Glycomimetic Langerin Ligands for Langerhans Cell Targeting

Inauguraldissertation
to obtain the academic degree
Dr. rerum naturalium (*Dr. rer. nat.)*

Submitted to the Department of Biology, Chemistry and Pharmacy
of Freie Universität Berlin

by

Eike-Christian Wamhoff
from Bassum, Germany

Berlin, June 2017

The work presented in this dissertation was conducted between April 2013 and May 2017 at the Max Planck Institute of Colloids and Interfaces under the supervision of Dr. Christoph Rademacher.

1st Reviewer: Dr. Christoph Rademacher
2nd Reviewer: Prof. Dr. Christian Freund

Date of Defense:

Acknowledgements

Foremost, I would like to express my sincere gratitude to Dr. Christoph Rademacher for his mentorship, his trust and his patience. His passion for science has been inspirational and his integrity exemplary. It was an honor to work on my dissertation under Dr. Christoph Rademacher's supervision and I am certain that he will be an outstanding mentor for future generations of Ph. D. students.

I would like to thank Prof. Peter H. Seeberger for providing an excellent interdisciplinary and international scientific environment at the Max Planck Institute of Colloids and Interfaces. Moreover, I am grateful for his career advice and support during application processes.

I deeply acknowledge the friendship, the professionalism and the lasting memories I have shared with my fellow Ph. D. students and co-workers in the Structural Glycobiology Group. Thank you, Dr. Jonas Aretz, Hannes Baukmann, Dr. Jonas Hanske, Jessica Schulze and Dr. Robert Wawrzinek! The last years have been an exciting time and would not have been the same without you.

My gratitude furthermore extends to the ever-motivated and talented students I had the opportunity to supervise. Felix Fuchsberger, David Hartmann, Lena Kilian, Julia Mastouri, Mareike Rentzsch, Lennart Schnirch and Christian Sommereisen have conducted important experiments and data analyses. Moreover, the help provided by the administrative and technical staff has been invaluable and I would like to specifically thank Dorothee Böhme, René Genz, Eva Settels, Felix Hentschel and Olaf Niemeyer.

Many collaborators have contributed substantially to the work presented in this dissertation. Dr. Oliver Schwardt and Prof. Beat Ernst have generously hosted me in their laboratory at the Universität Basel and their advice has been instrumental for the design and synthesis of several promising glycomimetics. Additionally, Prof. Beat Ernst and Dr. Alexander Titz have provided a focused glycomimetic library for screening against Langerin. Dr. Andrea Volkamer provided valuable insight into *in silico* druggability analysis. Gunnar Bachem and Prof. Oliver Seitz as well as Kira Neuhaus and Prof. Laura Hartmann have designed and synthesized multivalent glycomimetics. Here, I particularly acknowledge the productive discussions with Gunnar Bachem and his excellent synthetic work to enable the utilization of glycomimetics in cell-based *in vitro* and *ex vivo* experiments. These experiments were devised and conducted by Jessica Schulze, Lydia Bellmann, Dr. Martin Hermann and Assoc. Prof. Patrizia Stoitzner and lend this dissertation much of its relevance in the context of cancer immunotherapy. Finally, Maurice Grube and Dr. Daniel Varón Silva have patiently introduced me to the field of carbohydrate chemistry during my first year at the Max Planck Institute of Colloids and Interfaces.

Most importantly, I would like to thank my parents and my brother for their everlasting support and love. I greatly appreciate your continued encouragement and the sacrifices you endured to enable my education. Last but certainly not least, I would like express my deepest gratitude to Anna Behrens, who motivates me to be the best version of myself and gives me the opportunity to be happy. Thank you for your patience over the last months - I am excited to begin a new chapter in our lives.

Summary

Immune evasion represents an important hallmark of cancer progression. In this context, the induction of tumor-specific cytotoxic T cell immunity to overcome immunological tolerance has become a focal point of cancer immunotherapy. Langerhans cells (LCs) constitute a dendritic cell (DC) subset residing in the epidermis of the human skin. They have been recognized for their capacity to endocytose and cross-present exogenous antigens *via* MHC-I to efficiently prime naïve $CD8^+$ T cells. Langerin, an endocytic C-type lectin receptor (CLR) involved in the Ca^{2+}-dependent recognition of both pathogen- and self-associated glycans, displays an expression profile highly restricted to LCs. Hence, the targeted delivery of tumor-associated antigens (TAAs) to Langerin represents an intriguing approach to develop novel vaccination strategies.

Over the last decades, liposomes have emerged as versatile delivery platforms that can be targeted to LCs *via* the conjugation to Langerin ligands. As glycan interactions with CLRs are typically weak and highly promiscuous, liposomal targeting required the discovery of potent and specific glycomimetic ligands. However, the onerous synthesis of carbohydrate analogs as well as the hydrophilicity and high solvent exposure of carbohydrate binding sites render glycomimetic ligand design challenging. The structure-based *in silico* analysis of 21 X-ray structures presented in this dissertation corroborated the classification of CLRs as undruggable or challenging targets. Druggable secondary binding pockets adjacent to the carbohydrate binding site were exclusively identified for CLRs of limited therapeutic relevance.

Several strategies were employed to address the challenges outlined above and to discover potent carbohydrate analogs for Langerin. The structure-based *in silico* screening of substituents in C2 of the mannose (Man) scaffold served to design an initial focused glycomimetic library. The structure-activity relationship (SAR) of this scaffold was further elucidated *via* determination of affinities for an existing library of Man analogs derivatized in C1 and C6. Alternatively, a structure-based design strategy guided the exploration of substituents in C2 of glucsoamine-2-sulfate (GlcNS). These investigations ultimately led to the discovery of potent Man (K_I = 0.25±0.07 mM) and GlcNS (K_I = 0.24±0.03 mM) analogs displaying a 40- to 42-fold affinity increase over naturally occurring carbohydrate ligands.

The multivalent organization of carbohydrates or their synthetic analogs to match the geometry of the Langerin trimer represents an attractive strategy to optimize their specificity and potency. The GlcNS analog was conjugated to nucleic acid scaffolds to design divalent glycomimetics (IC_{50} = 23±2 µM) resulting in an additional distance-dependent 12-fold avidity increase. In an alternative approach, trivalent Man-bearing glycoclusters (IC_{50} = 0.20±0.08 mM) displayed an 80-fold avidity increase over naturally occurring carbohydrate ligands.

The development of a sensitive ^{19}F R_2-filtered nuclear magnetic resonance (NMR) assay enabled the determination of K_I and IC_{50} values for Langerin. Importantly, the optimization of the assay setup with respect to throughput and material consumption proved instrumental for the discovery of potent carbohydrate analogs and multivalent glycomimetics. Moreover, the implementation of an explorative ^{19}F R_2-filtered NMR fragment screening led to the identification of the first non-carbohydrate inhibitor reported for Langerin. The assay was successfully transferred to the CLR DC-SIGN to evaluate the specificity of designed carbohydrate analogs.

While ^{19}F R_2-filtered NMR experiments served as the primary screening and characterization assay, affinities were validated *via* saturation transfer difference and ^{15}N heteronuclear single quantum coherence NMR. Furthermore, an integrated strategy combining these NMR experiments with molecular docking studies was implemented to analyze the Ca^{2+}-dependent binding mode of the designed Man and GlcNS analogs. These investigations enabled the development of suitable conjugation strategies for liposomal formulations.

The GlcNS analog displayed remarkable specificity against DC-SIGN in ^{19}F R_2-filtered NMR experiments and was thus utilized for the preparation of targeted liposomes. Flow cytometry studies were employed to optimize liposomal formulations and to validate the binding of these liposomes to Langerin$^+$ model cells *in vitro*. Finally, *ex vivo* experiments demonstrated their capacity to specifically target LCs in the human skin. The liposomes were efficiently endocytosed and thus represent a promising TAA delivery platform.

In conclusion, the integration of carbohydrate chemistry, structure-based *in silico* methods and NMR experiments enabled the discovery of carbohydrate analogs and multivalent glycomimetics as potent and specific ligands of the endocytic CLR Langerin. These ligands were demonstrated to specifically target liposomes to LCs in the human skin and to promote endocytosis. Consequently, the findings presented in this dissertation constitute an important advancement for the research field of DC immunology and the development of novel cancer immunotherapies.

Publications

Parts of this dissertation have been published or will be included in journal articles:

Wamhoff, E.-C., Schulze, J., Bellmann, L., Bachem, G., Seitz, O., Stoitzner, P., and Rademacher, C. (2017) A glycomimetic Langerin ligand for Langerhans cell targeting. (in preparation).

Hanske, J., Wawrzinek, R., Geissner, A., **Wamhoff, E.-C.**, Sellrie, K., Schmidt, H., Seeberger, P. H., and Rademacher, C. (2017) Calcium-independent activation of an allosteric network in Langerin by heparin oligosaccharides. *ChemBioChem* (accepted).

Wamhoff, E.-C., Hanske, J., Schnirch, L., Aretz, J., Grube, M., Varón Silva, D., and Rademacher, C. (2016) ^{19}F NMR-guided design of glycomimetic Langerin ligands. *ACS Chem. Biol.* 11 (9), 2407-13.

Aretz, J., **Wamhoff, E.-C.**, Hanske, J., Heymann, D., and Rademacher, C. (2014) Computational and experimental prediction of human C-type lectin receptor druggability. *Front. Immunol.* 5 (323).

Other publications:

Glas, A.*, **Wamhoff, E.-C.***, Krüger, D., Rademacher, C., and Grossmann, T. N. (2017) Increased conformational flexibility of a macrocycle–receptor complex contributes to reduced dissociation rates. *Chem. Commun.* (submitted).

*These authors contributed equally.

Spiliotopoulos, D., **Wamhoff, E.-C.**, Lolli, G., Rademacher, C., and Caflisch, A. (2017) Discovery of BAZ2A bromodomain ligands. *J. Med. Chem.* (under revision).

Spiliotopoulos, D., Zhu, J., **Wamhoff, E.-C.**, Deerain, N., Marchand, J.-C., Aretz, J., Rademacher, C., and Caflisch, A. (2017) Virtual screen to NMR (VS2NMR): discovery of fragment hits for the CBP bromodomain. *Bioorg. Med. Chem. Lett.* (accepted).

Möginger, U., Resemann, A., Martin, C. E., Parameswarappa, S., Govindan, S., **Wamhoff, E.-C.**, Broecker, F., Suckau, D., Pereira, C. L., Anish, C., Seeberger, P. H., and Kolarich, D. (2016) Cross Reactive Material 197 glycoconjugate vaccines contain privileged conjugation sites. *Sci. Rep.* 6 (20488).

Aretz, J., Wratil, P. R., **Wamhoff, E.-C.**, Nguyen, H. G., Reutter, W., and Rademacher, C. (2016) Fragment screening of N-Acetylmannosamine Kinase reveals non-carbohydrate inhibitors. *Can. J. Chem.* 94 (11).

List of Abbreviations

Ac	Acetyl
APC	Antigen presenting cell
ATR-FTIR	Attenuated total reflection Fourier transform infrared spectroscopy
Boc	*Tert*-butyloxycarbonyl
Cbz	Carboxybenzyl
CLR	C-type lectin receptor
CORCEMA-ST	Complete relaxation and conformational exchange matrix saturation transfer
CPMG	Carr-Purcell-Meiboom-Gill
CRD	Carbohydrate recognition domain
CSA	Chemical shift anisotropy
CSP	Chemical shift perturbation
CTLD	C-type lectin-like domain
DNA	Deoxyribonucleic acid
DC	Dendritic cell
EBV	Epstein-Barr virus
ECD	Extracellular domain
ELLA	Enzyme-linked lectin assay
ESI MS	Electron spray ionization mass spectrometry
FBLD	Fragment-based lead discovery
FcR	Fc receptor
Fuc	*L*-fucose
GAG	Glycosaminoglycan
Gal	Galactose
Gal-6-OS	Galactose-6-sulfate
GE	Group efficiency
Glc	Glucose
GlcN	Glucosamine
GlcNAc	*N*-acetylglucosamine
GlcNS-6-OS	Glucosamine-2-sulfate-6-sulfate
GPI	Glycan-protein interaction
GBP	Glycan-binding protein
GBSA	Generalized Born surface accessibility
HADDOCK	High ambiguity driven protein-protein docking
HIV	Human Immunodeficiency Virus

HR	High resolution
HSQC	Heteronuclear single quantum coherence
HTS	High-throughput screening
ILR	I-type lectin receptor
ITAM	Immunoreceptor tyrosine-based activation motif
ITC	Isothermal titration calorimetry
ITIM	Immunoreceptor tyrosine-based inhibition motif
LE	Ligand efficiency
Le^X	$Lewis^X$
Man	Mannose
Man-6-OP	Mannose-6-phosphate
Man-6-OS	Mannose-6-sulfate
ManNAc	*N*-acetylmannosamine
MHC-I/II	Major histocompatibility complex-I/II
MC	Monte Carlo
MD	Molecular dynamics
MM	Molecular mechanics
MR	Mannose Receptor
NeuAc	*N*-acetylneuraminic acid
NK cells	Natural killer cells
NMR	Nuclear magnetic resonance
NOE	Nuclear Overhauser effect
PEG	Polyethylene glycol
Phth	Phthalimide
PNA	Peptide nucleic acid
PRR	Pattern recognition receptor
PSA	Polarizable surface area
Rha	*L*-rhamnose
Sial	Sialoside
sLe^A	Sialyl $Lewis^A$
sLe^X	Sialyl $Lewis^X$
SNR	Signal-to-noise ratio
SPR	Surface plasmon resonance
SR	Scavenger receptor
TAA	Tumor-associated antigen
TLC	Thin layer chromatography
Troc	Trichloroehtyl

Trt	Trityl
Ts	Tosyl
TSA	Thermal shift assay
TCR	T cell receptor
TLR	Toll-like receptor
UrA	Uronic acid
trNOE	Transferred nuclear Overhauser effect

Table of Contents

1. Introduction

1.1. Cancer Immunotherapy and Dendritic Cells

Introduction to Cancer Immunotherapy

Cancer immunotherapy attempts to harness the immune system to treat tumors and has led to prolonged survival and remission in many patients. Tumor-specific immune responses can be enhanced *via* the administration of antibodies or cytokines. Prominent examples include Ipilimumab and Nivolumab, two antibodies approved for the treatment of melanoma (Sharma and Allison, 2015). They target the check point inhibitors CTLA-4 and PD-1, respectively, thereby promoting the activation of cytotoxic $CD8^+$ T cells. Utilizing an alternative mechanism, Rituximab engages the B cell maker CD20 to induce complement-mediated as well as cell-mediated cytotoxicity and is *e.g.* administered to patients with non-Hodgkin lymphoma (Plosker and Figgitt, 2003). The cytokine interferon-α has proven efficacious for several cancers including melanoma and chromic myeloid leukemia (Talpaz *et al.*, 2013).

The discovery of tumor-associated antigens (TAAs) has enabled the development of antigen-specific immunotherapies (Palucka and Banchereau, 2012). The immune system has the capacity to recognize TAAs and eliminate tumor cells during the initial stages of cancer progression, a process that is assumed to occur continuously in healthy humans. The activation of antigen-specific cytotoxic T cells has been identified as a primary mechanism of tumor elimination (Albert *et al.*, 1998). Cytotoxicity is initiated by the recognition of TAAs presented on the major histocompatibility complex-I (MHC-I) of tumor cells *via* the T cell receptor (TCR). Subsequently, cytotoxic T cells eliminate tumor cells, predominantly *via* the induction of apoptosis (Martinez-Lostao *et al.*, 2015). This process is typically mediated by the secretion of perforins and granzymes or the engagement of Fas by FasL.

However, the presentation of TAAs by tumor cells is limited and as tumorigenesis lacks an acute inflammatory phase, the induction of T cell responses tends to be inefficient (Zou, 2005). On the contrary, chronic inflammation impairs the T cell memory in cancer patients leading to the development of anergic cytotoxic T cells (Freeman *et al.*, 2006). Additionally, active immune evasion has been recognized as a hallmark of cancer progression (Hanahan and Weinberg, 2011). For instance, the tumor microenvironment conveys tolerance *via* the secretion of TGF-β to deactivate infiltrating cytotoxic T cells or the recruitment of regulatory $FOXP3^+$ T cells (Ostrand-Rosenberg and Sinha, 2009; Yang *et al.*, 2010). Moreover, some tumors express PDL-1 and its engagement of the checkpoint inhibitor PD-1 has been demonstrated to suppresses T cell immunity (Dong *et al.*, 2002). Accordingly, the efficient induction of cytotoxic T cells responses against TAAs has become a focal

point of cancer immunotherapy (Palucka and Banchereau, 2013). TAAs have been targeted using autologous cytotoxic T cells that are expanded and activated *ex vivo* and re-administered into patients (Restifo *et al.*, 2012). Antigen-specific cells can be isolated from tumors or engineered genetically to improve efficacy. Despite recent advances, adoptive T cell-based therapies have not been approved for treatment.

Dendritic Cells in Cancer Immunotherapy

Alternatively, active immunization represents a promising strategy for cancer immunotherapy. First described by Steinmann *et al.*, dendritic cells (DCs) are a population of antigen presenting cells (APCs) coordinating the interplay between innate and adaptive immunity (Banchereau and Steinman, 1998). Several DC subsets have been described in humans and can be classified according to their tissue distribution. Plasmacytoid, $BDCA1^{+}$-$CD1c^{+}$ and $BDCA^{3+}$-$CD141^{+}$ DCs are present in the blood or lymphoid tissue, while $CD1a^{intermediate}$-$CD1c^{+}$ and $CD141^{high}$-$XCR1^{+}$ dermal DCs as well as $Langerin^{+}$-$CD1a^{high}$ Langerhans cells (LCs) can be found in the skin (Clausen and Stoitzner, 2015). Additionally, the *in situ* differentiation of monocytes to monocyte-derived DCs has been observed under inflammatory conditions (Leon and Ardavin, 2008). Some of these subsets have been recognized for their capacity to cross-present exogeneous TAAs *via* MHC-I and efficiently prime naïve $CD8^{+}$ T cells (Regnault *et al.*, 1999; Rodriguez *et al.*, 1999). This has rendered DCs attractive targets for cancer immunotherapy (Palucka and Banchereau, 2013).

While the capacity of individual DC subsets to elicit tumor-specific T cell immunity is still under debate, recent findings suggest that the efficiency of this process is additionally determined by external parameters (Fehres *et al.*, 2014b). First, DCs have been described to exist in two distinct states (Palucka and Banchereau, 2012). Under steady state conditions, tissue-resident DCs are considered immature and induce antigen-specific tolerance *via* the deletion of cytotoxic T cells or the expansion of regulatory T cells (Idoyaga *et al.*, 2013). When DCs encounter inflammatory signals, they mature and typically migrate to lymph nodes to prime naïve $CD8^{+}$ T cells. For therapeutic purposes, adjuvants such as Toll-like receptor (TLR) agonists are required to elicit the desired T cell responses and often display differential effects for individual DC subsets (Fehres *et al.*, 2014a; Jongbloed *et al.*, 2010). Additionally, TAAs may either be internalized by phagocytosis or by receptor-mediated endocytosis and the endocytic route has been demonstrated to affect the efficiency of MHC-I-dependent presentation by DCs (Burgdorf *et al.*, 2007; Chatterjee *et al.*, 2012). Lastly, the differentiation and expansion of cytotoxic T cells depends on the canonical presentation of TAAs *via* MHC-II and the induction of a favorable helper T cell response (Gerner *et al.*, 2008). Notably, the immunological properties of several DC subsets differ from their murine orthologs which has compromised the transferability of observations from animal studies and represents a major challenge for the field of DC immunology (Clausen and Stoitzner, 2015).

The cell biology of cross-presentation by DCs has been studied extensively over the last decades (Rock *et al.*, 2016). Two distinct pathways have been proposed to enable the cross-presentation of exogenous TAAs *via* MHC-I. On the one hand, endocytosed antigens may be transferred from phago- or endosomes to the cytoplasm and thereby enter the canonical proteasome- and TAP-dependent pathway of MHC-I presentation (Kovacsovics-Bankowski and Rock, 1995). On the other hand, MHC-I has been observed to migrate to the endosomal compartment, where it encounters antigens processed by endosomal- or lysosomal proteases (Burgdorf *et al.*, 2008; Gromme *et al.*, 1999). More recently, a hybrid compartment formed by the fusion of the endoplasmic reticulum and phagosomes has been reported (Guermonprez *et al.*, 2003). While the mechanistic details have been elucidated for neither pathway, one central tendency has emerged for DCs: The receptor-dependent endocytosis and subsequent routing of TAAs to early or late endosomal compartments displaying limited antigen processing at relatively high pH values promotes cross-presentation compared to other APCs (Accapezzato *et al.*, 2005; Belizaire and Unanue, 2009; Chatterjee *et al.*, 2012; Delamarre *et al.*, 2005; van Montfoort *et al.*, 2009). Upon DC maturation, the MHC-I-antigen complex migrates to the plasma membrane and can be recognized by the TCR of $CD8^+$ T cells. Additionally, co-stimulatory signals from both DCs and helper T cells are required for efficient cross-priming (Palucka and Banchereau, 2013). DCs secrete cytokines such as IL-15 and express CD80 as well as other co-stimulatory receptors while the secretion of IL-21 and interferon-γ by helper T cells promotes the activation of both DCs and cytotoxic T cells.

Adoptive Dendritic Cell Therapy and Vaccination Strategies

The therapeutic potential of DCs for the active immunization against tumors has been successfully realized. Sipuleucel-T, an adoptive DC-based therapy, has been approved for the treatment of prostate cancer (Kantoff *et al.*, 2010). The treatment involves the isolation of autologous DC precursors from the blood followed by their *ex vivo* differentiation, expansion and maturation. Subsequently, the DCs are activated with a fusion protein consisting of the TAA PAP and the cytokine GM-CSF which promotes DC-dependent priming of naïve $CD8^+$ T cells (Lutz, 2012). Upon re-administration, clinical trials demonstrated prolonged survival and a reduced risk of death. While Sipuleucel-T represents the proof-of-principle for DC-based cancer immunotherapy, adoptive DC-based therapies are expensive and laborious. Moreover, only a fraction of re-administered DCs migrate to the lymph nodes and only selected DC subsets potentially displaying non-optimal immunogenicity can be isolated in sufficient amounts (Tacken *et al.*, 2007).

Encouraged by these findings, various strategies have been explored to address the limitations of adoptive DC-based therapies. These strategies primarily focus on the efficient induction of T cell immunity without the requirement to isolate and re-administer autologous DCs for individual patients. Recently, the design of synthetic DCs has been reported (Mandal *et al.*, 2013; Mandal *et al.*, 2015). Preliminary investigations demonstrated that the multivalent organization of anti-CD3 and anti-CD38

antibodies on synthetic scaffold structures increases the activation of T cells *in vitro*. By comparison, vaccination strategies targeting DCs *in vivo* are considerably more advanced and are generally recognized as an intriguing alternative to *ex vivo* therapies (Palucka and Banchereau, 2013). Ideally, these strategies will both restore anergic $CD8^+$ memory T cells and prime naïve $CD8^+$ cells. Over the last decades, three critical determinants for the clinical efficacy of DC-based vaccines have emerged. First, the potency of the generated cytotoxic T cells is essential. Important parameters include the avidity of TCRs for MHC-I, the amount of secreted granzymes and perforins and the migratory capacity (Appay *et al.*, 2008). Secondly, the quality of the induced helper T cell response is important for the selection of potent cytotoxic T cells and should actively promote tumor elimination by cytotoxic T cells (Schietinger *et al.*, 2010). Additionally, the generation of $CD8^+$ memory T cells is required to prevent relapse (Palucka and Banchereau, 2013). Third, the challenges posed by the tolerogenic tumor microenvironment must be addressed to ensure sustained T cell immunity. Here, combination therapy represents a promising approach and antibodies such as Ipilimumab and Nivolumab might be utilized to target the checkpoint inhibitors CTLA-4 and PD-1 (Sharma and Allison, 2015). Recently, a population of suppressed tumor-resident DCs has been identified for several cancers and its re-activation represents an intriguing approach to enhance T cell immunity (Gardner and Ruffell, 2016).

Initial studies have extensively explored the utilization of non-targeted TAAs and the corresponding vaccination strategies displayed promising clinical efficacy (Palucka and Banchereau, 2013). Importantly, these studies have substantially advanced the design of optimized antigens (Quakkelaar and Melief, 2012). However, non-targeted TAAs display critical disadvantages with respect to the considerations described above. They require the administration of high antigen doses and provide limited control over the targeted DC subset or the endocytic route. This lack of control translates into challenges for the optimization of vaccines *e.g. via* the co-administration of adjuvants. Consequently, the targeted delivery of TAAs to DCs has become an attractive alternative strategy. Intriguingly, DCs express several families of endocytic receptors that potentially promote antigen cross-presentation of TAAs including Fc receptors (FcR), scavenger receptors (SR) and C-type lectin receptors (CLRs) (Fehres *et al.*, 2014b). CLRs display restricted expression profiles that might be leveraged to target individual DC subsets. Additionally, CLR-dependent endocytosis is efficient and potentially routs antigens to early or late endosomal compartments. These characteristic render CLRs excellent target receptors for the controlled delivery of TAAs to DCs and the development of novel vaccination strategies.

1.2. Targeted Antigen Delivery *via* C-Type Lectin Receptors

C-Type Lectin Receptors in Innate and Adaptive Immunity

CLRs constitute a large family of glycan-binding proteins (GBPs) classified into various groups based on domain architecture and phylogenetic analyses (Groups I to XVII) (Varki, 2009). They are found in all organisms and share a conserved domain, termed C-type lectin-like domain (CTLD), which displays a characteristic double-loop fold stabilized by two disulfide bonds. The domain typically, but not exclusively binds glycans and is referred to as the carbohydrate recognition domain (CRD) for the corresponding CLRs (Drickamer and Taylor, 2015). Overall, the domain harbors up to four Ca^{2+} ions of which only the canonical binding site 2 (Ca^{2+}-2) is associated with glycan recognition (Zelensky and Gready, 2005). Depending on their domain architecture, CLRs can be secreted or membrane-associated. Many CLRs consist of additional domains that mediate signaling and the formation of homo- as well as heterooligomers.

Human CLRs are involved in various aspects of immune regulation (Varki, 2009). They act as adhesion receptors, recognize pathogens and contribute to self versus non-self discrimination. Some CLRs are effectors of innate immune responses while others are endocytic receptors promoting antigen processing and presentation to T cells. Endocytic CLRs are *i.a.* expressed on APCs like macrophages and DCs where they act as pattern recognition receptors (PRRs). Notably, the expression profiles of some receptors including Langerin, Clec9a and BDCA-2 are restricted to individual DCs, which renders them attractive targets for the delivery of TAAs (Clausen and Stoitzner, 2015; Ebner *et al.*, 2004; Huysamen *et al.*, 2008). Endocytic CLRs exert their physiological function primarily by two complementary mechanisms.

On the one hand, the recognition of pathogens and self-antigens might induce signaling *via* the cytoplasmic domain and thereby activate APCs and modulate T cell responses (Geijtenbeek and Gringhuis, 2009). While this process remains poorly described for many CLRs, several examples highlight the involvement of tyrosine kinase Syk-dependent signaling pathways. Furthermore, CARD9 has been recognized as an important downstream adaptor protein for these pathways. Dectin-1 recruits Syk directly *via* an immunoreceptor tyrosine-based activation-like motif (ITAM) while the recruitment for Mincle is mediated by FcRγ (Gringhuis *et al.*, 2009b; Yamasaki *et al.*, 2008). For both CLRs, Syk-dependent signaling results in APC activation and the secretion of pro-inflammatory cytokines such as IL-6 and TNF-α. By contrast, the immunoreceptor tyrosine-based inhibition motif (ITIM) of DCIR is associated with anti-inflammatory signaling (Meyer-Wentrup *et al.*, 2009). Alternative pathways have been identified for other CLRs. DC-SIGN-mediated signaling is transmitted *via* serine/threonine protease Raf1 and was observed to modulate signals from other PRR such as TLR-4 (Gringhuis *et al.*, 2009a). Strikingly, Geijtenbeek *et al.* demonstrated that mannose (Man)-type glycans present on the

surface of the Human Immunodeficiency Virus (HIV) elicit signaling pathways distinct from those induced by the recognition of *L*-fucose (Fuc)-type glycans present on *Helicobacter pylori* (Gringhuis *et al.*, 2014). These findings indicate that structural aspects of glycan recognition potentially modulate CLR-mediated signaling.

On the other hand, recognized pathogens and self-antigens are internalized *via* CLR-dependent endocytosis. Most CLRs including DC-SIGN, MGL and SRCL utilize clathrin-mediated endocytic routes (Coombs *et al.*, 2005; Engering *et al.*, 2002; van Vliet *et al.*, 2007). Notably, Langerin has been proposed to additionally induce caveolin-mediated pathways (Ng *et al.*, 2016; van den Berg *et al.*, 2014). The subsequent lysosomal degradation of pathogens potentially contributes to innate immunity as exemplified by Langerin-dependent neutralization of HIV (de Witte *et al.*, 2007). Additionally, CLRs are essential for the induction of adaptive immune responses as they mediate antigen processing and presentation to both $CD4^+$ and $CD8^+$ T cells (Fehres *et al.*, 2014b). First described for the Mannose Receptor (MR), some CLRs recycle constitutively between the plasma membrane and different endosomal compartments (Stahl *et al.*, 1980). Receptor recycling enables the efficient endocytosis of antigens with kinetics not limited by biosynthesis (Engering *et al.*, 1997). Importantly, this also requires the release of antigens in the endosomal compartment upon decreased pH values and Ca^{2+} concentrations (Mullin *et al.*, 1994; Wragg and Drickamer, 1999). As highlighted in the previous section, these characteristics simultaneously promote cross-presentation and the priming of $CD8^+$ T cells (Accapezzato *et al.*, 2005; Belizaire and Unanue, 2009; Delamarre *et al.*, 2005; van Montfoort *et al.*, 2009). Notably, CLR-mediated signaling has been reported to further enhance endocytosis and cross-presentation by DCs (Geijtenbeek and Gringhuis, 2009).

Glycan Recognition by C-Type Lectins

Glycan interactions are prevalent for endocytic CLRs and mediate the recognition of pathogens and self-antigens (Drickamer and Taylor, 2015). The interactions are typically dependent on the coordination of Ca^{2+}-2 by vicinal hydroxyl groups of terminal monosaccharides (Weis *et al.*, 1992). Secondary hydrogen bonds formed with hydrophilic receptor residues convey specificity. Here, two highly conserved structural motives have been identified (Drickamer, 1992). The EPN motif is found in CLRs binding to Man- or Fuc-type glycans while the presence of the QPD motif is indicative of galactose (Gal)-type glycan binding. However, exceptions to this canonical model have been reported, exemplified by the recognition of Gal and galatose-6-sulfate (Gal-6-OS) by BDCA-2 and Langerin, respectively (Feinberg *et al.*, 2011; Jegouzo *et al.*, 2015). Furthermore, CH-π interactions with aromatic residues have recently been proposed to play a pivotal role in glycan recognition (Hudson *et al.*, 2015). Importantly, the characteristics of glycan recognition by CLRs contribute to the regulation of antigen release in the endosomal compartment and thereby promote cross-presentation by DCs. Decreased Ca^{2+} concentrations observed upon endosomal maturation result in the dissociation of glycans and Ca^{2+} ions from the receptor. The corresponding CLR apo state might be stabilized

kinetically *via* the *cis-trans* isomerization of the proline residue associated with the EPN and QPD motives as observed for MBP-C (Ng *et al.*, 1998). For some CLRs, the stabilization of the double-loop fold by the presence of Ca^{2+} ions in the adjacent binding sites 1 and 3 (Ca^{2+}-1 and -3) is required for glycan recognition (Furukawa *et al.*, 2013; Onizuka *et al.*, 2012). Consequently, allosteric cooperativity effects have been proposed to contribute to antigen release at decreased Ca^{2+} concentrations. The discussed regulatory mechanisms are enhanced by decreased pH values encountered along the endocytic pathway. For instance, pH-dependent Ca^{2+} affinities have recently been demonstrated for ASGPR and Langerin and might also be relevant for other CLRs including DC-SIGN or DC-SIGNR (Feinberg *et al.*, 2001; Hanske *et al.*, 2016; Onizuka *et al.*, 2012; Stambach and Taylor, 2003).

The carbohydrate binding sites of CLRs are highly solvent exposed and hydrophilic (Drickamer and Taylor, 2015). Consequently, interactions with mono- and oligosaccharides are typically characterized by low affinities in the millimolar range (Weis and Drickamer, 1996). Furthermore, the recognition process is highly promiscuous as individual CLRs bind several mono- or oligosaccharides and *vice versa*. Notably, secondary interactions and favorable entropic profiles potentially result in specificity and increased affinities for the interactions between oligosaccharides and the CRDs of individual CLRs. The tetrasaccharide sialyl Lewis^X (sLe^X) displays low conformational flexibility in the free state and readily adopts the bioactive conformation recognized by E-Selectin (Binder *et al.*, 2012). The recognition process involves the displacement of several structural H_2O molecules by secondary hydrogen bonds which results in favorable entropic contributions and submillimolar affinities. This effect is potentially enhanced by the preorganization of the carbohydrate binding site as demonstrated for the interaction between the Le^X trisaccharide and SRCL (Taylor and Drickamer, 2009). Similarly, the formation of secondary interactions was observed to promote the recognition of larger, more flexible Man-type oligosaccharides by DC-SIGN and DC-SIGNR (Feinberg *et al.*, 2005).

First described comprehensively by Whitesides *et al.*, multivalency represents an important concept to rationalize the physiological and pathophysiological functions of CLRs (Mammen *et al.*, 1998). The multivalent organization of CLRs and glycans is thought to provide the specificity and increased avidities required to control these functions. This is achieved on two hierarchical levels. On the receptor level, CRDs are often organized multivalently. For instance, the monomeric MR comprises two or more CRDs and the resulting avidity effects have been proposed to confer specificity and enable efficient endocytosis (Taylor *et al.*, 1992). Other CLRs form hetero- or homooligomers enabling the distinct spatial organizations of CRDs. Prominently, the ASGPR heterotetramer has been demonstrated to recognize Gal-type glycans displaying matching geometries with considerably increased avidities (Bider *et al.*, 1996; Lee *et al.*, 1983). Moreover, the formation of heterodimers between Mincle and MCL facilitates the recognition of the *M. tuberculosis* virulence factor TDM and thereby modulates signaling processes (Richardson and Williams, 2014). Examples for

homooligomeric CLRs include DC-SIGN and Langerin (Mitchell *et al.*, 2001; Stambach and Taylor, 2003). On the cell level, endocytic CLRs are multivalently presented on the plasma membrane. This results in additional avidity effects and the distribution of CLRs as well as diffusion properties regulate glycan recognition. Prominently, the formation of nanoclusters reported for DC-SIGN is thought to provide specificity for viruses over bacteria and fungi (Manzo *et al.*, 2012). Another intriguing example is the multivalent organization of monomeric L-Selectins to enable leukocyte adhesion and extravasation under shear flow conditions in human blood vessels (Schwarz and Alon, 2004).

Because of their involvement in various aspects of immune regulation, CLRs are considered attractive target receptors for drug discovery (Ernst and Magnani, 2009). E-Selectin and DC-SIGN are prominent examples for the design of glycomimetic ligands to modulate CLR function. Under inflammatory conditions, E-Selectin is expressed on endothelial cells and mediates leukocytes adhesion followed by extravasation (Ley, 2003). In patients with Sickle-cell anemia, aberrant erythrocytes also adhere to the CLR and potentially cause the occlusion of blood vessels. Hence, E-Selectin inhibitors generally display anti-inflammatory properties and have entered clinical trials for the treatment of Sickle-cell anemia (Chang *et al.*, 2010). DC-SIGN inhibitors, in turn, are protective in *in vitro* HIV transmission models (Mangold *et al.*, 2012). DC-SIGN interacts with Man-type glycans present on the HIV surface and promotes the internalization of the virus by DCs. Interestingly, the virus has evolved to escape DC-SIGN-dependent lysosomal degradation and thereby infects DCs (Geijtenbeek *et al.*, 2000). The transmission of HIV from DCs to T cells *in vivo* is recognized as an essential process for the systemic progression of HIV infections. Other examples for CLRs in drug discovery include the utilization of glycomimetic ligands to induce CLR-mediated signaling (Lang *et al.*, 2011). Here, the design of novel adjuvants targeting Mincle has recently been reported (Decout *et al.*, 2017).

C-Type Lectin Receptors as Target Receptors for Antigen Delivery

Importantly, endocytic CLRs also represent promising target receptors for cancer immunotherapy. They display restricted expression profiles that enable the delivery of TAAs to individual DC subsets. Furthermore, recycling CLRs are potent mediators of endocytosis and have been reported to efficiently release glycan-associated antigens in the endosomal compartment in presence of decreased Ca^{2+} concentrations and pH values. This process is highly regulated by the structural properties of CLR-glycan interactions and serves to circumvent lysosomal degradation of antigens. Consequently, antigen delivery to CLRs potentially promotes cross-presentation by DCs and the priming of naïve $CD8^+$ T cells. CLR-dependent signaling simultaneously activates DCs and might be leveraged to modulate T cell responses for optimal therapeutic efficacy. Also, targeting individual CLRs on defined DC subsets contributes to basic research and will eventually facilitate the design of improved cancer immunotherapies.

Pioneering work by Steinman *et al.* established the utility of anti-CLR antibody-antigen conjugates to elicit cytotoxic T cell immunity (Bonifaz *et al.*, 2002). Antigen targeting to DEC-205 resulted in substantially increased antigen delivery to and cross-presentation by murine $CD11c^{+}$-$CD8\alpha^{+}$ DCs compared to the administration of non-targeted antigens. However, the development of peripheral tolerance was observed in absence of co-stimulatory signals and the co-administration of the pro-inflammatory anti-CD40 antibody was required to induce an immunogenic cytotoxic T cell response. Over the following decade, additional adjuvants such as agonists of TLR-3, -7 and -8 were found to promote DC maturation and the activation of cytotoxic T cells in this experimental system (Steinman, 2012). The vaccination of non-human primates with an anti-DEC-205 fused to the HIV protein Gag in a prime-boost approach further demonstrated robust T cell immunity and this strategy is currently explored in clinical trials (Flynn *et al.*, 2011; Tacken and Figdor, 2011). Importantly, targeting the TAA NYESO-1 to DEC-205 also results in tumor regression for patients with various cancers including melanoma (Dhodapkar *et al.*, 2014). The clinical trial involved the co-administration of a TLR-3 agonist and future efforts are aimed at combination therapy by targeting checkpoint inhibitors such as CTLA-4 and PD-1.

Another notable example for targeted antigen delivery to CLRs is MR. In analogy to DEC-205-based therapies, the TAA hCG-β was fused to an antibody and clinical efficacy was demonstrated in patients with *e.g.* breast and pancreatic cancer (Morse *et al.*, 2011). The co-administration of GM-CSF as well as TLR-3, -7 and -8 agonists was required to induce robust cytotoxic T cell responses. In an alternative approach, oxidized mannan was conjugated to the TAA MUC1 (Karanikas *et al.*, 1997). While most clinical trials are conducted at late stages of cancer, this vaccine was administered to early stage breast cancer patients and resulted in reduced relapse rates (Apostolopoulos *et al.*, 2006). Notably, no adjuvants or cytokines were utilized and the development of both humoral and T cell immunity was observed.

The clinical trials conducted for DEC-205 and MR provide the proof-of-principle for the improved efficacy of targeted antigen delivery to CLRs. This approach is not only applicable to cancer immunotherapy, it also bears considerable potential for the treatment of infectious and autoimmune diseases. However, these initial vaccination strategies display several limitations with respect to the targeted DC subsets. DEC-205 is expressed on monocyte-derived, $BDCA1^{+}$-$CD1c^{+}$ and $BDCA^{3+}$-$CD141^{+}$ DCs in blood and lymphoid tissue as well as epidermal LCs (Tacken and Figdor, 2011). Additionally, the CLR can be found on other immune cells such as macrophages, B cells and natural killer (NK) cells. While the expression profile of MR is restricted to $CD1a^{intermediate}$-$CD1c^{+}$ dermal DCs, monocyte-derived DCs and macrophages, other CLRs are expected to bind Man-type glycans. Consequently, it is unclear which DC subsets contributed to the development of anti-tumor T cell immunity. As individual subsets display unique TLR patterns, the co-administration of a given set of

adjuvants might induce tolerance or inefficient DC activation (Idoyaga *et al.*, 2013). Moreover, the targeted CLR will influence DC signaling and the efficiency of cross-presentation.

Over the last decades, other CLRs such as BDCA-2, Clec9a and Langerin displaying more favorable expression profiles have been discovered (Tacken and Figdor, 2011). BDCA-2 is exclusively expressed on plasmacytoid DCs and was shown to mediate cross-presentation *in vitro* (Dzionek *et al.*, 2001). However, the antibody engagement of BDCA-2 also inhibits TLR-9 signaling and the secretion of interferon-α and -β by DCs (Jahn *et al.*, 2010). This signaling profile leads to the development of regulatory T cell-dependent tolerance and renders BDCA-2 an attractive target receptor for the treatment of autoimmune diseases such as lupus erythematosus rather than cancer immunotherapy (Chappell *et al.*, 2014). In contrast, Clec9a, present exclusively on $BDCA^{3+}$-$CD141^{+}$ DCs and has been successfully targeted *in vivo* by antibodies to elicit TAA-specific helper and cytotoxic T cell responses (Sancho *et al.*, 2008; Tullett *et al.*, 2016). As observed for DEC-205, the development of robust T cell immunity required anti-CD40 or TLR-3 agonists. Notably, Clec9a does not bind glycans but recognizes exposed Actin filaments of mechanically disrupted cells (Zhang *et al.*, 2012).

While BDCA-2 and Clec9a are present in the blood and lymphoid tissue, Langerin is predominantly expressed on LCs in the epidermis of the skin (Clausen and Stoitzner, 2015). This enables the topical or dermal administration of cancer vaccines potentially facilitating TAA delivery and improving patient care (Flacher *et al.*, 2009; Flacher *et al.*, 2010). Hence, Langerin represents an attractive targeted for TAA delivery to LCs and has the capacity to promote antigen-specific T cell immunity *in vivo* (Idoyaga *et al.*, 2011). However, essential parameters for the development of vaccination strategies remain to be investigated and the properties of Langerin and LCs will be addressed in detail in the following section.

1.3. Targeting Liposomes to Langerin

The C-type Lectin Receptor Langerin

Langerin is a homotrimeric type II transmembrane receptor that comprises a cytoplasmic domain, a transmembrane region, an α-helical oligomerization domain and a CRD that harbors a single Ca^{2+}-2 ion (Figure 1a) (Feinberg *et al.*, 2010; Feinberg *et al.*, 2011). The CRD displays a canonical CTLD fold with a double-looped, two-stranded antiparallel β-sheet formed by the N- and C-terminal residues which is connected *via* two α-helices and a three-stranded antiparallel β-sheet. The fold is stabilized by two disulfide bonds. The extended loop located in proximity to the Ca^{2+}-2 ion is segmented by additional secondary structures and the individual regions will be referred to as the long and short loop region throughout this dissertation (Figure 1a). Langerin contains an EPN motif that entails a preference for the recognition of Man-type and Fuc-type glycans (Drickamer, 1992). The cytoplasmic region contains a proline-rich sequence associated with SH3 domain-dependent recruitment of *e.g.* the tyrosine kinase Src (Li, 2005; Valladeau *et al.*, 2000). However, Langerin-dependent signaling has not been described.

Langerin exerts its physiological function as an endocytic PRR and is expressed on LCs and $Langerin^{intermediate}$-$CD11c^{+}$ dermal DCs, two DC subsets residing in the human skin (Bigley *et al.*, 2015; Clausen and Stoitzner, 2015). It has been implicated in the formation of Birbeck granules, a tubular organelle connected to the endosomal network unique to LCs (Valladeau *et al.*, 2000). Similar to MR, receptor recycling between these Birbeck granules, the endosomal compartment and the plasma membrane has been observed (Gidon *et al.*, 2012; McDermott *et al.*, 2002). Clathrin- as well as caveolin-dependent endocytic routes have been reported for the internalization of different pathogens *via* Langerin (Ng *et al.*, 2016; van den Berg *et al.*, 2014).

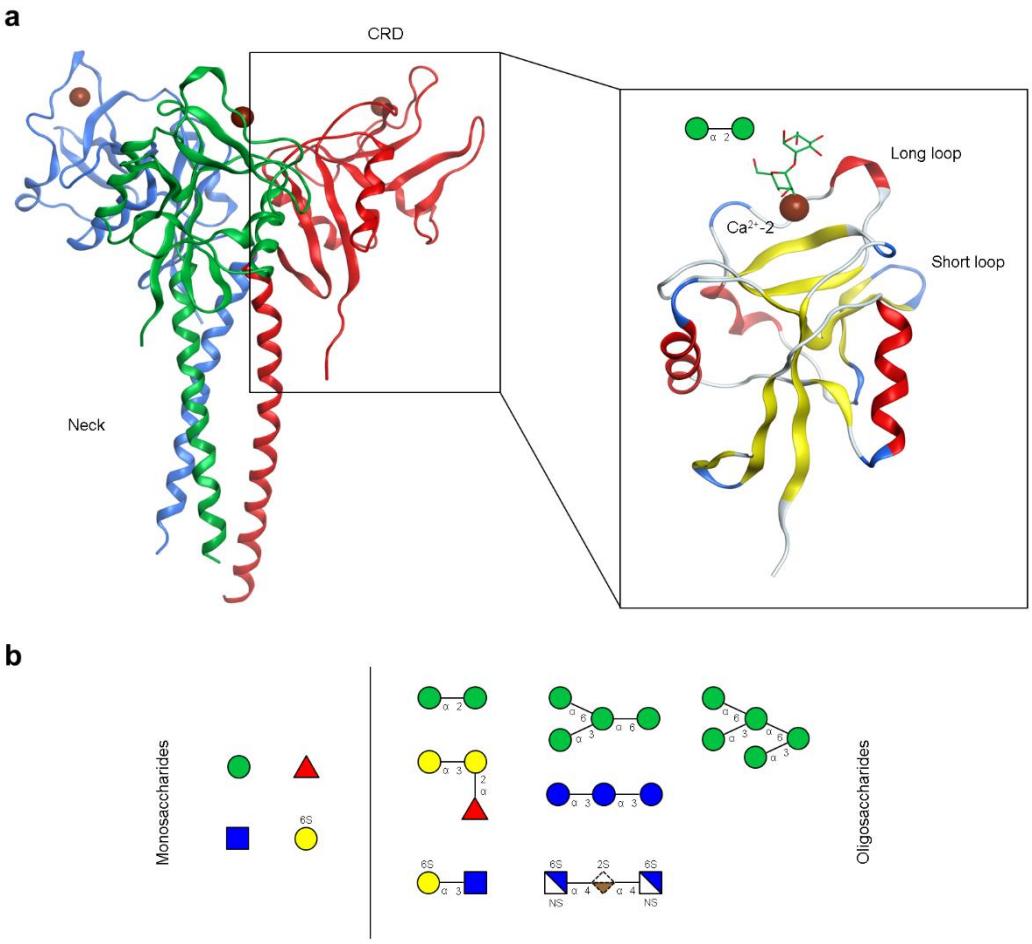


Figure 1. X-ray structures of Langerin and overview of naturally occurring mono- and oligosaccharide ligands.

a. The trimeric Langerin ECD comprises an α-helical oligomerization domain, herein termed neck region and a CRD (PDB code: 3KQG) (Feinberg *et al.*, 2010). The CRD displays a canonical CTLD fold and harbors a single Ca^{2+}-2 essential for glycan recognition (PDB code: 3P5F) (Feinberg *et al.*, 2011). The extended loop in proximity to the Ca^{2+}-2 ion is segmented by additional secondary structures and the individual segments are herein termed long and short loop. **b.** Glycan microarray studies have revealed the oligosaccharide specificity of Langerin. Recognized glycans contain either internal Man-α-(1-2)-Man units or terminal Man, Fuc, GlcNAc and Gal-6-OS, (left). Additionally, binding to several GAGs has been observed irrespective of these features. The binding mode of several oligosaccharides has been elucidated by X-ray crystallography or STD and trNOE NMR experiments (right). The ribbon representation for the CRD is colored according to secondary structure elements (red: α-helices; yellow: β-strands; blue: turn regions).

Langerin is involved in the recognition of both self- and pathogen-associated glycans. Recognized pathogens include viruses such as HIV, Measles Virus, Herpes Simplex Virus-2 and Influenza A Virus (de Jong *et al.*, 2010a; de Witte *et al.*, 2007; Ng *et al.*, 2016; van der Vlist *et al.*, 2011). Prominently, the Langerin-dependent internalization of HIV has been demonstrated to result in its degradation in Birbeck granules and thereby contributes to the innate immune response (van den Berg *et al.*, 2014). Additionally, Langerin binds to bacteria including *Yersinia pestis* and *Mycobacterium leprae* as well as fungi such as *Candida albicans* (de Jong *et al.*, 2010b; Hunger *et al.*, 2004; Kim *et al.*, 2015; Yang *et al.*, 2015). Self-associated glycans recognized by Langerin include blood group antigens as well as various glycosaminoglycans (GAGs) (Chabrol *et al.*, 2012; Feinberg *et al.*, 2011; Hanske *et al.*, 2017b; Munoz-Garcia *et al.*, 2015; van den Berg *et al.*, 2015; Zhao *et al.*, 2016). These interactions

might be involved in the physiological maintenance of tolerance or, in case of dermatan and heparan sulfate, contribute to the homeostasis and migration of LCs in the skin.

Glycan Recognition by Langerin

Over the last decade, glycan microarray studies served to elucidate the oligosaccharide specificity of Langerin (Figure 1b) (Feinberg *et al.*, 2013; Feinberg *et al.*, 2011; Hanske *et al.*, 2017a; Hanske *et al.*, 2017b; Holla and Skerra, 2011; Tateno *et al.*, 2010). Recognized glycans contain either internal Man-α-(1-2)-Man units or terminal Man, Fuc, *N*-acetylglucosamine (GlcNAc) and Gal-6-OS. Additionally, binding to several GAGs including heparin and heparan sulfate has been observed irrespective of these features. X-ray crystallography has revealed the binding mode for several of these oligosaccharides (Feinberg *et al.*, 2013; Feinberg *et al.*, 2011). Interestingly, glycan-Langerin interactions are predominantly confined to a single monosaccharide and dominated by the coordination of the Ca^{2+} ion by two vicinal, equatorial hydroxyl groups. These hydroxyl groups are part of an extended hydrogen bond network formed between the monosaccharide, the Ca^{2+} ion and receptor residues including E285, E293, N297 and N307. For Man and Fuc an additional hydrogen bond formed between the corresponding axial hydroxyl groups and K299 was observed. Notably, the obtained electron density map suggests the existence of an alternative orientation involving the 180° rotation of the Man scaffold that would not allow for the formation this interaction. Even for larger Man-type oligosaccharides few secondary interactions are formed and only a weak contact between the hydrophobic face of the adjacent Man and A289 was observed. By comparison, binding of the blood group antigen B trisaccharide results in the formation of several secondary interactions between Gal and the receptor surface in addition to the Ca^{2+} coordination by Fuc. GlcNAc displays a binding mode comparable to that of Man and Fuc with one prominent distinction: The equatorial acetamido group interacts with K299 *via* a structural H_2O molecule and a weak hydrophobic contact of the methyl group with P310.

Binding of Gal-6-OS to an EPN-type CLR is unexpected as Gal lacks vicinal equatorial hydroxyl groups. The recognition process is dominated by the formation of salt bridges between the sulfate group and K299 as well as K313. Additional hydrogen bonds are formed with N297 and N307. These interactions result in an altered orientation for the Gal scaffold that enables Ca^{2+} coordination by vicinal axial and equatorial hydroxyl groups. Interestingly, the common polymorphism K313I has been shown to result in a reduced affinity for Gal-6-OS and shifting the binding specificity to glycans with terminal GlcNAc (Feinberg *et al.*, 2013). Moreover, the mutation is statistically linked to the N288D polymorphism, destabilizing the Ca^{2+}-dependent carbohydrate binding site of Langerin and thereby potentially affecting the susceptibility to infectious diseases.

For GAGs, particularly heparin, Ca^{2+}-independent interactions have been observed (Chabrol *et al.*, 2012; Munoz-Garcia *et al.*, 2015; Zhao *et al.*, 2016). *In situ* conjugation studies have identified a

secondary binding pocket for heparin-derived hexasaccharides at the CRD interface of the Langerin trimer (Munoz-Garcia *et al.*, 2015). By contrast, the binding of heparin-derived trisaccharides is Ca^{2+}-dependent (Figure 1b). Saturation transfer difference (STD) and transferred nuclear Overhauser effect (trNOE) nuclear magnetic resonance (NMR) experiments served to determine the binding epitope and the conformation of these trisaccharides. Based on these experiments, molecular docking studies and molecular dynamics (MD) simulations were implemented to elucidate the binding mode (Munoz-Garcia *et al.*, 2015). In this model, the Ca^{2+} ion is coordinated by the vicinal, equatorial hydroxyl groups of glucosamine-2-sulfate-6-sulfate (GlcNS-6-OS). The adjacent iduronic acid-2-sulfate and the second GlcNS-6-OS are oriented towards K299 and K313. This orientation was proposed to result in the formation of an extended network of hydrogen bonds and salt bridges to promote the recognition process.

The affinities of several monosaccharides and smaller oligosaccharides for Langerin have been quantified. Isothermal titration calorimetry (ITC) experiments served to determine the K_D value of Man (K_D = ca. 6 mM) and a Man-type disaccharide (K_D = ca. 4 mM) while ^{15}N heteronuclear single quantum coherence (HSQC) NMR experiments revealed the affinity of a Man-type pentasaccharide (K_D = ca. 7 mM) (Hanske, 2016; Holla and Skerra, 2011). Consistent with the absence of strong secondary interactions, the observed affinity increase was negligible. Furthermore, heterogenous multi-well plate-based assays served to compare the affinities of monosaccharides (Feinberg *et al.*, 2013; Stambach and Taylor, 2003). Here, Man ($K_{I,app}$ = ca. 4 mM), GlcNAc ($K_{I,app}$ = ca 3 mM) and Gal-6-OS ($K_{I,app}$ = ca 3 mM) displayed the highest $K_{I,app}$ values. By comparison, statistical effects have been proposed to result in the observed avidity increase for a complex Man-type undecasaccharide ($K_{I,app}$ = ca 0.2 mM) (Feinberg *et al.*, 2011). Lastly, a submillimolar affinity was also determined for a heparin-derived trisaccharide (K_D = ca. 0.5 mM) in STD NMR experiments (Munoz-Garcia *et al.*, 2015).

Recently, the characteristics of Ca^{2+} recognition by Langerin were investigated *via* NMR and ITC experiments as well as MD simulations (Hanske *et al.*, 2016). The Ca^{2+} ion is coordinated by E285, N287, E293, N307 and D308 in a pentagonal pyramidal geometry and forms two additional coordination bonds with recognized monosaccharides (Feinberg *et al.*, 2011). Langerin's affinity for Ca^{2+} (K_D = ca. 100 µM) is pH-dependent and decreases approximately 6-fold from pH 7 to pH 6 (Hanske *et al.*, 2016). Notably, these pH values are representative of the physiological conditions at the plasma membrane and in the early endosome, respectively. This pH-dependency promotes the release of glycan ligands at decreased Ca^{2+} concentration upon endosomal maturation and thus potentially contributes to the cross-presentation of antigens by LCs. Interestingly, the observed mechanism seems to be regulated by an allosteric network that affects the conformational dynamics of the short and the long loop region (Hanske *et al.*, 2016). The loops are connected by interactions between K257 and H294, the latter of which is essential for the pH-dependent regulation of Ca^{2+}

recognition. In contrast to MBP-C, the kinetic stabilization of the apo state *via cis-trans* isomerization was not observed (Ng *et al.*, 1998).

Targeted Antigen Delivery to Langerin$^+$ Dendritic Cells

The utility of Langerin for the development of cancer vaccines was first investigated by Steinman *et al.* (Idoyaga *et al.*, 2011; Idoyaga *et al.*, 2009). The antibody-mediated delivery of antigens to Langerin$^+$ DCs was observed to elicit both helper and cytotoxic T cell responses *in vivo*. As expected, the development of robust T cell immunity required the co-administration of adjuvants, *i.e.* anti-CD40 and a TLR-3 agonist. However, these studies were conducted in mice and the induced T cell response was dependent on CD8α$^+$ DCs residing in lymphoid tissue. This DC subset does not exist in humans and while murine LCs are capable of cross-presentation, they have been reported to promote skin homeostasis and the development of peripheral tolerance (Flacher *et al.*, 2014; Idoyaga *et al.*, 2013; Romani *et al.*, 2010). The role of murine LCs in immune regulation has been under debate as other studies suggested their involvement the development of anti-tumor immunity *in vivo* (Stoitzner *et al.*, 2008).

Human LCs, by contrast, have been shown to efficiently activate cytotoxic T cells and induce helper T cell immunity in presence anti-CD40, TLR-4 agonists or *Candida albicans* (Klechevsky *et al.*, 2008; Seneschal *et al.*, 2012). Furthermore, optimal activation of LCs was observed for TLR-3 agonists resulting in superior cross-presentation efficiency compared to dermal DCs and macrophages (Fehres *et al.*, 2015a). These findings are consistent with the conservation of cross-priming-associated expression modules between murine CD8α$^+$ DCs and human LCs and their absence in murine LCs (Artyomov *et al.*, 2015). LCs and Langerin^{intermediate}-CD11c$^+$ dermal DCs are resident in the human skin (Clausen and Stoitzner, 2015). This enables the topical or dermal administration of TAAs potentially facilitating antigen delivery and improving patient care (Flacher *et al.*, 2009; Flacher *et al.*, 2010). Importantly, the topical administration of a TLR-7 agonist was observed to activate skin DCs and to promote cross-priming (Fehres *et al.*, 2014a). Recently, different strategies have been explored to target antigens to LCs isolated from healthy humans (Fehres *et al.*, 2015c). While liposomal formulations failed to induce the activation of CD8$^+$ T cells, the conjugation of antigens to an anti-Langerin antibody or the blood group antigen Le^Y promoted cross priming *in vitro*. T cell responses were further enhanced in presence of a TLR-3 agonist. In contrast to the utilization of an anti-Dectin-1 antibody, targeting Langerin resulted in antigen routing to early endosomes facilitating cross-presentation by LCs (Fehres *et al.*, 2015b). In summary, these characteristics render human LCs and Langerin *via*ble targets for cancer immunotherapy. However, available experimental evidence has been exclusively derived from *ex vivo* or in *vitro* studies and the identification of a suitable animal model will be essential for future investigations.

Glycomimetic Langerin Ligands and Targeted Liposomes

It can be argued that the carbohydrate binding site of Langerin has evolved to promote the activation of cytotoxic T cells upon recognition of pathogen-associated glycans. Ca^{2+}-dependent interactions are abrogated at decreased Ca^{2+} concentration and pH values, routing antigens to the early or late endosomal compartment (Hanske *et al.*, 2016). Thereby lysosomal degradation is avoided and antigens are efficiently presented to both $CD4^+$ and $CD8^+$ T cells (Accapezzato *et al.*, 2005; Belizaire and Unanue, 2009; Chatterjee *et al.*, 2012; Delamarre *et al.*, 2005; van Montfoort *et al.*, 2009). As Langerin is a recycling CLR, endosomal antigen release additionally enhances the endocytic capacity of LCs (Gidon *et al.*, 2012; McDermott *et al.*, 2002). In this context, the utilization of glycans as targeting ligands provides critical advantages over antibody-antigen conjugates. For example, antibodies engaging the GBP Siglec-2 were demonstrated to recycle to the plasma membrane which decreases the effective antigen concentration in the endosomal compartment (O'Reilly *et al.*, 2011). Similarly, antigen delivery *via* an antibody engaging the α-helical oligomerization domain of DC-SIGN enhanced cross-presentation compared to an anti-CRD antibody (Tacken *et al.*, 2011). This observation might be explained by the pH-dependent dissociation of the DC-SIGN tetramer and the ensuing release of antigens in the endosomal compartment (Guo *et al.*, 2004; Tabarani *et al.*, 2009). Despite recent advances in antibody engineering including humanization and defined glycosylation profiles, the administration of antibodies involves a considerable risk of adverse immune responses in patients (Harding *et al.*, 2010). Moreover, antibodies are prone to denaturation and pharmacokinetically less stable than glycan-antigen conjugates (Elliott *et al.*, 2003).

However, glycans are typically recognized by several CLRs or other GBPs and do not provide the specificity required to target individual DC subsets *in vivo* (Weis and Drickamer, 1996). Additionally, glycan-Langerin interactions display low affinities that might not be sufficient to promote the endocytosis of antigens (Weis and Drickamer, 1996). The design of potent and specific glycomimetic ligands represents an attractive strategy to address these limitations. Glycomimetics, particularly carbohydrate analogs, will share a subset of the interactions formed by pathogen-associated glycans. Hence, they are likely to meet the structural requirements to promote T cell immunity, *i.e.* to efficiently induce Langerin-dependent signaling, endocytosis and antigen routing. Recently, the design of glycomimetic ligands has enabled the development of targeted delivery approaches for Siglec-1 and 2, two GBPs expressed on macrophages and B cells, respectively (Kawasaki *et al.*, 2013; Macauley *et al.*, 2013). The designed sialoside (Sial) analogs displayed nanomolar to micromolar IC_{50} values corresponding to affinity increases of up to 3000-fold over naturally occurring oligosaccharides (Collins *et al.*, 2006; Nycholat *et al.*, 2012). For Siglec-1, the glycomimetics were utilized to target antigens to macrophages to prime NK cells *in vivo* which are for example required for tumor elimination (Bouwer *et al.*, 2014; Kawasaki *et al.*, 2013). Siglec-2-mediated antigen delivery to B cells

in vivo, in turn, induced antigen-specific apoptosis and conveyed tolerance that might be leveraged for the treatment of autoimmune diseases (Macauley *et al.*, 2013).

Over the last decades, nanotechnology has attracted considerable attention in cancer immunotherapy (Peer *et al.*, 2007). Advances in materials science have enabled the preparation of novel classes of nanoparticles such as quantum dots or nanoworms. By comparison, the utilization of liposomes in therapeutic applications is well-established and has been approved *e.g.* for the delivery of Doxorubicin to treat Kaposi's sarcoma (Barenholz, 2012). PEGylation has been demonstrated to confers steric stability and reduces immunogenicity (Sercombe *et al.*, 2015). In both studies presented above, liposomes served to organize targeting ligands and similar approaches were successfully implemented to target DC-SIGN *in vitro* (Fehres *et al.*, 2015c; Fehres *et al.*, 2015d). Accordingly, the conjugation of mono- and oligosaccharide is readily applicable to target liposomes to Langerin.

Liposomal formulations represent considerably more versatile delivery platforms than glycan-, glycomimetic- or antibody-antigen conjugates. The multivalent organization of targeting ligands allows for the adjustment of the avidity and specificity of targeted liposomes. Consequently, DC subsets with overlapping CLR expression profiles might be selectively targeted based on the expression levels of individual CLRs. Moreover, the conjugation of glycomimetics to the extended polyethylene glycol (PEG) linker structures potentially improves the accessibility for recognition by Langerin. As glycolipids are able diffuse through the lipid bilayer, the clustering of glycomimetics might further increase the avidity of targeted liposomes.

Furthermore, the conjugation of antigens to *e.g.* antibodies might alter the conformation of the antigen and compromise immunogenicity (Kastenmuller *et al.*, 2014). By contrast, hydrophilic TAAs can be encapsulated in the lumen of liposomes without chemical modifications. The co-administration of adjuvants has been identified as an important parameter for LC-mediated cross-priming of T cells (Fehres *et al.*, 2015a). Here, liposomes allow for both the encapsulation of hydrophilic adjuvants as well as the insertion of lipophilic adjuvants in the lipid bilayer (Boks *et al.*, 2015; Kawasaki *et al.*, 2013). Additionally, co-stimulatory molecules can be conjugated to lipids *via* the PEG linker structures. In summary, adjuvants can be targeted to the desired DC subset reducing the required doses and minimizing adverse effects.

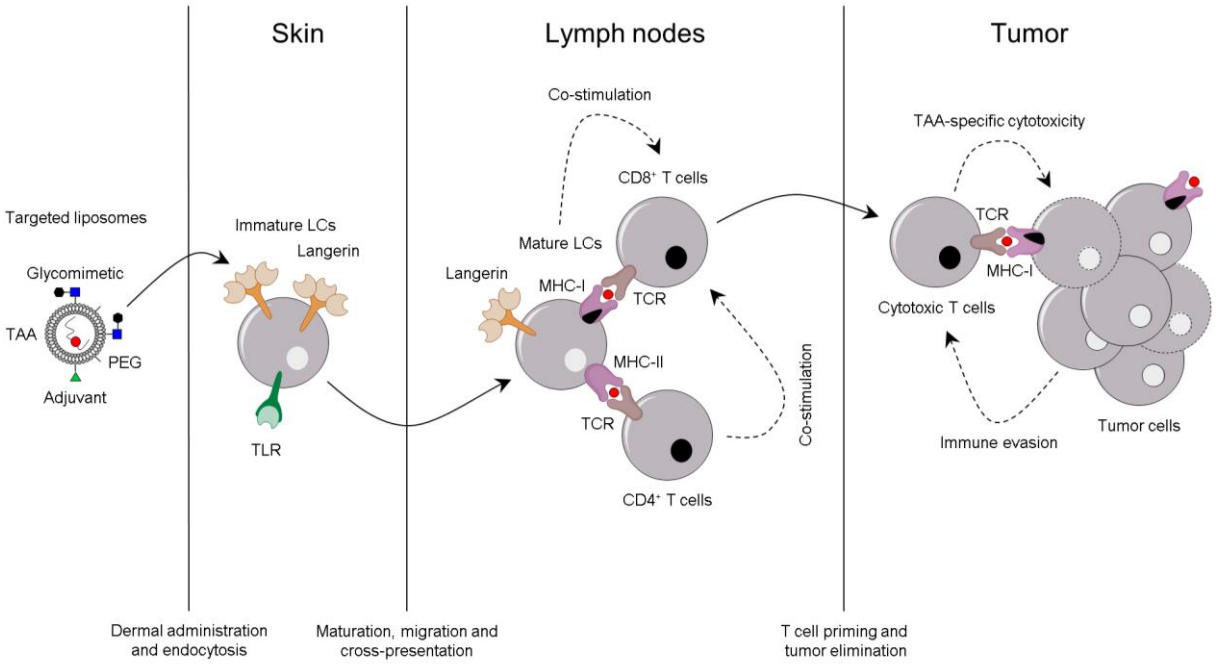


Figure 2. Targeted antigen delivery to LCs for cancer immunotherapy.

Targeting liposomes to LCs represents an intriguing strategy for the development for novel cancer vaccines. The designed liposomes present specific and potent glycomimetic Langerin ligands and are administered to the human skin, either topically or intradermally (left). Upon recognition by Langerin, the liposomes will be endocytosed. The co-formulation of adjuvants promotes LC maturation and migration to peripheral lymph nodes (middle left). Subsequently, delivered TAAs are processed and presented to both $CD4^+$ and $CD8^+$ T cells. The efficient activation of cytotoxic T cells depends on the expression or secretion of costimulatory molecules by LCs such as CD80 and IL-15 (Palucka and Banchereau, 2013). Additionally, a robust helper T cell response is required to promote T cell immunity, *i.a. via* the secretion of interferon-γ and IL-21 (middle right) (Palucka and Banchereau, 2013). Activated TAA-specific cytotoxic T cell disseminate and infiltrate the tumor. Upon recognition of the TAA, tumor cells are eliminated *via* the secretion of perforins and granzymes or the Fas-FasL pathway (right). The tolerogenic tumor microenvironment can be addressed by combination therapy, *e.g. via* the co-administration of anti-CTLA4 or anti-PD1 to target checkpoint inhibitors.

In conclusion, targeting liposomes to LCs *via* glycomimetic Langerin ligands represents an attractive cancer vaccination strategy (Figure 2). This approach addresses most of the limitation encountered for the utilization of glycan- or antibody-antigen conjugates and is readily transferrable to other therapeutic applications such as the treatment of autoimmune or infectious diseases. Moreover, targeted liposomes for $Langerin^+$ DC subsets in the human skin represent valuable molecular probes to investigate DC immunology and the underlying mechanisms of cross-presentation. The implementation of the proposed strategy requires specific and potent glycomimetic Langerin ligands. Notably, no ligand design studies have been published for Langerin and the identification of suitable monosaccharide analogs represents the central aim of the work presented in this dissertation.

1.4. Design Strategies for Glycomimetic C-Type Lectin Receptor Ligands

Glycomimetic Ligand Design – Definitions, Challenges and Opportunities

Glycan-protein interactions (GPIs) are involved in many physiological and pathophysiological processes including immune regulation, intracellular routing and pathogen adhesion (Varki, 2009). Hence, GBPs represent attractive targets for the development of novel therapies *e.g.* against infectious and autoimmune diseases or cancer (Ernst and Magnani, 2009; Palucka and Banchereau, 2013). Synthetic molecules able to modulate GPIs are termed glycomimetics and can be further classified according to their binding mode. Ligands that mimic a subset of the interactions formed between naturally occurring carbohydrates and GBPs can be defined as structural glycomimetics. For this dissertation, this definition includes both carbohydrate analogs and non-carbohydrate molecules. By contrast, functional glycomimetics include synthetic ligands that interact with secondary binding pockets on the GBP surface and modulate GPIs *via* steric hindrance or allosteric mechanisms.

Despite the considerable therapeutic potential associated with the modulation of GPIs, glycomimetics have been largely neglected by the drug discovery field (Ernst and Magnani, 2009). Glycans are complex molecules displaying diverse stereo- and regioisomerism as well as a high density of hydroxyl groups. Hence, the synthesis of carbohydrate analogs and other structural glycomimetics requires elaborate protecting group strategies and is generally considered onerous (Hahm *et al.*, 2017; Seeberger and Werz, 2007). The chemical space explored by structural glycomimetics has been limited and diverse libraries are typically not available for screening (Lenci *et al.*, 2016). Moreover, glycans are hydrophilic and display considerable conformational flexibility. These physicochemical properties have impeded the elucidation of the structural and thermodynamic determinants governing GPIs (DeMarco and Woods, 2008). As highlighted in the previous chapters, CLRs as well as other GBPs typically bind mono- or oligosaccharides with low affinities in the micro- to millimolar range (Chapter 1.2 und 1.3) (Weis and Drickamer, 1996). Carbohydrate binding sites have evolved to accommodate glycans and are thus hydrophilic and typically display high solvent exposure. In summary, the unique characteristics of glycans and GBPs render the design of potent glycomimetics challenging, particularly for traditional medicinal chemistry approaches (Ernst and Magnani, 2009).

The hydrophilicity of carbohydrates and their synthetic analogs additionally results in unfavorable pharmacokinetic profiles. Major limitations arise from rapid renal clearance and limited adsorption in the human intestine (Ernst and Magnani, 2009). Notably, these considerations are of limited concern for the design of glycomimetic targeting ligands. Upon their conjugation to nanoparticles, the pharmacokinetic profile is predominantly determined by the physicochemical properties of the

delivery platform. Moreover, the multivalent organization of the targeting ligands reduces the potency requirements for the design process. In contrast to the development of orally available GPI inhibitors, the design of glycomimetic targeting ligands involves the identification of suitable conjugation strategies. For liposomal LC targeting, the primary route of administration is the skin, circumventing the need for oral availability. Furthermore, the blood group antigen Le^X (K_D = ca. 1 mM) has been successfully utilized to target liposomes to DC-SIGN on DCs to activate T cells *in vitro* (Fehres *et al.*, 2015c; Pederson *et al.*, 2014). This observation indicates that micro- to millimolar affinities are likely sufficient to efficiently mediate Langerin-dependent endocytosis by LCs.

Recent advances have addressed many of the challenges of glycomimetic ligand design and led to the discovery of potent GPI inhibitors or targeting ligands for several classes of GBPs including CLRs, I-type lectin receptors (ILRs) and bacterial adhesins (Chang *et al.*, 2010; Mangold *et al.*, 2012; Mydock-McGrane *et al.*, 2016; Nycholat *et al.*, 2012; Rillahan *et al.*, 2012). This chapter aims to provide an overview of modern concepts and strategies that guide the design process. Here, the development of E-Selectin inhibitors represents an ideal case study as these inhibitors represent the first example of glycomimetic CLR ligands explored in clinical trials (Chang *et al.*, 2010). This case study is complemented by an introduction to strategies facilitating the design, synthesis and screening of focused mono- or oligosaccharide analog libraries. Lastly, the multivalent organization of glycomimetics represents an attractive strategy to target oligomeric CLRs and the underlying concepts are outlined briefly.

The Design of E-Selectin Inhibitors – A Case Study

As presented in the previous chapter, endothelial cells express E-Selectin in presence of inflammatory cytokines and thereby promote the adhesion and extravasation of leukocytes (Chapter 1.2) (Ley, 2003). The adhesion process is mediated by E-Selectin binding to the glycoproteins ESL-1 or PSL-1 which present the blood group antigens sLe^X and sLe^A (Lenter *et al.*, 1994; Steegmaier *et al.*, 1995). In patients with Sickle-cell anemia, aberrant erythrocytes also adhere to the CLR potentially causing the occlusion of blood vessels. Consequently, E-Selectin inhibitors generally display anti-inflammatory properties and have entered clinical trials for the treatment of Sickle-cell anemia (Chang *et al.*, 2010).

During the initial stages of glycomimetic ligand design, the elucidation of the mono- and oligosaccharide specificity of the targeted CLR is essential (Figure 3). To minimize the synthetic efforts for the ensuing design process it is desirable do define the minimal glycan epitope recognized. For E-Selectin, the sLe^X tetrasaccharide was identified as an optimal scaffold (Nelson *et al.*, 1993; Ramphal *et al.*, 1994). These initial experiments utilized heterogeneous fluorescence-based competitive binding assays to evaluate the affinities of individual oligosaccharides. By comparison, glycan microarrays provide optimized throughput and material consumption. Hence, the emergence of

this technology over the last decades has greatly facilitated the initial stages of glycomimetic ligand design (Broecker and Seeberger, 2017).

In absence of an X-ray structure for E-Selectin, trNOE and *trans*-glycoside J-coupling NMR experiments as well as Monte Carlo (MC) and molecular dynamics (MD) simulations served to determine the bioactive conformation of sLe^X (Harris *et al.*, 1999; Peters *et al.*, 1995; Scheffler *et al.*, 1997). The tetrasaccharide displays low conformational flexibility and readily adopts the bioactive conformation in the free state. This preorganization minimizes the decrease in conformational entropy upon E-Selectin binding and contributes to an entropically driven recognition process that is based on the replacement of structural H_2O molecules bound to the carbohydrate binding site (Binder *et al.*, 2012). Interestingly, the bioactive conformation is stabilized *via* hydrophobic interactions and the formation of a non-canonical, intramolecular CH-O hydrogen bonds between Gal and Fuc (Zierke *et al.*, 2013).

Extensive structure-activity relationship (SAR) studies revealed that only the carboxyl group, the hydroxyl group of Fuc and the hydroxyl groups in C4 and C6 of Gal form essential interactions with E-Selectin (Kolb and Ernst, 1997). Based on this SAR, an *in silico* approach utilizing MC simulations served to design a reduced scaffold preorganized in the previously determined bioactive conformation of sLe^X (Figure 3) (Kolb and Ernst, 1997). The *N*-acetylneuraminic acid (NeuAc) and the GlcNAc were substituted with cyclohexyl rings and the corresponding glycomimetic displayed a 15-fold affinity increase over sLe^X (Binder *et al.*, 2012). trNOE NMR experiments and MC simulations validated the efficient sampling of the bioactive conformation for the free state of the reduced scaffold (Jahnke *et al.*, 1997). The retrospective thermodynamic analysis of the interactions between the glycomimetic and E-Selectin revealed substantial enthalpic contributions to the affinity increase (Binder *et al.*, 2012). These contributions can be attributed to the absence of unfavorable solvation effects associated with the hydrophilicity of NeuAc and GlcNAc in sLe^X. Notably, the importance of solvation effects has also been observed for other GBPs such as bacterial adhesins and the concept of scaffold reduction has been transferred to monosaccharides, *e.g.* for the design of LecB inhibitors (Navarra *et al.*, 2017; Sommer *et al.*, 2014).

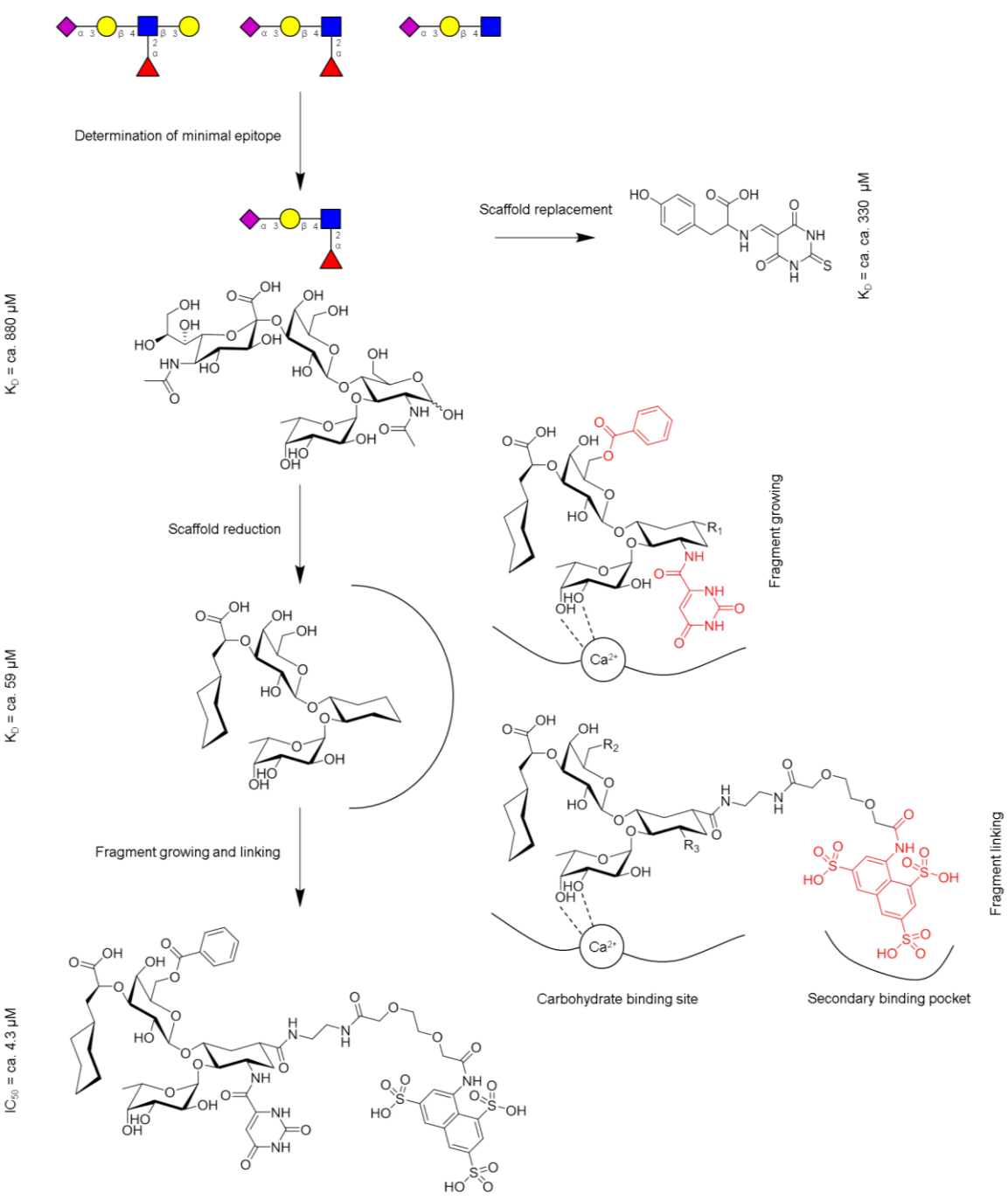


Figure 3. Design of glycomimetic E-Selectin ligands.

The initial stages of glycomimetic ligand design involved the identification of the sLe^X tetrasaccharide as an optimal scaffold (Nelson *et al.*, 1993; Ramphal *et al.*, 1994). The reduction of the sLe^X scaffold was guided by extensive SAR studies, trNOE NMR experiments as well as MC simulations and resulted in a 15-fold affinity increase (Binder *et al.*, 2012; Kolb and Ernst, 1997). Fragment growing led to the discovery of favorable substituents in C2 of the GlcNAc scaffold and C6 of the Gal scaffold that likely form secondary interactions with the carbohydrate binding site (Chang *et al.*, 2010; Thoma *et al.*, 2001a). Additionally, an R_2-filtered NMR screening served to identify fragments binding to a secondary binding pocket (Egger *et al.*, 2013). These fragments were linked to the reduced sLe^X scaffold and combined with substituents identified *via* fragment growing to afford an additional 14-fold affinity increase. The corresponding glycomimetic lead structure is currently explored in clinical trials for its efficacy in patients with Sickle cell anemia (Chang *et al.*, 2010). In an alternative strategy, an *in silico* screening served to replace the sLe^X scaffold with a non-carbohydrate inhibitor displaying a 3-fold affinity increase over sLe^X (Barra *et al.*, 2016).

Subsequently, the reduced sLe^X scaffold was derivatized to explore the carbohydrate binding site of E-Selectin for secondary interactions (Figure 3). The synthesis of focused glycomimetic libraries led to the discovery of favorable substituents in C6 of Gal and C2 of GlcNAc (Thoma *et al.*, 2001a; Thoma *et al.*, 1999; Thoma *et al.*, 2001b). Interestingly, modification in both positions were observed to also contribute to the preorganization of the reduced sLe^X scaffold (Schwizer *et al.*, 2012; Thoma *et al.*, 2001a). Additionally, an elegant R_2-filtered NMR screening strategy served to identify fragments binding to a secondary binding pocket adjacent to the carbohydrate binding site (Figure 3) (Egger *et al.*, 2013). A primary screening detected fragment interactions with the E-Selectin surface, irrespective of the location of the binding pocket. Next, a glycomimetic reporter molecule bearing a stable nitroxide spin label was designed to identify fragments binding in proximity of the carbohydrate binding site of E-Selectin. The utilized spin label induces distance-dependent paramagnetic relaxation enhancement and allowed for the determination of the orientation of selected fragments relative to the glycomimetic reporter molecule. Based on this structural information, fragments were conjugated to C6 of one of the cyclohexyl rings utilizing a library of linker structures. Strikingly, the identification of a suitable linker structure resulted in a 48-fold affinity increase for the designed glycomimetics. In a similar approach, the linker structure was optimized *in situ* to discover potent Sial analogs targeting Siglec-4 (Shelke *et al.*, 2010).

The substituents identified in these initial experiments were further optimized to improve both the affinity and the pharmacokinetic profile of the designed glycomimetics. Finally, a glycomimetic lead structure displaying a 210-fold affinity increase over sLe^X was identified (Figure 3). The corresponding molecule bears substituents targeting the carbohydrate site as well as the adjacent secondary binding pocket and has entered clinical trials for the treatment of Sickle cell anemia (Chang *et al.*, 2010). These final optimization studies were conducted by the company Glycomimetics and neither the corresponding SARs nor the formed interactions have been disclosed. Nevertheless, examples for other GBPs such as DC-SIGN or the bacterial adhesin FimH corroborate the feasibility of the derivatization of optimized mono- or oligosaccharide scaffolds to explore the carbohydrate binding site for favorable secondary interactions (Garber *et al.*, 2010; Mydock-McGrane *et al.*, 2016; Pang *et al.*, 2012; Varga *et al.*, 2013).

In an alternative approach, the sLe^X scaffold was replaced by non-carbohydrate inhibitors of E-Selectin (Figure 3) (Barra *et al.*, 2016). An *in silico* screening integrating molecular docking and MD simulations led to the identification of a novel class of glycomimetics displaying a 3-fold affinity increase over sLe^X. While no competitive binding experiments were conducted and the Ca^{2+}-dependency of the interaction was not evaluated, STD NMR experiments and CORCEMA-ST simulations served to validate binding of the non-carbohydrate inhibitors to the carbohydrate binding site of E-Selectin. Accordingly, these non-carbohydrate inhibitors can be classified as structural glycomimetics. Functional GPI inhibitors, in turn, have been discovered *via* fragment screening for

e.g. DC-SIGN and the hyaluronic acid-binding GBP CD44 (Aretz *et al.*, 2017; Liu and Finzel, 2014). For DC-SIGN, these functional glycomimetics bind to remote secondary binding pockets and presumably modulate GPIs *via* an allosteric mechanism. Notably, some of the identified fragments represent substructures of previously reported, potent non-carbohydrate DC-SIGN inhibitors (Mangold *et al.*, 2012). By contrast, the glycomimetic fragments discovered for CD44 were found to interact with secondary binding pocket adjacent to the carbohydrate binding site and modulate GPIs *via* steric hindrance.

The design of non-carbohydrate ligands for CLRs and other GBPs is generally desirable because these glycomimetics typically display improved pharmacokinetic profiles and synthetic efforts are considerably reduced (Ernst and Magnani, 2009). In this context, it is noteworthy that the utilization of functional glycomimetics as targeting ligands for Langerin is potentially disadvantageous. As argued in the preceding chapter, structural glycomimetics share a subset of the interactions formed by pathogen-associated glycans and are thus likely to promote Langerin-dependent endocytosis as well as antigen cross-presentation (Chapter 1.3). Moreover, the carbohydrate binding site of Langerin is located on top of the homotrimer formed under physiologically conditions and thus accessible for ligand binding (Feinberg *et al.*, 2010). Neither consideration is necessarily true for secondary binding pockets in other regions of the receptor surface. Concerns regarding the binding pocket accessibility particularly extend to secondary pockets validated *in vitro* experiments with recombinantly expressed receptors as these pockets might be masked *in vivo* due to posttranslational modifications. Langerin does, for example, contain two potential *N*-glycosylation sites in the α-helical oligomerization domain (Valladeau *et al.*, 2000).

In conclusion, this case study highlights several modern concepts and strategies to guide the challenging design process for glycomimetic CLR ligands. First, the identification of the minimal carbohydrate epitope recognized by the targeted CLRs is essential to minimize the synthetic efforts. Secondly, unfavorable solvation effects typically reduce the affinity of carbohydrate scaffolds. Consequently, the reduction of mono- and oligosaccharide scaffolds represents an attractive approach to design potent glycomimetics. Hydrophilic functional groups or monosaccharides not involved in the formation of essential interactions with the carbohydrate binding site can be substituted with more hydrophobic structures to optimize the thermodynamic profile of GPIs. Third, the reduced carbohydrate scaffold should be preorganized in its bioactive conformation. Finally, the derivatization of the scaffold to explore the CLR surface for secondary interactions has been demonstrated to further increase the affinity of designed glycomimetics. Here, substituents might target the carbohydrate binding site or adjacent, secondary binding pockets. Alternatively, the replacement of the carbohydrate scaffold with non-carbohydrate molecules potentially improves the pharmacokinetic profile and facilitates the synthesis of glycomimetics. Remarkably, many of the structural and thermodynamic hypotheses proposed during the described design process have been validated retrospectively *i.a.* by

X-ray crystallography as well as STD NMR and ITC experiments (Binder *et al.*, 2012; Preston *et al.*, 2016; Rinnbauer *et al.*, 2003; Schwizer *et al.*, 2012; Zierke *et al.*, 2013).

Transferring Concepts from Fragment-Based Lead Discovery

It is noteworthy that some of the concepts and strategies presented above originate from fragment-based lead discovery (FBLD). FBLD has been generally recognized as a powerful approach to address challenging drug targets and the underlying principles are outlined briefly in this section to highlight its utility for glycomimetic ligand design (Erlanson *et al.*, 2016). Historically, fragments have been derived from drug-like molecules present in high-throughput screening (HTS) libraries of pharmaceutical companies. According to Lipinsky, orally available, drug-like molecules are smaller than 500 Da, do not bear more than five hydrogen bond donors nor more than ten hydrogen bond acceptors and display partition coefficients clogP that do not exceed values of 5 (Lipinski *et al.*, 2001).

Fragments display reduced molecular weights between 150 and 300 Da and are less complex than drug-like molecules (Erlanson *et al.*, 2016). As a result of this low complexity, fragments are more likely to form unique interactions with protein targets (Hann *et al.*, 2001). However, the maximal affinity expected for protein-ligand interactions correlates with the number of heavy atoms of the ligand (Kuntz *et al.*, 1999). Consequently, fragments typically bind proteins with low affinities in the micromolar to millimolar range and sensitive biophysical techniques such as NMR, surface plasmon resonance (SPR) or thermal shift assays (TSA) are required to successfully screen fragment libraries (Davis and Erlanson, 2013; Mashalidis *et al.*, 2013). Additionally, fragments need to be soluble at micro- to millimolar concentrations in aqueous buffer and thus display increased hydrophilicity (Erlanson *et al.*, 2016).

As the maximal contributions per heavy atom to the free energy of binding ΔG are considerably reduced for larger ligands, fragment screening typically yields hits that form highly efficient interactions compared to HTS (Kuntz *et al.*, 1999). This renders fragments favorable starting points for the lead discovery process (Hann, 2011). Over the last decades, several parameters have been developed to quantify the efficiency of interactions and to guide FBLD (Hopkins *et al.*, 2014). For instance, the ligand efficiency (LE) is calculated by normalizing ΔG values to the number of heavy atoms of a ligand whereas other measures additionally consider physicochemical properties such as the clogP values or the polarizable surface area (PSA).

The combinatorial evolution of fragment hits into lead structures is based on two thermodynamic principles. First, the contributions of individual interactions to ΔG are additive and this additivity can be leveraged *via* the conjugation of two fragments hits (Jencks, 1981). Secondly, the rotational and translational entropies in the free state of the synthesized molecule are comparable to those of each individual fragment, resulting in favorable entropic contributions to ΔG (Murray and Verdonk, 2002).

Strikingly, the conjugation of two fragments displaying K_D values of 1 mM can theoretically translate into the discovery of nanomolar ligands, even without the formation of additional interactions. This estimation, however, assumes the existence of an ideal linker structure that does not interfere with the recognition process. Based on these favorable thermodynamics characteristics, FBLD also facilitates the exploration of chemical space (Erlanson *et al.*, 2016). While typical HTS libraries consist of several million molecules, the screening of fragment libraries at sizes of several thousand molecules is considerable more efficient. Theoretical considerations suggest, that the combinatorial evolution of fragment hits nevertheless provides superior target-specific diversity compared to traditional HTS approaches.

Different strategies are pursued to evolve fragment hits into potent ligands (Erlanson *et al.*, 2016). Fragment linking approaches aim to identify optimal linker structures for the conjugation of hits interacting with adjacent binding pockets. The feasibility of this approach was first demonstrated by Fesik *et al.* in a study that instigated the emergence of FBLD (Shuker *et al.*, 1996). Important parameters include the geometry and conformational flexibility of the linker structure as well as solvation effects (Erlanson *et al.*, 2016). Moreover, structural information on the fragment binding mode and the location of secondary binding pockets is required to guide fragment linking approaches. Consequently, the optimization process is considered difficult and represents a major challenge for FBLD. Alternatively, fragment hits can be derivatized *via* short, rigid linker structures to explore protein surface in proximity of the primary binding pocket for favorable interactions (Wyatt *et al.*, 2008). In contrast to traditional medicinal chemistry approaches, these fragment growing strategies focus on the identification of efficient interactions, *i.e.* by avoiding the introduction of larger, hydrophobic substituents (Hann, 2011; Hopkins *et al.*, 2014). Structural information is not necessarily required for the evolution of hits into potent ligands *via* fragment growing. Nevertheless, STD and ^{15}N HSQC NMR experiments or X-ray crystallography enable the implementation of rational design approaches considerably facilitating the identification of lead structures (Harner *et al.*, 2013). Lastly, two hits displaying overlapping binding modes can be combined to leverage efficient interactions of both fragments (Hudson *et al.*, 2012). This strategy is termed fragment merging.

In conclusion, the screening of typical fragment libraries with sensitive biophysical techniques enables the reliable identification of hits, even for challenging drug target such as CLRs and other GBPs (Aretz *et al.*, 2017; Aretz *et al.*, 2014; Liu and Finzel, 2014). Moreover, identified hits typically display increased LE values compared to HTS hits and thus represent favorable starting points for the design of potent glycomimetics. As fragments are more hydrophilic than drug-like molecules, they are potentially privileged to interact with hydrophilic and solvent exposed carbohydrate binding sites. These characteristics render fragment screening an intriguing approach to replace carbohydrate scaffolds with non-carbohydrate glycomimetics (Figure 3). As highlighted by the E-Selectin case study, carbohydrate scaffolds can be derivatized to discovery potent CLR ligands. This approach is

analogous to fragment linking and growing strategies implemented in FBLD (Figure 3). Hence, it can be argued that the utilization of fragments for the derivatization of carbohydrate scaffolds enables the efficient exploration of chemical space and facilitates the identification of favorable interactions. The thermodynamic principles of the combinatorial evolution of fragments are readily transferrable to other ligand classes. As carbohydrates typically bind GBPs with low affinities, the application of fragment growing and linking approaches are ideally suited to design potent glycomimetics (Egger *et al.*, 2013; Shelke *et al.*, 2010). The design process can be guided *via* the calculation of group efficiencies (GE) to monitor efficiency of secondary interactions formed by introduced substituents (Verdonk and Rees, 2008). In analogy to the LE, this parameter normalizes affinity increase over the carbohydrate scaffold, *i.e.* $\Delta\Delta G$ values, to the number of heavy atoms of the fragment.

Binding Pocket Identification and Druggability Analysis

Importantly, the feasibility of fragment growing and linking approaches to modulate GPIs depends on the existence of suitable secondary binding pockets, ideally in proximity of the carbohydrate binding site. Aside from proof-of-principle SAR studies for the target GBP, structural biology tools have been established to identify binding pockets. Prominent examples include the screening of solvent molecules and fragments *via* X-ray crystallography or protein-observed NMR experiments (Cheng *et al.*, 2007; Hajduk *et al.*, 2005a; Mattos and Ringe, 1996). Alternatively, several *in silico* methods are available to predict binding pockets and can be classified as energy-, geometry or evolutionary-based approaches. Compared to experimental techniques, these methods require limited resources and are thus particularly attractive for systematic evaluations of protein families, *e.g.* CLRs. Over three decades ago, Goodford reported the first energy-based approach utilizing small molecular probes to predict interaction potentials on the protein surface (Goodford, 1985). Geometry-based approaches typically identify binding pockets *via* grid-based algorithms that might be improved *via* the utilization of additional geometric descriptors such as difference of Gaussian (DoG) filters or α spheres (Le Guilloux *et al.*, 2009; Volkamer *et al.*, 2010). In contrast to energy- and geometry-based algorithms, evolutionary-based approaches do not rely on the availability of X-ray structures. Here, multiple sequence alignments and homology modeling serve to identify conserved residues and binding pockets (Armon *et al.*, 2001). Notably, conformational flexibility represents a major challenge for *in silico* binding pocket predictions. Recent advances have addressed this challenge *via* the integration of MD simulations (Bakan *et al.*, 2012; Cimermancic *et al.*, 2016).

The scientific problem of binding pocket identification is closely related to the concept of druggability. A therapeutically relevant protein family is defined as druggable if the function of at least one family member can be efficiently modulated by drug-like molecules (Hopkins and Groom, 2002). This requires the existence of at least one binding pocket with the capacity to bind drug-like molecules with nano- to micromolar affinities. By this definition, approximately 10% of the human genome have been estimated to encode for druggable target proteins (Hopkins and Groom, 2002).

Strikingly, approximately 60% of drug discovery projects fail during lead optimization as target protein are found to be undruggable (Brown and Superti-Furga, 2003). Considering the substantial expenses involved, methods to predict the druggability of protein families have gained attention over the last decades (Egner and Hillig, 2008). Experimentally determined hit rates from fragment screening have been recognized as robust predictors of druggability (Edfeldt *et al.*, 2011; Hajduk *et al.*, 2005a). Both ligand-observed and protein-observed NMR techniques have been extensively utilized as they display low false positive and false negative rates at sufficient throughput (Edfeldt *et al.*, 2011). Generally, alternative methods complying with these criteria can also be employed (Aretz *et al.*, 2016; Edfeldt *et al.*, 2011). Additionally, several integrated *in silico* methods have been developed to predict the druggability of identified binding pockets (Halgren, 2009; Schmidtke and Barril, 2010; Volkamer *et al.*, 2012a; Volkamer *et al.*, 2012b). The corresponding algorithms are trained *via* machine learning on available druggability data sets and typically account for geometry, physicochemical properties as well as the evolutionary conservation of binding pockets (Schmidtke and Barril, 2010). For many protein families, calculated druggability scores correlate well with experimentally determined hit rates (Chen and Hubbard, 2009; Hajduk *et al.*, 2005a; Jordan *et al.*, 2012). Prior to the work presented in this dissertation, the druggability of GBPs was evaluated in two *in silico* studies, both reporting low scores (Hajduk *et al.*, 2005b; Schmidtke and Barril, 2010). Yet, no detailed information on identified binding pockets or on the druggability of CLRs was available.

Strategies to Discover Potent Mono- and Oligosaccharide Analogs

As highlighted above, the derivatization of carbohydrate scaffolds to explore the CLR surface for favorable interactions represents a powerful approach to design potent glycomimetics. Fragment growing- or linking-inspired strategies are ideally suited to guide the design process. However, the onerous synthesis of mono- or oligosaccharide analogs impedes glycomimetic ligand design considerably. Recent advances in the automated or enzymatic assembly of oligosaccharides as well as novel diversity-oriented strategies for the preparation of structural glycomimetics have addressed these limitations (Danby and Withers, 2016; Hahm *et al.*, 2017; Lenci *et al.*, 2016; Seeberger and Werz, 2007). Prominently, a shikimic acid-based synthetic route for Man and Fuc analogs enabled the efficient identification of potent DC-SIGN inhibitors (Garber *et al.*, 2010; Grim *et al.*, 2011). Nevertheless, large and diverse libraries of structural glycomimetics are typically not available for screening.

Target-oriented synthesis thus represents a feasible alternative for glycomimetic ligand design. Focused libraries of carbohydrate analogs are typically prepared following traditional medicinal chemistry approaches to target individual GBPs (Pang *et al.*, 2012; Varga *et al.*, 2013). Over the last decade, several elegant strategies have been devised to further reduce the synthetic efforts associated with the screening of focused glycomimetic libraries. As the successful utilization of fragment linking to discover E-Selectin and Siglec-4 inhibitors has been described in detail above, this section focuses

on strategies that facilitate the application of fragment growing to modulate GPIs (Egger *et al.*, 2013; Shelke *et al.*, 2010).

Based on X-ray structures of GBP-carbohydrate complexes, structure-based *in silico* methods enable the generation of binding hypotheses to guide the design of focused glycomimetic libraries. Consequently, the number of synthesized mono- or oligosaccharide analogs required to identify potent ligands can be reduced substantially. This approach is of particular value during the initial stages of the design process when only limited SAR information is available. However, the prediction of structural and thermodynamic determinants of GPIs remains challenging, particularly for oligosaccharides and GBPs displaying high conformational flexibility (DeMarco and Woods, 2008). Current limitations arise *i.a.* from the selection of suitable scoring functions, force field parametrization as well as the impact of structural H_2O molecules and solvation effects (Binder *et al.*, 2012; Eid *et al.*, 2014; Huang and Shoichet, 2008; Kirschner *et al.*, 2008). In principle, MD simulations enable the estimation of ΔG values for GPIs *via* thermodynamic integration or free energy perturbation calculations (Sommer *et al.*, 2014; Wang *et al.*, 2015). Yet, both methods are computationally demanding and permit only a limited set chemical transformations. Hence, they are not suitable for the virtual screening of larger carbohydrate analog libraries. Molecular docking is less accurate with respect to the prediction of both binding modes and ΔG values but provides improved throughput and can be readily implemented to design focused glycomimetic libraries.

Modern molecular docking procedures typically involve five stages. Initially, a conformation database for the evaluated carbohydrate analogs is generated. Conformations are enumerated *via* stochastic and evolutionary algorithms or predicted from low mode MD simulations (Ferguson and Raber, 1989; Labute, 2010; Morris *et al.*, 2009). These conformations are placed into the carbohydrate binding site either utilizing a pharmacophore model or various available grid-based algorithms (Chemical Computing Group, 2016; Friesner *et al.*, 2004; Morris *et al.*, 2009). Subsequently, the generated docking poses are scored *via* heuristic scoring functions and highly scored poses are refined in molecular mechanics (MM) or MD simulations. Finally, refined docking poses are rescored to estimate ΔG values. Generated poses can additionally be filtered *via* a pharmacophore model and are typically evaluated visually.

As stated above, the selection of suitable scoring functions is critical for the efficient design of glycomimetic libraries (Eid *et al.*, 2014). The availability of X-ray structures of GBPs in complex with carbohydrate analogs enables the evaluation of different scoring functions to guide this selection. For instance, the program FRED was observed to be superior at reproducing the binding mode of Man analogs for DC-SIGN compared to several alternative molecular docking programs (Jug *et al.*, 2015). Additionally, the utilized consensus scoring *via* the Chemgauss and Chemscore functions robustly discriminated between active and inactive glycomimetics (Eldridge *et al.*, 1997; Verkhivker *et al.*, 2000). During the initial stages of glycomimetic ligand design, only limited SAR information and no

X-ray structures for carbohydrate analogs are available. Hence, the evaluation of different scoring functions and placement algorithms is not feasible. Under these conditions, the orientation of the carbohydrate scaffold in the binding site can be constrained by *e.g.* pharmacophore models (Nycholat *et al.*, 2012). This reduces the degrees of freedom during the generation and refinement of docking poses and circumvents the challenging prediction of GPIs. Generally, the utilization of sophisticated scoring functions *e.g.* based on MM/generalized Born surface accessibility (GBSA) simulations further improves the robustness affinity estimations (Massova and Kollman, 2000).

In silico methods have been instrumental for the design of potent Sial analogs for several ILRs including Siglec-1, -4 and -7 (Mesch *et al.*, 2010; Nycholat *et al.*, 2012; Rillahan *et al.*, 2013). For Siglec-1, a virtual screening of approximately 8000 commercially available carboxylic acid substituents in C9 of NeuAc led to the discovery of a nanomolar ligand with improved selectivity against other ILRs (Nycholat *et al.*, 2012). Based on an X-ray structures of Siglec-1 in complex with a previously reported Sial analog, the molecular docking procedure utilized a pharmacophore model to constrain the orientation of the carbohydrate scaffold and the placement of aromatic substituents (Zaccai *et al.*, 2003). The virtual *in situ* conjugation of the carboxylic acid substituents was followed by the grid-based generation of docking poses and scoring *via* the proprietary London ΔG function (Chemical Computing Group, 2016). This initial docking procedure was implemented *via* Molecular Operating Environment (MOE) and 3000 highly scored poses were subsequently re-evaluated in AutoDock (Morris *et al.*, 2009). Finally, a set of only eight Sial analogs was selected for synthesis and affinity characterizations. The most potent glycomimetic (IC_{50} = ca. 0.4 μM) was able to selectively target liposomes to Siglec-1^{+} macrophages *in vivo* (Nycholat *et al.*, 2012).

For Siglec-4 and -7, molecular docking procedures served to rationalize SAR information obtained for previously synthesized Sial analogs (Mesch *et al.*, 2010; Rillahan *et al.*, 2013). In the absence of X-ray structures for the corresponding glycomimetics, these studies provide valuable structural information for the rational design of next generation Sial analogs. Aside from the preliminary investigation for DC-SIGN, structure-based *in silico* methods have not been applied to discover favorable substituents on mono- or oligosaccharide scaffolds to target CLRs (Jug *et al.*, 2015). However, STD and trNOE NMR experiments were combined with molecular docking to predict the binding mode of the heparin-derived trisaccharides to Langerin (Munoz-Garcia *et al.*, 2015).

In addition to *in silico* methods, experimental approaches integrating solid-phase synthesis with the with the *in situ* screening of focused libraries of carbohydrate analogs have been successfully applied to discover potent glycomimetics for various GBPs including the ILRs Siglec-7 and Siglec-9 as well as Shiga-like Toxin-2 (Dasgupta *et al.*, 2014; Rillahan *et al.*, 2012; Rillahan *et al.*, 2013). These approaches provide an optimized throughput and reduced synthetic efforts as they circumvent the laborious purification of individual ligands. For the ILRs, Sial scaffolds bearing alkyne or azido groups in C9 or C5 of NeuAc were printed on NHS-activated microarrays (Rillahan *et al.*, 2012;

Rillahan *et al.*, 2013). Subsequently, these scaffolds were derivatized with up to 100 substituents *via* Cu(I)-catalyzed Huisgen 1,3-dipolar cycloaddition (Rostovtsev *et al.*, 2002). The incubation of the microarrays with fluorescently labeled ILRs led to the identification of potent glycomimetics, that were re-synthesized in solution phase and successfully utilized to target liposomes to Siglec-7^+ model cells or Siglec-9^+ monocytes *in vitro*.

In a similar approach, Bundle *et al.* conjugated a P^k trisaccharide scaffold bearing a Boc protected amino group in C2 of Gal to a vinyl polymer by Cu(I)-catalyzed Huisgen 1,3-dipolar cycloaddition (Dasgupta *et al.*, 2014; Rostovtsev *et al.*, 2002). The polymer was immobilized on multi-well plates, followed by the deprotection of the amino group and derivatization of the P^k trisaccharide scaffold with approximately 100 substituents bearing *e.g.* acyl chlorides or sulfonyl chlorides. Subsequently, the affinity of Shiga-like Toxin 2 for the synthesized carbohydrate analogs was evaluated *in situ* utilizing an enzyme-linked lectin assay (ELLA). The most promising hits were re-synthesized in solution phase revealing a 30-fold affinity increases over the P^k trisaccharide for the most potent glycomimetic (IC_{50} = ca. 0.2 mM). Notably, the utilized screening assay provides excellent sensitivity based on the multivalent interactions formed by the pentameric Shiga-like Toxin 2 and is thus ideally suited for the initial stages of glycomimetic ligand design.

In summary, the presented examples highlight recent advances to facilitate the discovery of potent carbohydrate analogs. The generation of binding hypotheses *via* structure-based *in silico* methods potentially reduces the number of synthesized glycomimetics during the design process. Alternatively, the combination of solid-phase synthesis and *in situ* screening circumvents the laborious purification of individual mono- or oligosaccharide analogs. Ultimately, these approaches might be integrated with the rational optimization of carbohydrate scaffolds and concepts from FBLD to overcome the fundamental challenges of glycomimetic ligand design and to efficiently target CLRs (Ernst and Magnani, 2009).

Multivalent Organization of Glycomimetic Ligands

The multivalent organization of mono- and oligosaccharides or glycomimetics on synthetic scaffolds structures can be leveraged to address the challenges encountered for the design of potent and specific CLR ligands outlined above (Cecioni *et al.*, 2015). Multivalent interactions display increased avidities compared to the corresponding monovalent recognition processes. The magnitude of these avidity increases are determined by statistical and chelate cooperativity effects (Fasting *et al.*, 2012). Statistical effects generally contribute positively to multivalent interactions and correlate with the valency of the system. In part, these effects can be attributed to increased effective concentrations and rebinding events. By contrast, chelate cooperativity might be positive or negative, depending on the scaffold structure as well as the thermodynamic and kinetic properties of the recognition process. In the context of this dissertation, chelate cooperativity is defined as the tendency of multivalent

glycomimetics to simultaneously engage multiple carbohydrate binding sites of an individual CLR. This definition is closely related to the additivity of ΔG values originally proposed by Jencks and outlined above in the context of FBLD (Jencks, 1981)

For the design of potent multivalent glycomimetics, statistical effects are typically optimized by increasing the valency of a scaffold structure (Fasting *et al.*, 2012). Notably, this approach potentially amplifies the avidity for other GBPs if the monovalent interactions are not specific for the targeted CLR. By contrast, the design of scaffold structures that match the geometry of oligomeric CLRs represents a powerful strategy to leverage chelate cooperativity effects and will ideally provide specificity (Fasting *et al.*, 2012). This concept can, in principle, also be applied to target monomeric CLRs based on their spatial distribution in the plasma membrane. However, the corresponding multi-state equilibria of are highly complex and the quantification of thermodynamic or kinetic parameters is non-trivial (Hunter and Anderson, 2009; Kitov and Bundle, 2003). Consequently, the design of multivalent glycomimetics is typically guided by empirical parameters such as apparent K_D or IC_{50} values to approximate avidity increases.

Over the last decades, the development of synthetic strategies for novel scaffold structures including glycoclusters, glycodendrimers, glycopolymers and glyconanoparticles has enabled the discovery of potent multivalent glycomimetic GBP ligands (Cecioni *et al.*, 2015). Other strategies involve the presentation of monovalent ligands on biomolecules such as proteins or nucleic acids. In a hallmark study, Bundle *et al.* reported the structure-based design of a decavalent glycocluster targeting the Shiga-like Toxin-1 (Kitov *et al.*, 2000). Remarkably, the inhibitor displayed a more than 10^6-fold avidity increase compared the interactions formed by the corresponding P^k trisaccharide. Alternatively, the toxin has been targeted with P^k trisaccharide-functionalized gold nanoparticles that bear potential as synthetic biomarkers for *E. coli* infections (Chien *et al.*, 2008). Recently, the application of solid-phase synthesis has enabled the design of monodisperse, polymeric glycoclusters displaying defined valency (Ponader *et al.*, 2014). The prepared heteromultivalent glycomimetics were utilized to target oligomeric plant lectins and led to the discovery of shielding effects for multivalent interactions.

Examples for CLRs include the synthesis of glycodendrimers and glycoclusters to inhibit DC-SIGN-mediated HIV and Ebola virus infections *in vitro* (Luczkowiak *et al.*, 2011; Ordanini *et al.*, 2015). While the regioselective functionalization of protein scaffolds remains challenging, recent advances enabled the utilization of virus capsid scaffold in similar approaches (Ribeiro-*Via*na *et al.*, 2012). Prominently, ASGPR-targeted glycoclusters are currently explored in clinical trials for their therapeutic potential against liver cancer (Huang *et al.*, 2017).

Nucleic acids represent an attractive class of scaffold structures as they provide excellent control over the valency and spatial orientation of organized ligands. The unique fidelity of Watson-Crick base pairing enables the programmed self-assembly of monodisperse deoxyribonucleic acid (DNA)-DNA

duplexes. Compared to these duplexes, the utilization of peptide nucleic acid (PNA) provides optimized thermodynamic and enzymatic stability (Scheibe *et al.*, 2011). Available NMR structures of PNA-DNA duplexes and MD simulations guide the design of these scaffolds (Eriksson and Nielsen, 1996; Scheibe *et al.*, 2011). Importantly, PNA-DNA duplexes were found to be rigid at lengths above 100 Å and the spatial organization of ligands can thus be predicted from the sequence of the DNA template. The flexibility required to optimally orient mono- and oligosaccharides or monovalent glycomimetics in the carbohydrate binding site is conveyed by the introduction of nicks between PNA fragments and by the linker structures utilized for ligand organization (Scheibe *et al.*, 2011).

Programmed nucleic acid scaffolds have been utilized to discover and organize divalent glycomimetics targeting DC-SIGN on DCs *iv vitro*. Yet, this design approach did not explicitly account for the geometry of the tetrameric DC-SIGN ECD, partly due to extended and branched linker structures (Ciobanu *et al.*, 2011). Recently, approaches providing optimized control over the spatial organization of mono- and oligosaccharides have successfully been applied to map the geometry of several of oligomeric plant lectins (Scheibe *et al.*, 2011; Scheibe *et al.*, 2013). Additionally, the designed PNA-DNA duplexes enabled the evaluation of the spacing between monomeric L-Selectin in the plasma membrane of a model cells *in vitro* (Scheibe *et al.*, 2013).

1.5. Characterizing C-Type Lectin Receptor-Ligand Interactions by NMR

Assay Development for Glycomimetic Ligand Design

The availability of robust assays providing suitable throughput and material consumption is essential for glycomimetic ligand design. However, the low affinities observed for GPIs translate into sensitivity problems and a lack of reporter molecules, particularly during the initial stages of the design process (Ernst and Magnani, 2009). These limitations have been compensated for by the multivalent organization of carbohydrates and CLRs in heterogeneous assays (Cecioni *et al.*, 2015). Several approaches leveraging avidity effects have been explored to detect GPIs or to quantify the affinity of glycomimetics.

Over the last decades, glycan microarrays have emerged as an efficient technology to evaluate oligosaccharide specificity of GBPs and were successfully employed to screen glycomimetic libraries against Siglec-7 (Broecker and Seeberger, 2017; Rillahan *et al.*, 2013). However, the technology does not enable the quantification of affinities due to *i.a.* the impact of the utilized linker structures (Grant *et al.*, 2014). Hence, glycan microarrays are not suitable for the elucidation of SARs in the context of glycomimetic ligand design. Alternatively, the immobilization of GBPs or carbohydrates on multi-well plates has been utilized to determine EC_{50} and IC_{50} values *e.g.* in enzyme-linked lectin assays (ELLA) or *via* fluorescence detection (Hanske *et al.*, 2017a; Scharenberg *et al.*, 2012). These competitive binding assays served to quantify the Langerin affinity for several mono- and oligosaccharides or to guide the design of glycomimetic DC-SIGN inhibitors (Feinberg *et al.*, 2013; Garber *et al.*, 2010; Stambach and Taylor, 2003).

The utilization of micro- and nanoparticles in similar approaches might be envisioned and has been demonstrated to be compatible with cell-based assays (Sprokholt *et al.*, 2016; Zhang *et al.*, 2015). The characterization of GPIs in these assays potentially involves flow cytometry experiments or the detection of aggregation *e.g. via* dynamic light scattering (DLS). Alternative cell-based approaches include HIV transmission inhibition assays to discover glycomimetic DC-SIGN inhibitors and the screening of glycomimetic libraries against Siglec-9^+ model cells *via* glycan microarray technology (Ordanini *et al.*, 2015; Rillahan *et al.*, 2012).

Furthermore, heterogeneous SPR experiments integrate the excellent sensitivity of biophysical techniques with avidity effects. These experiments provide optimal throughput and material consumption and were applied to guide the design of glycomimetic DC-SIGN or E-Selectin inhibitors (Andreini *et al.*, 2011; Aretz *et al.*, 2017; Egger *et al.*, 2013). For Langerin, SPR served to characterize interactions with several GAGs (Chabrol *et al.*, 2012; Zhao *et al.*, 2016). Notably, sensor chips

typically utilize dextran matrices that have been observed to interact with CLRs including Langerin, potentially introducing artifacts for the determination of IC_{50} or K_D values (Andreini *et al.*, 2011).

The presented examples highlight the utility of heterogenous and multivalent assays for glycomimetic ligand design. However, these methods introduce complex multi-state equilibria and surface phenomena, complicating data analysis as well as critically limiting the transferability of determined affinities between assay formats (Kitov and Bundle, 2003). Alternatively, sensitive biophysical techniques enable the development of homogeneous and monovalent assays. Prominently, ITC experiments can be employed to determine the thermodynamic profile of GPIs and have proven instrumental *e.g.* for the discovery of glycomimetic inhibitors of E-Selectin or the bacterial adhesin FimH (Binder *et al.*, 2012; Pang *et al.*, 2012). For Langerin, ITC served to determine K_D values for several mono- and oligosaccharides (Holla and Skerra, 2011). As these experiments display non-optimal throughput and material consumption, they are of limited utility during the initial stages of glycomimetic ligand design. By contrast, fluorescent polarization (FP)-based competitive binding assays provide high throughput and require considerably less material. These characteristics have rendered FP an attractive approach for the design of glycomimetic FimH inhibitors (Han *et al.*, 2010). However, suitable reporter molecules displaying nano- to micromolar affinities are required for the assay development.

Additionally, several ligand- and protein-observed NMR techniques can be applied to characterize GPIs and are readily transferrable to glycomimetic ligand design (Angulo and Nieto, 2011; Dalvit, 2007; Williamson, 2013). The integration of these techniques to establish both competitive binding experiments and the determination of K_D values is ideally suited to exclude false-positive mechanisms such as the aggregation of glycomimetics or the denaturation of CLRs (Gossert and Jahnke, 2016). Additionally, NMR experiments enable the elucidation of the conformation of oligosaccharides and represent an attractive approach to analyze the binding mode of glycomimetics. In the following section, important concepts for the implementation of ^{19}F R_2-filtered NMR experiments are outlined. Notably, this technique constitutes an integral part of the design strategy for glycomimetic Langerin ligands presented in this dissertation. Finally, the chapter concludes with a brief introduction to other relevant NMR techniques focusing on STD and ^{15}N HSQC NMR experiments.

^{19}F R_2-filtered NMR Competitive Binding Assays

Transversal relaxation rates R_2 obtained from ligand-observed NMR experiments are ideally suited to characterize weak protein-ligand interactions (Fielding, 2007). The determination of K_D values in the micro- to millimolar range is well-established and R_2-filtered NMR experiments are routinely utilized to screen fragment libraries (Hajduk *et al.*, 1997; Verdier *et al.*, 2000). Modern NMR techniques are based on the excitation of nuclear spins *via* radiofrequency pulses B_1 to generate phase coherence and detectable transversal magnetization M_{xy}. In this context, relaxation describes the decay of the

detectable magnetization during the relaxation time T (Keeler, 2011). Generally, two main relaxation processes are distinguished: Longitudinal R_1 and transversal R_2 relaxation. R_1 relaxation is defined as the return of the excited system to its thermal equilibrium and the regeneration of longitudinal magnetization M_z along the static magnetic field B_0.

Transversal relaxation rates R_2 can be determined *via* the Carr-Purcell-Meiboom-Gill (CPMG) pulse sequence (Carr H. Y. and M., 1954; Meiboom and Gill, 1958) (Figure 4). Following an initial $90°_x$ pulse, a train of $180°_y$ pulses is applied to refocus the magnetization vector along the y-axis with respect to field inhomogeneities and chemical shift differences $\Delta\delta$. The number of $180°_y$ pulses n and the interpulse delay τ determine the relaxation time T. During this time, thermal molecular motions contribute to the degeneration of phase coherence and enhance the decay of detectable magnetization for transversal relaxation. Integrals I obtained from NMR spectra acquired at different T values can be fitted to Equation 1 to calculate observed R_2 values.

$$I = I_0 e^{-R_{2,obs}T}$$

Equation 1

Thermal molecular motions can be described by the correlation time τ_c which is proportional to the molecular weight. These motions generate local magnetic fields B_2 that induce the relaxation processes outlined above (Keeler, 2011). The time dependency of these fields is characterized by the correlation function which depends on τ_c. Its Fourier transform, the spectral density function reveals the molecular motions present at the frequency ω. The analysis of these functions reveals two important characteristics of the transverse relaxation rate R_2. First, R_2 values are proportional to τ_c, *i.e.* the molecular weight. Secondly, R_2 depends on the spectral density at both the Larmor frequency ω_0 and at an ω value of 0 Hz. For nuclei displaying spin numbers s of 1/2 such as ^{1}H and ^{19}F, relaxation processes are induced by different mechanisms including dipolar interactions, chemical shift anisotropy (CSA) and spin rotation.

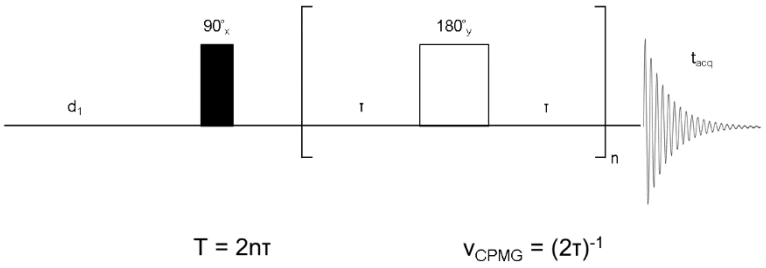


Figure 4. CPMG pulse sequence for the determination of transversal relaxation rates R_2.

An initial 90°$_x$ pulse generates coherent transversal magnetization M_{xy} along the y-axis. Subsequently, the magnetization vector is refocused over the relaxation time T utilizing a train of 180°$_y$ pulses. The value of T is determined by the number of 180°$_y$ pulses and the interpulse delay τ. The frequency these pulses ν_{CPMG} is given by $(2\tau)^{-1}$.

For protein-ligand interaction equilibria, ligands display distinct transversal relaxation rates $R_{2,f}$ and $R_{2,b}$ corresponding to their free and their bound state, respectively. Due the long correlation time τ_c of protein-ligand complexes and potential contributions from dipolar interactions or the CSA, $R_{2,b}$ values are increased compared to $R_{2,f}$ (Fielding, 2007). In the fast or fast-to-intermediate exchange regime, ligand resonances display population-averaged transversal relaxation rates R_{2obs} that are determined by the bound fraction of the ligand p_b. In practice, increased $R_{2,obs}$ values are observed in the presence of protein while $R_{2,f}$ values can be obtained from independent experiments. $R_{2,obs}$ additionally depends on the chemical exchange contribution $R_{2,ex}$ and its dependency on p_b is described by Equation 2 (Dalvit, 2007).

$$R_{2,obs} = p_f R_{2,f} + p_b R_{2,b} + R_{2,ex}$$

with

$$R_{2,ex} = \frac{4\pi^2 p_b p_f^2 \Delta\delta^2}{k_{off}} \left(1 - \frac{p_f k_{off} tanh\left(\frac{k_{off}\tau}{p_f}\right)}{\tau} \right)$$

Equation 2

Importantly, Equation 2 enables the determination of K_D values from titrations experiments. Ligands characterized in these experiments can subsequently be utilized as reporter molecules for the development of competitive binding assays (Dalvit, 2007; Roehrl *et al.*, 2004). Suitable reporter molecules should display high $R_{2,b}$ values that translate into an optimal dynamic range for the determination of K_I values. However, the quantification of affinities is based on analytical expressions only valid in the absence of chemical exchange contributions $R_{2,ex}$. In addition to p_b, $R_{2,ex}$ depends on the chemical shift difference $\Delta\delta$ between the free and the bound state, the dissociation rate k_{off} as well as the interpulse delay τ of the CPMG pulse sequence. This has two important implications for the development of competitive binding assays. First, suitable reporter molecules should display fast

dissociation rates compared to Δδ. Notably, this requirement is typically satisfied for affinities in the micro- to millimolar range and thus compatible with glycomimetic ligand design. Secondly, the chemical exchange contribution $R_{2,ex}$ can be minimized at short interpulse delays. By contrast, fragment screening strategies leverage high τ values to increase the sensitivity of R_2-filtered NMR experiments (Spiliotopoulos *et al.*, 2017). Moreover, the systematic variation of τ in relaxation dispersion experiments can be utilized to obtain k_{off} values for protein-ligand interactions (Moschen *et al.*, 2016).

The concepts outlined above are generally applicable to nuclei displaying spin number s of 1/2. In this context, ^{19}F R_2-filtered NMR provides considerable advantages over ^{1}H-detected experiments, particularly for the development of competitive binding assays (Dalvit, 2007). The high natural abundance and gyromagnetic ratio γ_n of the ^{19}F isotope result in an intrinsic sensitivity of approximately 83% compared to the detection of the ^{1}H isotope. In practice, the sensitivity of detection is superior because ^{19}F NMR experiments are not compromised by baseline distortions and the presence of background signals from solvents, additives and buffers. As no deuterated solvents are required, optimal buffer conditions for protein stability can be utilized. Moreover, the absence of spectral overlap with the protein or the competitor facilitates data analysis. Finally, the transversal relaxation rate R_2 of ^{19}F nuclei represents a highly sensitive parameter for the detection of protein-ligand interactions (Dalvit, 2007). On the one hand, the high CSA and ^{1}H-^{19}F dipolar interaction contributions to the spectral density at an ω value of 0 Hz translate into substantially increased $R_{2,b}$ values. On the other hand, ^{19}F nuclei display a large chemical shift range potentially resulting in a considerable chemical exchange contribution $R_{2,ex}$. While the latter property is not necessarily desirable for competitive binding assays, it increases the sensitivity for ^{19}F R_2-filtered NMR experiments for *e.g.* fragment screening.

The development of ^{19}F R_2-filtered NMR competitive binding assays for glycomimetic ligand design requires the synthesis of ^{19}F-labeled carbohydrate analogs. Here, various fluorination strategies for mono- and oligosaccharides have been established (Dumitrescu *et al.*, 2014; N'Go *et al.*, 2014; Unione *et al.*, 2017). Importantly, J coupling potentially introduces spectral artifacts compromising the determination of $R_{2,obs}$ values *via* the CPMG pulse sequence (Segawa and Bodenhausen, 2013). These artifacts can either be avoided by the utilization of more elaborate pulse sequences or the introduction of suitable trifluoromethyl groups such as trifluoroacetamido or trifluoromethylphenyl substituents. The absence of J coupling additionally translates into improved signal-to-noise (SNR) ratios for the detection of the reporter molecule in ^{19}F R_2-filtered NMR experiments, particularly in comparison to monofluorinated ligands (Dalvit *et al.*, 2005). Notably, ^{19}F NMR experiments have been extensively explored to characterize GPIs or to screen fragment libraries (Aretz *et al.*, 2014; Matei *et al.*, 2013; Ribeiro *et al.*, 2015). Yet, no ^{19}F R_2-filtered NMR competitive binding assays have been reported for CLRs.

Other NMR Techniques –^{15}N HSQC and STD NMR

Other NMR techniques ideally suited to guide glycomimetic ligand design include ^{15}N HSQC and STD NMR experiments (Angulo and Nieto, 2011; Williamson, 2013). Both techniques enable the determination of K_D values for GPIs and provide structural information on the recognition process (Munoz-Garcia *et al.*, 2015; Pederson *et al.*, 2014). Prominently, ^{15}N HSQC NMR experiments served to analyze the binding mode of non-carbohydrate inhibitors for DC-SIGN which led to the identification of secondary binding pockets (Aretz *et al.*, 2017). The technique was furthermore employed to characterize the recognition of carbohydrate analogs targeting *e.g.* the bacterial adhesin FimH or DC-SIGN (Fiege *et al.*, 2015; Prost *et al.*, 2012). The utility of STD NMR experiments for glycomimetic ligand design has been demonstrated for several CLRs including E-Selectin and DC-SIGN (Rinnbauer *et al.*, 2003; Thépaut *et al.*, 2013; Timpano *et al.*, 2008). Moreover, both techniques have been extensively utilized to screen fragment libraries (Meyer and Peters, 2003; Rademacher *et al.*, 2010; Shuker *et al.*, 1996).

In contrast to the ^{19}F R_2-filtered NMR experiments described above, ^{15}N HSQC NMR experiments involve the detection of protein resonances and require ^{15}N isotope-labeled samples (Williamson, 2013). The corresponding pulse sequence is based on the initial generation of coherent transversal magnetization M_{xy} for ^{1}H nuclei, followed by the INEPT magnetization transfer to ^{15}N nuclei (Bodenhausen and Ruben, 1980). After an incremented evolution time t_1, the magnetization is transferred back to ^{1}H nuclei. Consequently, the two-dimensional ^{15}N HSQC NMR spectra display correlated resonances corresponding to backbone amides and the side chains of residues bearing N-H bonds. The distribution of these resonances provides information on the integrity of the protein fold and thus represents a valuable quality control for glycomimetic ligand design. Due to the correlation between the molecular weight and the transversal relaxation rate $R_{2,}$ ^{15}N HSQC NMR experiments are limited to proteins smaller than approximately 25 kDa. Modern techniques including the deuteration of proteins and transversal relaxation-optimized NMR spectroscopy have addressed this limitation (Pervushin *et al.*, 1997).

Protein-ligand interaction equilibria potentially induce chemical shift perturbations (CSPs), line broadening $\Delta\nu_{0.5}$ or reduced integrals ΔV for potein resonances, depending on the chemical exchange regime (Williamson, 2013). CSPs observed in titration experiments are the result of changes in the chemical environment, either induced directly by the formation of protein-ligand interactions or indirectly due to the modulation of conformational dynamics. CSP values are typically calculated from empirically weighted Euclidean distances to protein resonances corresponding to the free state. For fast chemical exchange, these values can be utilized for the robust determination of K_D values (Williamson, 2013). Resonances in the fast-to-intermediate or intermediate-to-slow exchange regime are not suitable for the quantification of affinities. By contrast, slow chemical exchange enables the calculation of the bound fraction of the protein p_b from the integrals V_f and V_b corresponding to the

free and the bound state. This approach represents a robust alternative for the determination of K_D values (Fielding, 2007).

Furthermore, the assignment of ^{15}N HSQC NMR resonances to individual protein residues enables the mapping of CSP or ΔV values on available X-ray structures to visualize the three-dimensional pattern of induced perturbation. The structural information obtained from this approach can be leveraged to identify secondary binding pockets or to elucidate the interaction formed by glycomimetic ligands. Notably, the resonances assignment for the CRD of Langerin has recently been reported (Hanske *et al.*, 2016). CSP trajectories obtained from titration experiments, *i.e.* the vectors of induced perturbations, provide additional structural information as they depend on the chemical environment of the bound state (Chi *et al.*, 2015; Selvaratnam *et al.*, 2012) . Hence, the comparison of CSP trajectory for different glycomimetics potentially facilitates the analysis of binding modes.

Ligand-observed STD NMR experiments are based on the NOE to detect protein-ligand interactions (Mayer and Meyer, 1999). The corresponding pulse sequence involves the selective saturation of protein resonances during the saturation time τ_{sat}. The transferred magnetization is rapidly distributed throughout the protein *via* spin diffusion, *i.e.* intramolecular cross-relaxation. Upon ligand recognition, intermolecular cross-relaxation induces the distance-dependent saturation of ligand resonances. Subsequently, a $90°_x$ pulse is applied to generate coherent transversal magnetization M_{xy} for non-saturated ^1H nuclei. Fast chemical exchange results in the saturation of a detectable fraction of ligands during $\tau_{sat.}$ Consequently, difference spectra from on- and off-resonance saturation experiments display increased integrals I for ligand resonances corresponding to ^1H nuclei located in proximity of the protein surface.

STD effects can be quantified *via* the kinetic analysis of the saturation transfer at different saturation times. The calculation of initial transfer rates STD_0' from STD build-up curves accounts for R_1 relaxation and enables the elucidation of distance-dependent binding epitopes (Mayer and Meyer, 2001). Additionally, initial transfer rates obtained from titration experiments can be utilized to determine K_D values (Angulo *et al.*, 2010). As cross-relaxation rates σ correlate with τ_c, *i.e.* the molecular weight, STD effects are more pronounced for larger proteins. Moreover, the requirement for fast chemical exchange facilitates the detection of weak interactions. Hence, STD NMR experiments are ideally suited to characterize GPIs. Notably, more sophisticated approaches utilizing differential equations to describe the saturation transfer have been established (Jayalakshmi and Rama Krishna, 2004; Marcelo *et al.*, 2014). These approaches enable the determination of dissociation rates k_{off} or improve the spatial resolution of binding epitopes.

1.6. Aims and Objectives

The central aim of the work presented in this dissertation was the design of glycomimetic Langerin ligands. Langerin is an endocytic CLR predominantly expressed on LCs residing in the human skin. This DC subset has been recognized for its capacity to endocytose and cross-present exogeneous TAAs to efficiently activate cytotoxic T cells. Consequently, LCs and Langerin represent attractive targets for the development of novel cancer immunotherapies. The designed glycomimetics were envisioned to specifically target liposomes to LCs, enabling the efficient delivery of TAAs in the future. Six central objectives guided the challenging design of targeting ligands for Langerin.

First, the structure-based *in silico* analysis of 21 X-ray structures served to identify secondary binding pockets and to systematically evaluate the druggability of CLRs (Chapter 3.1). Secondly, ^{19}F R_2-filtered NMR experiments were explored to develop an efficient and robust competitive binding assay for Langerin (Chapter 3.2). This was imperative as available assays displayed considerable limitations with respect to throughput, material consumption and the robustness of affinity determination. The third objective was the discovery of potent monosaccharide analogs. To this end, the identification of suitable monosaccharide scaffolds was required. Utilizing these scaffolds, structure-based *in silico* design strategies or the screening of existing glycomimetic libraries facilitated the exploration of Langerin's carbohydrate binding site for favorable secondary interactions (Chapter 3.3 and 3.4). Fourthly, the binding mode of potent monosaccharide analogs was analyzed *via* ^{15}N HSQC and STD NMR experiments as well as molecular docking studies. This enabled the development of conjugation strategies for liposomal formulations (Chapter 3.3 and 3.4). The fifth objective represented the validation of the specificity of designed targeting ligands against other GBPs as well as the evaluation of the capacity of targeted liposomes to induce Langerin-mediated endocytosis by LCs (Chapter 3.4). Finally, the multivalent organization of monosaccharides and their synthetic analogs was explored to optimize the potency and specificity of the designed glycomimetics (Chapter 3.5).

2. Materials and Methods

2.1. Molecular Modelling and Bioinformatics

2.1.1. General Remarks

Unless stated otherwise, molecular modelling procedures were performed in MOE (Chemical Computing Group, 2016). Deviations from default options and parameters are noted. The AMBER10:EHT force field was selected for the refinement of docking poses and the hydrogen bond network while the MMFF94x force field was utilized for the generation conformers (Case *et al.*, 2008; Gerber and Müller, 1995; Halgren, 1996). Databases were processed in KNIME and tautomers were enumerated with ChemAxon's Calculator Plugin (Berthold *et al.*, 2007; ChemAxon, 2014). Receptor surfaces were visualized in Connolly representation (Connolly, 1993).

2.1.2. Structure-Based Multiple Sequence Alignment and Consensus Structure Analysis

The scope of structural data on human CLRs was assessed using the protein family (Pfam) database (accession code: PF00059) (Punta *et al.*, 2012). Natural killer (NK) cell lectin-like receptors were treated as a closely related, yet physiologically distinct subfamily according to the Pfam annotation and were not included in the analysis. Furthermore, CLRs crystallized as a domain swap dimer, namely blood BDCA-2 and MR, were omitted (Feinberg *et al.*, 2000; Nagae *et al.*, 2014). Murine Dectin-1 was included in the selection as it displays a unique Ca^{2+}-independent carbohydrate-binding mode and no structural information of the human orthologue is available (Brown *et al.*, 2007). All structures considered for the analysis are listed (Table 1). If available, a structure in complex with a carbohydrate ligand was selected. Prior to the calculations all structures were trimmed down to the respective CRD domain as inferred from the Pfam domain definition. A structure-based multiple sequence alignment was performed in MOE. RMSD values were determined for all pairs of C_α atoms unless a gap was found in one of the compared sequences. Next, a phylogenetic analysis based on the pairwise sequence similarities was conducted in R (Paradis *et al.*, 2004; R Core Team, 2014) Hierarchical clustering was performed based on the Manhattan metric and *via* the complete linkage criterion. To complement the phylogenetic analysis, MOE was used to predict a consensus structure of all CRDs. During model construction, up to 20 gaps and RMSD values of C_α up to 10 Å were allowed for a single position in the multiple sequence alignment.

Table 1. List of analyzed CLR X-ray structures.

Receptor	PDB code	Domain	Oligomerization	Reference
ASGPR	1DV8	CRD	Monomer	(Meier *et al.*, 2000)
CD23a	2H2T	CRD	Monomer	(Wurzburg *et al.*, 2006)
Clec1b	2C6U	CRD	Monomer	(Watson *et al.*, 2007)
Clec5a	2YHF	CRD	Monomer	(Watson *et al.*, 2011)
Clec9a	3VPP	CRD	Monomer	(Zhang *et al.*, 2012)
DC-SIGN	2XR6	CRD	Monomer	Unpublished
DC-SIGNR	1K9J	CRD	Monomer	(Feinberg *et al.*, 2001)
mDectin-1	2CL8	CRD	Monomer	(Brown *et al.*, 2007)
E-Selectin	1ESL	CRD-EGF	Monomer	(Graves *et al.*, 1994)
EMBP	1H8U	CRD	Monomer	(Swaminathan *et al.*, 2001)
Langerin	3P5F	CRD	Monomer	(Feinberg *et al.*, 2011)
	3KQG	ECD	Trimer	(Feinberg *et al.*, 2010)
L-Selectin	3CFW	CRD-EGF	Monomer	Unpublished
LOX-1	1YPO	ECD	Dimer	(Park *et al.*, 2005)
MBP-C	1HUP	ECD	Trimer	(Sheriff *et al.*, 1994)
MCL	3WHD	ECD	Monomer	(Furukawa *et al.*, 2013)
Mincle	3WH3	ECD	Monomer	(Furukawa *et al.*, 2013)
P-Selectin	1G1S	CRD-EGF	Monomer	(Somers *et al.*, 2000)
Reg3a	1UV0	CRD	Monomer	(Abergel *et al.*, 1999)
Reg1a	1QDD	CRD	Monomer	(Gerbaud *et al.*, 2000)
Collectin-12	2OX8	CRD	Monomer	(Feinberg *et al.*, 2000)
SP-D	3IKN	ECD	Trimer	(Shrive *et al.*, 2009)
Tetranectin	1HTN	CRD	Monomer	(Nielsen *et al.*, 1997)
	1TN3	ECD	Trimer	(Kastrup *et al.*, 1998)

2.1.3. Binding Pocket Identification and Druggability Analysis

Initially, CLR structures were superposed in MOE. For the superposition procedure and the subsequent druggability analysis, physiologically relevant oligomerization states were selected (Table 1). The EGF domains of Selectins structures were removed. The resulting files served as input data for binding pocket prediction with DoGSite (Volkamer *et al.*, 2010). The predicted binding pockets were mapped on the structure and classified into four categories following the reported nomenclature of secondary structure elements and Ca^{2+} binding pockets: I - Ca^{2+}-2-binding pockets, II - Ca^{2+}-associated binding pockets in the long loop region, III - Ca^{2+}-independent carbohydrate binding pockets, and IV - other binding pockets (Zelensky and Gready, 2005). A binding pocket was assigned to category I if the Ca^{2+}-2 ion was part of the predicted binding pocket. For category II, the criteria were less restrictive and all binding pockets with residues within a 6 Å radius of either Ca^{2+}-1, 2 or 3 were included. Binding pockets in category III are located in close proximity to experimentally determined, Ca^{2+}-independent carbohydrate binding sites. The druggability of all binding pockets was assessed with DoGSiteScorer (Volkamer *et al.*, 2012a). Finally, category I, II or III binding pockets that displayed the highest score for each receptor were selected and this selection served to determine a mean druggability score for CLRs.

2.1.4. Molecular Docking

Definition of the Pharmacophore Model and Preparation of the Langerin Complex

A structural alignment of Langerin carbohydrate binding sites in complex with different Man-type oligosaccharides was performed (PDB codes: 3P5D, 3P5E and 3P5F) (Feinberg *et al.*, 2011). Based on this visualization, a pharmacophore model was defined with features for O3, O4 and O5 of the Man scaffold. The spatial constraint on the O3 and O4 was defined by a sphere with a radius r of 0.5 Å while the position of O5 was constrained by a sphere with a radius r of 1.0 Å. Chain B of the Langerin CRD in complex with a Man-type disaccharide served as the structural basis for the performed *in silico* screening of *N*-acetylmannosamine (ManNAc) analogs **5** and the docking of Man analog **14.5** (PDB code: 3P5F). Of the two binding modes included in this model, the orientation for targeting the identified pockets in axial direction of C2 was selected. Additionally, an alternative conformation for K313 observed for the Langerin complex with Gal-6-OS was modeled and included into the analysis (Feinberg *et al.*, 2013). Overall model quality and protein geometry were evaluated in MolProbity and maintained utilized MOE's Structure Preparation (Chen *et al.*, 2010a). Next, protonation states and the hydrogen bond network of the complex were simulated with MOE's Protonate 3D followed by the removal of all solvent molecules. The definition of a pharmacophore model and the preparation of the receptor structure for the docking of GlcNS analog **41.2** was conducted in analogy to the procedure

described above utilizing Chain B of the Langerin CRD with GlcNAc (PDB code: 4N32) (Feinberg *et al.*, 2013).

For the *in silico* design of ManNAc analogs **5**, a propargyl group was modeled to the anomeric position of the mannose scaffold and the axial hydroxyl group in C2 was substituted with an acetamido group. The conformational space for the dihedral angle of the C2-N bond was explored *in situ via* Low Mode MD simulations assuming trans configuration of the acetamido group (Labute, 2010). Five energetically favorable rotamers corresponding to two different conformations of K313 were identified and served as the structural basis for the subsequent *in silico* screening of the carboxylic acid conformation database against Langerin.

Preparation of the Carboxylic Acid Conformation Database

Carboxylic acid conformations were generated from building block databases of selected manufacturers *via* MOE's Conformation Import (TimTec, TCI, Sigma Aldrich, Otava Chemicals, Life Chemicals, Focus Synthesis, Enamine, ChemDiv, ChemBridge and Asinex). Prior to this procedure, the database was processed with MOE's Wash and filtered to yield structures with a maximum of 23 heavy atoms and 6 rotatable bonds. Moreover, only molecules containing one carboxyl group and no amino, azido or alkyne group were retained. Next, this subset was filtered for reactive molecules followed by the generation of tautomers and protonation states with ChemAxon's Calculator Plugin (Oprea, 2000). Subsequently, the carboxyl group was removed and substituted by an annotated atom with MOE's Combinatorial Library that served as an annotation point for virtual conjugation to the modified Man scaffold *in situ*.

***In silico* Screening of the Conformation Database against Langerin**

The docking procedure was implemented with MOE's Combinatorial Builder. The carboxylic acids were conjugated to the modified Man scaffold by substituting the terminal methyl group *via* the virtual formation of a carbon-carbon bond. A grid-based placement method was utilized for the generation of docking poses by exploring the conformational space around this bond. The docking poses were scored *via* the proprietary London ΔG function and highly scored poses were selected for structure refinement *via* MM simulations. During refinement, the binding mode of the Man scaffold was constrained by the pharmacophore model described above. Conformational flexibility of the carbohydrate binding site was accounted for by introducing B-factor-derived tethers to side chain atoms. Refined docking poses were then filtered by the pharmacophore model, scored with the GBIV/WSA ΔG function and written into the output database (Corbeil *et al.*, 2012). Next, scores were referenced against **5.2** and calculated GBIV/WSA ΔΔG values served to determine the predicted group efficiency GE_{pred} as well as the predicted relative potency (Verdonk and Rees, 2008). Only poses with a GE_{pred} value higher than 0.15 $kJ \cdot mol^{-1}$ and an RMSD upon refinement lower than 2 Å were retained.

Highly scored ManNAc analogs **5** were evaluated visually and a focused library was composed. The composition of the library was guided by an attempt to maximize the diversity of pharmacophoric features and to ensure synthetic feasibility. Importantly, this first generation of analogs was selected to test basic binding hypotheses and to establish a structure activity relationship in axial direction of the C2.

Molecular Docking of Mannose and Glucosamine-2-Sulfate Analogs to Langerin

Conformations for Man analog **14.5** and GlcNS analog **41.2** were generated utilizing MOE's Conformation Import. A pharmacophore-based placement method was utilized to generate docking poses. Highly scored poses were refined utilizing MM simulations, rescored *via* the GBIV/WSA ΔG function, filtered *via* the pharmacophore model and written into the output database (Corbeil *et al.*, 2012). Conformational flexibility of the carbohydrate binding site was accounted for by introducing B-factor-derived tethers to side chain atoms. Refined docking poses were ranked according to their the GBIV/WSA ΔG score and evaluated visually.

2.2. Receptor Expression and Purification

2.2.1. General Remarks

Codon-optimized genes for the expression of Langerin and DC-SIGN in *E. coli* were purchased from GenScript and Life Technologies, respectively. All growth media or chemicals used for receptor expression and purification were purchased from Carl Roth if not stated otherwise.

2.2.2. Langerin Extracellular Domain

The truncated Langerin ECD (residues 148 to 328, forward primer: GGTGGTCATATGGCC TCGACGCTGAATGCCCAGATTCCGG, reverse primer: ACCACCAAGCTTTTATTTTTCAAAC TGCGGATG) was cloned with a C-terminal TEV cleavage site and a Strep-tag II into a pET30a expression vector (EMD Millipore) and expressed insolubly in *E. coli* BL21* (DE3) (Invitrogen).[1]

Precultures were incubated overnight in LB medium supplemented with 35 µg·ml^{-1} Kanamycin (50 ml) at 37° C and 220 rpm. The preculture was diluted to an OD_{600} of 0.1 into LB medium supplemented with 35 mg·ml^{-1} Kanamycin (500 ml). The culture was incubated at 37° C and 220 rpm and expression of the Langerin ECD was induced with 0.5 mM IPTG at an OD_{600} of 0.6 to 0.8. Cells were harvested 4 h after induction *via* centrifugation at 4000 g and 4° C for 20 min. Cell pellets were stored overnight at -20° C and subsequently resuspended in 50 mM Tris with 0.1% Triton X-100 and 10 mM $MgCl_2$ (20 ml) at pH 7.5. Catalytic quantities of lysozyme (Sigma Aldrich) was added and the sample was incubated for 3.5 h at 4° C. After the addition of catalytic quantities of DNase I (AppliChem) the sample was incubated for another 30 min at 4° C.

Inclusion bodies were harvested *via* centrifugation at 10000 g and 4° C for 10 min and washed three times with 25 mM Tris with 150 mM NaCl (20 ml) at pH 7.8. Inclusion body pellets were stored overnight at -20° C and subsequently solubilized overnight in 100 mM Tris with 6 M Gu-HCl and 1 mM DTT (20 ml) at pH 8.0 and 30° C. Following centrifugation at 15000 g and 4°C for 1.5 h, the Langerin ECD was refolded overnight *via* rapid dilution into 50 mM Tris with 0.4 M arginine, 20 mM NaCl, 0.8 mM KCl, 1 mM glutathione (AppliChem) and 0.2 mM glutathione disulfide (AppliChem) (200 ml) at pH 7.6 and 4°C. Next, the sample was dialyzed overnight against 25 mM Tris, 150 mM NaCl, 25 mM $CaCl_2$ at pH 7.8 and 4° C.

After centrifugation at 15000 g and 4° C for 2 h, the sample was purified *via* mannan-agarose (Sigma Aldrich, St. Louis, USA) affinity chromatography as previously published (Stambach and Taylor, 2003). The buffer was exchanged to 25 mM Tris with 150 mM NaCl at pH 7.8 *via* 7 kDa size-

[1] The procedure was devised and conducted by Dr. Jonas Hanske (Max Planck Institute of Colloids and Interfaces).

exclusion desalting columns (Thermo Scientific) and the concentration of Langerin ECD was determined *via* UV spectroscopy ($A_{280,\ 0.1\%}$ = 2.45) (Gasteiger *et al.*, 2005). Typical yields were in the range are of 10 $mg \cdot l^{-1}$ bacterial culture. Purity and monodispersity of Langerin ECD samples were analyzed *via* SDS PAGE and DLS, respectively (Appendix A.1).

For STD NMR experiments, Langerin ECD samples were dialyzed 5 times for at least 8 h against ultrapure H_2O. Subsequently, the H_2O was removed *via* lyophilization. The residue was stored at -80°C. Prior to conducting STD NMR experiments, the Langerin ECD was dissolved in in 25 mM Tris-d_{11} (Eurisotope) with 100% D_2O, 150 mM NaCl and 5 mM $CaCl_2$ at pH 7.8 and 25° C. The concentration of Langerin ECD was determined *via* UV spectroscopy and the monodispersity of the sample was evaluated *via* DLS (Gasteiger *et al.*, 2005).

2.2.3. Langerin Carbohydrate Recognition Domain

The Langerin CRD (residues 193 to 328, forward primer: GGTGGTCATATGGCCCAGGTGGTTAG CCAAGGCTGGAAATAC, reverse primer: ACCACCAAGCTTTTATTTTTCAAACTGCGGATG) was cloned with a C-terminal TEV cleavage site and a Strep-tag II into a pET30a expression vector (Invitrogen) and expressed insolubly in *E. coli* BL21* (DE3) (Invitrogen).[2]

Precultures were incubated overnight in M9 medium supplemented with 35 $\mu g \cdot ml^{-1}$ Kanamycin and 10 mM ^{15}N-labeled NH_4Cl (Sigma Aldrich) (50 ml) at 37° C and 220 rpm. The preculture was diluted to an OD_{600} of 0.1 into M9 medium supplemented with 35 $mg \cdot ml^{-1}$ Kanamycin and ^{15}N-labeled NH_4Cl (Sigma Aldrich) (500 ml). The culture was incubated at 37° C and 220 rpm and expression of the Langerin CRD was induced with 0.5 mM IPTG at an OD_{600} of 0.6 to 0.8. Cells were harvested 4 h after induction *via* centrifugation at 4000 g and 4° C for 20 min. Cell pellets were stored overnight at -20° C and subsequently resuspended in 50 mM Tris with 0.1% Triton X-100 and 10 mM $MgCl_2$ (20 ml) at pH 7.5. Catalytic quantities of lysozyme (Sigma Aldrich) was added and the sample was incubated for 3.5 h at 4° C. After the addition of catalytic quantities of DNase I (AppliChem) the sample was incubated for another 30 min at 4° C.

Inclusion bodies were harvested *via* centrifugation at 10000 g and 4° C for 10 min and washed three with 25 mM Tris with 150 mM NaCl (20 ml) at pH 7.8. Inclusion body pellets were stored overnight at -20° C and subsequently solubilized overnight in 100 mM Tris with 6 M Gu-HCl and 1 mM DTT (20 ml) at pH 8.0 and 30° C. Following centrifugation at 15000 g and 4°C for 1.5 h, the Langerin CRD was refolded overnight *via* rapid dilution into 50 mM Tris with 0.8 M arginine, 20 mM NaCl, 0.8 mM KCl, 1 mM glutathione (AppliChem) and 0.2 mM glutathione disulfide (AppliChem) (200 ml) at pH

[2] The procedure was devised and conducted by Dr. Jonas Hanske (Max Planck Institute of Colloids and Interfaces).

7.6 and 4°C. Next, the sample was dialyzed overnight against 50 mM Tris, 150 mM NaCl, 1 mM EDTA at pH 8.0 and 4° C.

After centrifugation at 15000 g and 4° C for 2 h, the sample was purified *via* StrepTactin affinity chromatography (Iba). The Langerin CRD was eluted with 50 mM Tris with 2.5 mM d-desthiobiotin, 150 mM NaCl at pH 7.5 and dialyzed against 25 mM MES with 40 mM NaCl at pH 6.0. After centrifugation at 15000 g and 4° C for 1.5 h, the buffer was exchanged to 25 mM HEPES with 150 mM NaCl at pH 7.0 *via* 7 kDa size-exclusion desalting columns (Thermo Scientific) and the concentration of Langerin CRD was determined *via* UV spectroscopy ($A_{280,\ 0.1\%}$ = 3.19) (Gasteiger *et al.*, 2005). Typical yields were in the range are of 5 $mg \cdot l^{-1}$ bacterial culture. Purity and monodispersity of Langerin CRD samples were analyzed *via* SDS PAGE and DLS, respectively (Appendix A.1).

2.2.4. DC-SIGN Carbohydrate Recognition Domain[3]

The DC-SIGN CRD (residues 205 to 404, forward primer: GCCGCCTCTAGAGGA GTAATACGACTCACTATAGGGACTAGAGAAAGAGGAGAAAACTAGATGGCTGAACGTCT GTGTCATCCGTG) was cloned with a C-terminal TEV cleavage site and a Strep-tag II into a pUC19 expression vector (Invitrogen) and expressed insolubly in *E. coli* BL21 (DE3) (Invitrogen).[4]

Precultures were incubated overnight in M9 medium supplemented with 50 $\mu g \cdot ml^{-1}$ Ampicillin and 10 mM ^{15}N-labeled NH_4Cl (Sigma Aldrich) (50 ml) at 37° C and 220 rpm. Cells were harvested *via* centrifugation at 4000 g and 4°C for 10 min. Cell pellets were resuspended with M9 medium supplemented with 50 $\mu g \cdot ml^{-1}$ Ampicillin and ^{15}N-labeled NH_4Cl (Sigma Aldrich) (50 ml). Next, the preculture was diluted to an OD_{600} of 0.1 into M9 medium supplemented with 50 $\mu g \cdot ml^{-1}$ Ampicillin and ^{15}N-labeled NH_4Cl (Sigma Aldrich) (500 ml). The culture was incubated at 37° C and 220 rpm and expression of the DC-SIGN CRD was induced with 0.5 mM IPTG at an OD_{600} of 0.6 to 0.8. Cells were harvested 4 h after induction *via* centrifugation at 4000 g and 4° C for 20 min. Cell pellets were stored overnight at -20° C and subsequently resuspended in 50 mM Tris with 0.1% Triton X-100 and 10 mM $MgCl_2$ (20 ml) at pH 7.5. Catalytic quantities of lysozyme (Sigma Aldrich) was added and the sample was incubated for 3.5 h at 4° C. After the addition of catalytic quantities of DNase I (AppliChem) the sample was incubated for another 30 min at 4° C.

Inclusion bodies were harvested *via* centrifugation at 10000 g and 4° C for 10 min and washed three with ultrapure H_2O. Inclusion body pellets were stored overnight at -20° C and subsequently solubilized overnight in 100 mM Tris with 6 M Gu-HCl and 1 mM DTT (20 ml) at pH 8.0 and 30° C. Following centrifugation at 15000 g and 4°C for 1.5 h, the DC-SIGN CRD was refolded overnight *via* rapid dilution into 100 mM Tris with 1.0 M arginine, 150 mM NaCl and 120 mM sucrose (100 ml) at

[3] The procedure was conducted by Christian Sommereisen (Freie Universität Berlin) under my supervision.
[4] The procedure was devised and conducted by Dr. Jonas Aretz (Max Planck Institute of Colloids and Interfaces).

pH 8.0 and 4°C. Next, the sample was dialyzed overnight against 50 mM Tris, 150 mM NaCl, 1 mM EDTA at pH 8.0 and 4° C.

After centrifugation at 15000 g and 4° C for 2 h, the sample was purified *via* StrepTactin affinity chromatography (Iba). The DC-SIGN CRD was eluted with 50 mM Tris with 2.5 mM d-desthiobiotin, 150 mM NaCl at pH 7.5 and dialyzed against 20 mM NaOAc with 50 mM NaCl at pH 4.5. After centrifugation at 15000 g and 4° C for 1.5 h, the buffer was exchanged to 25 mM HEPES with 150 mM NaCl at pH 7.4 *via* 7 kDa size-exclusion desalting columns (Thermo Scientific) and the concentration of DC-SIGN CRD was determined *via* UV spectroscopy ($A_{280,\ 0.1\%}$ = 2.98) (Gasteiger *et al.*, 2005). Typical yields were in the range are of 5 $mg \cdot l^{-1}$ bacterial culture. Purity and monodispersity of DC-SIGN CRD samples were analyzed *via* SDS PAGE and DLS, respectively (Appendix A.1).

2.3. Synthetic Chemistry

2.3.1. General Remarks

Reagents and solvents used were purchased from Sigma Aldrich unless indicated otherwise and used as supplied without any further purification. Anhydrous solvents were either purchased or taken from an anhydrous solvent system (JC-Meyer Solvent Systems).

Column chromatography was carried out using silica gel at a pore size from 40 to 60 Å (Machery Nagel). Reversed-phase column chromatography was carried out using Chromabond end-capped C_{18} columns at a pore size of 60 Å (Machery Nagel). MPLC and reversed-phase MPLC was conducted on a Combiflash R_f (Teledyne Isco) using RediSep and RediSep C_{18} columns (Teledyne Isco). Preparative reversed-phase HPLC was performed on a 1200 Series LC/MS using a semi-preparative HyperCarb column (Thermo Scientific).

Analytical thin layer chromatography (TLC) was performed on glass plates coated with silica gel at a pore size of 60 Å (Machery Nagel or Merck). Compounds were detected *via* 3-methoxyphenol reagent (0.2% 3-methoxyphenol in EtOH: 2 N sulfuric acid in EtOH (1:1)), ninhydrin reagent (1.5 g ninhydrin in 15 ml AcOH and 500 ml MeOH) or CAM reagent (1.0 g $Ce(SO_4)_2 \cdot 4H_2O$ and 2.5 g ammonium molybdate pentahydrate in 96 ml of H_2O and 6 ml of concentrated H_2SO4) upon heating or *via* UV adsorption ($\lambda = 254$ nm).

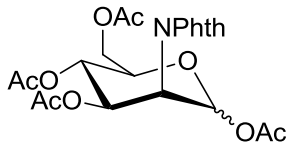
NMR experiments were conducted on a PremiumCompact 600 MHz or 400 MHz spectrometer (Agilent), an Avance III 400 MHz spectrometer (Bruker), an Avance DMX 500 MHz spectrometer (Bruker) or an Ascend 700 MHz spectrometer (Bruker) (Appendix A.2). Chemical shifts were referenced to the internal standards DMSO ($\delta(^1H) = 3.33$ ppm), $CHCl_3$ ($\delta(^1H) = 7.26$ ppm and $\delta(^{13}C) = 77.1$ ppm), H_2O ($\delta(^1H) = 7.26$ ppm), MeOH ($\delta(^1H) = 3.31$ ppm, $\delta(^{13}C) = 49.0$ ppm), TMS ($\delta(^1H) = 0.00$ ppm, $\delta(^{13}C) = 0.0$ ppm) and TFA ($\delta(^{19}F) = 76.55$ ppm). Coupling constants are reported in Hz and coupling patterns are indicated as s for singlet, d for doublets, dd for doublets of doublets, ddd for doublets of doublets of doublets, t for triplets, dt for doublets of triplets, td for triplet of doublets, q for quartets and m for multiplets. Resonances were assigned by means of COSY, TOCSY, ^{13}C HSQC and ^{13}C H2BC NMR experiments (Petersen *et al.*, 2006). Stereoselectivity at the anomeric position of the Man scaffold was analyzed by measuring $^1J_{C1,\,H1}$ coupling constants for **3**, **4** and **7** (Duus *et al.*, 2000). NMR spectra were processed in MestReNova (Mestrelab Research, 2016).

The specific optical rotation was determined using a Model 341 polarimeter (PerkinElmer). ESI-MS analysis was conducted using an 1100 Series LC/MS coupled to a G1946D ESI-Q spectrometer (Agilent) or Micromass ZQ spectrometer (Waters). HR ESI-MS analysis was conducted using a 6210

ESI-TOF spectrometer (Agilent) or an Acquity H-Class UPLC/MS coupled to a Xevo G2-S ESI-Q-TOF spectrometer (Waters). ATR-FTIR spectra were acquired using a Spectrum 100 FTIR spectrometer (PerkinElmer). Analytical reversed-phase HPLC was performed on a 1200 Series LC/MS coupled to a 6130 ESI-Q spectrometer (Agilent) using an analytical HyperCarb column (Thermo Scientific) or an analytical Atlantis T3 column (Waters) (Appendix A.3).

2.3.2. *N*-acetylmannosamine Analogs

1,3,4,6-tetra-O-acetyl-2-deoxy-2-phthalimido-D-mannopyranose



2

2 was prepared as previously published (Veselý *et al.*, 2003). Mannosamine hydrochloride (Dextra) (2.15 g, 10 mmol), phthalic anhydride (1.63 g, 11 mmol) and pyridine (2 ml, 25 mmol) were dissolved in a mixture of acetone:H_2O (1:1, 15 ml) and stirred at 50°C for 3 h. Progress of the reaction was monitored by analytical TLC (propan-1-ol:EtOAc:H_2O:25% aqueous ammonia (6:3:1:1)).

Solvents were evaporated *in vacuo* and acetic anhydride (14.1 ml, 150 mmol) and pyridine (40 ml) were added to the residue. The mixture was heated to 50°C and stirred for 5 h. Progress of the reaction was monitored by analytical TLC (toluene:EtOAc (4:1)). Solvents were evaporated *in vacuo* and the residue was taken up in chloroform (250 ml). The organic phase was extracted with 1 M HCl, saturated $NaHCO_3$ and H_2O. Subsequently, the organic phase was dried with $MgSO_4$. Solvents were evaporated *in vacuo* and the residue was purified *via* column chromatography (toluene:EtOAc (8:1)) to afford an α/β-anomer mixture of **2** (3.10 g, 6.50 mmol, 65 %) as a white solid.

^{1}H NMR (400.0 MHz, $CDCl_3$, α-anomer): δ = 7.92 - 7.75 ppm, m, 4 H (aromatic H of NPhth); δ = 6.59 ppm, d, 1 H (H1); δ = 5.55 ppm, dd, 1 H, J = 6.8, 8.2 Hz (H4); δ = 5.50 ppm, dd, 1 H, J = 5.3, 6.8 Hz (H3); δ = 4.89, dd, 1H, J = 3.9, 5.3 Hz (H2); δ = 4.45 ppm, dd, 1 H, J = 6.0, 12.2 Hz, (H6a); δ = 4.31, dd, 1 H, J = 3.3, 12.2 ppm (H6b); δ = 4.23 ppm, ddd, 1 H, J = 3.3, 6.0, 8.3 Hz, (H5);δ = 2.15 ppm, s, 3 H ($OCOCH_3$); δ = 2.13 ppm, s, 3 H ($OCOCH_3$); δ = 2.10 ppm, s, 3 H ($OCOCH_3$); δ = 1.96 ppm, s, 3 H ($OCOCH_3$).

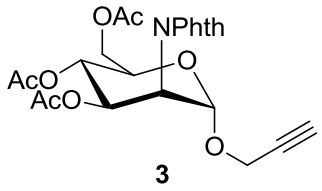
^{13}C NMR (100.6 MHz, $CDCl_3$, α-anomer): δ = 170.8 ppm, 1C ($OCOCH_3$); δ = 169.9, 1C ($OCOCH_3$); δ = 169.5 ppm ($OCOCH_3$); δ = 168.5 ppm, 1C ($OCOCH_3$); δ = 167.6 ppm, 2 C (carbonyl C of Phth); δ = 134.5 ppm, 2 C (aromatic C of Phth); δ = 131.3 ppm, 2 C (aromatic C of Phth); δ = 123.8 ppm, 2 C (aromatic C of Phth); δ = 90.4 ppm, 1 C (C1); δ = 71.1 ppm, 1 C (C5); δ = 69.0 ppm, 1 C (C3); δ = 67.8 ppm, 1 C (C4); δ = 62.5 ppm, 1 C (C6); δ = 50.5, 1 C (C2); δ = 21.0 ppm, 1 C ($OCOCH_3$); δ = 20.8 ppm, 2 C (two times $OCOCH_3$); δ = 20.7 ppm, 1 C ($OCOCH_3$).

^{1}H NMR (400.0 MHz, $CDCl_3$, β-anomer): δ = 7.92 - 7.75 ppm, m, 4 H (aromatic H of NPhth); δ = 6.06 ppm, t, 1 H, J = 9.5, 9.5 Hz (H4); δ = 5.99 ppm, d, 1 H, J = 2.8 Hz (H1); δ = 5.37 ppm, dd, 1 H, J = 6.8, 9.4 Hz (H3); δ = 5.06 ppm, dd, J = 2.8, 6.8 Hz (H2); δ = 4.48 ppm, dd, 1 H, J = 6.0 ppm, 12.2 (H6a); δ = 4.27, dd, 1 H, J = 2.3, 12.2 ppm (H6b); δ = 3.93 ppm, ddd, 1 H, J = 2.3, 6.0, 9.7 (H5); δ = 2.16 ppm, s, 3 H ($OCOCH_3$); δ = 2.07 ppm, s, 3 H ($OCOCH_3$); δ = 1.97 ppm, s, 3 H ($OCOCH_3$); δ = 1.91 ppm, s, 3 H ($OCOCH_3$).

^{13}C NMR (100.6 MHz, $CDCl_3$, β-anomer): δ = 170.7 ppm, 1 C ($OCOCH_3$); δ = 169.7, 1 C ($OCOCH_3$); δ = 169.5 ppm, 1 C (OCOCH3); δ = 168.4 ppm, 1 C ($OCOCH_3$); δ = 168.0 ppm, 2 C (carbonyl C of Phth); δ = 134.3 ppm, 2 C (aromatic C of Phth); δ = 131.2 ppm, 2 C (aromatic C of Phth); δ = 123.6 ppm, 2 C (aromatic C of Phth); δ = 90.3 ppm, 1 C (C1); δ = 73.9 ppm, 1 C (C5); δ = 69.8 ppm, 1 C (C3); δ = 66.7 ppm, 1 C (C4); δ = 62.5 ppm, 1 C (C6); δ = 49.9, 1 C (C2); δ = 20.7 ppm, 2 C (two times $OCOCH_3$); δ = 20.5 ppm, 1 C ($OCOCH_3$); δ = 20.4 ppm, 1 C ($OCOCH_3$).

R_f = 0.35 with toluene:EtOAc (2:1).

Propargyl-3,4,6-tri-O-acetyl-2-deoxy-2-phthalimido-α-D-mannopyranoside



3

2 (1.40 g, 2.9 mmol) was dissolved in anhydrous DCM:ether (29 ml, 2:1). The mixture was stirred at 0°C and kept under argon. Propargyl alcohol (700 µl, 11.7 mmol) and $BF_3{\cdot}OEt_2$ (740 µl, 5.9 mmol) were added, the reaction was allowed to heat up to room temperature and stirred for 64 h. Progress of the reaction was monitored by analytical TLC (toluene:EtOAc (2:1)). The mixture was diluted in DCM (200 ml) and the organic phase was extracted with saturated $NaHCO_3$ and H_2O. Subsequently, the organic phase was dried with $MgSO_4$. Solvents were evaporated *in vacuo* and the residue was purified *via* column chromatography (toluene:EtOAc (8:1)) to yield **3** (752 mg, 1.57 mmol, 54%) as a light yellow resin. A stereoselectivity of 10 to 1 favoring the α-anomer was determined *via* ^{1}H-NMR experiments of the crude mixture. Starting material **2** that was not converted was recovered.

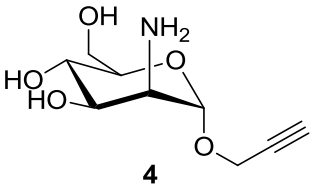
^{1}H NMR (400.0 MHz, $CDCl_3$): δ = 7.88 - 7.72 ppm, m, 4 H (aromatic H of NPhth); δ = 5.58 ppm, d, 1H, J = 4.1 Hz (H1); δ = 5.48 ppm, dd, 1 H, J = 6.5, 8.0 Hz (H4); δ = 5.45 ppm, dd, 1 H, J = 4.9, 6.4 Hz (H3); δ = 4.87 ppm, dd, 1 H, J = 4.0 Hz, 4.9 Hz (H2); δ = 4.46 ppm, dd, 1 H, J = 6.1, 12.1 Hz (H6a); δ = 4.28 ppm, dd, 2 H, J = 0.5, 2.4 Hz (OC*H*$_2$CCH); δ = 4.27 ppm, dd, 1 H, J = 3.0, 12.1 Hz (H6b); δ = 4.17 ppm, ddd, 1 H, J = 3.0, 6.2, 8,0 Hz (H5); δ = 2.41, t, 1 H, J = 2.4 Hz (OCH$_2$CC*H*); δ = 2.16 ppm, s, 3 H (OCOC*H*$_3$); δ = 2.09 ppm, s, 3 H (OCOC*H*$_3$); δ = 1.95 ppm, s, 3 H (OCOC*H*$_3$).

^{13}C NMR (100.6 MHz, $CDCl_3$): δ = 170.9 ppm, 1 C (O*C*OCH$_3$); δ = 170.0 ppm, 1 C (O*C*OCH$_3$); δ = 169.8 ppm, 1 C (O*C*OCH$_3$); δ = 168.0 ppm, 2 C (carbonyl C of NPhth); δ = 134.5 ppm, 2 C (aromatic C of NPhth); δ = 131.5 ppm, 2 C (aromatic C of NPhth); δ = 123.7 ppm, 2 C (aromatic C of NPhth); δ = 95.5 ppm, 1 C (C1); δ = 78.5 ppm, 1 C (OCH$_2$*C*CH); δ = 75.4 ppm, 1 C (OCH$_2$C*C*H); δ = 69.6 ppm, 1 C (C5); δ = 69.5 ppm, 1 C (C3); δ = 68.3 ppm, 1 C (C4); δ = 62.8 ppm, 1 C (C6); δ = 54.9 ppm, 1 C (O*C*H$_2$CCH); δ = 51.6 ppm, 1 C (C2); δ = 21.0 ppm, 2 C (two times OCO*C*H$_3$); δ = 20.9 ppm, 1 C (OCO*C*H$_3$).

R_f = 0.47 with toluene:EtOAc (2:1).

ESI-MS for $C_{23}H_{23}NO_{10}$: $m{\cdot}z^{-1}(M+Na^+)_{calc}$ = 496.1; $m{\cdot}z^{-1}(M+Na^+)_{obs}$ = 496.0; $m{\cdot}z^{-1}(M+NH_4^+)_{calc}$ = 491.2; $m{\cdot}z^{-1}$ $(M+NH_4^+)_{obs}$ = 491.2.

Propargyl-2-deoxy-2-amino-α-D-mannopyranoside



4

3 (2.70 g, 5.6 mmol) was dissolved in EtOH containing 33% $MeNH_2$ (110 ml) at room temperature. The mixture was stirred overnight and progress of the reaction was monitored by analytical TLC (20% MeOH in DCM). Solvents were evaporated *in vacuo* and the residue was purified *via* column chromatography (gradient: hexane, hexane:DCM (1:1), DCM, 1% MeOH in DCM, 5% MeOH in DCM and elution with 20% MeOH in DCM). Silica gel particles were removed by filtration in MeOH with a cellulose acetate membrane at a pore size 0.2 µm to yield **4** (950 mg, 4.40 mmol, 78%) as a white solid.

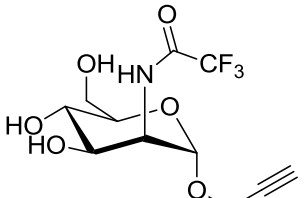
^{1}H NMR (400.0 MHz, MeOD): δ = 4.95 ppm, d, 1 H, J = 1.3 Hz (H1); δ = 4.26 ppm, d, 2 H, J = 2.4 Hz (OC*H*$_2$CCH); δ = 3.79 ppm, dd, 1 H, J = 2.4, 11.9 Hz (H6a); δ = 3.76 ppm, dd, 1 H, J = 4.5, 9.3 Hz (H3); δ= 3.73 ppm, dd, 1 H, J = 4.7, 11.8 Hz (H6b); δ = 3.57 ppm, m, 1 H (H4); δ = 3.50 ppm, ddd, 1 H, J = 2.3, 4.8, 9.8 Hz (H5), δ = 3.05 ppm, dd, 1 H, J = 1.4, 4.3 Hz (H2); δ = 2.85 ppm, t, 1 H, J = 2.5 Hz (OCH$_2$CC*H*).

^{13}C NMR (100.6 MHz, MeOD): δ = 100.0 ppm, 1 C (C1); δ = 80.0 ppm, 1 C (OCH$_2$*C*CH); δ = 76.0 ppm, 1 C (OCH$_2$C*C*H); δ = 74.7 ppm, 1 C (C5); δ = 71.8 ppm, 1 C (C3); δ = 67.7 ppm, 1 C (C4); δ = 62.3 ppm, 1 C (C6); δ = 55.7 ppm, 1 C (C 2); δ = 54.9 ppm, 1 C (O*C*H$_2$CCH).

R_f = 0.22 with 20% MeOH in DCM.

HR ESI-MS for $C_9H_{15}NO_5$: $m{\cdot}z^{-1}(M+Na^+)_{calc}$ = 240.085; $m{\cdot}z^{-1}(M+Na^+)_{obs}$ = 240.085.

Propargyl-2-deoxy-2-2',2',2'-(trifluoro)acetamido-α-D-mannopyranoside



5.1

4 (255 mg, 1.2 mmol) was dissolved in DMF (2.4 ml) and ethyl 2,2,2-trifluoroacetate (155 µl, 1.3 mmol) was added at room temperature. Progress of the reaction was monitored by analytical TLC (10% MeOH in DCM). After 6 h, solvents were evaporated *in vacuo* and the residue was purified *via*

column chromatography (gradient: DCM and elution with ml 5% MeOH in DCM). Silica gel particles were removed by filtration in MeOH with a cellulose acetate membrane at a pore size 0.2 µm to yield **5.1** (248 mg, 790 µmol, 67%) as a white solid.

^{1}H NMR (600.0 MHz, MeOD)**:** δ = 4.94 ppm, d, 1 H, J = 0.8 Hz (H1); δ = 4.39 ppm, dd, 1 H, J = 0.8, 4.9 Hz (H2); δ = 4.27 ppm, m, 2 H (OC*H*$_2$CCH); δ = 3.96 ppm, dd, 1 H, J = 4.8, 9.7 Hz (H3); δ = 3.87 ppm, dd, 2 H, J = 4.0, 11.7 Hz (H6a); δ = 3.79 ppm, dd, 2 H, J = 2.4, 11.7 Hz (H6b); δ = 3.65 ppm, t, 1 H, J = 9.8 Hz (H4); δ = 3.57 ppm, ddd, 1 H, J = 2.2, 3.8, 9.9 Hz (H5); δ = 2.88 ppm, t, 1 H, J = 2.5 Hz (OCH$_2$CC*H*).

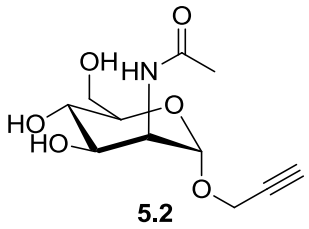
^{13}C NMR (100.6 MHz, MeOD): δ = 159.6 ppm, q, 1C, J = 37.6 Hz (NH*C*OCF$_3$); δ = 117.3 ppm, q, 1C, J = 286.6 Hz (NHCO*C*F$_3$); δ = 98.5 ppm, 1 C (C1); δ = 79.6 ppm, 1 C (OCH$_2$*C*CH); δ = 76.4 ppm, 1 C (OCH$_2$C*C*H); δ = 74.5 ppm, 1 C (C5); δ = 70.5 ppm, 1 C (C3); δ = 67.5 ppm, 1 C (C4); δ = 61.5 ppm, 1 C (C6); δ = 55.3 ppm, 1 C (O*C*H$_2$CCH); δ = 54.8 ppm, 1 C (C2).

^{19}F NMR (376.0 MHz, D$_2$O): δ = -75.8 ppm, s, 3 F (C*F*$_3$).

R_f = 0.39 with 10% in MeOH in DCM.

HR ESI-MS for $C_{11}H_{14}F_3NO_6$: $m{\cdot}z^{-1}(M+Na^+)_{calc}$ = 336.067; $m{\cdot}z^{-1}(M+Na^+)_{obs}$ = 336.070.

Propargyl-2-deoxy-2-acetamido-α-D-mannopyranoside



5.2

4 (59 mg, 270 µmol) was dissolved in MeOH (5 ml) and acetic anhydride (194 mg, 1.9 mmol) was added. The mixture was stirred at room temperature overnight. Progress of the reaction was monitored by analytical TLC (5% MeOH in DCM). Solvents were evaporated *in vacuo* and the residue was purified *via* column chromatography (gradient: hexane, hexane:DCM (1:1), DCM, 5% MeOH in DCM and elution with 20% MeOH in DCM). Silica gel particles were removed by filtration in MeOH with a cellulose acetate membrane at a pore size 0.2 µm to yield **5.2** (57 mg, 220 µmol, 81%) as a white solid.

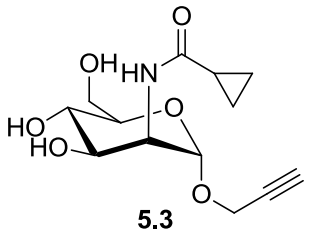
^{1}H NMR (400.0 MHz, D_2O): δ = 5.00 ppm, d, 1 H, J = 1.1 Hz (H1); δ = 4.37 ppm, dd, 1 H, J = 1.1, 4.8 Hz (H2); δ = 4.34 ppm, m, 2 H (OC*H*$_2$CCH); δ = 4.02 ppm, dd, 1 H, J = 4.8, 9.4 Hz (H3); δ = 3.88 ppm, m, 2 H (H6a/b); δ = 3.71 ppm, dt, 1 H, J = 3.5, 10.0 Hz (H 5); δ = 3.66 ppm, m, 1 H (H4); δ = 2.93 ppm, t, 1 H, J = 2.5 Hz (OCH$_2$CC*H*); δ = 2.07 ppm, s, 3 H (NHCOC*H*$_3$).

^{13}C NMR (100.6 MHz, D_2O): δ = 175.4 ppm, 1 C (NH*C*OCH$_3$); δ = 98.4 ppm, 1 C (C1); δ = 78.8 ppm, 1 C (OCH$_2$*C*CH); δ = 76.2 ppm, 1 C (OCH$_2$C*C*H); δ = 73.3 ppm, 1 C (C5); δ = 69.6 ppm, 1 C (C3); δ = 67.1 ppm, 1 C (C4); δ = 60.8 ppm, 1 C (C6); δ = 55.3 ppm, 1 C (O*C*H$_2$CCH); δ = 53.0 ppm, 1 C (C2); δ = 22.5 ppm, 1 C (NHCO*C*H$_3$).

R_f = 0.27 with 10% MeOH in DCM.

HR ESI-MS for $C_{11}H_{17}NO_6$: $m{\cdot}z^{-1}(M+Na^+)_{calc}$ = 282.095; $m{\cdot}z^{-1}(M+Na^+)_{obs}$ = 282.095.

Propargyl-2-deoxy-2-cyclopropanecarboxmido-α-D-mannopyranoside



5.3

4 (48 mg, 220 µmol) was dissolved in pyridine (5 ml) and cyclopropanecarbonyl chloride (160 µl, 1.8 mmol, Acros Organics) was added at 0°C. The mixture was allowed to heat up to room temperature and progress of the reaction was monitored by analytical TLC (3% MeOH in DCM). After 7 h EtOH containing 33% $MeNH_2$ (10 ml) was added and the mixture was stirred overnight at room temperature. Solvents were evaporated *in vacuo* and the residue was purified *via* column chromatography (gradient: hexane, hexane:DCM (1:1), DCM and elution with ml 1% MeOH in DCM) to yield **5.3** (27 mg, 95 µmol, 43%) as a yellow solid.

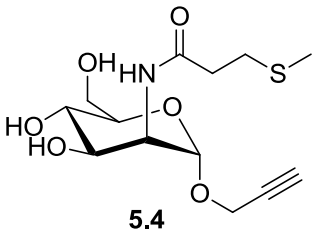
^{1}H NMR (400.0 MHz, D_2O): δ = 5.00 ppm, d, 1 H, J = 1.3 Hz (H1); δ = 4.40 ppm, dd, 1 H, J = 1.4, 4.8 Hz (H2); δ = 3.34 ppm, m, 2 H (OC*H2*CCH); δ = 4.02 ppm, m, 1 H (H3); δ = 3.89 ppm, m, 2 H (H6a/b); δ = 3.72, m, 2 H (H4, H5); δ = 2.93 ppm, t, J = 2.3 Hz (OCH_2CC*H*); δ = 1.73 ppm, m, 1 H ($NHCOC$*H*C_2H_4); δ = 0.89 ppm, m, 4 H ($NHCOCHC_2$*H*$_4$).

^{13}C NMR (100.6 MHz, D_2O): δ = 178.4 ppm, 1C ($NHCOCHC_2H_4$); δ = 98.3 ppm, 1 C (C1); δ = 79.1 ppm, 1 C (OCH_2*C*CH); δ = 76.7 ppm, 1 C (OCH_2CC*H*); δ = 73.2 ppm, 1 C (C5); δ = 69.5, 1 C (C3); δ = 67.0 ppm, 1 C (C4); δ = 60.7 ppm, 1 C (C6); δ = 55.2, 1 C (OCH_2CCH); δ = 53.0 ppm, 1 C (C2); δ = 14.4 ppm, 1 C ($NHCOCHC_2H_4$); δ = 7.7 ppm, 1 C ($NHCOCHC_2H_4$); δ = 7.6 ppm, 1 C ($NHCOCHC_2H_4$).

R_f = 0.51 with 3% MeOH in DCM

HR ESI-MS for $C_{13}H_{19}NO_6$: $m{\cdot}z^{-1}(M+Na^+)_{calc}$ = 308.111; $m{\cdot}z^{-1}(M+Na^+)_{obs}$ = 308.111.

Propargyl-2-deoxy-2-3'-(methylthio)propanamido-α-D-mannopyranoside



5.4

4 (3.03 mg, 150 µmol) was dissolved in pyridine (760 µl) and 3-(methylthio)propanoyl chloride (21 µl, 170 µmol, TCI) was added at 0°C. The reaction mixture was allowed to heat up to room temperature and after 4 h more 3-(methylthio)propanoyl chloride (21 µl, 170 µmol, TCI) was added. Progress of the reaction was monitored by analytical TLC (10% MeOH in DCM) and after 7 h EtOH containing 33% $MeNH_2$ (5 ml) was added at room temperature. The mixture was stirred overnight, solvents were evaporated *in vacuo* and the residue was purified *via* column chromatography (gradient: hexane, DCM, 1% MeOH in DCM, 5% MeOH in DCM and elution with 10% MeOH in DCM). Residual impurities were removed *via* reversed-phase column chromatography (gradient: H_2O, 1% MeOH in H_2O and elution with 5% MeOH in H_2O) to yield **5.4** (4.8 mg, 15 µmol, 10%) as a white solid.

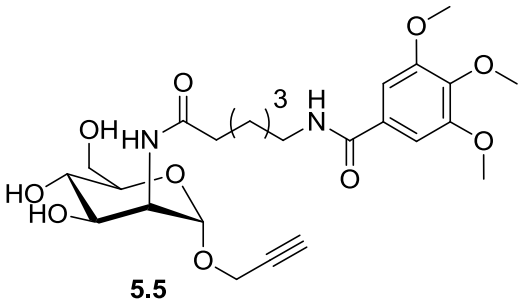
^{1}H NMR (400.0 MHz, MeOD): δ = 4.90 ppm, s, 1 H (H1); δ = 4.33 ppm, d, 1 H, J = 4.6 Hz (H2); δ = 4.26 ppm, m, 2 H (OC*H2*CCH); δ = 3.91 ppm, dd, J = 4.7, 9.1 Hz, 1 H (H3); δ = 3.81 ppm, d, 2 H, J = 3.2 Hz (H6a/b); δ = 3.59, m, 1 H (H4); δ = 3.54, dt, 1 H, J = 3.1, 10.0 Hz (H5); δ = 2.87 ppm, t, J = 2.2 Hz (OCH_2CCH); δ = 2.75 ppm, m, 2 H ($NHCOCH_2CH_2SCH_3$); δ = 2.58 ppm, m, 2 H ($NHCOCH_2CH_2SCH_3$); δ = 2.12 ppm, s, 3 H ($NHCOCH_2CH_2SCH_3$).

^{13}C NMR (100.6 MHz, MeOD): δ = 175.0 ppm, 1C ($NHCOCH_2CH_2SCH_3$); δ = 99.1 ppm, 1 C (C1); δ = 79.8 ppm, 1 C (OCH_2CCH); δ = 76.2 ppm, 1 C (OCH_2CCH); δ = 74.6 ppm, 1 C (C5); δ = 70.7, 1 C (C3); δ = 68.2 ppm, 1 C (C4); δ = 62.1 ppm, 1 C (C6); δ = 55.1, 1 C (OCH_2CCH); δ = 54.1 ppm, 1 C (C2); δ = 36.8 ppm, 1 C ($NHCOCH_2CH_2SCH_3$); δ = 30.8 ppm, 1 C ($NHCOCH_2CH_2SCH_3$); δ = 15.3 ppm, 1 C ($NHCOCH_2CH_2SCH_3$).

R_f = 0.18 with 10% MeOH in DCM.

HR ESI-MS for $C_{13}H_{21}NO_6S$: $m{\cdot}z^{-1}(M+Na^+)_{calc}$ = 342.099; $m{\cdot}z^{-1}(M+Na^+)_{obs}$ = 342.103.

Propargyl-2-deoxy-2-6'-(3'',4'',5''-(trimethoxy)benzamido)hexanamido-α-D-mannopyranoside



5.5

6-(3,4,5-Trimethoxybenzamido)hexanoic acid (90 mg, 0.28 mmol, Vitas-M Laboratory) was dissolved in DMF (700 µl). Subsequently, PyBOP (150 mg, 0.28 mmol) and DIPEA (100 µl, 0.55 mmol) were added and the mixture was stirred for 10 min at room temperature. **4** (30 mg, 0.14 mmol) was added and the mixture was stirred overnight at room temperature. Progress of the reaction was monitored by analytical TLC (10% MeOH in DCM). The reaction was quenched with MeOH (1 ml) and after addition of 1 M NaOH (830 µl, 0.83 mmol) solvents were removed *in vacuo.* The residue was purified *via* reversed-phase column chromatography (gradient: H_2O, 10 % MeOH in H_2O, 20 % MeOH in H_2O, 30 % MeOH in H_2O and elution with 40 % MeOH in H_2O) to yield **5.5** (51 mg, 97 µmol, 70 % yield) as a white solid.

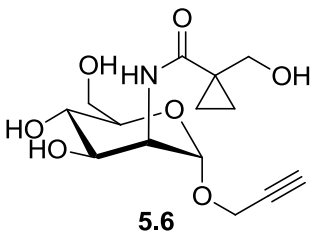
^{1}H NMR (400.0 MHz, MeOD): δ = 7.17 ppm, s, 2 H (aromatic H of 3',4',5'-trimethoxybenzamide); δ = 4.85 ppm, s, 1 H (H1); δ = 4.31 ppm, d, 1 H, J = 5.1 Hz (H2); δ = 4.23 ppm, m, 2 H (OC*H*$_2$CCH); δ = 3.91 ppm, m, 1 H (H3); δ = 3.89 ppm, s, 6 H (two times CH_3 of 3',4',5'-trimethoxybenzamide); δ = 3.81 ppm, m, 2 H, (H6a/b); δ = 3.81 ppm, s, 3 H (CH_3 of 3',4',5'-trimethoxybenzamide); δ = 3.60 ppm, m, 1 H (H4); δ = 3.53 ppm, dt, 1 H, J = 3.2, 10.1 Hz (H5); δ = 3.38 ppm, m, 2 H (NHCOCH$_2$CH$_2$CH$_2$CH$_2$C*H*$_2$NHCO); δ = 2.85 ppm, t, J = 2.4 Hz (OCH$_2$CC*H*); δ = 2.31 ppm, m, 2 H (NHCOC*H*$_2$CH$_2$CH$_2$CH$_2$CH$_2$NHCO); δ = 1.67 ppm, m, 4 H (NHCOCH$_2$C*H*$_2$CH$_2$C*H*$_2$CH$_2$NHCO); δ = 1.43 ppm, m, 2 H (NHCOCH$_2$CH$_2$C*H*$_2$CH$_2$CH$_2$NHCO).

^{13}C NMR (100.6 MHz, MeOD): δ = 175.0 ppm, 1C (NH*C*OCH$_2$CH$_2$CH$_2$CH$_2$CH$_2$NHCO); δ = 169.5 ppm, 1C (NHCOCH$_2$CH$_2$CH$_2$CH$_2$CH$_2$NH*C*O); δ = 154.4 ppm, 2C (C3' and C5' of 3',4',5'-trimethoxybenzamide); δ = 141.9 ppm, 1C (C4' of 3',4',5'-trimethoxybenzamide); δ = 131.1 ppm, 1C (C1' of 3',4',5'-trimethoxybenzamide); δ = 105.9 ppm, 2C (C2' and C6' of 3',4',5'-trimethoxybenzamide); δ = 99.2 ppm, 1 C (C1); δ = 79.80 ppm, 1 C (OCH$_2$*C*CH); δ = 76.2 ppm, 1 C (OCH$_2$C*C*H); δ = 74.6 ppm, 1 C (C5); δ = 70.69, 1 C (C3); δ = 68.2 ppm, 1 C (C4); δ = 62.1 ppm, 1 C (C6); δ = 61.1 ppm, 1 C (CH$_3$ of 3',4',5'-trimethoxybenzamide); δ = 56.7 ppm, 2 C (two times CH$_3$ of 3',4',5'-trimethoxybenzamide); δ = 55.1, 1 C (O*C*H$_2$CCH); δ = 54.0 ppm, 1 C (C2); δ = 41.0 ppm, 1 C (NHCOCH$_2$CH$_2$CH$_2$CH$_2$*C*H$_2$NHCO); δ = 36.7 ppm, 1 C (NHCO*C*H$_2$CH$_2$CH$_2$CH$_2$CH$_2$NHCO); δ = 30.2 ppm, 1 C (NHCOCH$_2$CH$_2$CH$_2$*C*H$_2$CH$_2$NHCO); δ = 27.6 ppm, 1 C (NHCOCH$_2$CH$_2$*C*H$_2$CH$_2$CH$_2$NHCO); δ = 26.6 ppm, 1 C (NHCOCH$_2$*C*H$_2$CH$_2$CH$_2$CH$_2$NHCO).

R_f = 0.23 with 10% MeOH in DCM.

HR ESI-MS for $C_{25}H_{36}N_2O_{10}$: $m{\cdot}z^{-1}(M+Na^+)_{calc}$ = 547.227; $m{\cdot}z^{-1}(M+Na^+)_{obs}$ = 547.234.

Propargyl-2-deoxy-2-1'-(hydroxymethyl)cyclopropanecarboxamide-α-D-mannopyranoside



5.6

1-hydroxycyclopropanecarboxylic acid (37 mg, 0.32 mmol, ChemBridge) was dissolved in DMF (800 µl). Subsequently, PyBOP (170 mg, 0.32 mmol) and DIPEA (120 µl, 0.65 mmol) were added and the mixture was stirred for 10 min at room temperature. **4** (35 mg, 0.16 mmol) was added, the mixture was stirred overnight at room temperature. Progress of the reaction was monitored by analytical TLC (15% MeOH in DCM). The reaction was quenched with MeOH (1 ml) and after addition of 1 M NaOH (1000 µl, 1.0 mmol) solvents were removed *in vacuo*. The residue was purified *via* preparative HPLC (gradient: H_2O for 10 min, from 0 to 20% acetonitrile in H_2O in 30 min, from 20 to 50% acetonitrile in H_2O in 10 min, from 50 to 100% acetonitrile in H_2O in 5 min and acetonitrile for 5 min at 3.2 $ml{\cdot}min^{-1}$) to yield **5.6** (16 mg, 49 µmol, 31%) as a white solid.

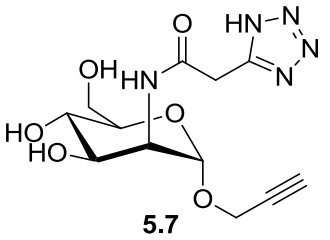
^{1}H NMR (400.0 MHz, MeOD): δ = 4.98 ppm, s, 1 H (H1); δ = 4.25 ppm, m, 2 H (OC*H2*CCH); δ = 4.23 ppm, m, 1 H (H2); δ = 3.95 ppm, dd, 1 H, J = 4.7, 9.3 Hz (H3); δ = 3.85 ppm, dd, 1 H, J = 1.9, 11.9 Hz (H6a); δ = 3.71 ppm, m, 1 H (H6b); δ = 3.64 ppm, m, 2 H (NHCOC(CH_2CH_2)CH_2OH); δ = 3.55, m, 1 H (H5); δ = 3.59, m, 1 H (H4); δ = 2.87 ppm, m, 1 H (OCH_2CC*H*); δ = 1.12 ppm, m, 2 H (NHCOC(C$H_2$$CH_2$)$CH_2$OH); δ = 0.71 ppm, m, 2 H (NHCOC(CH_2CH_2)CH_2OH).

^{13}C NMR (100.6 MHz, MeOD): δ = 177.0 ppm, 1C (NH*C*OC(CH_2CH_2)CH_2OH); δ = 98.5 ppm, 1 C (C1); δ = 79.8 ppm, 1 C (OCH_2*C*CH); δ = 76.2 ppm, 1 C (OCH_2C*C*H); δ = 74.7 ppm, 1 C (C5); δ = 70.4, 1 C (C3); δ = 68.9 ppm, 1 C (C4); δ = 66.7 ppm, 1 C (NHCOC(CH_2CH_2)*C*H_2OH); δ = 62.6 ppm, 1 C (C6); δ = 55.1 ppm, 1 C (O*C*H_2CCH); δ = 54.4 ppm, 1 C (C2); δ = 26.4 ppm, 1 C (NHCO*C*(CH_2CH_2)CH_2OH); δ = 13.4 ppm, 1 C (NHCOC(*C*$H_2$$CH_2$)$CH_2$OH); δ = 13.2 ppm, 1 C (NHCOC(CH_2*C*H_2)CH_2OH).

R_f = 0.33 with 15% MeOH in DCM.

HR ESI-MS for $C_{14}H_{21}NO_7$: $m{\cdot}z^{-1}(M+Na^+)_{calc}$ = 338.122; $m{\cdot}z^{-1}(M+Na^+)_{obs}$ = 338.121.

Propargyl-2-deoxy-2-2'-(*1H*-tetrazol-5'-yl)acetamido--α-D-mannopyranoside



5.7

2-(tetrazolidin-5-yl)acetic acid (59 mg, 0.46 mmol, Santa Cruz Biotechnology) was dissolved in DMF (1.2 ml). Subsequently, PyBOP (240 mg, 0.46 mmol) and DIPEA (160 µl, 0.92 mmol) were added and mixture was stirred for 10 min at room temperature. **4** (50 mg, 0.230 mmol) was added and the mixture was stirred overnight at room temperature. Progress of the reaction was monitored by analytical TLC (20% MeOH in DCM). Following quenching with MeOH (1 ml), solvents were removed *in vacuo*. The residue was purified *via* reversed-phase column chromatography (gradient: elution with H_2O). Residual impurities were removed *via* preparative HPLC (gradient: H_2O for 10 min, from 0 to 20% acetonitrile in H_2O in 30 min, from 20 to 50% acetonitrile in H_2O in 10 min, from 50 to 100% acetonitrile in H_2O in 5 min and acetonitrile for 5 min at 3.2 $ml{\cdot}min^{-1}$) to yield **5.7** as a white solid (16 mg, 49 µmol, 21 %).

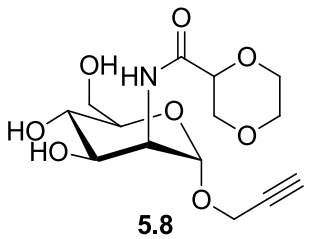
^{1}H NMR (400.0 MHz, MeOD): δ = 8.11 ppm, s, 1 H (N*H*COCH$_2$N$_4$H); δ = 4.94 ppm, s, 1 H (H1); δ = 4.25 ppm, d, 1 H, J = 4.7 Hz (H2); δ = 4.27 ppm, m, 2 H (OC*H2*CCH); δ = 4.03 ppm, m, 2 H (NHCOC*H$_2$*N$_4$H); δ = 3.93 ppm, dd, 1 H, J = 4.7, 8.9 Hz (H3); δ = 3.83 ppm, m, 2 H (H6a/b); δ = 3.61 ppm, m, 1 H (H4); δ = 3.58 ppm, m, 1 H (H5); δ = 2.87 ppm, m, 1 H (OCH$_2$CC*H*).

^{13}C NMR (100.6 MHz, MeOD): δ = 169.5 ppm, 1C (NH*C*OCH$_2$N$_4$H); δ = 98.9 ppm, 1 C (C1); δ = 79.7 ppm, 1 C (OCH$_2$*C*CH); δ = 76.2 ppm, 1 C (OCH$_2$C*C*H); δ = 74.7 ppm, 1 C (C5); δ = 70.7 ppm, 1 C (C3); δ = 68.3 ppm, 1 C (C4); δ = 62.2 ppm, 1 C (C6); δ = 55.1, 1 C (O*C*H$_2$CCH); δ = 54.4 ppm, 1 C (C2); δ = 31.0 ppm, 1 C (NHCO*C*H$_2$N$_4$H).

R_f = 0.06 with 20% MeOH in DCM.

HR ESI-MS for $C_{12}H_{17}N_5O_6$: m·z^{-1}(M+Na$^+$)$_{calc}$ = 350.108; m·z^{-1}(M+Na$^+$)$_{obs}$ = 350.109.

Propargyl-2-deoxy-2-1',4'-dioxane-2'-carboxamido-α-D-mannopyranoside



5.8

1,4-dioxane-2-carboxylic acid (40 mg, 306 µmol, Santa Cruz Biotechnology) was dissolved in DMF (760 µl). Subsequently PyBOP (158 mg, 306 µmol) and DIPEA (110 µl, 612 µmol) were added and the mixture was stirred for 10 min at room temperature. **4** (33 mg, 152 µmol) was added and the mixture was stirred overnight at room temperature. Progress of the reaction was monitored by analytical TLC (10% MeOH in DCM). Solvents were removed *in vacuo* and the residue was purified *via* column chromatography (gradient: hexane, DCM, 5 % MeOH in DCM and elution with 10 % MeOH in DCM). Residual impurities were removed *via* revers-phase column chromatography (gradient: H_2O, 5% MeOH in H_2O and elution at 10% MeOH in H_2O) to yield **5.8** as a white solid (20 mg, 60 µmol, 40%).

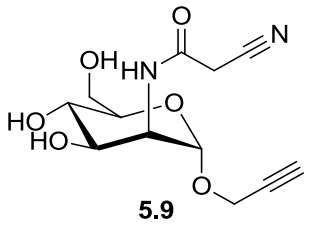
^{1}H NMR (400 MHz, MeOD): δ = 4.95 ppm, dd, 1 H, J = 1.0, 8.4 Hz (H1); δ = 4.27 ppm, m, 2H (OC*H$_2$*CCH); δ = 4.25 ppm, m, 1H (H2); δ = 4.15 ppm, m, 1 H (NHCOC*H*CH$_2$OCH$_2$CH$_2$O); δ = 3.99 ppm, m, 1 H (NHCOCHC*H$_2$*OCH$_2$CH$_2$O); δ = 3.94 ppm, m, 1 H (H3); δ = 3.90 ppm, m, 1 H (NHCOCHCH$_2$OC*H$_2$*CH$_2$O); δ = 3.79 ppm, m, 3 H (H6a/b, NHCOCHCH$_2$OC*H$_2$*CH$_2$O); δ = 3.72 ppm, m, 1 H (NHCOCHCH$_2$OCH$_2$C*H$_2$*O); δ = 3.62 ppm, m, 1H (NHCOCHCH$_2$OCH$_2$C*H$_2$*O); δ = 3.54 ppm, m, 1H (H5); δ = 3.52 ppm, m, 1 H (H4); δ = 3.47 ppm, m, 1 H (NHCOCHC*H$_2$*OCH$_2$CH$_2$O); δ = 2.87 ppm, m, 1 H (OCH$_2$CC*H*)

^{13}C NMR (100.6 MHz, MeOD): δ = 171.5 ppm, 1 C (NH*C*OCHCH$_2$OCH$_2$CH$_2$O) ; δ = 98.8 ppm, d, 1 C, J = 16.5 Hz (C1); δ = 79.7 ppm, 1 C (OCH$_2$*C*CH); δ = 76.3 ppm, m, 1 C (NHCO*C*HCH$_2$OCH$_2$CH$_2$O); δ = 76.2 ppm, 1 C (OCH$_2$C*C*H); δ = 74.6 ppm, d, 1C, J = 4.2 Hz (C5); δ = 70.3 ppm, d, J = 20.0 Hz (C3); δ = 69.4 ppm, d, 1C, J = 8.9 Hz (NHCOCH*C*H$_2$OCH$_2$CH$_2$O); δ = 68.2 ppm, 1C (C4); δ = 67.6 ppm, d, 1C, J = 3.9 Hz (NHCOCHCH$_2$OC*H*$_2$CH$_2$O); δ = 67.3 ppm, d, 1C, J = 1.6 Hz (NHCOCHCH$_2$OCH$_2$C*H*$_2$O); δ = 62.0 ppm, d, 1C, J = 1.3 Hz (C6); δ = 55.2 ppm, 1C (O*C*H$_2$CCH); δ = 53.81 ppm, d, 1C, J = 4.9 Hz (C2)

R_f = 0.19 with 10% in MeOH in DCM.

HR ESI-MS for $C_{14}H_{21}NO_8$: m·z^{-1}(M+Na$^+$)$_{calc}$ = 354.117; m·z^{-1}(M+Na$^+$)$_{obs}$ = 354.114.

Propargyl-2-deoxy-2-2'-cyanoacetamido-α-D-mannopyranoside



5.9

2-cyanoacetic acid (39.0 mg, 0.46 mmol) was dissolved in DMF (1.2 ml). Subsequently PyBOP (240 mg, 0.46 mmol) and DIPEA (160 µl, 0.92 mmol) were added and the mixture was stirred for 10 min at room temperature. **4** (50 mg, 0.230 mmol) was added and the mixture was stirred overnight at room temperature. Progress of the reaction was monitored by analytical TLC (10% MeOH in DCM). Solvents were removed *in vacuo* and the residue was purified *via* column chromatography (gradient: hexane, DCM, 5 % MeOH in DCM and elution with 20 % MeOH in DCM). Residual impurities were removed *via* preparative HPLC (gradient: H_2O for 10 min, from 0 to 20% acetonitrile in H_2O in 30 min, from 20 to 50% acetonitrile in H_2O in 10 min, from 50 to 100% acetonitrile in H_2O in 5 min and acetonitrile for 5 min at 3.2 ml·min^{-1}) to yield **5.9** as a white solid (10.2 mg, 36 µmol, 16%).

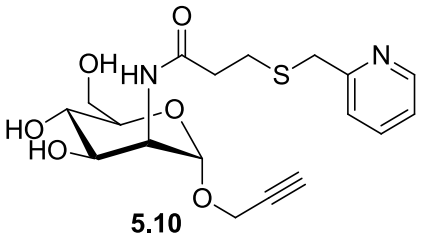
^{1}H NMR (600.0 MHz, MeOD): δ = 8.16 ppm, s, 1 H (N*H*COCH$_2$CN); δ = 4.91 ppm, d, 1 H, J = 1.1 Hz (H1); δ = 4.29 ppm, m, 1 H (H2); δ = 4.27 ppm, m, 2 H (OC*H2*CCH); δ = 3.92 ppm, dd, 1 H, J = 4.8, 9.1 Hz (H3); δ = 3.81 ppm, m, 2 H (H6a/b); δ = 3.62 ppm, m, 2 H (NHCOC*H*$_2$CN); δ = 3.54 ppm, m, 2 H (H4, H5); δ = 2.87 ppm, t, 1 H, J = 2.3 Hz (OCH$_2$CC*H*).

^{13}C NMR (100.6 MHz, MeOD): δ = 165.3 ppm, 1 C (NH*C*OCH$_2$CN); δ = 116.0 ppm, 1 C (NHCOCH$_2$*C*N); δ = 98.7 ppm, 1 C (C1); δ = 79.7 ppm, 1 C (OCH$_2$*C*CH); δ = 76.3 ppm, 1 C (OCH$_2$C*C*H); δ = 74.6 ppm, 1 C (C5); δ = 70.47, 1 C (C3); δ = 68.2 ppm, 1 C (C4); δ = 62.1 ppm, 1 C (C6); δ = 55.1, 1 C (O*C*H$_2$CCH); δ = 54.7 ppm, 1 C (C2); δ = 26.0 ppm, 1 C (NHCO*C*H$_2$CN).

R_f = 0.25 with 10% in MeOH in DCM.

HR ESI-MS for $C_{12}H_{16}N_2O_6$: $m \cdot z^{-1}(M+Na^+)_{calc}$ = 307.091; $m \cdot z^{-1}(M+Na^+)_{obs}$ = 307.096.

Propargyl-2-deoxy-2-3'-((pyridin-2'-ylmethyl)thio)-propanamido-α-D-mannopyranoside



5.10

3-((Pyridin-2-ylmethyl)thio)propanoic acid (42.0 mg, 210 µmol, Enamine) was dissolved in DMF (600 µl). Subsequently PyBOP (110 mg, 210 µmol) and DIPEA (80 µl, 430 µmol) were added and the mixture was stirred for 10 min at room temperature. **4** (23 mg, 110 µmol) was added and the mixture was stirred overnight at room temperature. Progress of the reaction was monitored by analytical TLC (15% MeOH in DCM). The reaction was quenched with MeOH (1 ml) and after addition of 1 M NaOH (600 µl, 600 µmol) solvents were removed *in vacuo*. The residue was purified *via* preparative HPLC (gradient: H_2O for 10 min, from 0 to 20% acetonitrile in H_2O in 30 min, from 20 to 50% acetonitrile in H_2O in 10 min, from 50 to 100% acetonitrile in H_2O in 5 min and acetonitrile for 5 min at 3.2 $ml \cdot min^{-1}$) to yield **5.10** as a white solid (26.1 mg, 66 µmol, 62%).

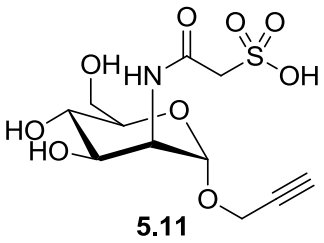
^{1}H NMR (400.0 MHz, MeOD): δ = 8.46 ppm, d, 1 H, J = 4.8 Hz ($SCH_2CCHCHCHCH\mathit{H}N$); δ = 8.12 ppm, s, 1 H ($N\mathit{H}COCH_2CH_2S$); δ = 7.82 ppm, td, 1 H, J = 1.5, 7.7 Hz ($SCH_2CCHC\mathit{H}CHCHN$); δ = 7.52 ppm, d, 1 H, J = 7.9 Hz ($SCH_2CC\mathit{H}CHCHCHN$); δ = 7.31 ppm, dd, 1 H, J = 5.2, 7.4 Hz ($SCH_2CCHCHC\mathit{H}CHN$); δ = 4.89 ppm, s, 1 H (H1); δ = 4.32 ppm, d, 1 H, J = 5.0 Hz (H2); δ = 4.26 ppm, m, 2 H ($OC\mathit{H2}CCH$); δ = 3.91 ppm, dd, 1 H, J = 4.8, 9.3 Hz (H3); δ = 3.87 ppm, s, 2 H ($SC\mathit{H}_2CCHCHCHCHN$); δ = 3.81 ppm, m, 2 H (H6a/b); δ = 3.61 ppm, m, 1 H (H4); δ = 3.54 ppm, m, 1 H (H5); δ = 2.86 ppm, t, 1 H, J = 2.3 Hz ($OCH_2CC\mathit{H}$); δ = 2.74 ppm, m, 2 H ($NHCOC\mathit{H}_2CH_2S$); δ = 2.58 ppm, m, 2 H ($NHCOCH_2C\mathit{H}_2S$).

^{13}C NMR (100.6 MHz, MeOD): δ = 174.7 ppm, 1 C ($NHCOCH_2CH_2S$); δ = 159.8 ppm, 1 C ($SCH_2CCHCHCHCHN$); δ = 149.6 ppm, 1 C ($SCH_2CCHCHCHCHN$); δ = 139.1 ppm, 1 C ($SCH_2CCHCHCHCHN$); δ = 125.1 ppm, 1 C ($SCH_2CCHCHCHCHN$); δ = 123.7 ppm, 1 C ($SCH_2CCHCHCHCHN$); δ = 99.0 ppm, 1 C (C1); δ = 79.8 ppm, 1 C (OCH_2CCH); δ = 76.2 ppm, 1 C (OCH_2CCH); δ = 74.6 ppm, 1 C (C5); δ = 70.62, 1 C (C3); δ = 68.2 ppm, 1 C (C4); δ = 62.2 ppm, 1 C (C6); δ = 55.1 ppm, 1 C (OCH_2CCH); δ = 54.1 ppm, 1 C (C2); δ = 37.9 ppm, 1 C ($SCH_2CCHCHCHCHN$); δ = 36.8 ppm, 1 C ($NHCOCH_2CH_2S$); δ = 28.2 ppm, 1 C ($NHCOCH_2CH_2S$).

R_f = 0.67 with 15% in MeOH in DCM.

HR ESI-MS for $C_{18}H_{24}N_2O_6S$: $m \cdot z^{-1}(M+Na^+)_{calc}$ = 419.125; $m \cdot z^{-1}(M+Na^+)_{obs}$ = 419.124.

Propargyl-2-deoxy-2-2'-sulfoacetamido-α-D-mannopyranoside



5.11

2-Sulfoacetic acid (65 mg, 460 µmol) was dissolved in DMF (1.2 ml). Subsequently, PyBOP (240 mg, 460 µmol) and DIPEA (160 µl, 920 µmol) were added and the mixture was stirred for 10 min at room temperature. **4** (45 mg, 230 µmol) was added and the mixture was stirred overnight at room temperature. Progress of the reaction was monitored by analytical TLC (20% MeOH in DCM). After the addition of 1 M NaOH (1.4 ml, 1.4 mmol) solvents were removed *in vacuo.* The residue was purified *via* reversed-phase column chromatography (elution with H_2O). Residual impurities were removed *via* preparative HPLC (gradient: H_2O for 10 min, from 0 to 100% acetonitrile in H_2O in 40 min, and acetonitrile for 10 min at 1.0 $ml \cdot min^{-1}$) to yield **5.11** as a white solid (18 mg, 53 µmol, 23%). A purity of more than 90% purity was determined *via* analytical HPLC and ^1H NMR using TSP as an internal reference.

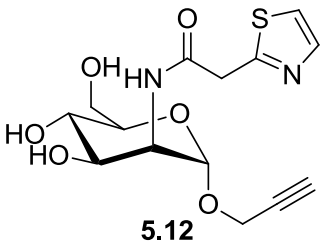
^1H NMR (600.0 MHz, MeOD): δ = 8.09 ppm, s, 1 H (N*H*COCH$_2$SO$_3$H); δ = 4.98 ppm, s, 1 H (H1); δ = 4.35 ppm, d, 1 H, J = 4.2 Hz (H2); δ = 4.27 ppm, m, 2 H (OC*H2*CCH); δ = 3.89 ppm, dd, 1 H, J = 4.5, 9.2 Hz (H3); δ = 3.78 ppm, m, 3 H (NHCOC*H$_2$*SO$_3$H, H6a/b); δ = 3.55 ppm, m, 2 H (H4, H5); δ = 2.85 ppm, t, 1 H, J = 2.4 Hz (OCH$_2$CC*H*).

^{13}C NMR (100.6 MHz, MeOD): δ = 168.2 ppm, 1 C (NH*C*OCH$_2$SO$_3$H); δ = 98.7 ppm, 1 C (C1); δ = 79.8 ppm, 1 C (OCH$_2$*C*CH); δ = 76.2 ppm, 1 C (OCH$_2$C*C*H); δ = 74.8 ppm, 1 C (C5); δ = 71.1 ppm, 1 C (C3); δ = 68.6 ppm, 1 C (C4); δ = 62.5 ppm, 1 C (C6); δ = 58.1 ppm, 1 C (NHCO*C*H$_2$SO$_3$H); δ = 55.1 ppm, 1 C (O*C*H$_2$CCH); δ = 54.4 ppm, 1 C (C2).

R_f = 0.04 with 20% in MeOH in DCM.

HR ESI-MS for $C_{11}H_{17}NO_9S$: $m \cdot z^{-1}(M\text{-}H^+)_{calc}$ = 338.054; $m \cdot z^{-1}(M\text{-}H^+)_{obs}$ = 338.058.

Propargyl-2-deoxy-2-2'-(thiazol-2''-yl)acetamido-α-D-mannopyranoside



5.12

2-(Thiazol-2-yl)acetic acid (50.0 mg, 350 µmol) was dissolved in DMF (0.9 ml). Subsequently, PyBOP (180 mg, 350 µmol) and DIPEA (120 µl, 700 µmol) were added and the mixture was stirred for 10 min at room temperature. 16 h after the addition of **4** (38 mg, 180 µmol), more PyBOP (180 mg, 350 µmol), DIPEA (120 µl, 700 µmol) and 2-(thiazol-2-yl)acetic acid (50 mg, 350 µmol) were added. The mixture was stirred overnight at room temperature and progress of the reaction was monitored by analytical TLC (10% MeOH in DCM). After the addition of 1 M NaOH (1.4 ml, 1.4 mmol) solvents were removed *in vacuo*. The residue was purified *via* reversed-phase column chromatography (elution with H_2O). Residual impurities were removed *via* preparative HPLC (gradient: H_2O for 10 min, from 0 to 20% acetonitrile in H_2O in 30 min, from 20 to 50% acetonitrile in H_2O in 10 min, from 50 to 100% acetonitrile in H_2O in 5 min and acetonitrile for 5 min at 3.2 ml·min^{-1}) to yield **5.12** as a white solid (5.2 mg, 15 µmol, 9%).

^{1}H NMR (600.0 MHz, MeOD): δ = 4.97 ppm, d, 1 H, J = 1.2 Hz (H1); δ = 4.30 ppm, dd, 1 H, J = 1.3, 4.8 Hz (H2); δ = 4.28 ppm, m, 2 H (OC*H2*CCH); δ = 4.02 ppm, s, 2 H (NHCOC*H*$_2$CNCCS); δ = 3.96 ppm, dd, 1 H, J = 4.8, 8.8 Hz (H3); δ = 3.83 ppm, m, 1 H (H6a); δ = 3.75 ppm, m, 1 H (H6b); δ = 3.57 ppm, m, 1 H (H5); δ = 3.53 ppm, m, 1 H (H4); δ = 2.87 ppm, t, 1 H, J = 2.4 Hz (OCH$_2$CC*H*).

^{13}C NMR (100.6 MHz, MeOD): δ = 175.5 ppm, 1 C (NH*C*OCH$_2$CNCCS); δ = 98.6 ppm, 1 C (C1); δ = 79.7 ppm, 1 C (OCH$_2$*C*CH); δ = 76.2 ppm, 1 C (OCH$_2$C*C*H); δ = 74.6 ppm, 1 C (C5); δ = 70.4 ppm, 1 C (C3); δ = 68.4 ppm, 1 C (C4); δ = 62.6 ppm, 1 C (NHCO*C*H$_2$CNCCS); δ = 62.1 ppm, 1 C (C6); δ = 55.1 ppm, 1 C (O*C*H$_2$CCH); δ = 53.8 ppm, 1 C (C2).

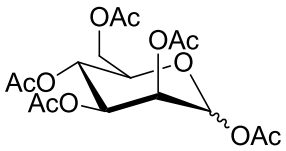
The aromatic carbon atoms of the thiazolyl were not detected in the conducted ^{13}C NMR experiments.

R_f = 0.24 with 10% in MeOH in DCM.

HR ESI-MS for $C_{14}H_{18}N_2O_6S$: m·z^{-1}(M+Na$^+$)$_{calc}$ = 365.078; m·z^{-1}(M+Na$^+$)$_{obs}$ = 365.076.

2.3.3. Mannose Analogs

1,2,3,4,6-penta-O-acetyl-D-mannopyranose



10

10 was prepared as published previously (Hasegawa *et al.*, 2007). Mannose (5.0 g, 28 mmol) and acetic anhydride (42 ml, 440 mmol) were dissolved in pyridine (100 ml). The reaction mixture was stirred over night at 50°C. Progress of the reaction was monitored by TLC in toluene:EtOAc (2:1). Solvents were evaporated *in vacuo* and the residue was taken up in chloroform (250 ml). The organic phase was extracted with 1 M HCl, saturated $NaHCO_3$ and H_2O. Subsequently, the organic phase was dried with $MgSO_4$. Solvents were evaporated in *in vacuo* and the residue was purified *via* column chromatography (toluene:EtOAc (6:1)) to afford an α/β-anomer mixture of **10** (7.20 g, 18.45 mmol, 80 %) as a white solid.

^{1}H NMR (400.0 MHz, $CDCl_3$, α-anomer): δ = 6.05 ppm, m, 1 H (H1); δ = 5.30 ppm, m, 2 H (H3, H4); δ = 5.22, m, 1H (H2); δ = 4.23 ppm, dd, 1 H, J = 4.7, 12.1 Hz (H6a); δ = 4.05, m, 1 H (H6b); δ = 4.01 ppm, m, 1 H (H5); δ = 2.17 - 1.96 ppm, m, 15 H (five times $OCOCH_3$).

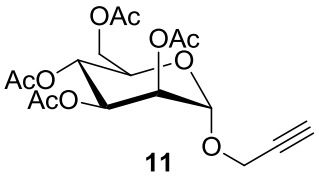
^{13}C NMR (100.6 MHz, $CDCl_3$, α-anomer): δ = 170.6 ppm, 1C ($OCOCH_3$); δ = 170.0, 1C ($OCOCH_3$); δ = 169.7 ppm ($OCOCH_3$); δ = 169.5 ppm, 1C ($OCOCH_3$); δ = 168.1 ppm, 1C ($OCOCH_3$); δ = 90.6 ppm, 1 C (C1); δ = 70.6 ppm, 1 C (C5); δ = 68.8 ppm, 1 C (C3); δ = 68.4 ppm, 1 C (C2); δ = 65.6 ppm, 1 C (C4); δ = 62.1, 1 C (C6); δ = 20.9 ppm, 1 C ($OCOCH_3$); δ = 20.8 ppm, 1 C ($OCOCH_3$); δ = 20.7 ppm, 3 C (three times $OCOCH_3$).

^{1}H NMR (400.0 MHz, $CDCl_3$, β-anomer): δ = 5.83 ppm, m, 1 H (H1); δ = 5.44 ppm, m, 1 H (H2); δ = 5.25 ppm, m, 1 H (H4); δ = 5.10, dd, 1H, J = 2.5, 10.2 Hz (H3); δ = 4.27 ppm, m, 1 H (H6a); δ = 4.05, m, 1 H (H6b); δ = 3.77, m, 1 H (H5); δ = 2.17 - 1.96 ppm, m, 15 H (five times $OCOCH_3$).

^{13}C NMR (100.6 MHz, $CDCl_3$, β-anomer): δ = 170.6 ppm, 1C ($OCOCH_3$); δ = 170.2, 1C ($OCOCH_3$); δ = 169.8 ppm ($OCOCH_3$); δ = 169.6 ppm, 1C ($OCOCH_3$); δ = 168.4 ppm, 1C ($OCOCH_3$); δ = 90.5 ppm, 1 C (C1); δ = 73.3 ppm, 1 C (C5); δ = 70.7 ppm, 1 C (C3); δ = 68.2 ppm, 1 C (C2); δ = 65.5 ppm, 1 C (C4); δ = 62.1, 1 C (C6); δ = 20.8 – 20.6 ppm, 5 C (five times $OCOCH_3$).

R_f = 0.43 with toluene:EtOAc (3:1).

Propargyl-2,3,4,6-tetra-O-acetyl-α-D-mannopyranoside



11

10 (7.20 g, 19 mmol) was dissolved in anhydrous DCM:ether (189 ml, 2:1). The mixture was stirred at 0°C and kept under argon. Propargyl alcohol (4.4 ml, 73.8 mmol) and $BF_3{\cdot}OEt_2$ (4.7 ml, 36.9 mmol) were added and the reaction was allowed to heat up to room temperature and stirred. After 24 h additional propargyl alcohol (4.4 ml, 73.8 mmol) and $BF_3{\cdot}OEt_2$ (4.7 ml, 36.9 mmol) were added at 0°C and the reaction was stirred for another 40 h. Progress of the reaction was monitored by analytical TLC (toluene:EtOAc (3:1)). The mixture was diluted in DCM (500 ml) and the organic phase was extracted with saturated $NaHCO_3$ and H_2O. Subsequently, the organic phase was dried with $MgSO_4$. Solvents were evaporated *in vacuo* and the residue was purified *via* column chromatography (toluene:EtOAc (8:1)) to yield **11** (188 mg, 487 µmol, 3%) as a light yellow resin. Starting material **10** that was not converted was recovered.

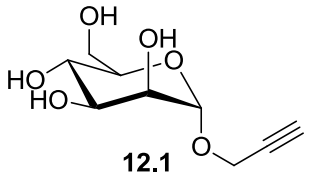
^{1}H NMR (400.0 MHz, $CDCl_3$): δ = 5.17 ppm, d, 1 H, J = 1.1 Hz (H3); δ = 5.15 ppm, m (H4); δ = 5.11 ppm, m, 1 H (H2); δ = 4.87 ppm, d, 1H, J = 1.3 Hz (H1); δ = 4.16 ppm, m, 1 H (H6a); δ = 4.11 ppm, d, 2 H, J = 2.4 Hz (OC*H*$_2$CCH); δ = 4.27 ppm, dd, 1 H, J = 2.3, 12.2 Hz (H6b); δ = 3.86 ppm, ddd, 1 H, J = 2.3, 5.3, 9.3 Hz (H5); δ = 2.32, t, 1 H, J = 2.4 5z (OCH$_2$CC*H*); δ = 2.00 ppm, s, 3 H (OCOC*H*$_3$); δ = 1.94 ppm, s, 3 H (OCOC*H*$_3$); δ = 1.88 ppm, s, 3 H (OCOC*H*$_3$); δ = 1.83 ppm, s, 3 H (OCOC*H*$_3$).

^{13}C NMR (100.6 MHz, CDCl3): δ = 170.7 ppm, 1 C (O*C*OCH$_3$); δ = 170.0 ppm, 1 C (O*C*OCH$_3$); δ = 169.9 ppm, 1 C (O*C*OCH$_3$); δ = 169.8 ppm, 1 C (O*C*OCH$_3$); δ = 95.5 ppm, 1 C (C1); δ = 96.3 ppm, 1 C (C1); δ = 78.0 ppm, 1 C (OCH$_2$*C*CH); δ = 75.7 ppm, 1 C (OCH$_2$C*C*H); δ = 69.5 ppm, 1 C (C2); δ = 69.1 ppm, 1 C (C5); δ = 69.0 ppm, 1 C (C3); δ = 66.1 ppm, 1 C (C4); δ = 62.4 ppm, 1 C (C6); δ = 55.1 ppm, 1 C (O*C*H$_2$CCH); δ = 20.9 ppm, 1 C (OCO*C*H$_3$); δ = 20.8 ppm, 2 C (three times OCO*C*H$_3$).

R_f = 0.38 with toluene:EtOAc (3:1).

ESI-MS for $C_{17}H_{22}O_{10}$: $m{\cdot}z^{-1}(M+Na^+)_{calc}$ = 409.1; $m{\cdot}z^{-1}(M+Na^+)_{obs}$ = 409.2; $m{\cdot}z^{-1}(M+NH_4^+)_{calc}$ = 404.2; $m{\cdot}z^{-1}$ $(M+NH_4^+)_{obs}$ = 404.2.

Propargyl-α-D-mannopyranoside



12.1

11 (188 mg, 487 µmol) was dissolved in EtOH containing 33% $MeNH_2$ (13 ml) at room temperature. The mixture was stirred overnight and progress of the reaction was monitored by analytical TLC (20% MeOH in DCM). Solvents were evaporated *in vacuo* and the residue was purified *via* column chromatography (gradient: hexane, hexane:DCM (1:1), DCM, 1% MeOH in DCM, 5% MeOH in DCM and elution with 20% MeOH in DCM). Silica gel particles were removed by filtration in MeOH with a cellulose acetate membrane at a pore size 0.2 µm to yield **12.1** (91 mg, 417 µmol, 86%) as a white solid.

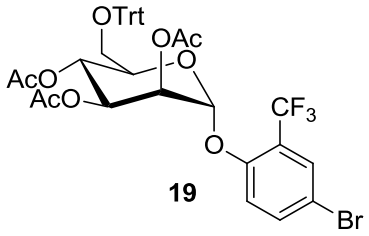
^{1}H NMR (400.0 MHz, MeOD): δ = 4.95 ppm, d, 1 H, J = 1.2 Hz (H1); δ = 4.26 ppm, d, 2 H, J = 2.6 Hz (OC*H*$_2$CCH); δ = 3.79 ppm, dd, 1 H, J = 2.2, 11.7 Hz (H6a); δ = 3.76 ppm, dd, 1 H, J = 1.6, 3.1 Hz (H2); δ = 3.70 ppm, m, 1 H (H6b); δ = 3.65 ppm, m, 1 H (H3); δ = 3.61 ppm, m, 1 H (H4); δ = 3.50 ppm, ddd, 1 H, J = 2.0, 5.8, 9.6 Hz (H5); δ = 2.85 ppm, t, 1 H, J = 2.4 Hz (OCH$_2$CC*H*).

^{13}C NMR (100.6 MHz, MeOD): δ = 99.8 ppm, 1 C (C1); δ = 80.0 ppm, 1 C (OCH$_2$*C*CH); δ = 76.0 ppm, 1 C (OCH$_2$C*C*H); δ = 75.1 ppm, 1 C (C5); δ = 72.5 ppm, 1 C (C3); δ = 72.0 ppm, 1 C (C2); δ = 68.5 ppm, 1 C (C4); δ = 62.8 ppm, 1 C (C 6); δ = 54.8 ppm, 1 C (O*C*H$_2$CCH).

R_f = 0.36 with 20% MeOH in DCM.

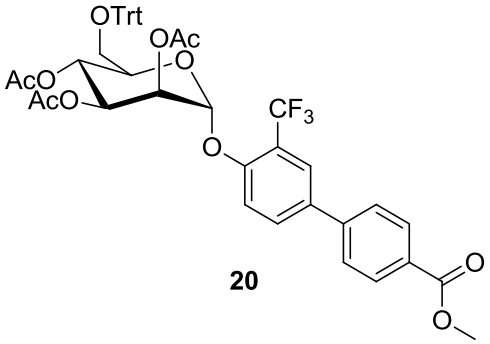
HR ESI-MS for $C_9H_{14}O_6$: $m{\cdot}z^{-1}(M+Na^+)_{calc}$ = 241.069; $m{\cdot}z^{-1}(M+Na^+)_{obs}$ = 241.068.

4'-bromo-2'-(trifluoromethyl)phenoxy-2,3,4-tetra-O-acetyl-O-trityl-α-D-mannopyranoside[5]



19

Methyl-4'-(2'',3'',4''-tetra-O-acetyl-6''-O-trityl-α-D-mannopyranosyloxy)-3'-trifluoromethylbiphenyl-4-carboxylate



20

A Schlenk tube was charged with **19** (500 mg, 648 µmol), (4-(methoxycarbonyl)phenyl)boronic acid (128 mg, 713 µmol), K_3PO_4 (206 mg, 972 µmol) and $Pd(dppf)Cl_2 \cdot CH_2Cl_2$ (16 mg, 19 µmol). The tube was sealed, evacuated and subsequently flushed with argon. This procedure was repeated twice. Next, anhydrous DMF (6 ml) was added under a stream of argon. The reaction mixture was degassed *via* ultrasonication, flushed with argon and stirred overnight at 80°C. Product formation was monitored *via* TLC (2:1 (PE:EtOAc). The reaction mixture was cooled to room temperature and diluted with EtOAc (50 ml). Next, the diluted reaction mixture was extracted with H_2O (50 ml) and brine (50 ml) and dried with Na_2SO_4. Solvents were removed *in vacuo* and the residue was purified *via* MPLC (gradient: 100% PE to 100% EtOAc in 20 min) to yield **20** as a white solid (464 mg, 561 µmol, 87%).

[5]The procedure was devised and conducted by Dr. Oliver Schwardt (Universität Basel).

^{1}H NMR (500.0 MHz, $CDCl_3$): δ = 8.14 – 8.10 ppm, m, 2 H (aromatic H of biphenyl); δ = 7.89 ppm, d, 1 H, J = 2.2 Hz (aromatic H of biphenyl); δ = 7.69 ppm, dd, 1 H, J = 9.0, 2.2 Hz (aromatic H of biphenyl); δ = 7.62 – 7.58 ppm, m, 2 H (aromatic H of biphenyl); δ = 7.55 ppm, d, 1 H, J = 8.9 Hz (aromatic H of biphenyl); δ = 7.39 – 7.33 ppm, m, 6 H, (aromatic H of OTrt); δ = 7.25 – 7.17 ppm, m, 9 H, (aromatic H of OTrt); δ = 5.70 – 5.72 ppm, m, 1 H (H1); δ = 5.52 – 5.47 ppm, m, 2 H (H2, H3); δ = 5.36 – 5.29 ppm, m, 1 H (H4); δ = 3.95 ppm, s, 3 H (OC*H*$_3$); δ = 3.92 ppm, ddd, 1 H, J = 10.2, 5.9, 2.0 Hz (H5); δ = 3.27 ppm, dd, 1 H, J = 10.8, 6.2 Hz (H6a); δ = 3.18 ppm, dd, 1 H, J = 10.9, 2.0 Hz (H6b); δ = 2.22 ppm, s, 3 H (OCOC*H*$_3$); δ = 2.01 ppm, s, 3 H (OCOC*H*$_3$); δ = 1.77 ppm, s, 3 H (OCOC*H*$_3$).

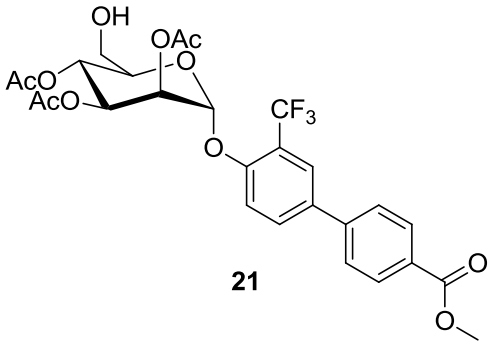
^{13}C NMR (125.8 MHz, $CDCl_3$): δ = 170.0 ppm, 1 C (O*C*OCH_3); δ = 169.7 ppm, 1 C (O*C*OCH_3); δ = 169.4 ppm, 1 C (O*C*OCH_3); δ = 166.8 ppm, 1 C (carbonyl C of biphenyl); δ = 153.2 ppm, 1 C (aromatic C of biphenyl); δ = 143.6 ppm, 3 C (aromatic C of OTrt); δ = 143.5 ppm, 1 C (aromatic C of biphenyl); δ = 134.3 ppm, 1 C (aromatic C of biphenyl); δ = 131.9 ppm, 1 C (aromatic C of biphenyl); δ = 130.3 ppm, 2 C (aromatic C of biphenyl); δ = 129.3 ppm, 1 C (aromatic C of biphenyl); δ = 128.6 ppm, 6 C (aromatic C of OTrt); δ = 127.8 ppm, 6 C (aromatic C of OTrt); δ = 127.0 ppm, 3 C (aromatic C of OTrt); δ = 126.8 ppm, 2 C (aromatic C of biphenyl); δ = 126.0 - 125.8 ppm, m, 1 C (aromatic C of biphenyl); δ = 123.2 ppm, q, 1 C, J = 272.4 Hz (*C*F_3); δ = 120.6 ppm, q, 1 C, J = 31.6 Hz (aromatic C of biphenyl); δ = 116.6 ppm, 1 C (aromatic C of biphenyl); δ = 95.6 ppm, 1 C (C1); δ = 86.8 ppm, 1 C (aliphatic C, OTrt); δ = 71.5 ppm, 1 C (C5); δ = 69.4 ppm, 1 C (C2); δ = 68.8 ppm, 1 C (C3); δ = 66.1 ppm, 1 C (C4); δ = 62.4 ppm, 1 C (C6); δ = 52.2 ppm, 1 C (O*C*H_3); δ = 20.9 ppm, 1 C (OCO*C*H_3); δ = 20.7 ppm, 1 C (OCO*C*H_3); δ = 20.5 ppm, 1 C (OCO*C*H_3).

R_f = 0.43 with PE:EtOAc (2:1).

$[\alpha]^{20}_D$ = +67.7° (c = 1.00, $CHCl_3$).

ESI-MS for $C_{46}H_{41}F_3O_{11}$: $m \cdot z^{-1}(M+Na^+)_{calc}$ = 849.80; $m \cdot z^{-1}(M+Na^+)_{obs}$ = 849.37.

Methyl-4'-(2'',3'',4''-tetra-O-acetyl-α-D-mannopyranosyloxy)-3'-trifluoromethylbiphenyl-4-carboxylate



21

20 (454 mg, 549 µmol) was dissolved in DCM (20 ml). Next, $FeCl_3$ (178 mg, 1.10 mmol) and H_2O (120 µl, 6.59 mmol) were added. The reaction mixture was stirred for 5 h at room temperature. Product formation was monitored *via* TLC (1:2 (PE:EtOAc). The reaction mixture was diluted with DCM (40 ml) and the organic phase was extracted with H_2O (40 ml) and dried with Na_2SO_4. Solvents were evaporated *in vacuo* and the residue was purified *via* MPLC (gradient: 100% PE to 100% EtOAc in 15 min) to yield **21** as a white solid (223 mg, 382 µmol, 69%).

^{1}H NMR (500.0 MHz, $CDCl_3$): δ = 8.14 – 8.10 ppm, m, 2 H (aromatic H of biphenyl); δ = 7.87 – 7.85 ppm, m, 1 H (aromatic H of biphenyl); δ = 7.77 – 7.72 ppm, m, 1 H (aromatic H of biphenyl); δ = 7.64 – 7.58 ppm, m, 2 H (aromatic H of biphenyl); δ = 7.34 ppm, d, 1 H, J = 9.0 Hz (aromatic H of biphenyl); δ = 5.77 – 5.73 ppm, m, 1 H (H1); δ = 5.64 ppm, dd, 1 H, J = 10.2, 3.4 Hz (H3); δ = 5.53 – 5.50 ppm, m, 1 H (H2); δ = 5.42 – 5.36 ppm, m, 1 H (H4); δ = 3.95 ppm, s, 3 H (OCH_3); δ = 3.88 – 3.83 ppm, m, 1 H (H5); δ = 3.73 – 3.66 ppm, m, 1 H (H6a); δ = 3.65 – 3.57 ppm, m, 1 H (H6b); δ = 2.21 ppm, s, 3 H (OCOCH_3); δ = 2.11 ppm, s, 3 H (OCOCH_3); δ = 2.06 ppm, s, 3 H (OCOCH_3).

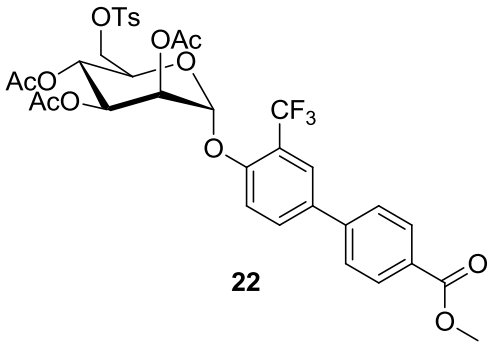
^{13}C NMR (125.8 MHz, $CDCl_3$): δ = 171.2 ppm, 1 C (O*C*OCH$_3$); δ = 170.9 ppm, 1 C (O*C*OCH$_3$); δ = 169.6 ppm, 1 C (O*C*OCH$_3$); δ = 166.8 ppm, 1 C (carbonyl C of biphenyl); δ = 152.9 ppm, 1 C (aromatic C of biphenyl); δ = 143.4 ppm, 1 C (aromatic C of biphenyl); δ = 134.4 ppm, 1 C (aromatic C of biphenyl); δ = 132.0 ppm, 1 C (aromatic C of biphenyl); δ = 130.3 ppm, 2 C (aromatic C of biphenyl); δ = 129.4 ppm, 1 C (aromatic C of biphenyl); δ = 126.8 ppm, 2 C (aromatic C of biphenyl); δ = 126.2 - 126.0 ppm, m, 1 C (aromatic C of biphenyl); δ = 123.1 ppm, q, 1 C, J = 272.5 Hz (*C*F$_3$); δ = 120.4 ppm, q, 1 C, J = 31.8 Hz (aromatic C of biphenyl); δ = 115.5 ppm, 1 C (aromatic C of biphenyl); δ = 95.6 ppm, 1 C (C1); δ = 72.0 ppm, 1 C (C5); δ = 69.2 ppm, 1 C (C2); δ = 68.3 ppm, 1 C (C3); δ = 65.9 ppm, 1 C (C4); δ = 60.9 ppm, 1 C (C6); δ = 52.2 ppm, 1 C (O*C*H$_3$); δ = 20.9 ppm, 1 C (OCO*C*H$_3$); δ = 20.8 ppm, 1 C (OCO*C*H$_3$); δ = 20.7 ppm, 1 C (OCO*C*H$_3$).

R_f = 0.53 with EtOAc:PE (2:1).

$[\alpha]^{20}_D$ = +62.1° (c = 1.00, $CHCl_3$).

ESI-MS for $C_{27}H_{27}F_3O_{11}$: $m \cdot z^{-1}(M+Na^+)_{calc}$ = 607.14; $m \cdot z^{-1}(M+Na^+)_{obs}$ = 607.10.

Methyl-4'-(2'',3'',4''-tetra-O-acetyl-6''-O-tosyl-α-D-mannopyranosyloxy)-3'-trifluoromethylbiphenyl-4-carboxylate



22

21 (216 mg, 370 µmol) was dissolved in dry pyridine (4 ml) and tosyl chloride (106 mg, 554 µmol) and a small amount of DMAP was added at 0°C. After 4 h more tosyl chloride (70.6 mg, 247 µmol) was added at 0°C. Next, the reaction mixture was stirred overnight at room termperature. Product formation was monitored *via* TLC (1:1 (PE:EtOAc)). Solvents were evaporated *in vacuo*. The residue was dissolved in DCM (50 ml) and the organic phase was extracted with 0.1 M HCl (50 ml), sat. $NaHCO_3$ (50 ml) and H2O (50 ml). Subsequently, the organic phase was dried with Na_2SO_4 and solvents were evaporated *in vacuo*. The residue was purified *via* MPLC (gradient: 100% PE to 100% EtOAc in 20 min) to yield **22** as a white solid (183 mg, 248 µmol, 67%).

^{1}H NMR (500.0 MHz, $CDCl_3$): δ = 8.15 – 8.11 ppm, m, 2 H (aromatic H of biphenyl); δ = 7.85 ppm, d, 1 H, J = 2.0 Hz (aromatic H of biphenyl); δ = 7.74 – 7.68 ppm, m, 3 H (aromatic H of OTs, biphenyl); δ = 7.65 – 7.59 ppm, m, 2 H (aromatic H of biphenyl); δ = 7.33 – 7.27 ppm, m, 3 H (aromatic H of OTs, biphenyl); δ = 5.60 ppm, d, 1 H, J = 1.6 Hz (H1); δ = 5.50 ppm, dd, 1 H, J = 10.0, 3.4 Hz (H3); δ = 5.53 – 5.50 ppm, dd, 1 H, J = 3.5, 1.9 Hz (H2); δ = 5.33 – 5.26 ppm, m, 1 H (H4); δ = 4.17 ppm, dd, 1 H, J = 11.3, 5.9 Hz (H6a); δ = 4.12 ppm, dd, 1 H, J = 11.3, 2.4 Hz (H6b); δ = 4.04 ppm, ddd, 1 H, J = 10.0, 5.9, 2.5 Hz (H5); δ = 3.95 ppm, s, 3 H (OC*H*$_3$); δ = 2.41 ppm, s, 3 H (aliphatic H, OTs); δ = 2.19 ppm, s, 3 H (OCOC*H*$_3$); δ = 2.03 ppm, s, 3 H (OCOC*H*$_3$); δ = 2.00 ppm, s, 3 H (OCOC*H*$_3$).

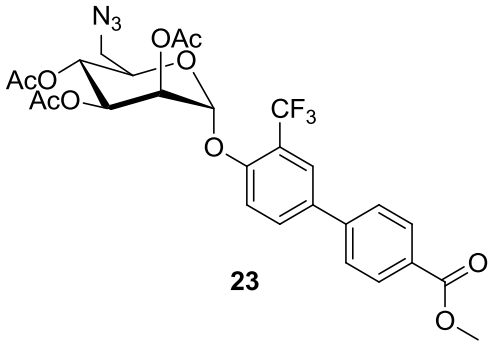
^{13}C NMR (125.8 MHz, $CDCl_3$): δ = 169.9 ppm, 1 C ($OCOCH_3$); δ = 169.6 ppm, 1 C ($OCOCH_3$); δ = 169.6 ppm, 1 C ($OCOCH_3$); δ = 166.8 ppm, 1 C (carbonyl C of biphenyl); δ = 152.9 ppm, 1 C (aromatic C of biphenyl); δ = 145.1 ppm, 1 C (aromatic C of OTs); δ = 143.4 ppm, 1 C (aromatic C of biphenyl); δ = 134.7 ppm, 1 C (aromatic C of biphenyl); δ = 132.7 ppm, 1 C (aromatic C of OTs); δ = 132.0 ppm, 1 C (aromatic C of biphenyl); δ = 130.3 ppm, 2 C (aromatic C of biphenyl); δ = 129.8 ppm, 2 C (aromatic C of OTs); δ = 129.4 ppm, 1 C (aromatic C of biphenyl); δ = 128.0 ppm, 2 C (aromatic C of OTs); δ = 126.8 ppm, 2 C (aromatic C of biphenyl); δ = 126.0 ppm, q, 1 C, J = 5.1 Hz (aromatic C of biphenyl); δ = 123.1 ppm, q, 1 C, J = 272.5 Hz (CF_3); δ = 120.5 ppm, q, 1 C, J = 31.8 Hz (aromatic C of biphenyl); δ = 116.1 ppm, 1 C (aromatic C of biphenyl); δ = 95.7 ppm, 1 C (C1); δ = 69.6 ppm, 1 C (C5); δ = 69.0 ppm, 1 C (C2); δ = 68.4 ppm, 1 C (C3); δ = 67.6 ppm, 1 C (C6); δ = 65.6 ppm, 1 C (C4); δ = 52.2 ppm, 1 C (OCH_3); δ = 21.6 ppm, 1 C (aliphatic C, OTs); δ = 20.8 ppm, 1 C ($OCOCH_3$); δ = 20.6 ppm, 1 C ($OCOCH_3$); δ = 20.6 ppm, 1 C ($OCOCH_3$).

R_f = 0.65 with EtOAc:PE (1:1).

$[\alpha]^{20}_D$ = +58.4° (c = 1.00, $CHCl_3$).

ESI-MS for $C_{34}H_{33}F_3O_{13}S$: $m{\cdot}z^{-1}(M+Na^+)_{calc}$ = 761.15; $m{\cdot}z^{-1}(M+Na^+)_{obs}$ = 761.17.

Methyl-4'-(2'',3'',4''-tetra-O-acetyl-6''-deoxy-6''-azido-α-D-mannopyranosyloxy)-3'-trifluoromethylbiphenyl-4-carboxylate



23

22 (30 mg, 41 µmol), sodium azide (14 mg, 203 µmol) and 1,4,7,10,13,16-hexaoxacyclooctadecane (6 mg, 20 µmol) were dissolved in anhydrous DMF (2 ml) at room temperature under argon atmosphere. The reaction mixture was stirred at 80°C overnight and subsequently diluted with diethyl ether (25 ml). The organic phase was extracted with H_2O (25 ml) and dried over Na_2SO_4. Solvents were evaporated *in vacuo* and the residue was purified *via* MPLC (gradient: 100% PE to 100% EtOAc in 20 min) to yield **23** (20.4 mg, 33 µmol, 82%) as a light orange solid.

^{1}H NMR (500.0 MHz, $CDCl_3$): δ = 8.15 – 8.09 ppm, m, 2 H (aromatic H of biphenyl); δ = 7.87 ppm, d, 1 H, J = 2.0 Hz (aromatic H of biphenyl); δ = 7.78 ppm, dd, 1 H, J = 8.9, 2.1 Hz (aromatic H of biphenyl); δ = 7.65 – 7.59 ppm, m, 2 H (aromatic H of biphenyl); δ = 7.37 ppm, d, 1 H, J = 8.8 Hz (aromatic H of biphenyl); δ = 5.72 ppm, d, 1 H, J = 1.7 Hz (H1); δ = 5.57 ppm, dd, 1 H, J = 10.9, 3.4 Hz (H3); δ = 5.50 ppm, dd, 1 H, J = 3.5, 1.9 Hz (H2); δ = 5.41 – 5.35 ppm, m, 1 H (H4); δ = 4.08 – 4.02 ppm, m, 1 H (H5); δ = 3.95 ppm, s, 3 H (OC*H*$_3$); δ = 3.39 ppm, dd, 1 H, J = 13.3, 6.6 Hz (H6a); δ = 3.27 ppm, dd, 1 H, J = 13.4, 2.5 Hz (H6b); δ = 2.23 ppm, s, 3 H (OCOC*H*$_3$); δ = 2.07 ppm, s, 3 H (OCOC*H*$_3$); δ = 2.05 ppm, s, 3 H (OCOC*H*$_3$).

^{13}C NMR (125.8 MHz, $CDCl_3$): δ = 169.9 ppm, 1 C (O*C*OCH$_3$); δ = 169.8 ppm, 1 C (O*C*OCH$_3$); δ = 169.6 ppm, 1 C (O*C*OCH$_3$); δ = 166.8 ppm, 1 C (carbonyl C of biphenyl); δ = 152.8 ppm, 1 C (aromatic C of biphenyl); δ = 143.4 ppm, 1 C (aromatic C of biphenyl); δ = 134.6 ppm, 1 C (aromatic C of biphenyl); δ = 132.0 ppm, 1 C (aromatic C of biphenyl); δ = 130.3 ppm, 2 C (aromatic C of biphenyl); δ = 129.4 ppm, 1 C (aromatic C of biphenyl); δ = 126.8 ppm, 2 C (aromatic C of biphenyl); δ = 126.1 ppm, q, 1 C, J = 5.1 Hz (aromatic C of biphenyl); δ = 123.2 ppm, q, 1 C, J = 272.8 Hz (*C*F$_3$); δ = 120.4 ppm, q, 1 C, J = 31.8 Hz (aromatic C of biphenyl); δ = 115.6 ppm, 1 C (aromatic C of biphenyl); δ = 95.3 ppm, 1 C (C1); δ = 71.5 ppm, 1 C (C5); δ = 69.1 ppm, 1 C (C2); δ = 68.3 ppm, 1 C (C3); δ = 66.6 ppm, 1 C (C4); δ = 52.2 ppm, 1 C (O*C*H$_3$); δ = 51.0 ppm, 1 C (C6); δ = 20.9 ppm, 1 C (OCO*C*H$_3$); δ = 20.7 ppm, 1 C (OCO*C*H$_3$); δ = 20.6 ppm, 1 C (OCO*C*H$_3$).

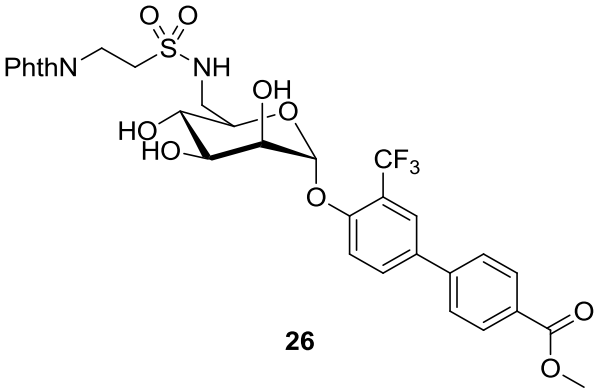
R_f = 0.56 with PE:EtOAc (2:1).

$[\alpha]^{20}_{D}$ = +35.8° (c = 1.00, $CHCl_3$).

ESI-MS for $C_{27}H_{26}F_3N_3O_{10}$: $m \cdot z^{-1}(M+Na^+)_{calc}$ = 632.15; $m \cdot z^{-1}(M+Na^+)_{obs}$ = 632.25.

ATR-FTIR (selected resonances): ν = 2104.6 cm^{-1} (azide stretching).

Methyl-4'-(2'',3'',4''-tetra-O-acetyl-6''-deoxy-6''-2'''-(phthalimido)ethylsulfonamido-α-D-mannopyranosyloxy)-3'-trifluoromethylbiphenyl-4-carboxylate



26

23 (101 mg, 166 µmol) was dissolved in anhydrous MeOH (5 ml) and 1 M MeONa (30 µl, 30 µmol). The reaction mixture was stirred for 4 h at room temperature under argon. Product formation was monitored *via* TLC (PE:EA (1:2)). The pH was adjusted to 6 to 7 using Amberlite IR120 (H^+) and solvents were removed *in vacuo*.

Next, the residue was dissolved in 1,4-dioxane (3 ml) and Pd/C (40 mg) was added under argon atmosphere in a twin-neck flask. The flask was flushed with H_2 at atmospheric pressure. The reaction mixture was stirred for overnight at room temperature. Product formation was monitored *via* TLC (2:1 (PE:EtOAc)) and ESI-MS. Pd/C was removed and solvents were removed *in vacuo*.

The residue was dissolved in anhydrous DMF (1.5 ml) and Et_3N (25 µl, 174 µmol) was added. 2-(phthalimido)ethanesulfonyl chloride (30 mg, 104 µmol) dissolved in DMF (750 µl) was slowly added under argon atmosphere at 0°C. The reaction mixture was stirred for 1 h at 0°C, allowed to heat up to room temperature and stirred for an additional 48 h. Solvents were evaporated *in vacuo* and the residue was purified *via* reversed-phase MPLC (gradient: 100% H2O to 100% MeOH, elution at 100% MeOH) to yield **26** (34 mg, 49 µmol, 41% over 3 steps) as a white solid.

^{1}H NMR (500.0 MHz, MeOD): δ = 8.03 – 7.98 ppm, m, 2 H (aromatic H of biphenyl); δ = 7.87 ppm, dd, 1 H, J = 8.6, 2.3 Hz (aromatic H of biphenyl); δ = 7.76 – 7.72 ppm, m, 5 H (aromatic H of biphenyl, NPhth); δ = 7.65 – 7.59 ppm, m, 3 H (aromatic H of biphenyl); δ = 5.73 ppm, d, 1 H, J = 1.5 Hz (H1); δ = 4.08 ppm, dd, 1 H, J = 3.5, 1.8 Hz (H2); δ = 4.03 – 3.87 ppm, m, 6 H (NCH_2CH$_2$S, H3, OCH_3); δ = 3.74 – 3.69 ppm, m, 1 H (H4); δ = 3.67 – 3.62 ppm, m, 1 H (H5); δ = 3.39 ppm, dd, 1 H, J = 14.1, 2.4 Hz (H6a); δ = 3.38 – 3.35 ppm, m, 3 H, J = 13.4, 2.5 Hz (H6b, NCH$_2$CH_2S).

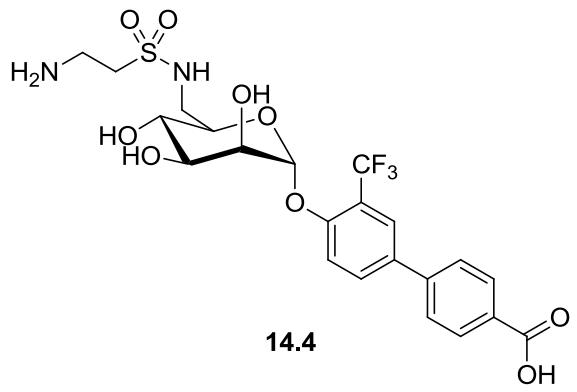
^{13}C NMR (125.8 MHz, MeOD): δ = 169.9 ppm, 2 C (carbonyl C of NPhth); δ = 168.2 ppm, 1 C (carbonyl C of biphenyl); δ = 155.1 ppm, 1 C (aromatic C of biphenyl); δ = 143.8 ppm, 1 C (aromatic C of biphenyl); δ = 135.4 ppm, 2 C (aromatic C of NPhth); δ = 134.6 ppm, 1 C (aromatic C of biphenyl); δ = 133.5 ppm, 1 C (aromatic C of biphenyl); δ = 133.2 ppm, 2 C (aromatic C of NPhth); δ = 131.2 ppm, 2 C (aromatic C of biphenyl); δ = 130.3 ppm, 1 C (aromatic C of biphenyl);); δ = 127.7 ppm, 2 C (aromatic C of biphenyl); δ = 126.4 ppm, q, 1 C, J = 5.2 Hz (aromatic C of biphenyl); δ = 125.0 ppm, q, 1 C, J = 272.2 Hz (*C*F_3); δ = 124.2 ppm, 2 C (aromatic C of NPhth); δ = 120.8 ppm, q, 1 C, J = 30.8 Hz (aromatic C of biphenyl); δ = 117.7 ppm, 1 C (aromatic C of biphenyl); δ = 99.8 ppm, 1 C (C1); δ = 74.7 ppm, 1 C (C5); δ = 71.9 ppm, 1 C (C3); δ = 71.6 ppm, 1 C (C2); δ = 69.1 ppm, 1 C (C4); δ = 52.7 ppm, 1 C (O*C*H_3); δ = 52.7 ppm, 1 C (NCH_2*C*H_2S); δ = 45.0 ppm, 1 C (C6); δ = 33.7 ppm, 1 C (N*C*$H_2$$CH_2$S).

R_f = 0.61 with 5% MeOH in DCM.

$[\alpha]^{20}_D$ = +57.8° (c = 1.00, MeOH).

ESI-MS for $C_{31}H_{29}F_3N_2O_{11}S$: $m{\cdot}z^{-1}(M\text{-}H^+)_{calc}$ = 693.14; $m{\cdot}z^{-1}(M\text{-}H^+)_{obs}$ = 693.31.

4'-(2'',3'',4''-tetra-O-acetyl-6''-deoxy-6''-2'''-aminoethylsulfonamido-α-D-mannopyranosyloxy)-3'-trifluoromethylbiphenyl-4-carboxylate



14.4

26 (33 mg, 48 µmol) was dissolved in MeOH (3.2 ml), hydrazine monohydrate (33 µl, 665 µmol) was added and the reaction mixture was stirred overnight at room temperature. Product formation was analyzed *via* MS and TLC (5% MeOH in DCM).

Next, 2 M aqueous NaOH (340 µl, 665 µmol) was added and the reaction mixture was stirred at room temperature for 4 h. More 2 M NaOH (340 µl, 665 µmol) was added and reaction mixture was stirred at room temperature overnight. Product formation was analyzed *via* ESI-MS. Solvents were evaporated *in vacuo* and the residue was purified *via* reversed-phase MPLC (gradient: 100% H_2O to 100% MeOH in 30 min) to yield **14.4** (12.3 mg, 22 µmol, 47% over 2 steps) as a white solid after lyophilization from H_2O.

^{1}H NMR (700.0 MHz, D_2O): δ = 8.04 – 8.00 ppm, m, 1 H (aromatic H of biphenyl); δ = 7.98 – 7.92 ppm, m, 3 H (aromatic H of biphenyl); δ = 7.76 – 7.70 ppm, m, 2 H (aromatic H of biphenyl); δ = 7.51 ppm, d, 1 H, J = 8.9 Hz (aromatic H of biphenyl); δ = 5.88 – 5.84 ppm, m, 1 H (H1); δ = 4.27 – 4.23 ppm, m, 1 H (H2); δ = 4.06 ppm, dd, 1 H, J = 9.5, 3.5 Hz (H3); δ = 3.71 – 3.66 ppm, m, 1 H (H4); δ = 3.65 – 3.60 ppm, m, 1 H (H5); δ = 3.53 ppm, dd, 1 H, J = 14.8, 2.1 Hz (H6a); δ = 3.25 ppm, dd, 1 H, J = 14.5, 8.1 Hz (H6b); δ = 3.12 – 2.98 ppm, m, 2 H ($NCH_2C\mathit{H}_2S$); δ = 2.86 – 2.80 ppm, m, 2 H ($NC\mathit{H}_2CH_2S$).

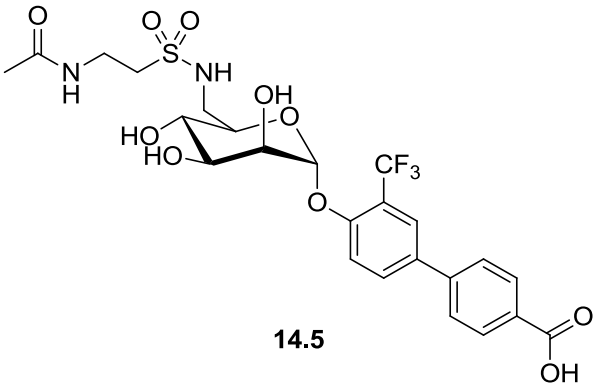
^{13}C NMR (176.0 MHz, D_2O): δ = 175.3 ppm, 1 C (carbonyl C of biphenyl); δ = 160.9 ppm, 1 C (aromatic C of biphenyl); δ = 152.1 ppm, 1 C (aromatic C of biphenyl); δ = 141.3 ppm, 1 C (aromatic C of biphenyl); δ = 135.4 ppm, 1 C (aromatic C of biphenyl); δ = 133.8 ppm, 1 C (aromatic C of biphenyl); δ = 132.1 ppm, 1 C (aromatic C of biphenyl); δ = 129.4 ppm, 2 C (aromatic C of biphenyl); δ = 126.5 ppm, 2 C (aromatic C of biphenyl); δ = 125.8 ppm, 1 C (aromatic C of biphenyl); δ = 123.5 ppm, q, 1 C, J = 272.4 Hz ($\mathit{C}F_3$); δ = 119.1 ppm, q, 1 C, J = 30.9 Hz (aromatic C of biphenyl); δ = 116.1 ppm, 1 C (aromatic C of biphenyl); δ = 96.7 ppm, 1 C (C1); δ = 72.7 ppm, 1 C (C5); δ = 70.1 ppm, 1 C (C3); δ = 69.6 ppm, 1 C (C2); δ = 67.8 ppm, 1 C (C4); δ = 54.2 ppm, 1 C ($NCH_2\mathit{C}H_2S$); δ = 43.1 ppm, 1 C (C6); δ = 35.3 ppm, 1 C ($N\mathit{C}H_2CH_2S$).

R_f = no migration with DCM:MeOH:aqueous NH_4OH (8:2:0.4).

$[\alpha]^{20}_D$ = +71.6° (c = 0.10, H_2O).

HR ESI-MS for $C_{22}H_{25}F_3N_2O_9S$: $m \cdot z^{-1}(M+H^+)_{calc}$ = 551.131; $m \cdot z^{-1}(M+H^+)_{obs}$ = 551.132.

4'-(2'',3'',4''-tetra-O-acetyl-6''-deoxy-6''-2'''-(acetamido)ethylsulfonamido-α-D-mannopyranosyloxy)-3'-trifluoromethylbiphenyl-4-carboxylate



14.5

14.4 (5.2 mg, 9.5 µmol) was dissolved in MeOH (1 ml) and acetic anhydride (5.4 µl, 57 µmol) was added dropwise at room temperature under argon atmosphere. The reaction mixture was stirred for 4 h. Product formation was monitored *via* ESI-MS and TLC (10% MeOH in DCM). Solvents were evaporated *in vacuo* and the residue was purified *via* reversed-phase MPLC (gradient 100% H2O to 100% MeOH in 35 min) to yield **14.5** (3.98 mg, 6.7 µmol, 70.7%) as a white solid after lyophilization from H_2O. A purity higher than 95% was demonstrated *via* analytical reversed-phase HPLC.

^{1}H NMR (500.0 MHz, D_2O): δ = 8.00 – 7.90 ppm, m, 4 H (aromatic H of biphenyl); δ = 7.74 – 7.68 ppm, m, 2 H (aromatic H of biphenyl); δ = 7.47 ppm, d, 1 H, J = 8.8 Hz (aromatic H of biphenyl); δ = 5.85 ppm, d, 1 H, J = 1.5 Hz (H1); δ = 4.24 ppm, dd, 1 H, J = 3.6, 1.8 Hz (H2); δ = 4.05 ppm, dd, 1 H, J = 9.6, 3.5 Hz (H3); δ = 2.67 – 2.61 ppm, m, 1 H (H4); δ = 3.60 – 3.53 ppm, m, 2 H (H5, H6a); δ = 3.23 ppm, dd, 1 H, J = 15.6, 9.3 Hz (H6b); δ = 3.19 – 3.00 ppm, m, 2 H (NC$H_2$$CH_2$S); δ = 3.09 – 2.86 ppm, m, 2 H ($NCH_2C$$H_2$S); δ = 1.70 ppm, s, 3 H (NCOCH_3).

^{13}C NMR (125.7 MHz, D_2O): δ = 175.2 ppm, 1 C (carbonyl C of biphenyl); δ = 173.7 ppm, 1 C (N*C*OCH_3); δ = 161.3 ppm, 1 C (aromatic C of biphenyl); δ = 151.8 ppm, 1 C (aromatic C of biphenyl); δ = 141.1 ppm, 1 C (aromatic C of biphenyl); δ = 135.3 ppm, 1 C (aromatic C of biphenyl); δ = 133.5 ppm, 1 C (aromatic C of biphenyl); δ = 131.9 ppm, 1 C (aromatic C of biphenyl); δ = 129.6 ppm, 2 C (aromatic C of biphenyl); δ = 126.3 ppm, 2 C (aromatic C of biphenyl); δ = 125.6 ppm, q, 1 C, J = 5.4 Hz (aromatic C of biphenyl); δ = 123.5 ppm, q, 1 C, J = 272.0 Hz (*C*F_3); δ = 118.9 ppm, q, 1 C, J = 30.9 Hz (aromatic C of biphenyl); δ = 115.8 ppm, 1 C (aromatic C of biphenyl); δ = 96.3 ppm, 1 C (C1); δ = 72.8 ppm, 1 C (C5); δ = 70.0 ppm, 1 C (C3); δ = 69.5 ppm, 1 C (C2); δ = 67.9 ppm, 1 C (C4); δ = 51.3 ppm, 1 C (NCH_2*C*H_2S); δ = 43.2 ppm, 1 C (C6); δ = 33.5 ppm, 1 C (N*C*$H_2$$CH_2$S) δ = 21.4 ppm, 1 C (NCO*C*H_3).

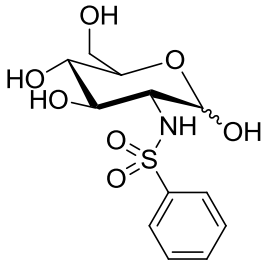
R_f = 0.15 with 10% MeOH in DCM.

$[\alpha]^{20}_D$ = +74.8° (c = 0.05, H_2O).

HR ESI-MS for $C_{24}H_{27}F_3N_2O_{10}S$: $m \cdot z^{-1}(M+Na^+)_{calc}$ = 615.124; $m \cdot z^{-1}(M+Na^+)_{obs}$ = 615.125.

2.3.4. Glucosamine-2-Sulfate Analogs

2-deoxy-2-phenylsulfonylamido-D-glucopyranose



31.1

Glucosamine hydrochloride (50 mg, 232 µmol) was dissolved in anhydrous MeOH (2.3 ml) *via* the addition of Et_3N (81 µl, 580 µmol). Benzenesulfonyl chloride (178 µl, 1.39 mmol) was added dropwise at 0°C under argon atmosphere. The reaction was allowed to heat up to room temperature and stirred for 2 h. More Et_3N (32 µl, 230 µmol) and benzenesulfonyl chloride (89 µl, 695 µmol) were added dropwise at 0°C. The reaction was again allowed to heat up to room temperature and stirred for another 2h. Solvents were evaporated i*n vacuo* and the residue was purified *via* reversed-phase column chromatography (gradient: H_2O, 1%, 5% MeOH in H_2O and elution with 10% MeOH in H_2O) and subsequently *via* column chromatography (gradient: DCM, 5% MeOH, 10% MeOH in DCM and elution with 20% MeOH in DCM). Silica particles were removed *via* filtration in MeOH with a cellulose acetate membrane at a pore size 0.2 µm to yield to yield an α/β-mixture of **31.1** (23.9 mg, 75 µmol, 32%) as a white solid.

The ratio of α- and β-anomer was determined to be 5:1 *via* ^{1}H NMR. Here, only chemical shifts corresponding to the β-anomer are documented. As a result of the low signal intensities for the α-anomer the corresponding resonances were not assigned.

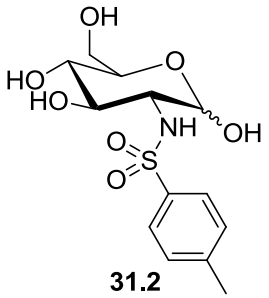
^{1}H NMR (400.0 MHz, MeOD, β-anomer): δ = 7.96 – 7.87 ppm, m, 2 H (aromatic H); δ = 7.64 – 7.45 ppm, m, 3 H (aromatic H); δ = 4.78 ppm, d, 1 H , J = 3.5 Hz (H1); δ = 3.76 – 3.68 ppm, m, 2 H (H6a, H5); δ = 3.69 – 3.55 ppm, m, 2 H (H6a, H3); δ = 3.30 – 3.24 ppm, m, 1 H (H4); δ = 3.13 ppm, dd, 1 H, J = 10.1, 3.5 Hz (H2).

^{13}C NMR (100.6 MHz, MeOD, β-anomer): δ = 143.4 ppm, 1 C (aromatic C); δ = 133.4 ppm, 1 C (aromatic C); δ = 130.0 ppm, 2 C (aromatic C); δ = 128.0 ppm, 2 C (aromatic C); δ = 92.9 ppm, 1 C (C1); δ = 72.8 ppm, 1 C (C5); δ = 72.5 ppm, 1 C (C3); δ = 72.2 ppm, 1 C (C4); δ = 62.6 ppm, 1 C (C6); δ = 59.9 ppm, 1 C (C2).

R_f = 0.50 with 20% MeOH in DCM.

HR ESI-MS for $C_{12}H_{17}NO_7S$: $m \cdot z^{-1}(M+Na^+)_{calc}$ = 342.062; $m \cdot z^{-1}(M+Na^+)_{obs}$ = 342.063.

2-deoxy-2-O-tosyl-D-glucopyranose[6]



31.2

Glucosamine hydrochloride (3.0 g, 13.9 mmol) were dissolved in acetone (15 ml). Tosyl chloride (2.6 g, 13.9 mmol) was added and reaction was stirred for 4 h at room temperature. Solvents were removed *in vacuo* and the residue was purified *via* MPLC (15% MeOH in DCM) to yield an α/β-mixture of **31.2** (4.1 g, 12.5 mmol, 90%) as a white solid.

The ratio of α- and β-anomer was determined to be 10:1 *via* ^{1}H NMR. Here, only chemical shifts corresponding to the β-anomer are documented. As a result of the low signal intensities for the α-anomer the corresponding resonances were not assigned.

^{1}H NMR (400.0 MHz, MeOD, β-anomer): δ = 7.82 – 7.77 ppm, m, 2 H (aromatic H); δ = 7.37 – 7.32 ppm, m, 2 H (aromatic H); δ = 4.77 ppm, d, 1 H , J = 3.5 Hz (H1); δ = 3.76 – 3.69 ppm, m, 2 H (H6a, H5); δ = 3.68 – 3.55 ppm, m, 2 H (H6a, H3); δ = 3.30 – 3.24 ppm, m, 1 H (H4); δ = 3.09 ppm, dd, 1 H, J = 10.1, 3.6 Hz (H2); δ = 2.41 ppm, s, 3 H (CH_3) .

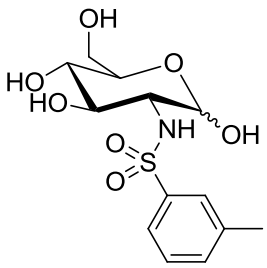
^{13}C NMR (100.6 MHz, MeOD, β-anomer): δ = 144.4 ppm, 1 C (aromatic C); δ = 140.4 ppm, 1 C (aromatic C); δ = 130.6 ppm, 2 C (aromatic C); δ = 128.1 ppm, 2 C (aromatic C); δ = 92.9 ppm, 1 C (C1); δ = 72.8 ppm, 1 C (C5); δ = 72.5 ppm, 1 C (C3); δ = 72.2 ppm, 1 C (C4); δ = 62.6 ppm, 1 C (C6); δ = 59.8 ppm, 1 C (C2); δ = 21.5 ppm, 1 C (CH_3).

[6] The procedure was devised and conducted by Gunnar Bachem (Humboldt-Universität zu Berlin).

R_f = 0.38 with 10% MeOH in DCM.

HR ESI-MS for $C_{13}H_{19}NO_7S$: $m \cdot z^{-1}(M+Na^+)_{calc}$ = 356.078; $m \cdot z^{-1}(M+Na^+)_{obs}$ = 356.077.

2-deoxy-2-3'-(methyl)phenylsulfonylamido-D-glucopyranose[7]



31.3

Glucosamine hydrochloride (49 mg, 227 µmol) was dissolved in DMSO (4.6 ml) *via* addition of DIPEA (240 µl, 1.39 mmol). Next, 3-methylphenylsulfonyl chloride (84 µl, 580 µmol) was slowly added and the reaction mixture was stirred for 2 h at room temperature. The reaction was quenched with MeOH and solvents were removed *in vacuo*. Residual DMSO was removed *via* lyophilization. The residue was purified *via* column chromatography (gradient: DCM, 5% MeOH in DCM and elution with 10% MeOH in DCM) and reversed-phase column chromatography (gradient: H_2O, 1%, 5% MeOH in H_2O and elution with 10% MeOH in H_2O) to yield **31.3** (12.9 mg, 39 µmol, 17%) as a white solid after lyophilization from H_2O.

The ratio of α- and β-anomer was determined to be 10:1 *via* ^{1}H NMR. Here, only chemical shifts corresponding to the β-anomer are documented. As a result of the low signal intensities for the α-anomer the corresponding resonances were not assigned.

^{1}H NMR (400.0 MHz, MeOD, β-anomer): δ = 7.78 – 7.66 ppm, m, 2 H (aromatic H); δ = 7.45 – 7.38 ppm, m, 2 H (aromatic H); δ = 4.78 ppm, d, 1 H , J = 3.6 Hz (H1); δ = 3.76 – 3.68 ppm, m, 2 H (H6a, H5); δ = 3.68 – 3.56 ppm, m, 2 H (H6a, H3); δ = 3.30 – 3.24 ppm, m, 1 H (H4); δ = 3.12 ppm, dd, 1 H, J = 10.5, 3.6 Hz (H2); δ = 2.42 ppm, s, 3 H (CH_3) .

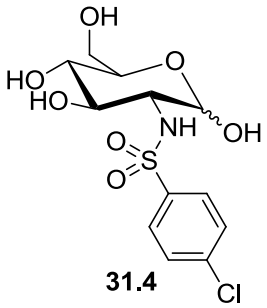
^{13}C NMR (100.6 MHz, MeOD, β-anomer): δ = 143.2 ppm, 1 C (aromatic C); δ = 140.4 ppm, 1 C (aromatic C); δ = 134.1 ppm, 1 C (aromatic C); δ = 129.9 ppm, 1 C (aromatic C); δ = 128.3 ppm, 1 C (aromatic C); δ = 125.1 ppm, 1 C (aromatic C); δ = 92.9 ppm, 1 C (C1); δ = 72.8 ppm, 1 C (C5); δ = 72.6 ppm, 1 C (C3); δ = 72.2 ppm, 1 C (C4); δ = 62.7 ppm, 1 C (C6); δ = 59.9 ppm, 1 C (C2); δ = 21.3 ppm, 1 C (CH_3).

[7] The procedure was conducted by David Hartmann (University of Cambridge) under my supervision.

R_f = 0.34 with 10% MeOH in DCM.

HR ESI-MS for $C_{12}H_{19}NO_7S$: $m \cdot z^{-1}(M+Na^+)_{calc}$ = 356.078; $m \cdot z^{-1}(M+Na^+)_{obs}$ = 356.079.

2-deoxy-2-4'-(chloro)phenylsulfonylamido-D-glucopyranose[8]



31.4

Glucosamine hydrochloride (51 mg, 234 mmol) was dissolved in anhydrous MeOH (2.3 ml) *via* the addition of DIPEA (240 µl, 1.39 mmol). 4-chlorophenylsulfonyl chloride (196 mg, 928 µmol) was dissolved in anhydrous MeOH (1.0 ml) and slowly added. The reaction mixture was stirred 4 h at room temperature. Solvents were removed *in vacuo* and the residue was purified *via* reversed-phase column chromatography (gradient: H_2O, 1% MeOH in H_2O and elution with 5% MeOH in H_2O) to yield **31.4** (35 mg, 99 µmol, 42%) as a white solid after lyophilization from H_2O.

The ratio of α- and β-anomer was determined to be 4:1 *via* ^{1}H NMR. Here, only chemical shifts corresponding to the β-anomer are documented. As a result of the low signal intensities for the α-anomer the corresponding resonances were not assigned.

^{1}H NMR (400.0 MHz, MeOD, β-anomer): δ = 7.92 – 7.87 ppm, m, 2 H (aromatic H); δ = 7.56 – 7.51 ppm, m, 2 H (aromatic H); δ = 4.86 ppm, d, 1 H , J = 3.6 Hz (H1); δ = 3.77 – 3.69 ppm, m, 2 H (H6a, H5); δ = 3.68 – 3.55 ppm, m, 2 H (H6a, H3); δ = 3.29 – 3.23 ppm, m, 1 H (H4); δ = 3.14 ppm, dd, 1 H, J = 10.2, 3.5 Hz (H2).

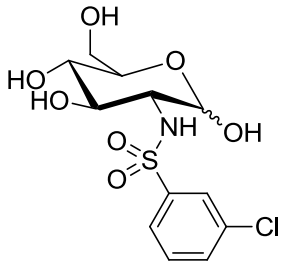
^{13}C NMR (100.6 MHz, MeOD, β-anomer): δ = 142.3 ppm, 1 C (aromatic C); δ = 139.4 ppm, 1 C (aromatic C); δ = 130.1 ppm, 2 C (aromatic C); δ = 129.8 ppm, 2 C (aromatic C); δ = 93.2 ppm, 1 C (C1); δ = 72.8 ppm, 1 C (C5); δ = 72.5 ppm, 1 C (C3); δ = 72.3 ppm, 1 C (C4); δ = 62.6 ppm, 1 C (C6); δ = 59.9 ppm, 1 C (C2).

R_f = 0.30 with 10% MeOH in DCM.

HR ESI-MS for $C_{12}H_{16}ClNO_7S$: $m \cdot z^{-1}(M+Na^+)_{calc}$ = 376.023; $m \cdot z^{-1}(M+Na^+)_{obs}$ = 376.024.

[8] The procedure was conducted by David Hartmann (University of Cambridge) under my supervision.

2-deoxy-2-3'-(chloro)phenylsulfonylamido-D-glucopyranose[9]



31.5

Glucosamine hydrochloride (50 mg, 234 µmol) was dissolved in anhydrous MeOH (2.3 ml) *via* addition of DIPEA (240 µl, 1.39 mmol). 3-chlorophenylsulfonyl chloride (130 µl, 930 µmol) was slowly added and the reaction mixture was stirred for 4 h at room temperature. Solvents were removed *in vacuo* and the residue was purified *via* reversed-phase column chromatography (gradient: H_2O, 1% MeOH in H_2O and elution with 5% MeOH in H_2O) to yield **31.5** (6.0 mg, 17 µmol, 7%) as a white solid after lyophilization from H_2O.

The ratio of α- and β-anomer was determined to be 8:1 *via* ^{1}H NMR. Here, only chemical shifts corresponding to the β-anomer are documented. As a result of the low signal intensities for the α-anomer the corresponding resonances were not assigned.

^{1}H NMR (400.0 MHz, MeOD, β-anomer): δ = 7.97 – 7.79 ppm, m, 2 H (aromatic H); δ = 7.63 – 7.47 ppm, m, 2 H (aromatic H); δ = 4.87 ppm, m, 1 H (H1); δ = 3.78 – 3.69 ppm, m, 2 H (H6a, H5); δ = 3.69 – 3.55 ppm, m, 2 H (H6a, H3); δ = 3.30 – 3.24 ppm, m, 1 H (H4); δ = 3.12 ppm, dd, 1 H, J = 10.6, 3.5 Hz (H2).

^{13}C NMR (100.6 MHz, MeOD, β-anomer): δ = 145.5 ppm, 1 C (aromatic C); δ = 135.8 ppm, 1 C (aromatic C); δ = 133.2 ppm, 1 C (aromatic C); δ = 131.6 ppm, 1 C (aromatic C); δ = 128.0 ppm, 1 C (aromatic C); δ = 126.4 ppm, 1 C (aromatic C); δ = 93.2 ppm, 1 C (C1); δ = 72.8 ppm, 1 C (C5); δ = 72.5 ppm, 1 C (C3); δ = 72.3 ppm, 1 C (C4); δ = 62.6 ppm, 1 C (C6); δ = 60.0 ppm, 1 C (C2).

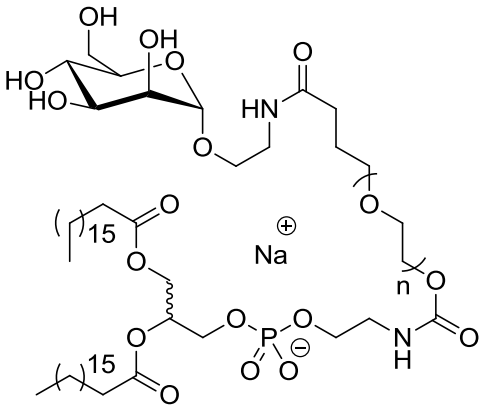
R_f = 0.51 with 10% MeOH in DCM.

HR ESI-MS for $C_{12}H_{16}ClNO_7S$: $m \cdot z^{-1}(M+Na^+)_{calc}$ = 376.023; $m \cdot z^{-1}(M+Na^+)_{obs}$ = 376.024.

[9] The procedure was conducted by David Hartmann (University of Cambridge) under my supervision.

2.3.5. Glycolipids

12.29-PEG-DSPE[10]



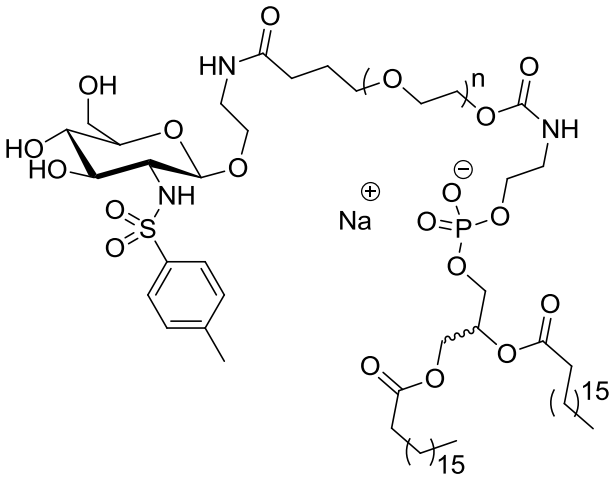
42

NHS-activated PEG-DSPE (3.1 kDa, NOF) (2.5 mg, 810 nmol) was dissolved in DMF (130 µl) and added to **12.29** (1.5 mg, 6.70 µmol) dissolved in 0.1 M $NaCO_3H$ at pH 8.4 (1.1 ml). The reaction mixture was stirred overnight at room temperature. Solvents were removed *in vacuo* and the residue was dissolved in 0.1 M $NaCO_3H$ at pH 8.4 (2.0 ml). Next, the reaction mixture was dialyzed three times against 0.1 M $NaCO_3H$ at pH 8.4 (0.5 ml) and three times against H_2O (0.5 ml). After lyophilization from H_2O, **42** (2.1 mg, 530 nmol, 66%) was obtained as a white solid. Product formation was analyzed *via* ^{1}H NMR and the yield was calculated *via* the integration of characteristic resonances.

^{1}H NMR (400.0 MHz, DMSO-d_6, characteristic resonances): δ = 4.61 – 4.58 ppm, m, 1 H (H1, **12.29**): δ = 0.85 ppm, t, 3 H (two times $CH_2C\mathit{H}_3$).

[10] The procedure was devised by Jessica Schulze (Max Planck Institute of Colloids and Interfaces) and me and conducted by Jessica Schulze.

41.1-PEG-DSPE[11]



43

NHS-activated PEG-DSPE (3.1 kDa, NOF) (2.8 mg, 900 nmol) was dissolved in DMF (140 µl) and added to **41.1** (2.7 mg, 7.20 µmol) dissolved in 0.1 M $NaCO_3H$ at pH 8.4 (1.2 ml). The reaction mixture was stirred overnight at room temperature. Solvents were removed *in vacuo* and the residue was dissolved in 0.1 M $NaCO_3H$ at pH 8.4 (2.0 ml). Next, the reaction mixture was dialyzed three times against 0.1 M $NaCO_3H$ at pH 8.4 (0.5 ml) and three times against H_2O (0.5 ml). After lyophilization from H_2O, **43** (2.4 mg, 710 nmol, 79%) was obtained as a white solid. Product formation was analyzed *via* ^{1}H NMR and the yield was calculated *via* the integration of characteristic resonances.

^{1}H NMR (400.0 MHz, DMSO-d_6, characteristic resonances): δ = 7.30 – 7.10 ppm, m, 2 H (aromatic H, **41.1**); δ = 0.85 ppm, t, 3 H (two times CH_2CH_3).

[11] The procedure was devised by Jessica Schulze (Max Planck Institute of Colloids and Interfaces) and me and conducted by Jessica Schulze.

2.4. ^{19}F R_2-filtered NMR

2.4.1. General Remarks

^{19}F NMR and ^{19}F R_2-filtered NMR experiments were conducted on a PremiumCompact 600 MHz spectrometer (Agilent). Spectra were processed in MestReNova and data analysis was performed with OriginPro (Mestrelab Research, 2016; OriginLab, 2015). Unless stated otherwise, experiments utilizing the Langerin ECD were performed at a receptor concentration of 50 µM in 25 mM Tris with 10% D_2O, 150 mM NaCl and 5 mM $CaCl_2$ at pH 7.8 and 25° C. Experiments utilizing the Langerin CRD or the DC-SIGN CRD were performed at a receptor concentration of 50 µM in 25 mM HEPES with 10% D_2O, 150 mM NaCl and 5 mM $CaCl_2$ at pH 7.0 and 25°C. TFA served as an internal reference at a concentration of 50 or 100 µM. For each spectrum 128 scans were recorded in 3 mm sample tubes at sample volumes of 150 µl. Relaxation rates $R_{2,obs}$ were determined with the CPMG pulse sequence by fitting Equation 1 to integrals of the ^{19}F resonance of **5.1** (Carr H. Y. and M., 1954; Meiboom and Gill, 1958). T represents the relaxation time and I_0 is the integral at a T value of 0 s. The relaxation delay d_1 was set to 2.0 s, the acquisition time t_{acq} was set to 0.8 s and the frequency of 180° pulses ν_{CPMG} was set to 500 Hz.

$$I = I_0 e^{-R_{2,obs}T}$$

Equation 1

2.4.2. K_D Determination

K_D values were determined in titration experiments at five ligand concentrations $[L]_T$. Samples were prepared *via* serial dilution. The K_D and the $R_{2,b}$ value of the reporter molecule **5.1** were derived from Equation 3 by detection of ^{19}F NMR relaxation rates $R_{2,obs}$ in a two parameter fit (Dalvit, 2007; Roehrl *et al.*, 2004). $R_{2,b}$ represents the relaxation rate in bound state of the ligand and p_b is the bound fraction of the ligand while $[P]_T$ represents the concentration of the receptor. The relaxation rate of the free ligand $R_{2,f}$ was measured at 0.1 mM **5.1** in absence of the receptor. An EDTA control experiment was conducted at 12.5 mM **5.1** for the Langerin ECD and CRD and at 0.1 mM **5.1** for the DC-SIGN CRD. To ensure the validity of Equation 3, the chemical exchange contribution $R_{2,ex}$ was estimated by ^{19}F NMR relaxation dispersion experiments at 0.1 mM **5.1** in presence of receptor. For the Langerin ECD and the DC-SIGN CRD the K_D and the $R_{2,b}$ value of the reporter molecule **5.1** was determined from three independent titration. Standard errors for the Langerin CRD were derived directly from the fitting procedure.

$$R_{2,obs} = R_{2,f} + (R_{2,b} - R_{2,f})p_b$$

with

$$p_b = \frac{[P]_T + [L]_T + K_D - \sqrt{([P]_T + [L]_T + K_D)^2 - 4[P]_T[L]_T}}{2[L]_T}$$

Equation 3

2.4.3. K_I Determination

K_I values were determined in competitive binding experiments *via* the detection of binding of 0.1 mM **5.1** to either the Langerin ECD or the DC-SIGN CRD at five or more competitor concentrations $[I]_T$. Samples were prepared *via* serial dilution. Equation 4 served to derive $[P]_T$ and K_I values from $R_{2,obs}$ values in a two parameter fit (Roehrl *et al.*, 2004). This procedure is herein termed Setup 1.

$$R_{2,obs} = R_{2,f} + (R_{2,b} - R_{2,f})p_b$$

with

$$p_b = \frac{2cos\left(\frac{\theta}{3}\right)\sqrt{a^2 - 3b} - a}{3K_D + 2cos\left(\frac{\theta}{3}\right)\sqrt{a^2 - 3b} - a}$$

and

$$\theta = cos^{-1}\left(\frac{-2a^3 + 9ab - 27c}{2\sqrt{(a^2 - 3b)^3}}\right), \qquad a = K_D + K_I + [L]_T + [I]_T - [P]_T,$$

$$b = ([I]_T - [P]_T)K_D + ([L]_T - [P]_T)K_D + K_I K_D, \qquad c = -K_I K_D [P]_T$$

Equation 4

The effects of 10% DMSO and 0.01% Tween-20 on the determined K_I values were evaluated in titration experiments with Man using the Langerin ECD. Titration experiments with naturally occurring monosaccharide derivative (Dextra Laboratories), ManNAc analogs **5**, Man analog **12.30** and GlcNS analogs **31** and **41** were conducted in the absence of these additives. In contrast, K_I values for Man analogs **12.2**, **12.5**, **12.11**[12], **12.24**, **14.1** and **14.5** were determined in presence of 10% DMSO and 0.01% Tween-20. For the acids Man-6-OS, Galc-6-OS, GlcNS, GlcNAc-6-OS, GlcNS-6-OS,

[12] ^{19}F R_2-filtered NMR experiments to evaluate changes $R_{2,obs}$ for the trifluoromethyl group of **12.11** in presence of the Langerin ECD were conducted under the same experimental conditions.

5.11, **12.5**, **12.11** and **14.5** the pH values of the NMR samples were monitored. The pH value of stock solutions for **5.11** was adjusted to 7.8 using 1 M NaOH. K_I values of Man and ManNAc for the Langerin ECD were determined from thee independent titrations. Due to the limited availability of other molecules, standard errors were derived directly from the fitting procedures.

2.4.4. Screening Setup

Estimated K_I values were determined in competitive binding experiments *via* the detection of binding of 0.1 mM **5.1** to the Langerin ECD at a single competitor concentration. Here, $[P]_T$ values were directly calculated from $R_{2,max}$, the data point at 0.1 mM **5.1** and in absence of competitor, *via* Equation 3. Subsequently, K_I values were estimated from Equation 4 in a one parameter fit. This procedure is herein termed Setup 2.

The assay performance in Setup 2 was evaluated in a retrospective analysis of the titration experiments with Man, ManNAc and ManNAc analogs **5**. Data points were selected for this evaluation if the competitor concentrations fell within one order of magnitude of the K_I value determined in Setup 1. De*via*tions from Setup 1 were quantified *via* the calculation of ΔK_I values. Setup 2 was found to be well suited for affinity ranking in screening applications.

The fragment screening was conducted in presence of 10% DMSO and 0.01% Tween-20. The mean $\mu_{Reference}$ and standard de*via*tion σ of $R_{2,obs}$ values in absence of competitor were estimated from nine independent experiments. Overall, 290 fragments (Key Organics) were randomly selected from our in-house library (Appendix A.5). These fragments were binned into mixtures of five or six and screened at concentrations of 0.5 mM. A 3σ-threshold was utilized to define screening hits and estimated K_I values for Man and fragment **8** were determined in Setup 2 were determined from three independent experiments.

A library of Man analogs **12**[13] and **14**[14] was screened in Setup 2 in presence of 0% to 10% DMSO at concentration between 0.1 and 10.0 mM (Hauck *et al.*, 2013; Pang *et al.*, 2012) (Appendix B.4). Overall, 26 Man analogs were screened. Estimated K_I values were determined for all Man analogs that displayed a decreased $R_{2,obs}$ value. The solubility of the Man analogs was assessed either visually or in selected cases *via* ^{1}H NMR experiments using TSP-d_6 as an internal reference at 0.1 mM.

2.4.5. IC_{50} Determination for Multivalent Glycomimetics

Due to the complex equilibria and cooperativity effects present in competitive binding experiments with multivalent glycomimetics, IC_{50} values rather than K_I values were utilized to quantify affinities.

[13] These structures were provided by Prof. Dr. Beat Ernst (Universität Basel).
[14] These structures were provided by Dr. Alexander Titz (Helmholtz Center for Infection Research).

IC_{50} values were determined in competitive binding experiments *via* the detection of binding of 0.1 mM **5.1** to either the Langerin ECD at six competitor concentrations. Samples were prepared *via* serial dilution. Equation 5 served to derive IC_{50} values and Hills factors p from $R_{2,obs}$ values in a two parameter fit (Stefan and Le Novère, 2013). Standard errors were derived directly from the fitting procedures. $R_{2,max}$ represents the relaxation rate at 0.1 mM **5.1** in presence of receptor and in absence of competitor.

$$R_{2,obs} = R_{2,f} + \left(\frac{R_{2,max} - R_{2,f}}{1 + \left(\frac{IC_{50}}{[I]_T} \right)^p} \right)$$

Equation 5

Titration experiments with divalent glycomimetics using DNA-based scaffold were conducted in presence of 10% DMSO.[15] To prepare these samples for ^{19}F R_2-filtered NMR experiments, stock solution of the divalent glycomimetics in 25 mM Tris with 10% D_2O, 10% DMSO, 150 mM NaCl and 5 mM $CaCl_2$ at pH 7.8 were heated to 90°C. Subsequently, the samples were slowly cooled to 25°C to allow for hybridization of the DNA template and with the functionalized PNA strands. ^{19}F R_2-filtered NMR experiments with second generation divalent glycomimetics were conducted at a decreased Langerin ECD concentration of 25 µM. This setup is herein termed Setup 3. Titration experiments with trivalent glycoclusters were conducted in absence of DMSO.[16]

[15] These structures were provided by Prof. Dr. Oliver Seitz (Humboldt-Universität zu Berlin).
[16] These structures were provided by Prof. Dr. Laura Hartmann (Heinrich-Heine-Universität Düsseldorf).

2.5. STD NMR

2.5.1. General Remarks

STD NMR experiments were conducted on a PremiumCompact 600 MHz spectrometer (Agilent) (Mayer and Meyer, 1999). Spectra were processed in MestReNova and data analysis was performed with OriginPro (Mestrelab Research, 2016; OriginLab, 2015). Experiments were conducted utilizing the Langerin ECD in 25 mM Tris-d_{11} (Eurisotope) with 100% D_2O, 150 mM NaCl and 5 mM $CaCl_2$ at pH 7.8 and 25° C. Experiments were repeated in absence of receptor to exclude STD effects due to direct saturation of fragments or monosaccharide analogs. Residual H_2O or TSP-d_6 at 0.1 mM served as an internal reference. Spectra were recorded in 5 mm sample tubes at sample volumes of 500 µl. Saturation was implemented *via* a train of 50 ms Gauss pulses at varying saturation times t_{sat}. The on-resonance irradiation frequency ν_{sat} was set to 0.0 ppm and the off-resonance irradiation frequency ν_{ref} was set to 80.0 ppm. The acquisition time t_{acq} was set to 2.0 s and the DPFGSE pulse sequence was utilized for solvent suppression (Hwang and Shaka, 1995). Receptor resonances were suppressed *via* a $T_{1,rho}$ filter at a relaxation time τ of 35 ms.

2.5.2. Screening Setup

Fragment mixtures displaying inhibition in the ^{19}F R_2-filtered NMR screening were deconvoluted *via* STD NMR. These mixtures were screened at fragment concentrations of 0.5 mM in presence of 25 µM Langerin ECD. For each spectrum 512 scans were recorded. The relaxation delay d_1 was set to 0 s and the saturation time t_{sat} was set to 4 s. STD spectra were generated *via* the internal subtraction of on- and off-resonance spectra during acquisition. Resonances of the analyzed fragment mixture were assigned by comparison to previously acquired ^1H NMR spectra of the individual fragments.

2.5.3. Epitope Mapping

Binding epitopes were determined for Man analogs **12.2**, **12.11**, **12.30** and **14.5** as well as GlcNS analog **41.2** at concentrations of 500 µM in presence of 50 µM Langerin ECD. Experiments with Man analogs **12.2**, **12.11**, **14.1** and **14.5** were conducted in presence of 10% DMSO. For each spectrum 512 scans were recorded. The relaxation delay d_1 was set to 6 s and spectra were recorded at 5 different saturation time t_{sat} varying from 0.25 to 6.00 s. Equation 6 served to derive the STD effect for each analyzed resonance from the corresponding on- and off-resonance spectra (Mayer and Meyer, 2001). I_0 represents the integral of a resonance in the off-resonance spectrum and I_{sat} represents the integral of a resonance in the on-resonance spectrum.

$$STD = \frac{I_0 - I_{sat}}{I_0}$$

Equation 6

The apparent saturation rate k_{sat} and the maximal STD effect STD_{max} were derived from Equation 7 in a two parameter fit (Angulo and Nieto, 2011). Standard errors were derived directly from the fitting procedures. These parameters were utilized to calculate the initial slope of the STD build-up curves STD'_0 *via* Equation 8. STD'_0 values were normalized and mapped on the corresponding Man analog structures. Resonances of the analyzed Man analogs were assigned by means of COSY, TOCSY, ^{13}C HSQC and ^{13}C H2BC NMR experiments (Petersen *et al.*, 2006). Only resonances for which at least part of a multiplet was isolated were considered for the epitope mapping. Here, H3 of Man analog **12.2** represents an exception. The individual parameters for H3 were derived from a linear combination of Equation 7 in a two parameter fit to the combined STD values for H3 and H6b. STD_{max} and k_{sat} for H6b were derived independently from the analysis of an isolated region of the H6b multiplet.

$$STD = STD_{max}\left(1 - e^{-k_{sat}t_{sat}}\right)$$

Equation 7

$$STD'_0 = STD_{max}k_{sat}$$

Equation 8

2.6. ^{15}N HSQC NMR

2.6.1. General Remarks

^{15}N HSQC NMR experiments were conducted on a PremiumCompact 600 MHz spectrometer (Agilent) or an Ascend 700 MHz spectrometer (Bruker) (Bodenhausen and Ruben, 1980). Spectra were processed in NMRPipe (Delaglio *et al.*, 1995). Data analysis was performed using CCPN Analysis, MatLab and OriginPro (MathWorks, 2016b; OriginLab, 2015; Vranken *et al.*, 2005). Experiments were conducted with the Langerin CRD in 25 mM HEPES with 10% D_2O, 150 mM NaCl and 5 mM $CaCl_2$ at pH 7.0 and 25°C. DSS-d_6 served as an internal reference at a concentration of 100 µM. Spectra were referenced *via* the internal spectrometer reference. On the PremiumCompact 600 MHz spectrometer (Agilent), spectra were acquired with 128 increments and 8 scans per increment using 500 µl samples in 5 mm sample tubes. Langerin CRD concentrations ranged from 160 to 200 µM. The relaxation delay d_1 was set to 1.5 s and the saturation time t_{sat} was set to 150 ms. This procedure was followed for titration experiments with Man, ManNAc and Man analogs **5.1** and **5.11**. On the Ascend 700 MHz spectrometer (Bruker), spectra were acquired with 128 increments and 32 scans per increments for 150 µl samples in 3 mm sample tubes. Langerin CRD concentrations ranged from 90 to 110 µM. The relaxation delay d_1 was set to 1.4 s and the saturation time t_{sat} was set to 100 ms. This procedure was followed for titration experiments with Man, Man analogs **12.2**, **12.11**, **12.30**, **14.1** and **14.5** and GlcNS analogs **31.2** and **41.2**. In both procedures, The W5 Watergate pulse sequence was utilized for solvent suppression (Liu *et al.*, 1998). The utilized resonance assignment for the Langerin CRD has been published previously (Hanske *et al.*, 2016). For experiments with Man, ManNAc and Man analogs **5.1** and **5.11** a preliminary version of this assignment was utilized. Titration experiments with Man analogs **12.2**, **12.11**, **14.1** and **14.5** were conducted in presence of 10% DMSO. Here, assignments were transferred from a reference spectrum in absence of DMSO to the nearest neighbor in the reference spectrum in presence of DMSO. In case this approach was ambiguous, the corresponding resonances were flagged during data processing and analysis. For the acids **5.11**, **12.11** and **14.5** the pH values of the NMR samples were monitored. The pH value of stock solutions for **5.11** was adjusted to 7.8 using 1 M NaOH.

2.6.2. K_D Determination

K_D values were determined in titration experiments at six ligand concentrations. For Man, Man analogs **12.2**, **12.11**, **12.30**, **14.1** and **14.5** and GlcNS analogs **31.2** and **41.2**, samples were prepared *via* serial dilution. Chemical shift perturbations CSP for Langerin CRD resonances in the fast or fast-to-intermediate exchange regime observed upon titration with ligand were calculated *via* Equation 9 (Williamson, 2013).

$$CSP = \sqrt{\frac{\delta(^{1}\mathrm{H})^2 + \left(0.15\delta(^{15}N)\right)^2}{2}}$$

Equation 9

A standard de*via*tion of σ = 0.02 ppm was previously determined for the measurement of chemical shifts in ^{15}N HSQC NMR experiments with the Langerin CRD (Hanske *et al.*, 2016). Accordingly, only assigned resonances that displayed CSP values higher than a threshold defined in multiples of σ at the highest ligand concentration were selected for the determination of K_D values *via* Equation 10 in a global two parameter fit (Williamson, 2013). Standard errors were derived directly from the fitting procedures. The threshold was set to 2σ with the following exceptions: The threshold for the titration with Man was set to 3σ and the threshold for the titrations with **12.2**, **12.11**, **14.1** and **14.5** was set to σ. Additionally, resonances that displayed line broadening $\Delta\nu_{0.5}$ larger than 10 Hz upon titration in either the ^{1}H or the ^{15}N dimension were not considered for the determination of K_D values. CSP_{max} represents the CSP value observed upon saturation of the binding site.

$$CSP = CSP_{max} p_b$$

with

$$p_b = \left(\frac{[P]_T + [L]_T + K_D - \sqrt{([P]_T + [L]_T + K_D)^2 - 4[P]_T[L]_T}}{2[P]_T} \right)$$

Equation 10

For resonances assumed to be in the slow exchange regime upon titration, $\mathbf{K_D}$ values were derived from integrals V_b and V_f corresponding to the bound and free state of the Langerin CRD, respectively. V values served to calculate the bound fraction of the receptor p_b *via* Equation 11. Integrals V were normalized *via* integral V of the N-terminal K347 and served to calculate the bound fraction of the receptor p_b *via* Equation 11. For these calculations, only resonances for which the bound state could be assigned were considered. Selected data points displaying a low SNR or issues with the baseline correction were treated as outliers and not considered for the determination of p_b values. Next, a one parameter fit of Equation 11 to mean p_b values served to determine K_D values.

$$\frac{V_b}{(V_b + V_f)} = p_b$$

with

$$p_b = \left(\frac{[P]_T + [L]_T + K_D - \sqrt{([P]_T + [L]_T + K_D)^2 - 4[P]_T[L]_T}}{2[P]_T} \right)$$

Equation 11

2.6.3. Binding Mode Analysis

CSP values, changes in the integral of the bound state ΔV_f as well as $\Delta\nu_{0.5}(^1H)$ and $\Delta\nu_{0.5}(^{15}N)$ values were mapped on the X-ray structure of the Langerin CRD (PDB code: 3P5F or 4N32) using Matlab's Bioinformatics Toolbox *via* substitution of the B-factor values (Feinberg *et al.*, 2013; Feinberg *et al.*, 2011; MathWorks, 2016a). The patterns obtained for these observables were visualized in MOE using Chain B (Chemical Computing Group, 2016). Model quality was maintained using MOE's Structure Preparation followed by the simulation of protonation states and the hydrogen bond network of the complex with MOE's Protonate 3D. Receptor surfaces were visualized in Connolly representation (Connolly, 1993). For CSPs, the values corresponding to the highest ligand concentration were mapped on the X-ray structure. For ΔV_f as well as $\Delta\nu_{0.5}(^1H)$ and $\Delta\nu_{0.5}(^{15}N)$, the highest modulus observed upon titration was mapped on the X-ray structure. In case of $\Delta\nu_{0.5}(^1H)$ and $\Delta\nu_{0.5}(^{15}N)$, only resonances that displayed ΔV_f values smaller than 0.9 and were not flagged due to spectral overlap or low spectra quality were considered. These flags were set based on visual assessment.

2.7. Lipid-ELLA

1 mg*l^{-1} of **42**, **43** or unconjugated PEG-DSPE (3.0 kDa, NOF) in 100 mM Tris (50 µl per well) at pH 8.9 and 4°C were added to a 384 well MaxiSorp plate (Nunc) and incubated overnight to immobilize the lipids. After removal of this solution, the wells were blocked with 2% BSA in 25 mM Tris with 150 mM NaCl and 0.1% Tween-20 (70 µl per well) at pH 7.6 and room temperature for 1 h. Next, the wells were washed three times using 25 mM HEPES with 150 mM NaCl, 5 mM $CaCl_2$ and 0.01% Tween-20 at pH 7.6 and room temperature (100 µl per well). The wells were incubated with Langerin ECD in 25 mM HEPES with 150 mM NaCl, 5 mM $CaCl_2$ and 0.01% Tween-20 (50 µl per well) at pH 7.6 and room temperature for 4 h at 10 different concentrations. Subsequently, the wells were washed and incubated with HRP conjugate and 2% BSA in 25 mM HEPES with 150 mM NaCl, 5 mM $CaCl_2$ and 0.01% Tween-20 (50 µl) at pH 7.6 and room temperature for 45 min. The wells were washed and developed with TMB solution (Rockland) (50 µl per well). The reaction was quenched after 5 min by addition of 0.18 M sulfuric acid (50 µl per well). Binding of the Langerin ECD was detected *via* absorbance A_{450} measurements at 450 nm using a SpectraMax spectrometer (Molecular Devices). EC_{50} values were derived from a modified expression of Equation 5 in a two parameter fit. EC_{50} values were determined from 3 independent titrations.

3. Results and Discussion

3.1. *In silico* Analysis of C-Type Lectin Receptor Druggability

Carbohydrate binding sites of CLRs are typically hydrophilic, harbor Ca^{2+} ions and display high solvent exposure (Ernst and Magnani, 2009). These characteristics impede glycomimetic ligand design (Chapter 1.4). As discussed above, the transfer of concepts from FBLD bears the potential to overcome these challenges (Chapter 1.4). Yet, fragment growing- and linking approaches for glycomimetics depend on the existence of suitable secondary binding pockets. A set of well-established experimental methods including the screening of solvent molecules and fragments *via* X-ray crystallography or protein-observed NMR spectroscopy may be employed to identify these pockets (Cheng *et al.*, 2007; Hajduk *et al.*, 2005a; Mattos and Ringe, 1996). Alternatively, a variety of *in silico* methods has been developed over the last decade (Halgren, 2009; Le Guilloux *et al.*, 2009; Schmidtke and Barril, 2010; Volkamer *et al.*, 2012a; Volkamer *et al.*, 2012b). Modern algorithms typically utilize X-ray crystallographic data and require limited resources compared to experimental approaches. In a two-step process, pockets on the receptor surface are first identified and then scored for their capacity to interact with drug-like ligands.

Here, a set of X-ray structures of 21 human CLRs was evaluated *in silico* to identify potential secondary binding pockets for glycomimetic ligand design. Additionally, murine Dectin-1 was included in the analysis as it represents a well-studied example of Ca^{2+}-independent carbohydrate recognition by a CLR and no structure of the human homolog was available (Brown *et al.*, 2007). The scientific problem of *in silico* binding pocket identification is closely related to the concept of druggability (Hopkins and Groom, 2002). Previously, two studies predicted low druggability scores for GBP (Hajduk *et al.*, 2005a; Schmidtke and Barril, 2010). Yet, no detailed information on identified binding pockets or on the druggability of CLRs was available. The work presented in this chapter was originally published as part of a study revisiting the druggability of CLRs (Aretz *et al.*, 2014). Accordingly, the results are discussed in the context of experimental druggability analyses conducted in our laboratory (Aretz, 2017; Aretz *et al.*, 2016; Wamhoff *et al.*, 2016).[17] Finally, the implications of these findings for the design of glycomimetic Langerin ligands are addressed.

[17] In part, these procedures were conducted by Dr. Jonas Aretz (Max Planck Institute of Colloids and Interfaces).

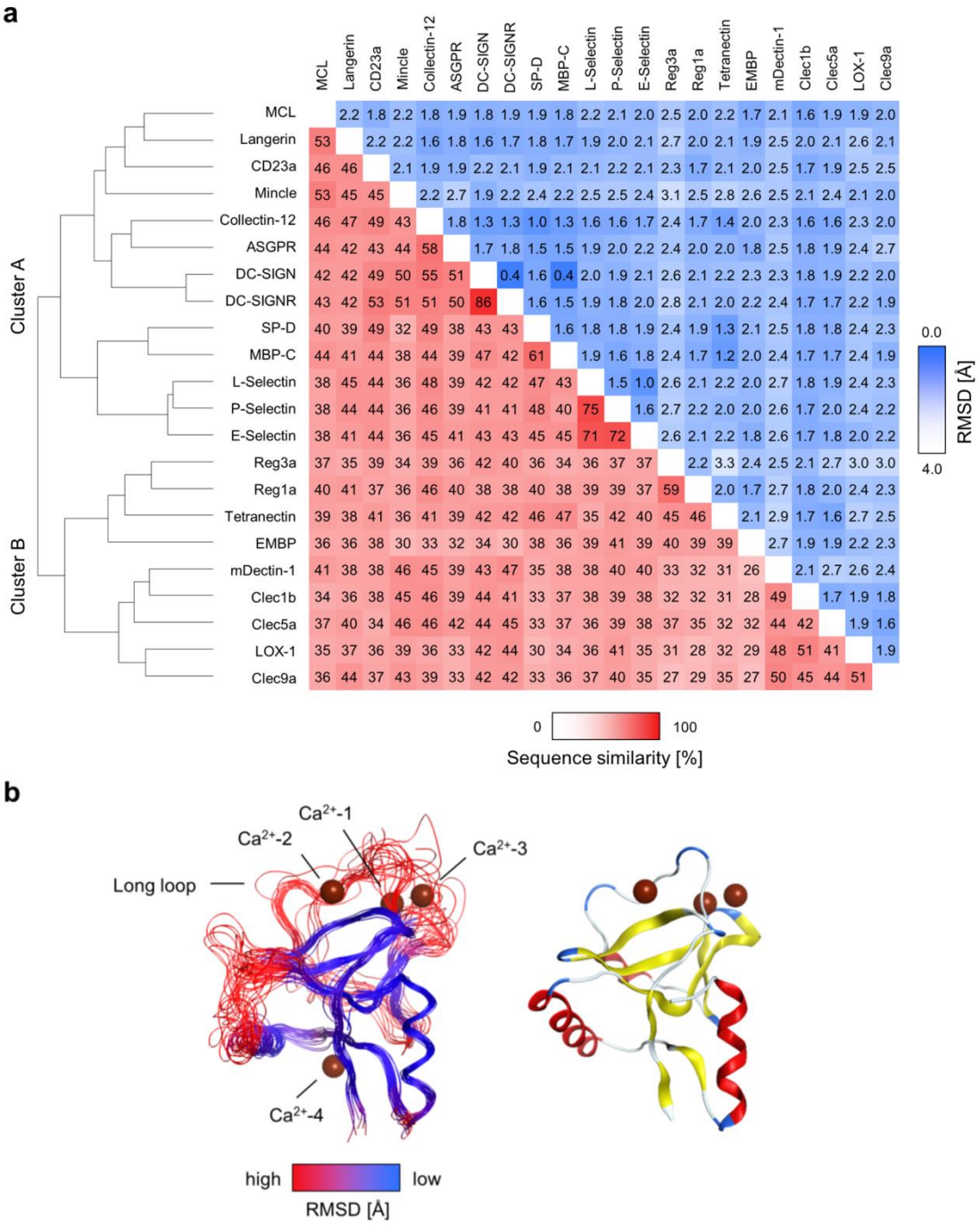


Figure 5. Structure-based sequence alignment and consensus structure of CLRs.

a. The dendrogram depicts the hierarchical clustering of CRDs based on a structure-based multiple sequence alignment. The two major branches are termed cluster A and B. Pairwise sequence similarities and RMSDs of C_α atoms are shown in matrix format. **b.** The locations of key structural features of the CTLD fold with respect to the predicted consensus structures are indicated (left). The peptide backbone is colored according to RMSD of C_α atoms. The CRD of DC-SIGN (PDB code: 2XR6) is shown for comparison (right). The ribbon representation is colored according to secondary structure elements (red: α-helices; yellow: β-strands; blue: turn regions). Adapted with permission from (Aretz *et al.*, 2014).

The implemented druggability prediction is based on a comparative structural analysis of the CRDs. Hence, a structure-based sequence alignment was conducted (Appendix B.1). While all CRDs share the same CTLD fold, the global sequence similarity within the set of receptors is low with an average value of 41% (Figure 5a). As the set contains closely related homologs such as DC-SIGN and DC-SIGNR, it additionally displays a high variance spanning from 26% to 86% sequence similarity. A phylogenetic analysis based on this alignment yields a classification that resembles the canonical nomenclature of CLRs, in particular with respect to the correct assignment of members of the Groups II, III, IV, V and VII (Varki, 2009). Collectin-12 is the only receptor deviating from the canonical nomenclature, as it is part of the Group II cluster. Moreover, Tetranectin and EMBP are the only representatives of Group IX and XII used in this study. Both display elevated distances to other branches. EMBP and Tetranectin as well as Clec9a, Lox-1, Clec1b, and Reg1a have been reported to interact with non-carbohydrate ligands and these CLRs were assigned to Cluster B by the hierarchical clustering algorithm. Strikingly, CRDs known to recognize carbohydrates *via* the Ca^{2+}-2 binding site are exclusively present in Cluster A (Figure 5a). These findings indicate that a comparative analysis of the global sequence similarity might be a good predictor for Ca^{2+}-dependent carbohydrate affinity of CLRs if the value is derived from structural information.

In contrast to the low global sequence similarity, the overall structure of the CTLD is highly conserved. RMSD values of C_α atoms obtained from the structure-based multiple sequence alignment are uniformly low and do not exceed 3.2 Å (Figure 5a). To visualize the conservation of the domain architecture a consensus structure was calculated (Figure 5b). While the core of the CTLD displays only minor deviations, a higher level of structural variability characterizes the two loop regions. This correlates with increased structural flexibility observed in experiments and MD simulations for selected CLRs (Hanske *et al.*, 2016; Pederson *et al.*, 2014; Probert *et al.*, 2013). The long loop is of particular interest as it harbors the Ca^{2+}-1, -2 and -3 sites and thus plays a fundamental role in Ca^{2+}-dependent carbohydrate recognition (Zelensky and Gready, 2005). Moreover, it is part of an allosteric network that modulates Ca^{2+} recognition by Langerin (Hanske *et al.*, 2016).

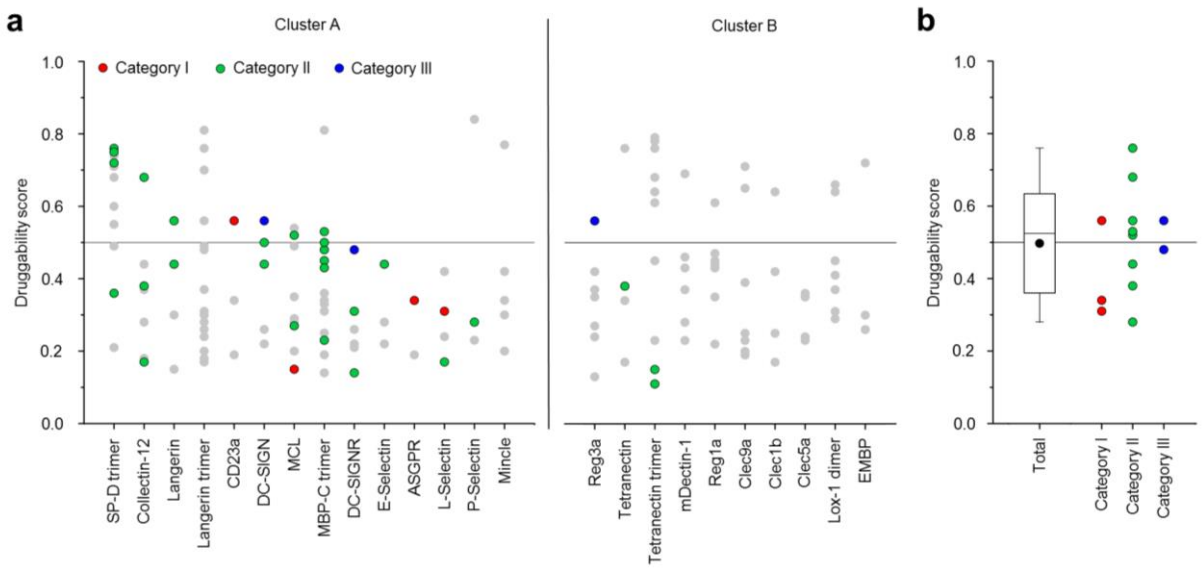


Figure 6. Computational druggability prediction for CLRs.

a. A comprehensive account of all binding sites predicted by DoGSite is given. CLRs were classified according to the structure-based multiple sequence alignment into Clusters A and B. A druggability score of less than 0.5 is indicative of an undruggable binding site. **b.** A boxplot of the highest druggability scores of binding site categories I to III of each CLR is shown. A mean druggability score of 0.50±0.04 was determined. Tukey-style representation was chosen. Adapted with permission from (Aretz *et al.*, 2014).

Next, binding pockets were predicted using the program DoGSite (Volkamer *et al.*, 2010; Volkamer *et al.*, 2012b). DoGSite evaluates the receptor surface geometry combining a grid-based algorithm with a difference of Gaussian (DoG) filter (Marr and Hildreth, 1980). Initially identified pockets are decomposed into subpockets to improve the accuracy of the approach. If X-ray structures of oligomeric CLR ECDs were available, these complexes were analyzed. Generally, X-ray structures corresponding to CLR-ligand complexes were prioritized. For Langerin, both the CRD and the trimeric ECD were analyzed as a CLR-ligand complex was only available for the former. Subsequently, DoGSiteScorer was applied to calculate druggability scores (Volkamer *et al.*, 2012a; Volkamer *et al.*, 2012b). These calculations are based on the geometric and physicochemical parameters of the identified subpockets including volume, depths and amino acid composition. In the applied scoring scheme, scores over 0.5 are indicative of a druggable binding pocket. Following this definition, at least one druggable pocket was found for the majority of the analyzed CLRs (Figure 6a).

It is *a priori* unknown whether a drug-like molecule binding to a pocket will exert a physiological effect, *e.g.* as an anti-infective. However, structural changes upon Ca^{2+} recognition, particularly for the long loop region have been reported from many CLRs (Abergel *et al.*, 1999; Blomhoff *et al.*, 1982; Furukawa *et al.*, 2013; Ng *et al.*, 1998; Nielbo *et al.*, 2004; Onizuka *et al.*, 2012). Recent publications on allosteric modulation of CLR function suggest that targeting pockets located in proximity to Ca^{2+} ions in the long loop regions might represent a feasible approach (Aretz *et al.*, 2017; Furukawa *et al.*, 2013; Hanske *et al.*, 2016). From the perspective of the design of carbohydrate analogs, targeting pockets remote from the carbohydrate binding site will complicate fragment growing- or linking

approaches. Hence, pockets were classified according to their location on the CLR surface: Category I pockets are in immediate proximity to Ca^{2+}-dependent carbohydrate binding sites, Category II pockets are found within 6 Å of either Ca^{2+}-1, Ca^{2+}-2 or Ca^{2+}-3 in the long loop region and Category III pockets are associated with Ca^{2+}-independent carbohydrate binding sites. The sites were experimentally identified for DC-SIGN, DC-SIGNR, Reg3a and murine Dectin-1 *via* X-ray crystallography (Brown *et al.*, 2007; Feinberg *et al.*, 2001; Lehotzky *et al.*, 2010).

Overall, three to nine pockets were identified for CRDs (Figure 6a). These values are in accordance with previous studies on other protein families (Sheridan *et al.*, 2010). For Langerin, MBP-C, Lox-1 and Tetranectin, X-ray structures of oligomeric ECDs were available. In these cases, nine to 19 pockets were identified with a considerable number of these pockets located at oligomer interfaces. Ca^{2+}-dependent carbohydrate binding sites corresponding to Category I were only identified for four of the 14 relevant CLRs. The calculated druggability scores are low on average with CD23a displaying a moderate score of 0.56. The latter is expressed on follicular B cells and involved in the regulation of IgE levels (Sukumar *et al.*, 2006). Targeting the Ca^{2+}-dependent carbohydrate binding site of CD23a with drug-like molecules or glycomimetics potentially bears therapeutic potential for the treatment of allergies and might be less challenging than for other CLRs in the analyzed data set. Ca^{2+}-independent carbohydrate binding sites were identified more reliably with a success rate of 3 out of 4. Notably, the existence of such a binding site in murine Dectin-1 has been controversial and the interaction observed in the X-ray structure might be artificial (Brown *et al.*, 2007). Category III pockets also received low to moderate druggability scores rendering the design of drug-like ligands challenging. For both, DC-SIGN and DC-SIGNR, the identified Ca^{2+}-independent binding sites are located in proximity to the primary carbohydrate binding site. The DoGSiteScorer analysis suggests that targeting these Category III pockets might facilitate glycomimetic ligand design or the identification of drug-like inhibitors as Ca^{2+}-dependent carbohydrate binding sites were not identified as pockets for most CLRs.

In summary, these findings reflect the limitations of typical grid-based *in silico* methods utilized in druggability evaluation with respect to the identification of solvent exposed and hydrophilic binding sites. These limitations can be attributed to the fact that algorithms are typically trained on data sets of binding pockets for drug-like molecules, thereby introducing a property bias and preventing the recognition of carbohydrate binding sites. Simultaneously, the results obtained for CLRs using DoGSiteScorer validate the assumption that targeting carbohydrate binding sites with drug-like molecules is challenging. Notably, alternative *in silico* approaches have recently been developed to specifically identify carbohydrate binding sites. These approaches include molecular docking and grid-based algorithms as well as evolutionary conservation scores (Frank, 2014; Zhao *et al.*, 2014).

Ca^{2+}-associated binding pockets of either Category I or II were identified for all CLRs harboring Ca^{2+}-2 ions except for Mincle and the trimeric Langerin ECD. A comparison with the Langerin CRD reveals that structural changes in the short loop region induced upon either oligomerization or ligand

binding potentially abrogate the recognition of Category II pockets by DoGSite (Figure 7a) (Feinberg *et al.*, 2010). As observed for Category I and III pockets, Category II pockets displaying high druggability scores are sparse. Collectin-12 and SP-D, members of the CLR Group III, represent notable exceptions. Both Collectins are essential mediators of innate immunity and do not represent well-established therapeutic targets for glycomimetics and drug-like molecules.

The existence of a single druggable pocket is sufficient to render a protein druggable (Hopkins and Groom, 2002). Following this definition, for each CLR the Category I, II and III pockets that displayed the highest druggability score was selected to calculate a mean druggability score of 0.50±0.04 (Figure 6b). This renders CLRs challenging or undruggable targets, a result consistent with predictions for GBP in general (Hajduk *et al.*, 2005b; Schmidtke and Barril, 2010; Volkamer *et al.*, 2012a). Individual receptors such as SP-D and Collectin-12 possess favorable pockets in the long loop region. Other targets such as E-Selectin or ASGPR display druggability values well below the mean. Notably, DogSite does not account for structural flexibility and will consequently not identify pockets potentially targeted *via* induced fit or conformational selection binding mechanisms. This limitation is generally recognized as a challenge for the development of improved algorithms and has recently been addressed (Cimermancic *et al.*, 2016).

Experimentally determined hit rates from fragment screening implicitly account for these binding mechanisms and are robust predictors for druggability (Aretz *et al.*, 2016; Edfeldt *et al.*, 2011; Hajduk *et al.*, 2005a). Consistent with the *in silico* analysis, screening approaches utilizing carbohydrate-derived reporter molecules revealed low hit rates below 2.1% for Langerin and DC-SIGN (Aretz, 2017; Aretz *et al.*, 2016; Wamhoff *et al.*, 2016). While these findings further validate the low druggability of carbohydrate binding sites, alternative screening approaches indicate that targeting the long loop region might be more feasible. In a ^{19}F NMR screening against Langerin, DC-SIGN and MCL increased hit rates ranging from 10.0% to 15.7% were observed for Ca^{2+}-dependent interactions (Aretz *et al.*, 2014). For DC-SIGN 47.0% of these hits were validated by SPR. Assuming these interactions might be leveraged to modulate CLR function, the hit rates indicate a moderate druggability for CLRs. Notably, allosteric inhibitors of DC-SIGN were recently identified in our laboratory (Aretz *et al.*, 2017). Moreover, following the definition for druggability, CLRs should be considered druggable as a drug-like DC-SIGN inhibitor with nano- to micromolar affinity has previously been identified (Hopkins and Groom, 2002; Mangold *et al.*, 2012). A comparison of these experimental data with the *in silico* analysis highlights the limitations of DoGSite and similar algorithms with respect to the structural flexibility of CLRs.

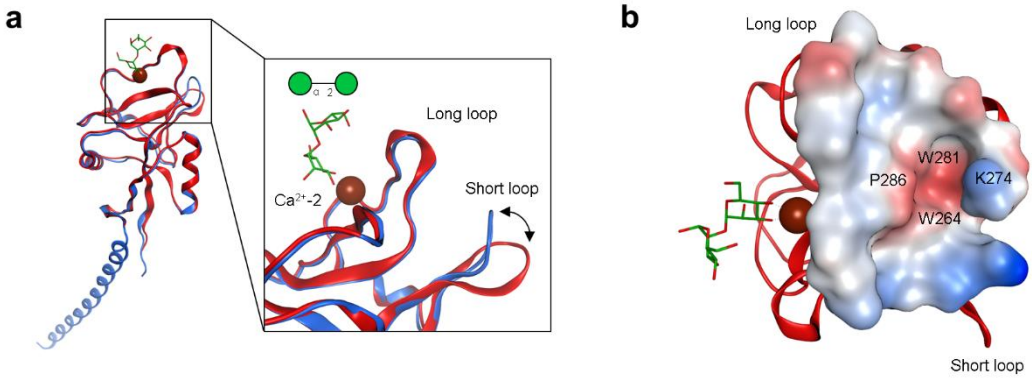


Figure 7. Pocket identification for Langerin.

a. A comparison between the X-ray structures of the Langerin ECD (blue) (PDB code: 3KQG) and the Langerin CRD in complex with Man_2 (red) (PDB code: 3P5F) reveals structural flexibility in the short loop region. The closing of the short loop abrogates the identification of the binding pocket predicted for the CRD (Feinberg *et al.*, 2010; Feinberg *et al.*, 2011). **b.** The Category II pocket displaying the highest druggability with a sore of 0.56 is located on the back surface of the Ca^{2+}-dependent carbohydrate binding site between the long and the short loop region. The lipophilic base of the pocket is formed by W281 and W264. The Connolly surface was calculated for receptor atoms assigned to the pocket by DogSite. Consequently, this surface representation might not be identical with the receptor surface. The pocket surface is colored according to its lipophilicity (lipophilic: red, hydrophilic: blue).

Shifting the focus from the global druggability prediction for CLRs back to the identification of secondary binding pockets for Langerin, the *in silico* analysis mostly corroborates the assumption that glycomimetic ligand design is challenging. The Ca^{2+}-dependent carbohydrate binding site of Langerin was not classified as a pocket by DogSite and both identified Category II pockets received low to moderate druggability scores. Interestingly, the pocket displaying the higher druggability score is located on the back surface of the carbohydrate binding site between the long and the short loop region (Figure 7b). The lipophilic base of the pocket is formed by W281 and W264 flanked by K274 and P286. Targeting this subpocket will potentially enable the design of glycomimetics *via* fragment linking. As highlighted above, this approach has been successfully applied for E-Selectin *via* a second site NMR screening (Egger *et al.*, 2013). The screening was implemented utilizing a spin-labeled glycomimetic inducing paramagnetic relaxation enhancement for fragments binding close to the carbohydrate binding site. Moreover, W264 was later observed to display CSPs in ^{15}N HSQC NMR titration experiments with Man and various glycomimetics (Chapter 3.2, 3.3 and 3.4). The role of W264 in the allosteric network previously described has not been elucidated, yet this observation suggests that targeting the identified secondary pocket might result in an allosteric inhibition of Langerin function.

Considering the low to moderate druggability score for this pocket and the fact that suitable structural biology tools were not established in our laboratory at the initial stages of the work presented in this dissertation, the secondary binding pocket described above was not validated experimentally and other approaches to glycomimetic ligand design were prioritized. These approaches will be discussed in the following chapters and predominantly focus on fragment growing strategies to explore the Ca^{2+}-binding site for favorable secondary interactions (Chapter 3.3, 3.4 and 3.5). This decision is clearly

justified by the compromised robustness of the *in silico* predictions for the Ca^{2+}-dependent carbohydrate binding sites of CLRs and reports on potent glycomimetics for other human CLRs (Borrok and Kiessling, 2007; Chang *et al.*, 2010). As suitable structural biology tools, *i.e.* potent glycomimetics potentially serving as reporter molecules for second site NMR screenings and ^{15}N HSQC NMR experiments are established in our laboratory at this stage, fragment linking strategies might be revisited in the future to optimize designed ligands.

3.2. ^{19}F R_2-filtered NMR-Guided Glycomimetic Ligand Design

3.2.1. Assay Development for Langerin

As a result of the hydrophilicity and high solvent exposure of Ca^{2+}-dependent CLR binding sites, monovalent interactions with naturally occurring mono- and oligosaccharides typically display low affinities (Ernst and Magnani, 2009; Weis and Drickamer, 1996). This does not only result in challenges for the glycomimetic ligand design process itself. Assay development is impeded by sensitivity problems and a lack of suitable reporter molecules, particularly at the initial stages of the design process. This has been compensated for *via* the multivalent organization of carbohydrates and CLRs in heterogeneous assays (Cecioni *et al.*, 2015). For Langerin, heterogeneous assays served to quantify its specificity for several mono- and oligosaccharides (Chabrol *et al.*, 2012; Feinberg *et al.*, 2013; Stambach and Taylor, 2003; Zhao *et al.*, 2016). Yet, these methods introduce complex multi-state equilibria and surface phenomena, complicating data analysis as well as limiting the transferability of determined affinities between assay formats (Kitov and Bundle, 2003). Alternatively, ITC and STD NMR experiments have enabled the determination of K_D values in the millimolar range (Holla and Skerra, 2011; Munoz-Garcia *et al.*, 2015). While these biophysical methods provide the required sensitivity, they are limited with respect to throughput and material consumption.

To address the limitations discussed above, a ^{19}F R_2-filtered NMR competitive binding assay for Langerin was developed during the initial stages of glycomimetic ligand design (Wamhoff *et al.*, 2016). R_2-filtered NMR represents a well-established method for the detection of ligand-receptor interactions (Hajduk *et al.*, 1997). In combination with the high contributions from the CSA and ^{1}H-^{19}F dipolar interactions observed for ^{19}F NMR, the determination of transversal relaxation rates R_2 provides excellent sensitivity, enabling the utilization of monovalent reporter molecules even at low affinities (Dalvit, 2007). Based on an analytical expression for the equilibrium of competitive binding experiments, the determination of K_I values directly yields information on the binding site and the stoichiometry of glycomimetic ligands (Roehrl *et al.*, 2004). The absence of background resonances from *e.g.* solvents, additives or the receptor in ^{19}F NMR generally improves data quality and facilitates sample preparation. In general, the approach is applicable to screen *e.g.* focused libraries, providing robust ranking of the identified hits.

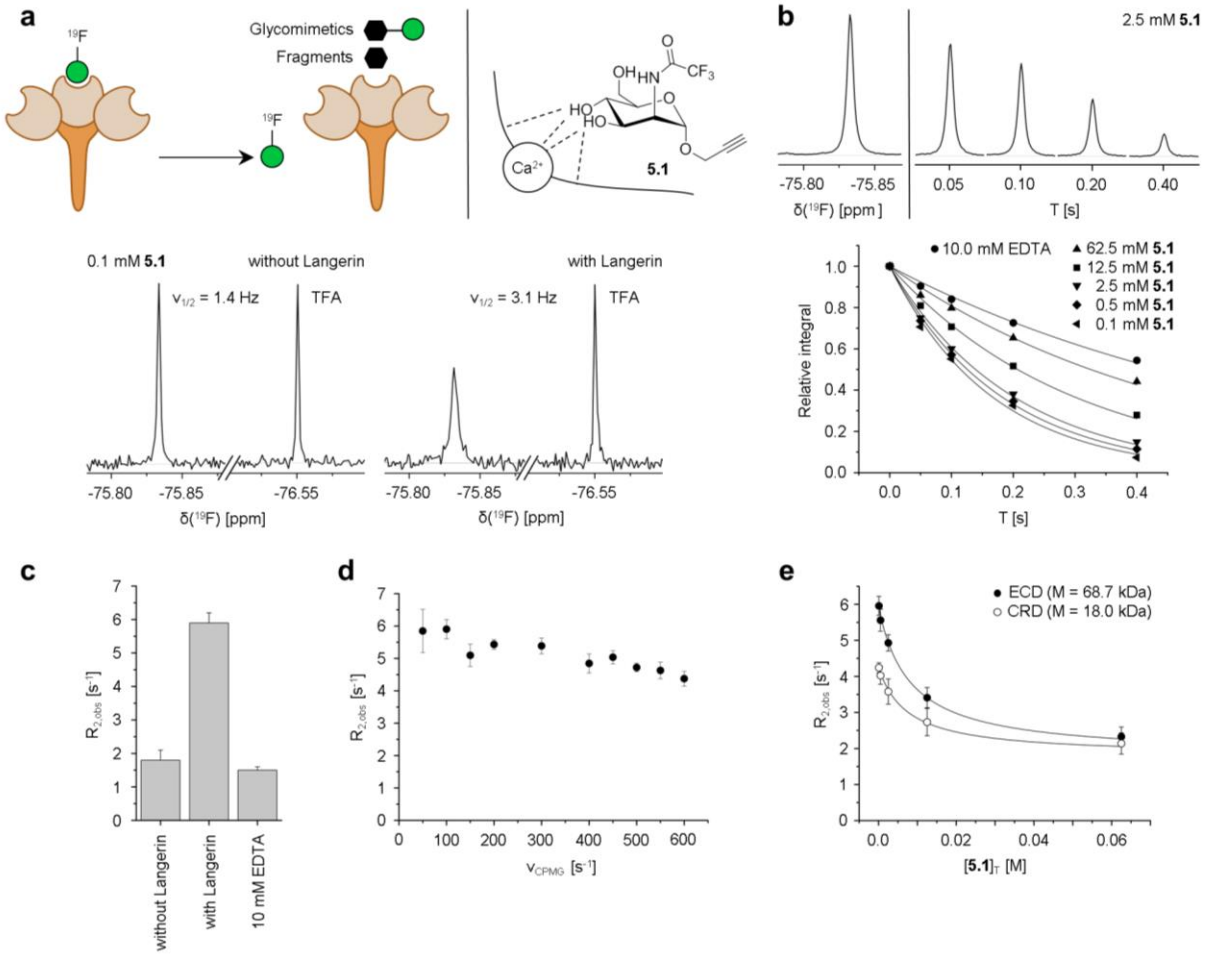
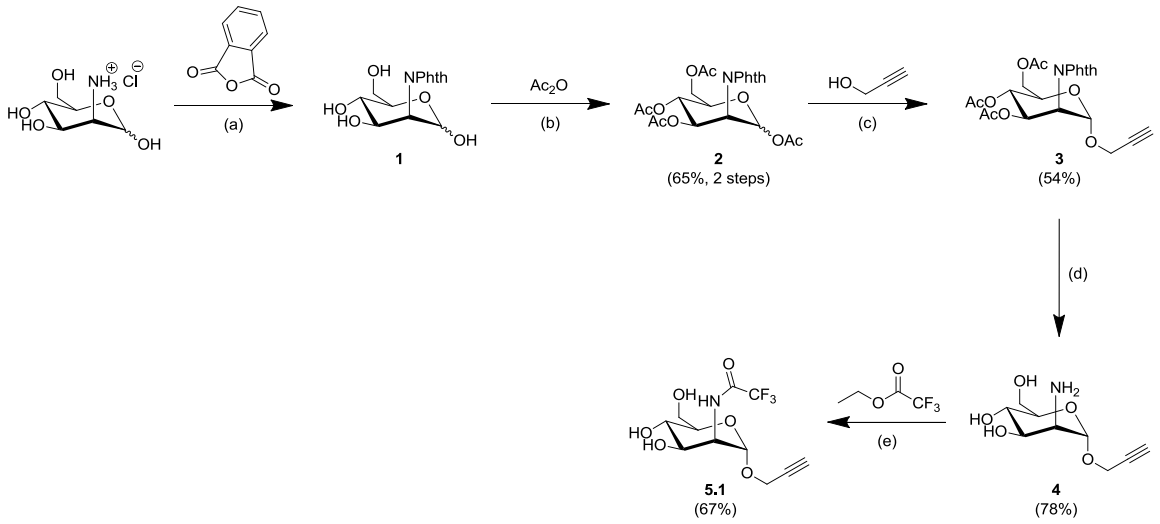


Figure 8. ^{19}F R_2-filtered NMR assay development.

a. In presence of the Langerin ECD, the ^{19}F NMR resonance of the trifluoroacetamido group of reporter molecule **5.1** displays line broadening $\Delta\nu_{0.5}$. This phenomenon was utilized to develop a competitive binding assay. **b.** The interaction between the reporter molecule and Langerin can be quantified *via* the transversal relaxation rate $R_{2,obs}$ using the CPMG pulse sequence (Carr H. Y. and M., 1954; Meiboom and Gill, 1958). Representative decay curves for the ECD are shown. **c.** The Ca^{2+}-dependency of the interaction was validated *via* addition of EDTA. The $R_{2,obs}$ value in absence of Langerin was determined from four independent experiments. The standard error for the experiments in presence of Langerin and EDTA is derived directly from the fitting procedure. **d.** Relaxation dispersion experiments for **5.1** in presence of the Langerin ECD indicate a negligible exchange contribution $R_{2,ex}$ at a ν_{CPMG} value of 500 Hz. **e.** Titration experiments revealed comparable K_D values of 7.9±0.7 mM and 7.3±1.0 mM for ECD and CRD, respectively. A significantly higher amplitude in $R_{2,obs}$ was observed for the ECD. $R_{2,obs}$ values for the ECD were determined from three independent titrations. Adapted with permission from (Wamhoff *et al.*, 2016).



Scheme 1. Synthesis of reporter molecule 5.1.

Intermediate **2** was prepared as previously published (Veselý *et al.*, 2003). Reaction conditions for the preparation of **5.1**: (a) Pyridine, acetone:H_2O (1:1), 50°C; (b) pyridine, 50°C; (c) $BF_3.OEt_2$, anhydrous DCM:ether (2:1), 0°C to room temperature; (d) $MeNH_2$, EtOH, room temperature; (e) DMF, room temperature.

To implement the ^{19}F R_2-filtered NMR assay, a reporter molecule was designed based on ManNAc ($K_{I,app}$ = 5.6±0.9 mM) which had previously been reported to bind Langerin with an affinity comparable to that of Man ($K_{I,app}$ = 2.3±0.1 mM) (Figure 8a) (Stambach and Taylor, 2003). ManNAc analog **5.1** was prepared over five steps with an overall yield of 18% (Scheme 1). The synthesis involved the stereospecific introduction of a propargyl group at the anomeric position to enable conjugation to *e.g.* nanoparticles or polymers. The ^{19}F NMR spectrum of **5.1** displays a single resonance for the trifluoroacetamido group, ensuring an optimal SNR. Moreover, the absence of J-coupling enables the utilization of the CPMG pulse sequence (Segawa and Bodenhausen, 2013). In presence of Langerin, the ^{19}F NMR resonance displayed considerable line broadening $\Delta\nu_{0.5}$ of 1.7 Hz (Figure 8a). This observation is indicative of binding and the simultaneous absence of CSPs highlights the excellent sensitivity of ^{19}F R_2-filtered NMR in experiments with large, oligomeric CLRs.

Table 2. Parameters for the ^{19}F R_2-filtered NMR assay.

Parameter	CRD	ECD
$R_{2,f}$ [s^{-1}]	1.8±0.3[a]	1.8±0.3[a]
$R_{2,b}$ [s^{-1}]	360±50	660±50[b]
K_D [mM] for **5.1**	7±1	7.9±0.7[b]

[a] This value was determined from four independent experiments *via* ^{19}F R_2-filtered NMR.
[b] This value was determined from three independent titration experiments *via* ^{19}F R_2-filtered NMR.

Next, the interaction between Langerin and **5.1** was quantified by measuring observed relaxation rates $R_{2,obs}$ in CPMG experiments (Figure 8b, Equation 1) (Carr H. Y. and M., 1954; Meiboom and Gill, 1958). The Ca^{2+}-dependency of the interaction was validated *via* the addition of EDTA (Figure 8c). The analytical expressions underlying the determination of K_D and K_I values are limited to the fast chemical exchange regime (Equation 3 and 4) (Dalvit, 2007). Hence, relaxation dispersion experiments were conducted to exclude a chemical exchange contribution $R_{2,ex}$ to $R_{2,obs}$ (Figure 8d). $R_{2,ex}$ is negligible at a frequency of 180° pulses ν_{CPMG} of 500 Hz and this frequency was selected for the establishment of the ^{19}F R_2-filtered NMR assay. Assuming a one-site binding model, titration experiments served to determine the affinity of both the CRD (K_D = 7.3±1.0 mM) and the trimeric ECD (K_D = 7.9±0.7 mM) for **5.1** (Figure 8e, Table 2). An increased transversal relaxation rate for the bound state $R_{2,b}$ with the trimeric ECD (M = 68.7 kDa) was observed, presumably due to an increased molecular weight over the CRD (M = 18.0 kDa). This increase translates into an improved dynamic range for competitive binding experiments.

The determination of K_I values by ^{19}F R_2-filtered NMR was established *via* competitive binding experiments with Man (K_I = 4.5±0.5 mM) and ManNAc (K_I = 24.9±0.9 mM) (Figure 9a to c, Table 3). In both cases, the binding isotherms are consistent with the assumed one-site binding model of the fitting procedure. For Man, complete competition with **5.1** binding was demonstrated and the determined affinity is in accordance with literature values (Feinberg *et al.*, 2013; Stambach and Taylor, 2003). In contrast, the affinity obtained for ManNAc de*via*tes significantly from the previously reported apparent K_I value (Stambach and Taylor, 2003).

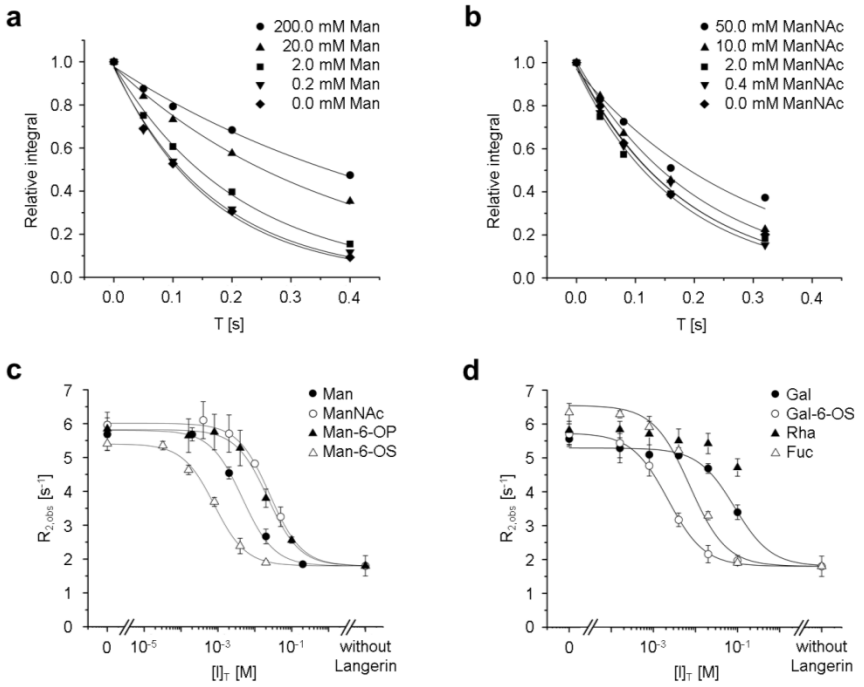


Figure 9. K_I determination for selected monosaccharides.

a. and b. Representative decay curves from the competitive binding experiments with Man and ManNAc are shown. **c.** Competitive binding experiments served to determine affinities for Man (K_I = 4.5±0.5 mM) and ManNAc (K_I = 25±1 mM). For these experiments, $R_{2,obs}$ values were derived from three independent titrations. While the sulfation in C6 of the mannose scaffold results in an affinity increase (K_I = 0.8±0.1 mM), a decreased affinity was observed for Man-6-OP (K_I = 22±1 mM). **d.** Gal (K_I = 80±20 mM) displays a low affinity close to the detection limit. The sulfation in C6 results in a decreased K_I value (K_I = 2.2±0.2 mM). Titration experiments with the deoxy monosaccharides Rha (K_I > 100 mM) and Fuc (K_I = 7±2 mM) complete the acquired data set. Adapted with permission from (Wamhoff *et al.*, 2016).

Consequently, orthogonal ^{15}N HSQC NMR titration experiments were conducted to validate the ^{19}F R_2-filtered NMR assay (Figure 10, Table 3). Here, CSPs of backbone amide resonances of the Langerin CRD induced upon ligand binding were utilized to determine K_D values (Williamson, 2013). As discussed for $R_{2,obs}$, analytical expressions for underlying K_D determinations are only valid in the fast exchange regime. Consequently, only those resonances that displayed $\Delta\nu_{0.5}$ values lower than 10 Hz were considered for analysis. Overall, the K_D values obtained for **5.1** (K_D = 5.8±0.5 mM), Man (K_D = 5.8±0.3 mM) and ManNAc (K_D = 21±3 mM) are consistent with affinities derived from ^{19}F R_2-filtered NMR experiments. Hence, the deviations from the literature value for ManNAc ($K_{I,app}$ = 5.6±0.9 mM) may reflect the limitations of heterogeneous assays discussed above (Stambach and Taylor, 2003).

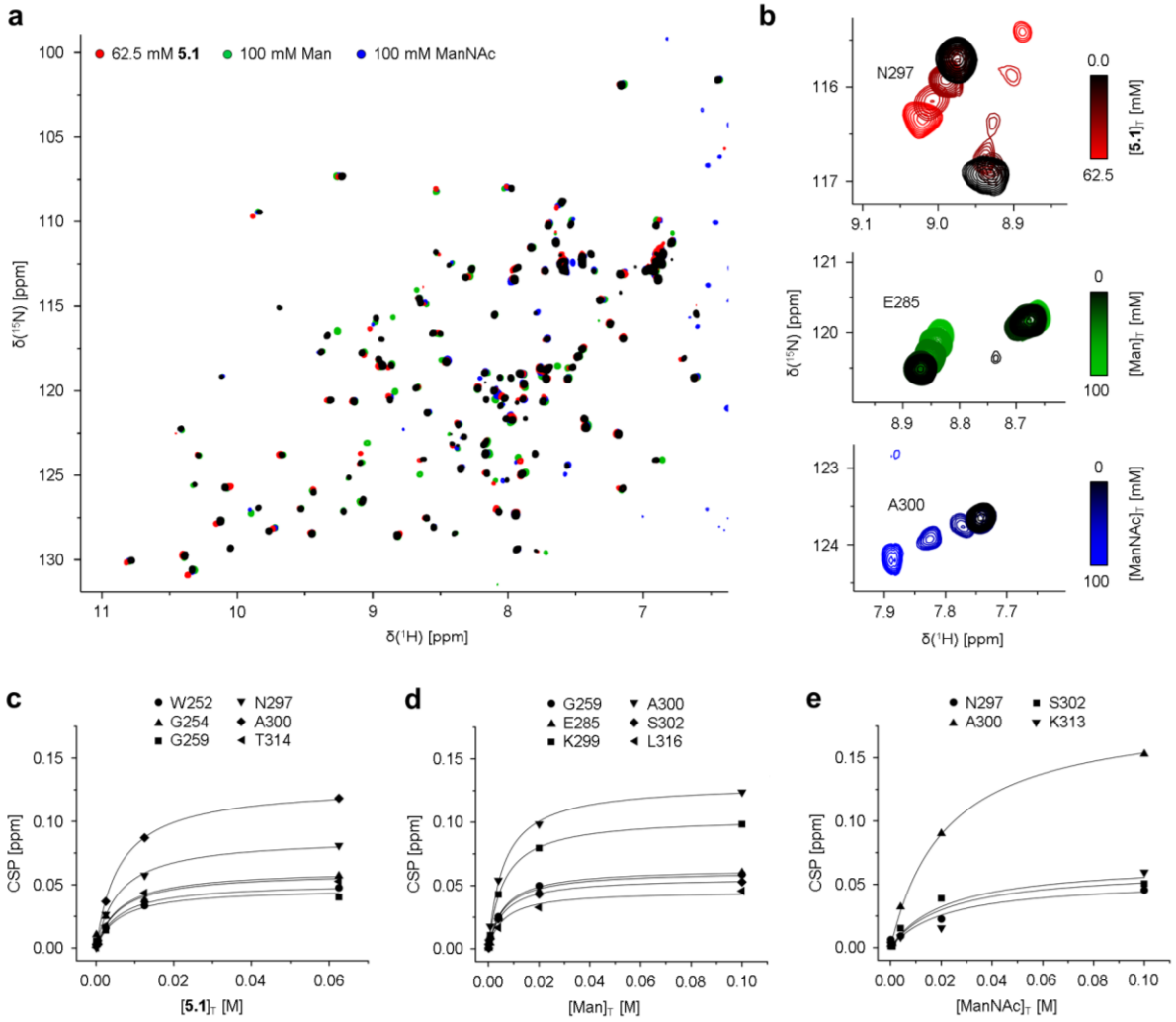


Figure 10. K_D determination for reporter molecule 5.1, Man and ManNAc.

a. Affinities determined for **5.1**, Man and ManNAc in the ^{19}F R_2-filtered NMR assay were validated *via* ^{15}N HSQC NMR titration experiments with the Langerin CRD. **b.** Assigned resonances displaying fast chemical exchange and CSPs larger than 0.04 ppm were selected for the determination of K_D values. Representative CSP trajectories are depicted. **c to e.** The K_D values obtained for **5.1** (K_D = 5.8±0.5 mM), Man (K_D = 5.8±0.3 mM) and ManNAc (K_D = 21±3 mM) are consistent with the results from the ^{19}F R_2-filtered NMR assay. Adapted with permission from (Wamhoff *et al.*, 2016).

Moreover, the previously reported assignment of backbone amide resonances for the Langerin CRD enabled the evaluation of the binding mode of reporter molecule **5.1**, Man and ManNAc (Hanske *et al.*, 2016). To this end, CSP values were mapped on the X-ray structure of the Langerin CRD (Figure 11) (Feinberg *et al.*, 2011). The data analysis in this chapter is based on a preliminary version of the resonance assignment. The discussion will focus on aspects relevant for the validation of the ^{19}F R_2-filtered NMR assay and a more elaborate analysis of the recognition of monosaccharides and glycomimetics by Langerin is included in the following chapters (Chapter 3.3 and 3.4).

For all three ligands, most CSPs are observed in proximity to the carbohydrate binding site. Of the residues observed to directly interact with Man or the Ca^{2+} ion in X-ray crystallography, only E285 and K299 were assigned (Feinberg *et al.*, 2011). Both residues display CSPs larger than 0.02 ppm upon titration with **5.1**, Man and ManNAc (Figure 11b). Importantly, this observation validates a Ca^{2+}-dependent binding mode for reporter molecule **5.1**. The CSP value of 0.02 ppm was previously determined as the standard deviation for the measurement of CSPs (Hanske *et al.*, 2016). Additional CSPs are observed in axial direction of C2 of the Man scaffold (Figure 11c). Prominent residues displaying CSPs include I250, W252, G254, N297, A300, K313 and T314 (Figure 11b and c). Man induces higher CSPs for K299 and K313 than both **5.1** and ManNAc, consistent with the formation of a hydrogen bond by the axial hydroxyl group in C2 (Feinberg *et al.*, 2011). The absence of this hydrogen bond for ManNAc potentially explains the decreased affinity determined *via* ^{19}F R_2-filtered NMR. For **5.1**, residues in proximity of N297, *e.g.* W252 and G254 display considerably higher CSP values than observed for Man and ManNAc. This observation can be attributed to favorable interactions formed by the trifluoromethyl group of **5.1** which potentially induces the observed affinity increase over ManNAc. Interestingly, observed CSPs are not confined to the proximity of the carbohydrate binding site. They do also extend to remote regions of the CTLD fold, in particular the short loop region (Figure 11c). Prominent residues include A258 and G259. This indicates a modulation of the previously described allosteric network involved in Ca^{2+} recognition by Langerin (Hanske *et al.*, 2016). While the observed distribution of CSPs complicates the interpretation of ^{15}N HSQC NMR experiments, a careful comparative analysis for residues in proximity of the carbohydrate binding site nevertheless yields robust information on the binding modes of glycomimetic Langerin ligands.

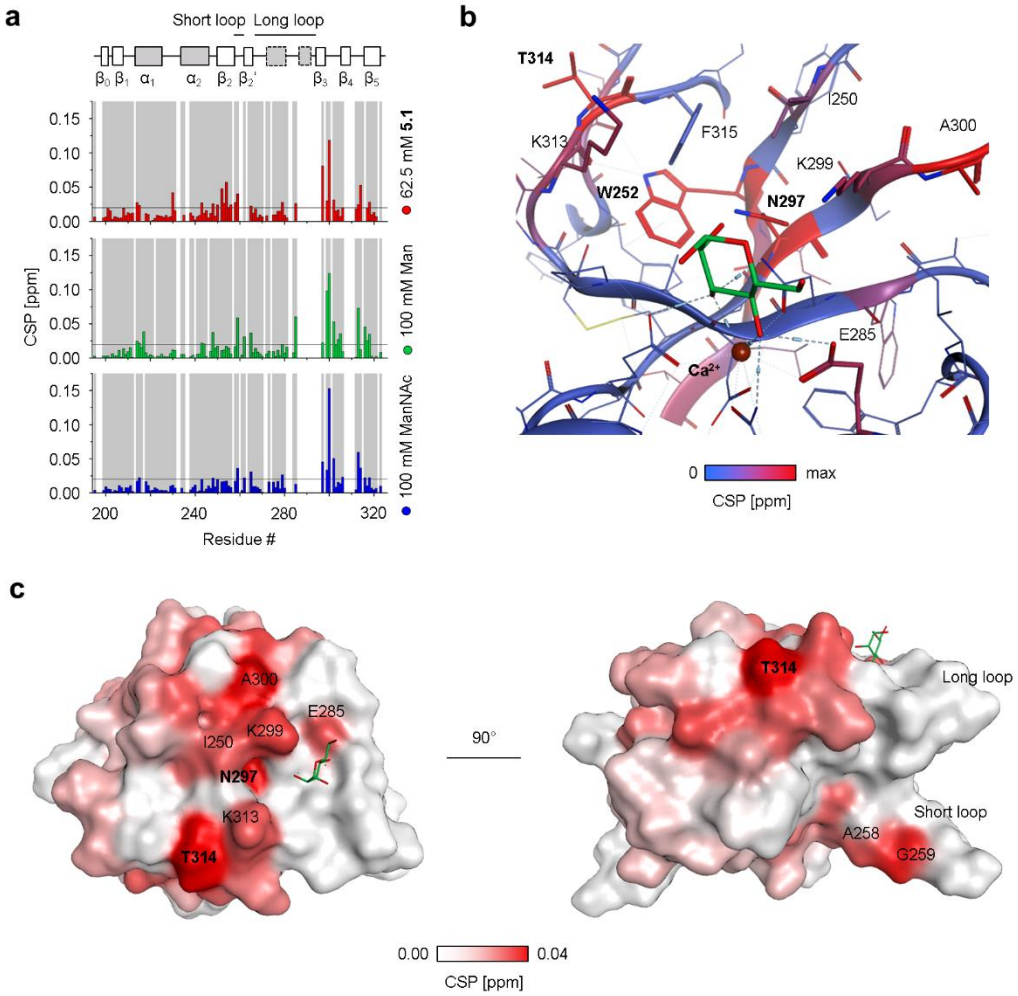


Figure 11. ^{15}N HSQC NMR binding mode analysis for reporter molecule 5.1, Man and ManNAc.

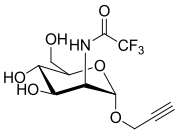
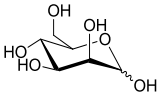
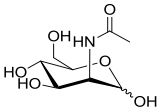
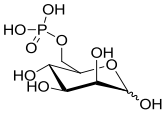
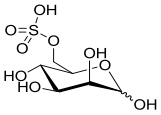
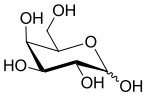
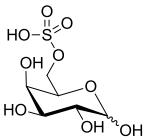
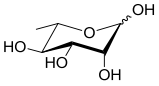
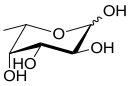
a. ^{15}N HSQC NMR experiments with the Langerin CRD revealed similar CSP patterns for **5.1**, Man and ManNAc. Assigned resonances detected in the reference spectrum are highlighted (grey). **b.** Mapping the CSPs on the X-ray structure of the Langerin CRD (PDB code: 3P5F) validated a Ca^{2+}-dependent binding mode for **5.1** as indicated by CSPs observed for E285 and K299 (Feinberg *et al.*, 2011). CSPs observed for *e.g.* N297 and W252 are potentially induced by interactions formed by the trifluoromethyl group of **5.1** and differ from the CSPs obtained for Man and ManNAc. **c.** Additional CSPs located in remote regions of the CTLD fold, particularly for G259 and A258 in the short loop region, indicate a modulation of the previously reported allosteric network upon **5.1** binding (Hanske *et al.*, 2016). This modulation was also observed for titration experiments with Man and ManNAc.

In summary, the results characterize ManNAc analog **5.1** as a suitable reporter molecule for the robust determination of K_I values for mono- and oligosaccharides as well as glycomimetics. Notably, the ^{19}F R_2-filtered NMR assay developed for Langerin represents the first instance, where this technique is applied to CLRs. In principal, the approach is transferrable to other CLRs and its excellent sensitivity and robustness potentially facilitates the challenging quantification of affinities in this research field. The applicability of the ^{19}F R_2-filtered NMR assay to screen larger compound libraries as well as its transfer to DC-SIGN will be demonstrated in the following sections (Chapter 3.2.2 and 3.2.3). In combination with ^{15}N HSQC and STD NMR experiments, the assay provides a powerful approach to characterize the interactions between designed glycomimetics and Langerin that is extensively utilized throughout the following chapters (Chapter 3.3 and 3.4). The combination of both protein- and ligand-observed experiments to detect GPIs, either directly or *via* competitive binding experiments, is ideally suited to exclude false-positive mechanisms such as the aggregation of glycomimetics or the denaturation of CLRs (Gossert and Jahnke, 2016). Moreover, ^{15}N HSQC NMR experiments are conducted with the monomeric CRD, while both the STD NMR and the ^{19}F R_2-filtered NMR experiments are conducted with the trimeric ECD. Hence, the characterization does not only provide information on the binding mode, it will additionally explore the effects of receptor oligomerization.

This section concludes with the K_I determination for an extended set of monosaccharide (Figure 9c and d, Table 3). These investigations aimed to both validate additional literature values and to identify novel monosaccharide ligands to aid glycomimetic ligand design. Apparent K_I values were previously reported for Fuc ($K_{I,app}$ = 2.6±0.5 mM), Gal-6-OS ($K_{I,app}$ = 3.1±0.1 mM) and Gal ($K_{I,app}$ = 27.6±0.1 mM) (Feinberg *et al.*, 2013; Stambach and Taylor, 2003). Interestingly, Gal (K_I = 80±20 mM) displayed a significantly higher K_I value in the ^{19}F R_2-filtered NMR assay while the affinity increase for Gal-6-OS (K_I = 2.2±0.1 mM) over Man was validated. The common polymorphism K313I has been shown to result in a reduced affinity for Gal-6-OS ($K_{I,app}$ = 10.2±0.4 mM) as recognition of this monosaccharide involves salt bridges formed with K299 and K313 (Feinberg *et al.*, 2013). Simultaneously, the affinity for GlcNAc ($K_{I,app}$ = n.d.) is increased. Moreover, the mutation is statistically linked to the N288D polymorphism which destabilizes the Ca^{2+}-dependent carbohydrate binding site and thereby potentially affects the susceptibility to infectious. As observed for Gal and ManNAc, the K_I value of the 6-deoxy monosaccharide Fuc (K_I = 7±2 mM) determined *via* ^{19}F R_2-filtered NMR is increased compared to literature values. X-ray crystallographic studies reveal that the alternative chair conformation adopted by Fuc still allows for a binding mode analogous to that of Man (Feinberg *et al.*, 2011). The Ca^{2+} ion is coordinated by two equatorial hydroxyl group and enabling the formation of a hydrogen bond between K299 and the adjacent axial hydroxyl group. For *L*-rhamnose (Rha), another 6-deoxy monosaccharide, this binding mode results in a different position for the ring oxygen. This potentially results in the abrogated affinity (K_I > 100 mM) for Langerin.

Interestingly, mannose-6-sulfate (Man-6-OS) (K_I = 0.8±0.1 mM) was identified as a novel, potent monosaccharide ligand of Langerin. The occurrence of Man-6-OS in glycan structures is poorly described, limiting interpretations with respect to the physiological relevance of the discovered interaction. The monosaccharide has been implicated in the sorting of lysosomal hydrolases in amoebozoa, similar to mannose-6-phosphate (Man-6-OP) (K_I = 22±1 mM), which itself displays a considerably lower affinity for Langerin (Feasley *et al.*, 2010; Freeze and Wolgast, 1986). While structural aspects of this particular interaction were not investigated further, the increased affinities of both Gal-6-OS and Man-6-OS suggest a propensity of the Langerin carbohydrate binding site to interact with sulfate groups and their bioisosteres. This hypothesis will be evaluated in the following chapters and eventually contribute substantially to the design of potent glycomimetic ligands for Langerin (Chapter 3.3 and 3.4).

Table 3. Affinity determination for reporter molecule 5.1 and selected monosaccharides.

Name	Structure	K_I [mM]	K_D [mM]	Relative potency[a]
5.1			7.9±0.7[b] 5.8±0.5[c]	0.57
Man		4.5±0.5[b]	5.8±0.3[c]	1.0
ManNAc		25±1[b]	21±3[c]	0.18
Man-6-OP		22±1		0.21
Man-6-OS		0.8±0.1		6.6
Gal		80±20		0.056
Gal-6-OS		2.2±0.1		2.1
Rha		>100		
Fuc		7±2		0.63

[a] The relative potency was calculated utilizing the K_I value determined for Man.
[b] This value was determined from three independent titration experiments *via* ^{19}F R_2-filtered NMR.
[c] This value was determined *via* ^{15}N HSQC NMR.

3.2.2. Assay Optimization - Screening Applications and Multivalent Ligands

The ^{19}F R_2-filtered NMR assay was demonstrated to enable the robust determination of K_I values for a set of monosaccharides. Yet, the efficient elucidation of SARs for focused libraries of glycomimetics or the identification of novel glycomimetic scaffolds from fragment screening requires an assay optimized for throughput and material consumption. Ideally, an optimized setup provides the required sensitivity for the affinity ranking of identified hits.

Along these lines, the applicability of the ^{19}F R_2-filtered NMR assay to the screening of larger libraries was explored. Initially, assay performance of the original setup utilized in the preceding section was evaluated (Setup 1) (Figure 12a). The accuracy of K_I determination is high as determined from experiments conducted in triplicate (relative S.E. = 0.09) (Table 3). Furthermore, an excellent Z-score of 0.69 was determined to quantify the assay performance (Zhang *et al.*, 1999). Generally, competitive binding experiments display an upper limit for accurate affinity quantification (Roehrl *et al.*, 2004). The simulation of binding isotherms at different K_I values indicated that satisfactory sensitivity is maintained for affinities up to 1 µM. Considering the low affinities determined for various monosaccharides and expected for fragment hits, this affinity range is suitable for the initial stages of glycomimetic ligand design for Langerin (Table 3). Notably, the K_I determination in Setup 1 involves the measurement of $R_{2,obs}$ values at least five competitor concentrations. For each $R_{2,obs}$ value, CPMG experiments are conducted at five relaxation times T. To ensure a satisfactory SNR for the ^{19}F NMR spectra, affinity determination requires 225 min on a regular 600 MHz NMR spectrometer. This experimental time, in combination with a receptor consumption m_P of approximately 1 mg, renders Setup 1 unsuitable for screening applications.

To optimize both material consumption and throughput, an alternative assay setup was explored (Setup 2) (Figure 12b). A retrospective analysis of previously acquired affinity data served to predict the accuracy of K_I value estimation at a single ligand concentration. The data set was composed from titration experiments with Man and ManNAc as well as the focused library of ManNAc analogs **5** designed and synthesized in parallel (Chapter 3.3.1). Data points were selected if the concentrations fell within one order of magnitude of the K_I value determined in Setup 1 (Figure 12b). The hereby estimated affinities fell within 46% of the real K_I value for 99% of the experiments (σ = 0.23). The experimental time for Setup 2 amounts to 45 min per estimated K_I value and the receptor consumption is reduced to 0.2 mg. Moreover, screening a library at a concentration around the K_I value of the parent scaffold reduces the ligand consumption by about one order of magnitude compared to complete titration experiments.

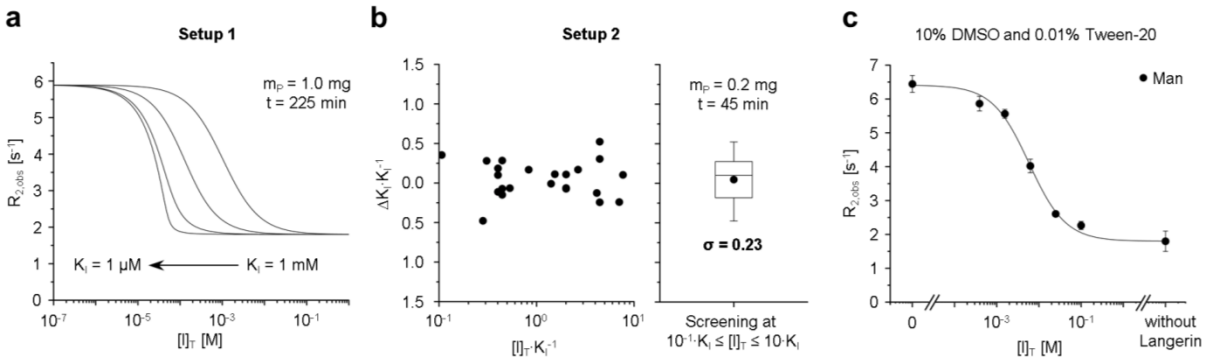


Figure 12. Assay optimization for screening applications.

a. The simulated binding isotherms indicate the upper limit of accurate affinity determination in Setup 1. The isotherms begin to converge at K_I values below 1 µM. **b.** Setup 2 involves the estimation of K_I values from single ligand concentrations and allows for the efficient elucidation of SARs or fragment screening. 23 data points selected from competitive binding experiments with **5.7**, **5.8**, **5.9**, **5.11**, Man and ManNAc served to simulate the accuracy of affinity determination (Chapter 3.3.1). **c.** The ^{19}F R_2-filtered NMR assay was demonstrated to yield robust K_I values for Man (K_I = 5.6±0.2 mM) in presence of 10% DMSO and 0.01% Tween-20. Adapted with permission from (Wamhoff *et al.*, 2016).

Finally, the robustness of K_I determination in presence of additives typically utilized in fragment screenings was evaluated (Figure 12c). A titration experiment with Man (K_I = 5.6±0.2 mM) in presence of 10% DMSO and 0.01% Tween-20 revealed an affinity comparable to the K_I value obtained in absence of these additives (Table 3). The detergent Tween-20 reduces the detection of false positives due to aggregation, while DMSO is improves the solubility of fragments and glycomimetics (Shoichet, 2006). Collectively, these features facilitate the elucidation of SARs for focused glycomimetic libraries and enable fragment screening *via* the ^{19}F R_2-fitlered NMR assay. Experiments at a single ligand concentration provide optimized throughput and material consumption while still allowing for the accurate estimation of K_I values. Notably, the optimization of the assay was essential for the efficient affinity determination of Man analog libraries **12** and **14**. This application is discussed in the following chapter (Chapter 3.3).

Furthermore, the utility of Setup 2 was demonstrated in an explorative fragment screening to identify novel glycomimetic scaffolds (Figure 13). To implement this screening, 290 fragments were randomly selected from our in-house library and screened at concentrations of 0.5 mM in presence of 10% DMSO and 0.01% Tween-20 (Appendix A.5). As low to moderate hit rates are generally expected for carbohydrate binding sites, five to six fragments were binned into one mixture to further optimize throughput and material consumption. Only four fragment mixtures displayed significantly reduced $R_{2,obs}$ values, corresponding to a low hit rate of 1.4% (Figure 13a). Moreover, none of these fragment mixtures displayed an apparent K_I value below 5 mM. Man, on the other hand, was reliably detected at the same concentration, highlighting the excellent sensitivity of ^{19}F R_2-filtered NMR assay. As discussed above, these observations further corroborate the classification of the Langerin carbohydrate binding site as a challenging target for drug discovery (Chapter 3.1).

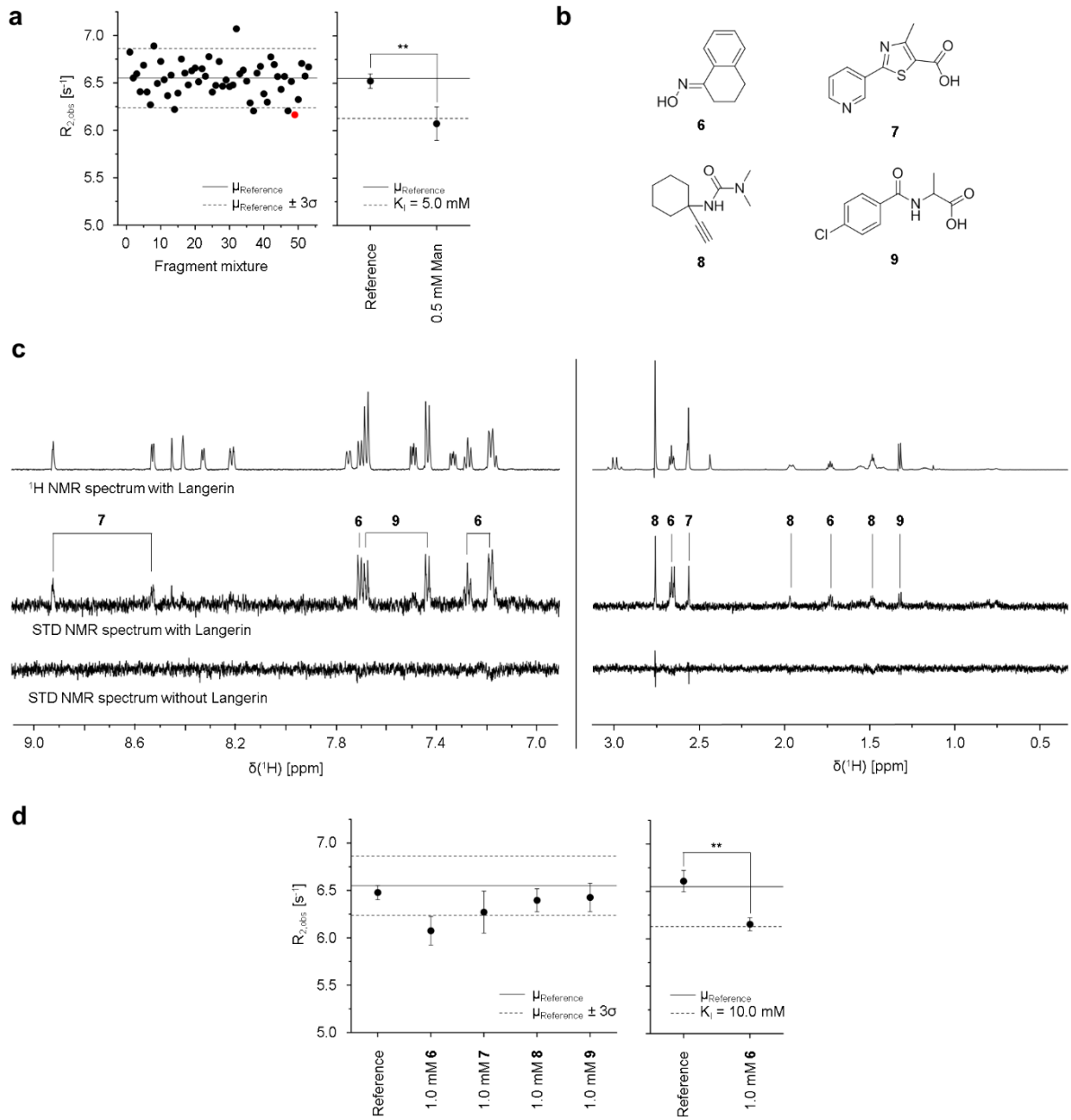


Figure 13. Fragment screening *via* ^{19}F R_2-filtered NMR.

a. Setup 2 of the ^{19}F R_2-filtered NMR assay was utilized in an explorative fragment screening. While only few hits were identified, Man was reliably detected at a concentration of 0.5 mM ($n = 3$ and $p < 0.01$, Student's t test). **b. and c.** The fragment mixture displaying the highest competition was analyzed by STD NMR at a saturation time t_{sat} of 4.0 s. Overall, four out of six fragments were found to interact with the Langerin ECD. STD NMR spectra are magnified 6.25-fold (left) or 25.00-fold (right). **d.** Upon deconvolution of this mixture, only fragment **6** was observed to compete with reporter molecule **5.1** ($K_{I,est}$ = 9±1 mM, $n = 3$ and $p < 0.01$, Student's t test). Adapted with permission from (Wamhoff *et al.*, 2016).

The fragment mixture displaying the highest competition ($\Delta R_{2,obs}$ = 0.39 s^{-1}) was subsequently analyzed *via* STD NMR (Figure 13b and c). While no STD effects were observed in absence of Langerin, binding was detected for four of the six fragments. The deconvolution of the fragment mixture at concentrations of 1 mM revealed fragment **6** ($K_{I,est}$ = 9.7±1.2 mM) as a potential glycomimetic (Figure 13d). Fragments **7**, **8** and **9** did not display competition with binding of reporter molecule **5.1**. The integration of ^{19}F R_2-filtered and STD NMR experiments generally reduces the experimental time required for the deconvolution of fragment mixtures and thus represents a powerful approach for the identification of glycomimetic fragments.

Only two examples for the identification of fragment inhibitors for GBPs have been reported. Functional glycomimetics targeting CD44 bind to a secondary pocket adjacent to the carbohydrate binding site and inhibit the recognition of hyaluronate *via* steric hindrance (Liu and Finzel, 2014). Additionally, several classes of allosteric DC-SIGN inhibitor were recently discovered in our laboratory (Aretz *et al.*, 2017). Fragment **6** bears only a single functional group, an oxime, capable of hydrogen bond formation. Hence, inhibition mechanisms similar to those listed above can be envisioned. Additionally, oxime groups are potent metal chelators and a direct interaction with the Ca^{2+} ion would represent an alternative mechanism (Smith *et al.*, 2003). While an interaction of fragment **6** with Langerin was observed in two orthogonal NMR techniques and in the presence of Tween-20, dose dependency could not be shown, potentially due to limited solubility. Currently, the explorative fragment screening is followed up by a comprehensive screening of our in-house library combining a cell-based competitive binding assay with ^{19}F R_2-filtered and ^{15}N HSQC NMR experiments. This screening will potentially result in the identification of additional glycomimetic scaffolds and SAR studies will explore the evolution of identified fragment hits, including fragment **6**, into potent glycomimetic ligands. Ideally, this fragment-based design approaches might be integrated with information gained from the design of monosaccharide analogs as glycomimetic Langerin ligands.

The multivalent presentation of glycomimetic ligands represents a powerful strategy to target oligomeric CLRs in which geometry-based avidity effects are leveraged to improve potency and specificity (Cecioni *et al.*, 2015). However, chelate cooperativity and statistical rebinding introduce complex equilibria that cannot be expressed analytically (Hunter and Anderson, 2009; Kitov and Bundle, 2003). Thus, affinities are typically quantified *via* IC_{50} values obtained from the empirical Hill equation (Setup 3) (Equation 5). While the design of multivalent glycomimetics for Langerin will be discussed in the following chapters, these considerations have important implications for the optimization of the ^{19}F R_2-filtered NMR assay, in particular with respect to the upper limit of accurate affinity determination (Chapter 3.5).

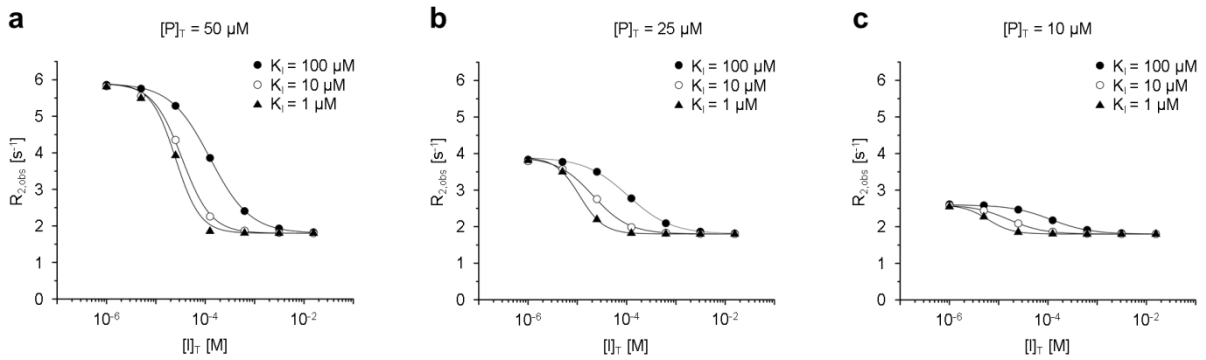


Figure 14. Assay optimization for multivalent ligands.

a to c. Due to complex equilibria and cooperativity effects in experiments with multivalent glycomimetics, IC_{50} value rather than K_I values are utilized to quantify affinities. The limits of accurate affinity quantification in Setup 3 were explored *via* the simulation of binding isotherms utilizing Equation 4 at the indicated Langerin ECD concentrations $[P]_T$ and apparent K_I values. IC_{50} values were determined by fitting Equation 5 to the generated data points and compared to the assumed K_I values. Conducting the ^{19}F R_2-filtered NMR assay in presence of 25 µM Langerin ECD enables the discrimination of K_I values of 1 µM and 10 µM while still providing a satisfactory dynamic range of 2.1.

To estimate the sensitivity of the assay in Setup 3, isotherms corresponding to a one-site binding model were simulated at different apparent K_I values and receptor concentrations $[P]_T$ (Figure 14, Equation 4). Subsequently, IC_{50} values and Hill factors p were determined *via* fitting the Hill equation to the generated data points (Table 4, Equation 5). Notably, these simulations are an approximation as they do *e.g.* not consider cooperativity effects explicitly. Nevertheless, a comparison of the apparent K_I and the IC_{50} values provides a measure for the relative sensitivity of the assay for multivalent high affinity ligands. Strikingly, and in contrast to Setup 1, Setup 3 does not enable the discrimination of apparent affinities below 10 µM when utilizing a receptor concentration of 50 µM. The relative sensitivity of the ^{19}F R_2-filtered NMR assay can be optimized by reducing the receptor concentration. This optimization will simultaneously reduce the dynamic range of the assay. At a $[P]_T$ value of 25 µM a satisfactory dynamic range of 2.1 is maintained while the relative sensitivity is improved by approximately 2-fold at an apparent affinity of 1 µM. These conditions still account for a systematic underestimation of the apparent affinity by a factor of approximately 10. Nevertheless, they enable the robust discrimination of apparent K_I values of 10 µM and 1 µM *via* the determination of IC_{50} values. This renders Setup 3, at reduced receptor concentrations, valuable for the design of multivalent glycomimetics even at apparent affinities of 1 µM to 10 µM.

Table 4. Assay optimization for multivalent ligands.

$[P]_T$ [µM][a]	K_I [µM][a]	Dynamic range	p[b]	IC_{50} [µM][b]	Relative sensitivity
50	100	3.3	1.1	130	0.76
50	10	3.3	1.5	34	0.29
50	1.0	3.3	1.7	25	0.04
25	100	2.1	1.0	110	0.91
25	10	2.1	1.2	21	0.47
25	1.0	2.1	1.8	11	0.09
10	100	1.4	1.1	120	0.83
10	10	1.4	1.2	16	0.63
10	1.0	1.4	1.8	6.0	0.17

[a] These values were utilized to simulate binding isotherms for competitive binding experiments from Equation 4
[b] These values were obtained *via* fitting Equation 5 to the simulated binding isotherms.

The findings discussed above ultimately reveal the limitations of reporter molecule **5.1** and potentially of the ^{19}F R_2-filtered NMR approach in general. Further optimization to accurately quantify affinities below 1 µM requires substantially reduced receptor concentrations. As **5.1** (K_D = 7.9±0.7 mM) displays an affinity in low millimolar range, this will necessarily result in a low dynamic range. The design of a more potent reporter molecule would resolve this problem and provide a satisfactory dynamic range, even at reduced receptor concentrations. However, a shift towards the intermediate exchange regime on the ^{19}F NMR timescale is expected for increased affinities and the resulting exchange contribution $R_{2,ex}$ to $R_{2,obs}$ would considerably complicate data analysis. In case glycomimetic ligands with affinities in the low micromolar range are available, fluorescence polarization represents a powerful alternative approach to quantify high affinity interactions for CLRs (Han *et al.*, 2010).

In addition to the decreasing relative sensitivity, the simulated data set indicates that determined Hill factors p do not reflect the cooperativity effects observed for analyzed multivalent interaction. As apparent K_I values approach the upper limit of accurate affinity quantification, the corresponding binding isotherms converge and display increasing slopes. When fitting a Hill equation to the corresponding data points, this phenomenon alone leads to increased p values (Table 4). Thus, mechanistic interpretations of Hill factors p obtained from ^{19}F R_2-filtered NMR should be treated cautiously.

In summary, the utility of the ^{19}F R_2-filtered NMR assay for screening applications and the design of multivalent glycomimetics was evaluated. Based on the robust determination of K_I values from titration experiments (Setup 1), a setup with optimized throughput and material consumption was developed (Setup 2). The setup provides the required accuracy for the estimation of K_I values from competitive binding experiments at a single competitor concentration. Consequently, the ^{19}F R_2-filtered NMR assay enables the efficient elucidation of SARs for focused glycomimetic libraries as well affinity ranking of hits identified from fragment screening. Notably, an explorative fragment screening resulted in the identification of a novel glycomimetic fragment for Langerin. Finally, the evaluation was expanded to explore the limitations of the assay with respect to the quantification of multivalent interactions (Setup 3). Overall, this section highlights the importance of assay development for CLRs. The optimization of the ^{19}F R_2-filtered NMR assay in particular was essential for the design of glycomimetic Langerin ligands presented in this dissertation. The application of the optimized setup to the identification of potent monosaccharide analogs will be discussed in the following chapters while this chapter will conclude with the transfer of the developed ^{19}F R_2-filtered NMR approach to DC-SIGN (Chapter 3.3 and 3.4).

3.2.3. Assay Transfer to DC-SIGN

The ^{19}F R_2-filterd NMR assay should in principal be transferable to other CLRs. The transfer can be implemented by either utilizing reporter molecule **5.1** or by designing a novel reporter molecule *via* the fluorination of other monosaccharide scaffolds. While various fluorination strategies for mono- and oligosaccharides have been proposed, the introduction of trifluoroacetamido groups provides optimal spectroscopic properties (Dumitrescu *et al.*, 2014; N'Go *et al.*, 2014; Unione *et al.*, 2017). Here, the absence of J coupling is of particular interest as it allows for the utilization of the CPMG pulse sequence to determine $R_{2,obs}$ values (Segawa and Bodenhausen, 2013). The utility of reporter molecule **5.1** depends on its affinity for the CLR of interest as well as the corresponding $R_{2,b}$ value. CLRs reported to recognize Man or ManNAc should thus be prioritized to explore assay development.

One of the CLRs reported to bind to Man and ManNAc is DC-SIGN (Garber *et al.*, 2010; Holla and Skerra, 2011). Similar to Langerin, DC-SIGN is a tetrameric, endocytic receptor involved in the recognition of self- and pathogen-associated carbohydrates and displays endocytic activity promoting antigen processing and presentation to T cells (Fehres *et al.*, 2015c; Geijtenbeek *et al.*, 2000). It is expressed on macrophages as well as DC subsets distinct from LCs and has been implicated in the transmission of HIV to T cells (Geijtenbeek *et al.*, 2000). This has rendered DC-SIGN a prominent target for drug discovery. As Langerin potentially displays protective properties in HIV infection, the utilization of ^{19}F R_2-filtered NMR to establish a specificity assay is of considerable interest for the design of anti-infectives (de Witte *et al.*, 2007). From the perspective of antigen delivery to Langerin, the availability of ^{19}F R_2-filtered NMR assays for other CLRs is generally desirable because it enables *in vitro* specificity screenings for targeting ligands. Obtained K_I values potentially complement *in vivo* data for targeted liposomes and provide information on the targeted DC subset.

The ^{19}F R_2-filtered NMR assay for DC-SIGN was implemented in analogy to the procedure presented for Langerin (Figure 15, Chapter 3.2.1). As available protocols for the expression and purification of the DC-SIGN ECD did not provide the required yields for ^{19}F R_2-filtered NMR experiments, the assay was only established for the CRD. Briefly, reporter molecule **5.1** displays substantial line broadening $\Delta\nu_{0.5}$ in presence of DC-SIGN (Figure 15a). As observed for Langerin, **5.1** did not display CSPs. The interaction between DC-SIGN and **5.1** was subsequently quantified *via* the determination of observed relaxation rates $R_{2,obs}$ from CPMG experiments (Figure 15b) (Carr H. Y. and M., 1954; Meiboom and Gill, 1958). The addition of EDTA served to validate the Ca^{2+}-dependency of **5.1** binding (Figure 15c).

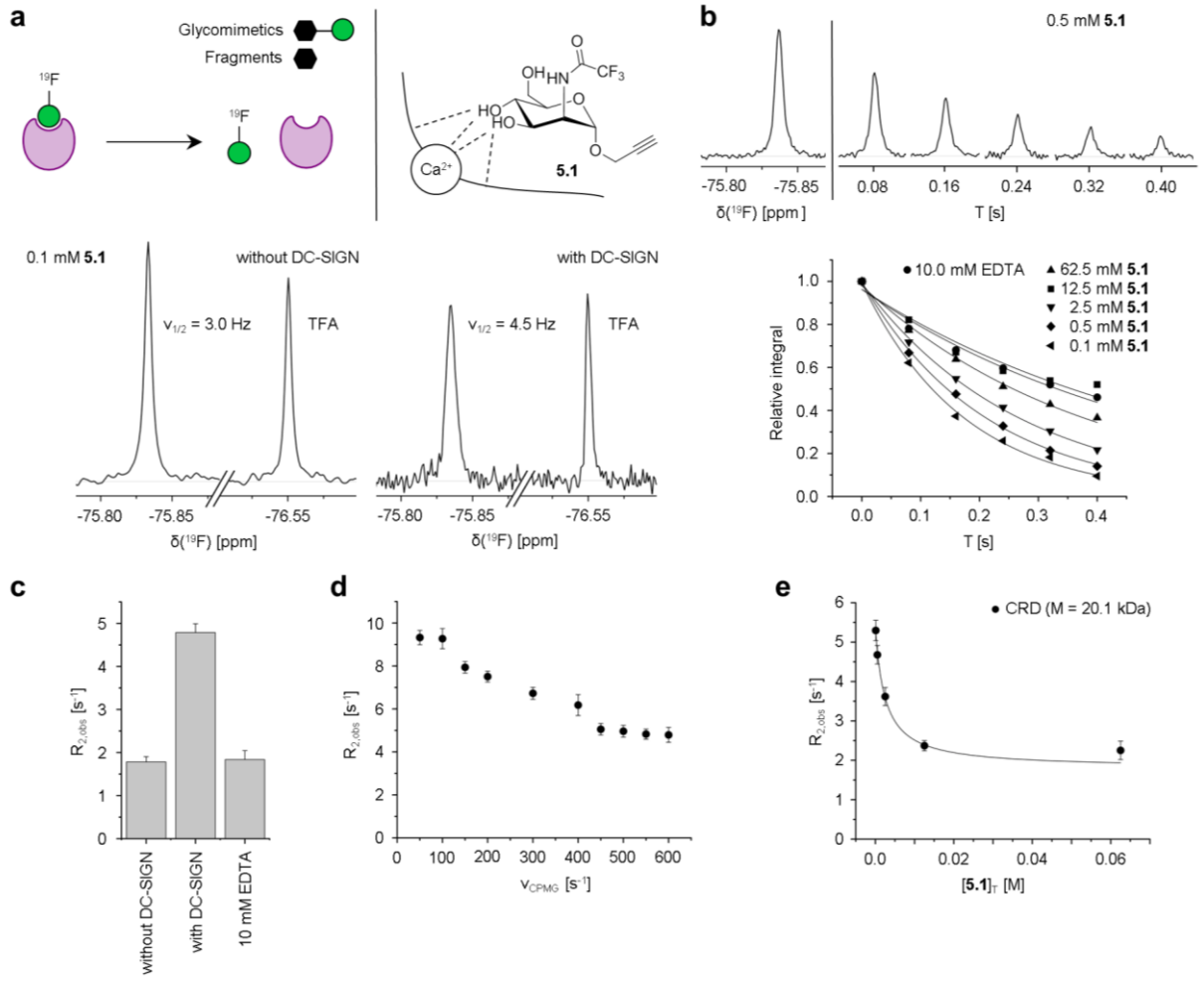


Figure 15. ^{19}F R_2-filtered NMR assay development with DC-SIGN.

a. In presence of the DC-SIGN CRD, the ^{19}F NMR resonance of the trifluoroacetamido group of reporter molecule **5.1** displays line broadening. This phenomenon was utilized to transfer the ^{19}F R2-filtered NMR assay from Langerin to DC-SIGN. The spectrum in absence of DC-SIGN is processed with an exponential apodization function at 1.8 Hz. **b.** As for Langerin, the interaction between the reporter molecule and DC-SIGN was quantified *via* the transversal relaxation rate $R_{2,obs}$ using the CPMG pulse sequence (Carr H. Y. and M., 1954; Meiboom and Gill, 1958). Representative decay curves for the CRD are shown. **c.** The Ca^{2+}-dependency of the interaction was validated *via* addition of EDTA. The $R_{2,obs}$ value in absence of DC-SIGN was determined from four independent experiments. The standard error for the experiments in presence of DC-SIGN and EDTA is derived directly from the fitting procedure. **d.** Relaxation dispersion experiments for **5.1** in presence of the DC-SIGN CRD indicate at negligible exchange contribution $R_{2,ex}$ at a ν_{CPMG} value of 500 Hz. In contrast to Langerin, a considerable exchange contribution is observed a lower ν_{CPMG} values. **e.** Titration experiments revealed a K_D value of 2.3±0.5 mM. A representative binding isotherm is shown. The K_D value was determined from three independent titrations.

Interestingly, relaxation dispersion experiments revealed a substantial chemical exchange contribution $R_{2,ex}$ at a frequency of 180° pulses ν_{CPMG} of 100 Hz ($R_{2,ex}$ = 4.3 s^{-1}) (Figure 15c). This observation indicates either an increased CSP induced upon binding to DC-SIGN or a decreased dissociation rate k_{off} compared to experiments with Langerin. However, at a ν_{CPMG} value of 500 Hz, the chemical exchange contribution $R_{2,ex}$ is negligible, enabling the determination of K_D and K_I values from titration experiments. Finally, the affinity of **5.1** for the DC-SIGN CRD (K_D = 2.3±0.5 mM) was determined *via* ^{19}F R_2-filtered NMR assuming a one-site binding model (Figure 15e, Table 5).

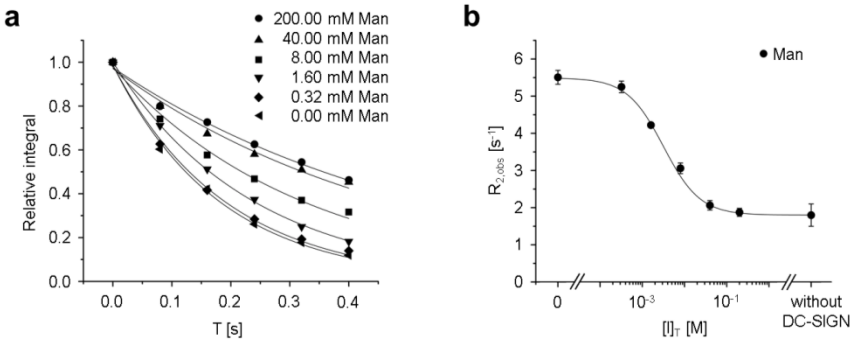


Figure 16. K_I determination for Man with DC-SIGN.

a. Representative decay curves from the competitive binding experiments with Man are shown. **b.** Competitive binding experiments yield a K_I value for Man (K_I = 3.0±0.3 mM), consistent with literature values (Holla and Skerra, 2011).

The affinity of **5.1** for DC-SIGN is higher than for Langerin. As the assay conditions are identical for both receptors, this results in an increased dynamic range of 3.1 when comparing the CRDs. Interestingly, the determined $R_{2,b}$ value is decreased in spite of the higher molecular weight of the DC-SIGN CRD (M = 20.1 kDa). While this observation might be explained by a decreased CSP compared Langerin, this explanation is contrasted by the increased chemical exchange contribution $R_{2,ex}$ and was not investigated further. Finally, the determination of K_I values by ^{19}F R_2-filtered NMR was established *via* competitive binding experiments with Man (K_I = 3.0±0.3 mM) (Figure 16, Table 5). The binding isotherm is consistent with the assumed one-site binding model and complete competition with **5.1** binding was observed. Importantly, the determined affinity is comparable to literature values (Holla and Skerra, 2011).

In contrast to Langerin, the interaction between **5.1** and DC-SIGN was not investigated *via* ^{15}N HSQC NMR. Yet, the ^{19}F R_2-filtered NMR data acquired for DC-SIGN are both consistent with the experiments conducted for Langerin and with the assumed one-site binding model. Moreover, the K_I value obtained for Man is comparable to literature values (Holla and Skerra, 2011). These observations and the extensive validation of the approach for Langerin justify the conclusion that ManNAc analog **5.1** represents a suitable reporter molecule for DC-SIGN and will enable the robust determination of K_I values for glycomimetics. In this dissertation, the ^{19}F R_2-filtered NMR assay will be utilized to evaluate the specificity of glycomimetic Langerin ligands versus DC-SIGN as discussed in the following chapters (Chapter 3.4). Yet, its relevance extends beyond this application as the assay will be pivotal for the fragment-based design of DC-SIGN inhibitors currently pursued in our laboratory (Aretz *et al.*, 2017). Moreover, it represents a valuable addition to available methods for the *in vitro* characterization of anti-infectives targeting *e.g.* HIV in general.

Table 5. Parameters for ^{19}F R_2-filtered NMR assay with DC-SIGN.

Parameter	CRD
$R_{2,f}$ [s^{-1}]	1.8±0.3[a]
$R_{2,b}$ [s^{-1}]	190±20[b]
K_D [mM] for **5.1**	2.3±0.5[b]
K_I [mM] for Man	3.0±0.3

[a] This value was determined from four independent experiments *via* ^{19}F R_2-filtered NMR.
[b] This value was determined from three independent titration experiments *via* ^{19}F R_2-filtered NMR.

3.3. Structure-Activity Relationship between the Mannose Scaffold and Langerin

3.3.1. Structure-Based *in silico* Design of *N*-Acetylmannosamine Analogs

Langerin is a trimeric CLR predominantly expressed on LCs that promotes antigen processing and cross-presentation to $CD8^+$ T cells (Clausen and Stoitzner, 2015; Idoyaga *et al.*, 2011; Klechevsky *et al.*, 2008). Hence, glycomimetic Langerin ligands are of considerable interest for the development of cancer immunotherapies as their conjugation to liposomes enables the specific and efficient delivery of TAAs (Stoitzner *et al.*, 2014). Glycomimetics provide critical advantages over alternative ligand classes such as naturally occurring mono- and oligosaccharides, antibodies or drug-like molecules targeting secondary binding pockets (Chapter 1.2). However, the carbohydrate binding site of Langerin and other CLRs is hydrophilic, harbors a Ca^{2+} ion and displays high solvent exposure (Feinberg *et al.*, 2011). These characteristics render glycomimetic ligand design challenging. Nevertheless, recent advances have demonstrated the feasibility of glycomimetic ligand design for other CLRs such as E-Selectin or DC-SIGN as well as several bacterial lectins (Chang *et al.*, 2010; Han *et al.*, 2010; Hauck *et al.*, 2013; Mangold *et al.*, 2012). No ligand design studies other than the findings presented in the dissertation have been reported for Langerin and limited SAR information on the recognition of mono- and oligosaccharides was available during the initial stages of the design process (Wamhoff *et al.*, 2016). In combination with the limited access to libraries of mono- and oligosaccharide analogs in general, this further impeded the design of glycomimetic Langerin ligands.

By contrast, X-ray crystallography provided detailed structural information on the Ca^{2+}-dependent recognition of terminal Man, Fuc, GlcNAc and Gal-6-OS (Feinberg *et al.*, 2013; Feinberg *et al.*, 2011). X-ray structures enable the implementation of structure-based *in silico* design approaches for the design of focused libraries of mono- or oligosaccharide analogs. As highlighted above, binding hypotheses generated from molecular docking potentially reduce the considerable synthetic efforts associated with the preparation of these libraries (Chapter 1.4). This is of particular value during the initial stages of the design process when only limited SAR information is available. While the prediction of GPIs remains challenging, *in silico* approaches have been successfully applied to identify potent carbohydrate analogs for the ILRs Siglec-1, -4 and -7 (Mesch *et al.*, 2010; Nycholat *et al.*, 2012; Rillahan *et al.*, 2013). The studies involved both the prediction of favorable substituents from virtual screenings and the retrospective evaluation of SARs.

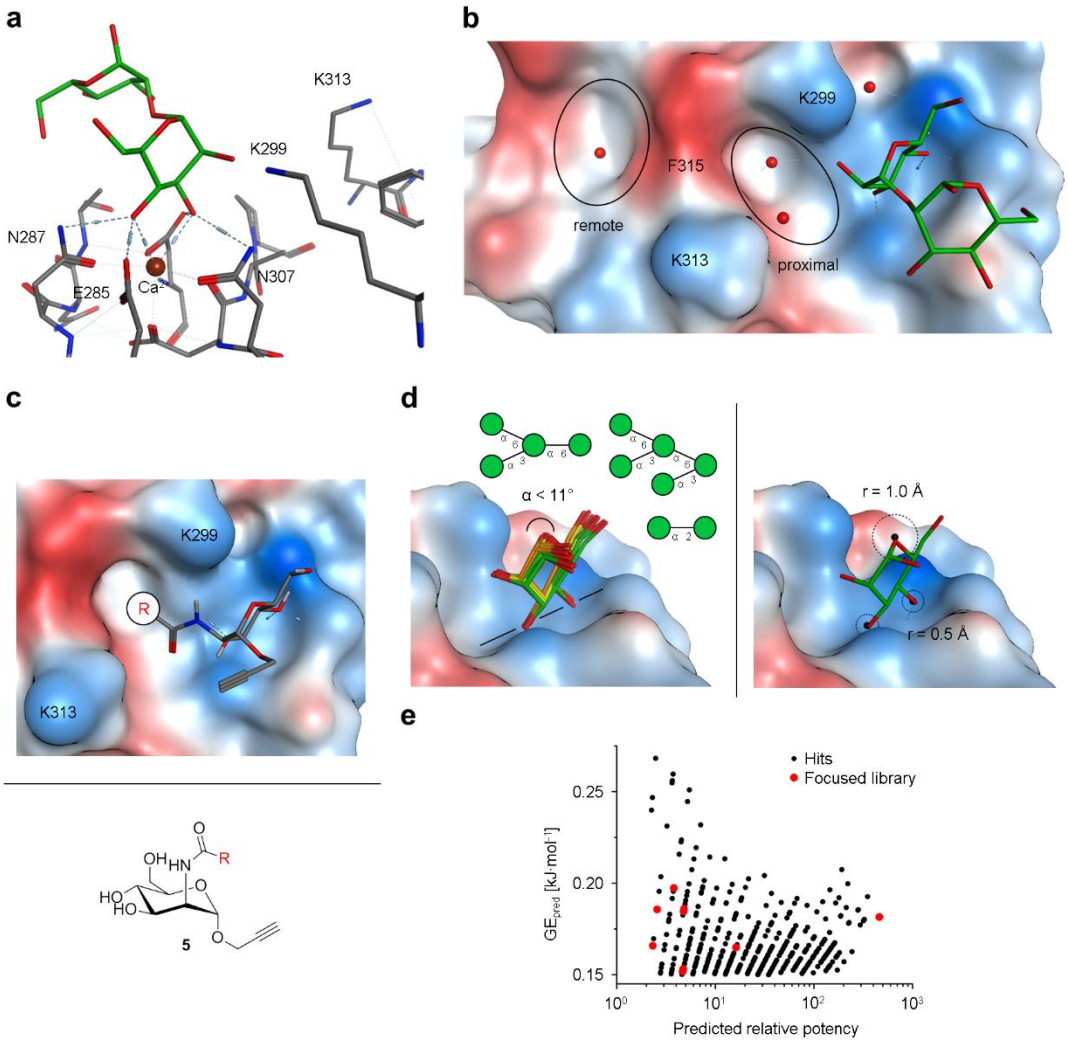
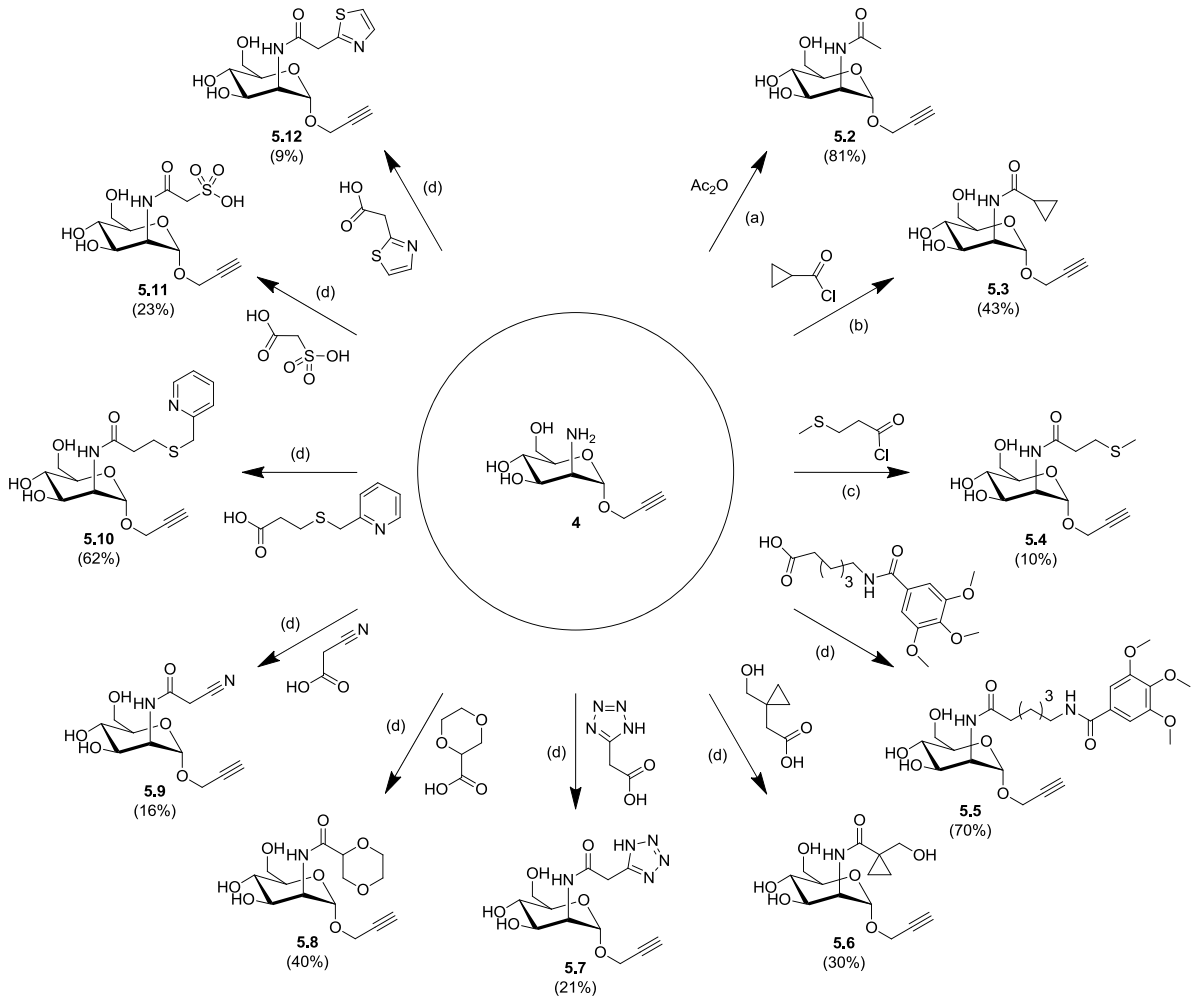


Figure 17. ***In silico*** **design strategy for ManNAc analogs 5.**

a. Man recognition by Langerin is dominated by the Ca^{2+} coordination *via* two equatorial hydroxyl groups. Only few secondary interactions are observed. One of these interactions is the weak hydrogen bond formed between K299 and the hydroxyl group in C2 (PDB code: 3P5F) (Feinberg *et al.*, 2011). **b.** A visual evaluation of the receptor surface reveals two pockets in axial direction of C2 of the Man scaffold. The pockets are flanked by K299, K313 and F315 and harbor several structural H_2O molecules. **c.** The pockets were targeted by the structure-based *in silico* design of ManNAc analogs **5**. **c.** To implement the *in silico* screening, carboxamide conformations were generated *via* low mode MD simulations and utilized for the *in situ* conjugation of commercially available carboxylic acids. Moreover, the alternative conformation of K313 observed for the complex with Gal-6S was accounted for in additional docking runs (Feinberg *et al.*, 2011). Here, one of the five initial states of the binding site is depicted (Appendix B.2). **d.** The structural alignment of the binding sites of available X-ray structures of Langerin in complex with different Man-type oligosaccharides is depicted (PDB codes: 3P5D, 3P5E and 3P5F) (Feinberg *et al.*, 2011). The orientation of the directly bound Man is highly conserved. Based on these observations, a pharmacophore model was defined to constrain the orientation of the Man scaffold during the MM simulation-based refinement of generated docking poses. The defined features require an oxygen atom to be positioned within the indicated spheres. **e.** The distribution of hits with respect to the predicted relative potency and the corresponding group efficiency GE_{pred} is depicted. Both values were calculated relative the score obtained for ManNAc analog **5.2** (Scheme 2). The focused library of ManNAc analogs **5** was selected from hits displaying a GE_{pred} value higher than 0.15 $kJ \cdot mol^{-1}$. The receptor surface is colored according to its lipophilicity (lipophilic: red, hydrophilic: blue). Adapted from (Wamhoff *et al.*, 2016).

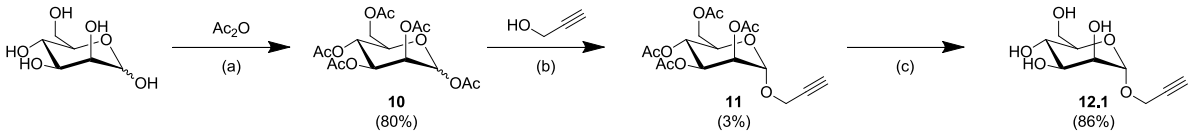
The developed *in silico* design strategy to identify glycomimetic ligands for Langerin integrates the availability of X-ray structures with the obtained SAR information and considerations regarding the synthetic tractability of designed glycomimetics. On the one hand, the Man scaffold displays increased affinities compared to *e.g.* Fuc or Gal (Table 3). Moreover, the derivatization of the latter is generally considered to be more laborious (Hauck *et al.*, 2013). On the other hand, the recognition of Man-type di- and oligosaccharides by Langerin in dominated by the Ca^{2+} coordination *via* its two equatorial hydroxyl groups (Figure 17a). Only few secondary interactions are observed with the weak hydrogen bond formed between K299 and the axial hydroxyl in C2 likely being responsible for the observed specificity versus ManNAc (Table 3) (Feinberg *et al.*, 2011). This is consistent with the marginal affinity increase observed for a Man disaccharide and other small oligosaccharide (Hanske, 2016; Holla and Skerra, 2011). Consequently, Man was identified as a promising scaffold for the *in silico* design of a focused library of monosaccharide analogs.

The visual evaluation of the Langerin surface in proximity of the carbohydrate binding site revealed two potential binding pockets in axial direction of C2 of the Man scaffold (Figure 17b). The pockets were not identified by DoGSite, likely due to their small size (Chapter 3.1). They are flanked by K299, K313 as well as F315 and harbor several structural H_2O molecules conserved over a set of three Langerin complexes with Man-type oligosaccharides (Feinberg *et al.*, 2011). Notably, the displacement of structural H_2O molecules represents a powerful design approach potential resulting in favorable entropic contributions (Binder *et al.*, 2012; Congreve *et al.*, 2011). Additionally, the interactions formed by trifluoromethyl group of ManNAc analog **5.1** result in a decreased K_I values compared to ManNAc (Table 3). Based on these observations, the identified pockets were targeted *via* an *in silico* screening of substituents in axial direction of C2 to compose a focused library of ManNAc analogs **5** (Figure 17c). This strategy was facilitated by the availability of intermediate **4**, enabling the introduction of substituents by the formation of carboxamides (Scheme 1).



Scheme 2. Synthesis of ManNAc analog library 5.

Reaction conditions for the preparation of library 5: (a) MeOH, room temperature; (b) pyridine, room temperature; (c) pyridine, 0°C to room temperature; $MeNH_2$, EtOH, room temperature; (d) PyBOP, DIPEA, DMF, room temperature.



Scheme 3. Synthesis of Man analog 12.1.

Intermediate **10** was prepared as previously published (Hasegawa *et al.*, 2007). Reaction conditions for the preparation of **12.1**: (a) pyridine, 50°C; (b) $BF_3.OEt_2$, anhydrous DCM:ether (2:1), 0°C to room temperature; (c) $MeNH_2$, EtOH, room temperature.

To implement the *in silico* screening, favorable conformations of the carboxamide were evaluated in low mode MD simulations (Figure 17c, Appendix B.2). As K313 adopts an alternative conformation in the Langerin complex with Gal-6-OS, two receptor states were considered for these simulations (Feinberg *et al.*, 2011). Overall, five initial states for the complex were selected for the *in situ* conjugation to previously generated conformers of 24,800 commercially available carboxylic acids (Appendix B.2). The subsequent docking procedure involved a grid-based placement of the carboxylic acid conformers and scoring of docking poses using the proprietary London ΔG function (Chemical Computing Group, 2016). Highly scored poses were refined utilizing MM simulations and rescored *via* the GBVI/WSA ΔG function (Corbeil *et al.*, 2012; Massova and Kollman, 2000). As the orientation of the Man coordinating the Ca^{2+} ion was observed to be conserved over a set of three Langerin complexes with Man-type oligosaccharides, a pharmacophore model was developed to constrain the orientation of the Man scaffold during the MM simulations and to filter generated docking poses (Figure 17d). This approach minimizes potential artifacts arising from the limitations for the prediction of GPIs discussed above (Chapter 1.4). Predicted ΔG values were normalized to the score obtained for ManNAc analog **5.2** and expressed as relative potencies (Scheme 2, Appendix B.2).

Importantly, the composition of the focused library was not solely based on these predicted relative potencies (Figure 17e, Appendix B.2). Generated docking poses were visually assessed and analogs were selected to systematically explore predicted interactions. Here, substituents displaying minimal complexity were prioritized to follow a fragment growing approach. Additionally, more complex substituents were selected to explore the remote pocket for favorable interactions as exemplified by ManNAc analogs **5.5** and **5.10** (Scheme 2, Appendix B.2).

Next, the focused library of ManNAc analogs **5** was synthesized over five steps and subsequently characterized using the ^{19}F R_2-filtered NMR assay (Figure 18, Scheme 2, Table 6, Appendix B.3). The synthesis involved either the utilization of activated carboxylic acid derivatives such as carboxylic anhydrides or acyl chlorides or the *in situ* activation *via* PyBOP (Coste *et al.*, 1990). Notably, a subset of prepared analogs, particularly **5.7** and **5.11**, is highly hydrophilic requiring purification by HPLC utilizing a semi-preparative HyperCarb column (Thermo Scientific). This resulted in yields ranging from 2% to 22%.

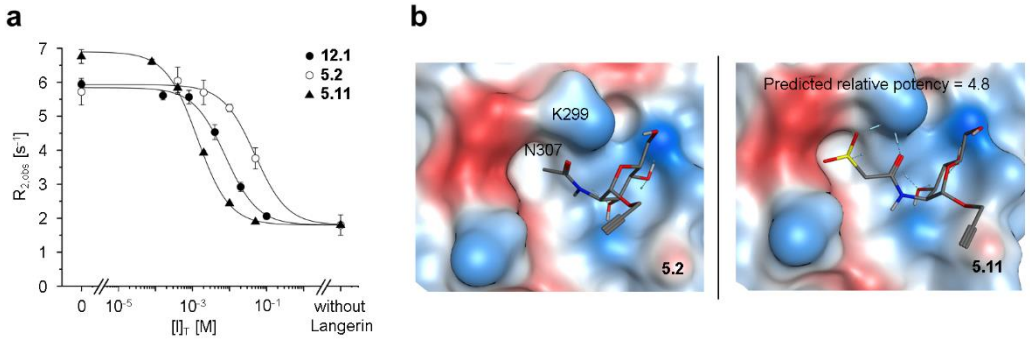


Figure 18. Docking poses and K_I determination for Man analog 12.1, ManNAc analogs 5.2 and 5.11.

a. The ^{19}F R_2-filtered NMR assay served to determine a 5.9-fold affinity increase over **12.1** and a 36-fold affinity increase over **5.2** for **5.11**. **b**. The docking pose generated for **5.11** reveals a set of hydrogen bonds formed between the sulfonate and K299 as well as N307. These interactions likely result in the observed affinity increase. The receptor surface is colored according to its lipophilicity (lipophilic: red, hydrophilic: blue). Adapted from (Wamhoff *et al.*, 2016).

Overall, the SAR is dominated by an unfavorable contribution of the carboxamide as observed for the comparison between Man and ManNAc (Chapter 3.2.1). The formation of an α-mannoside or the introduction of the propargyl group results in an additional affinity decrease reflected by the K_I value determined for **5.2** (K_I = 48±6 mM). No affinity increase was observed for substituents targeting the distal pocket. While this observation might be attributed to the absence of favorable interactions formed by **5.5** (K_I = 45±6 mM) and **5.10** (K_I = 46±6 mM), both substituents contain flexible linker structures. Hence, the formation of these interactions might be masked by unfavorable entropic contributions. By contrast, the SAR obtained for the proximal pocket revealed more potent analogs. The introduction of negatively charged substituents for **5.7** (K_I = 7.1±1.7 mM) and **5.11** (K_I = 1.3±0.1 mM) resulted in substantially improved K_I values.

With a 36-fold affinity increase over reference molecule **5.2**, **5.11** represents the most potent member of the focused library and displays an excellent group efficiency (GE = 2.23 $kJ \cdot mol^{-1}$) (Figure 18a) (Hopkins *et al.*, 2014). This affinity increase can be rationalized by the formation of salt bridges or hydrogen bonds between the sulfonate group and K299 as well as N307 (Figure 18b). Notably, interactions similar to those predicted for **5.11** have been observed for Gal-6-OS, validating the generated docking pose (Feinberg *et al.*, 2011). The propargyl group enables the conjugation of **5.11** to liposomes or fluorescent dyes by Cu(I)-catalyzed Huisgen 1,3-dipolar cycloaddition (Rostovtsev *et al.*, 2002). Consequently, the ManNAc analog bears potential as a targeting ligand for LCs or as a chemical probe to investigate DC immunology in general. The conjugation of naturally occurring mono- and oligosaccharides is typically achieved *via* the anomeric position. To demonstrate the affinity increase (relative potency = 5.9) in this context, Man analog **12.1** (K_I = 7.6±0.1 mM) was synthesized and characterized in the ^{19}F R_2-filtered NMR assay (Figure 18a, Scheme 3, Table 6).

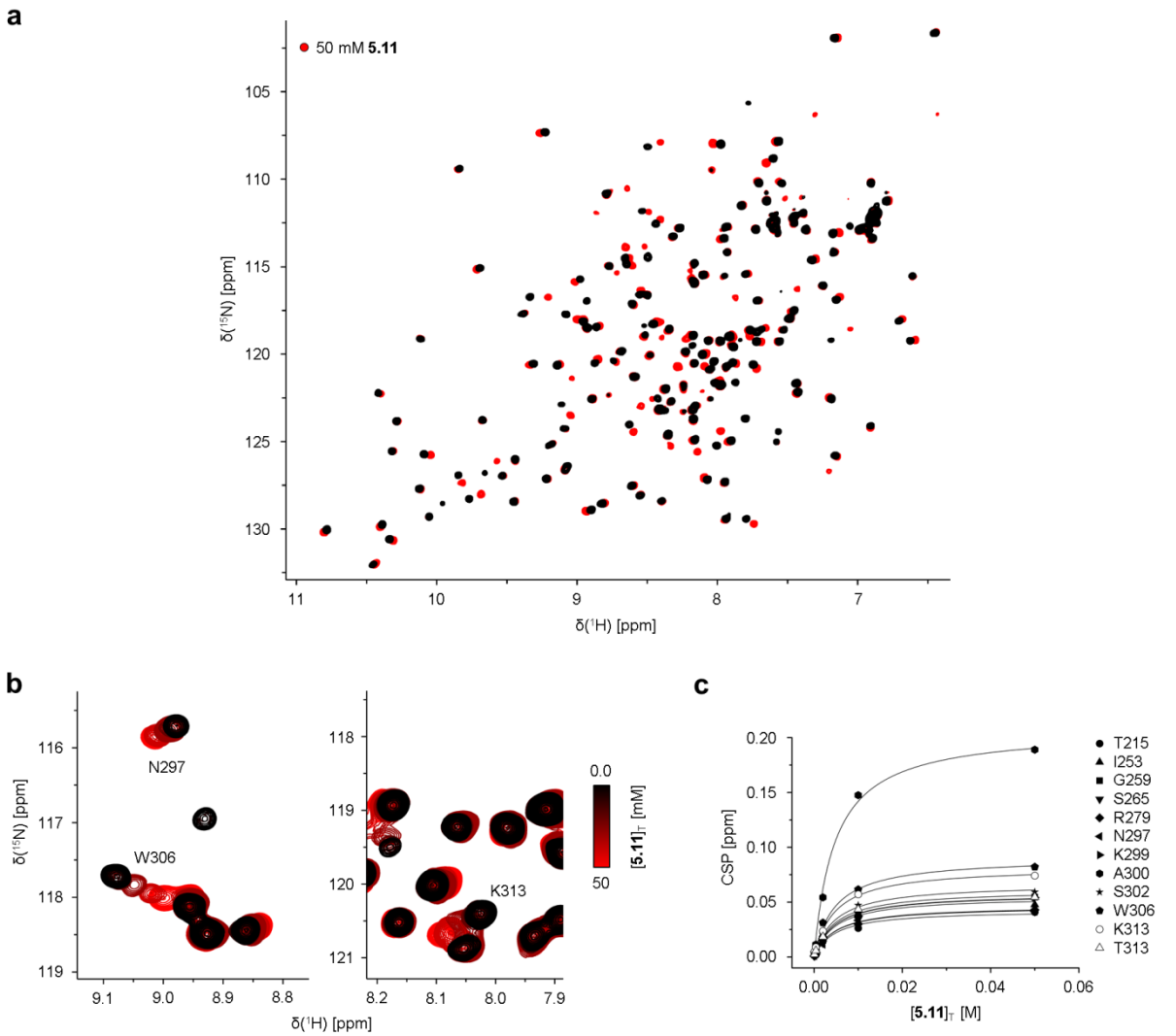
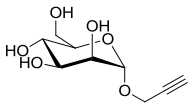
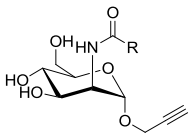
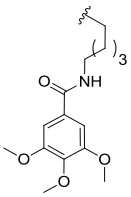
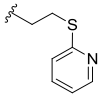


Figure 19. K_D determination for ManNAc analog 5.11.

a. The affinity determined for **5.11** in the ^{19}F R_2-filtered NMR assay were validated in ^{15}N HSQC NMR titration experiments with the Langerin CRD. **b.** Assigned resonances displaying fast chemical exchange and CSPs larger than 0.04 ppm were selected for the determination of the K_D value. Representative resonances corresponding to N297, W306 and K313 are displayed. **c.** The K_D value obtained for **5.11** (K_D = 4.3±0.3 mM) is consistent with the results from the ^{19}F R_2-filtered NMR assay. Adapted from (Wamhoff *et al.*, 2016).

The K_I value obtained for **5.11** was validated *via* ^{15}N HSQC NMR (Figure 19, Table 6). Aside from minor dev*ia*tions, the determined affinity increase over Man (K_D = 4.3±0.3 mM) is consistent with the ^{19}F R_2-filtered NMR experiments (Table 3). To evaluate the binding mode of **5.11**, observed CSPs were mapped on the X-ray structure of Langerin (Figure 20). Notably, the data analysis in this section is based on the preliminary version of the reported resonance assignment (Hanske *et al.*, 2016). This version was also utilized for experiments with Man, ManNAc and ManNAc analog **5.1** (Chapter 3.2.1).

Table 6. Structure-activity relationship of ManNAc analog library 5.

Name	Structure	R	K_I [mM]	K_D [mM]	GE [kJ*mol^{-1}][a]	Relative potency[a]
12.1			7.6±0.1			1.0
5.1				7.9±0.7[b] 7.4±0.0.6[c]	1.5	0.96
5.2			48±6[c]			0.16
5.3			37±7		0.33	0.21
5.4			40±3		0.16	0.19
5.5			45±4		0.010	0.17
5.6			40±4		0.12	0.19
5.7			7±2		1.0	1.1
5.8			24±2		0.020	0.32
5.9			19±4		1.2	0.40
5.10			46±6		0.010	0.17
5.11			1.3±0.1	4.3±0.3[c]	2.2	5.9
5.12			33±3		0.19	0.23

[a] The group efficiency was calculated relative to the K_I value determined for **5.2**.
[a] The relative potency was calculated relative to the K_I value determined for **12.1**.
[b] This value was determined from three independent titration experiments *via* ^{19}F R_2-filtered NMR.
[c] This value was determined *via* ^{15}N HSQC NMR.

As for Man and **5.1**, most CSPs are located proximal to the carbohydrate binding site. K299 and E285 directly interact with Man or the Ca^{2+} ion as observed *via* X-ray crystallography (Feinberg *et al.*, 2011). These interactions are also observed in the ^{15}N HSQC NMR experiments with **5.11** (Figure 20a and b). While the mapping of CSPs might be utilized to provide constraints for docking procedures, it additionally represent a powerful approach to validate independently generated docking poses (Dominguez *et al.*, 2003). The pose generated for **5.11** predicts the formation of hydrogen bonds with K299 and N307. In comparison to the CSP pattern obtained with **5.1** for which these interactions are likely absent, increased CSP values are observed for both K299 and W306, the residue adjacent to N307 (Figure 20a and b). It is noteworthy that the resonance corresponding to N307 was not assigned. In comparison to titration experiments with Man, which forms a hydrogen bond with K299, only W306 displays an increased CSPs in titration experiments with **5.11**.

Moreover, additional residues in proximity to both K299 and K313, most prominently A300 and T314, display increased CSPs, an effect likely induced by the sulfonate group of **5.11** (Figure 20b and c). Notably, the formation of a hydrogen bond with K313 was not predicted in the docking pose corresponding to the highest GBIV/WSA ΔG score. Yet, alternative docking poses generated with the alternative conformation for K313 account for this interaction (data not shown). This is consistent with the binding mode determined for Gal-6-OS by X-ray crystallography, indicating conformational flexibility for the carbohydrate binding site of Langerin (Feinberg *et al.*, 2001). The hypothesis of conformational flexibility is further supported by the relatively high number of residues displaying CSPs upon **5.11** binding. In distinction from the proposed modulation of the allosteric network also observed for **5.11**, these effects are spatially associated with carbohydrate binding site (Figure 20c). Notably, the receptor backbone was observed to be predominantly rigid in MD simulations suggesting that conformational flexibility is likely conveyed *via* side chain dynamics (Hanske *et al.*, 2016). Indeed, alternative states for the side chains of several residues associated with the carbohydrate binding site were identified including I250, I282, K299 and K313.[18] Overall, these observations highlight the potential importance of induced fit or conformational selection mechanisms for the recognition of glycomimetic Langerin ligands. As observed for Man and **5.1**, CSPs are also observed in remote regions of the CTLD fold, particularly for G259 and A258 in the short loop region. This suggests a modulation of the previously reported allosteric network upon **5.11** binding and is in agreement with the proposed conformation flexibility (Hanske *et al.*, 2016; Hanske *et al.*, 2017b).

[18] The corresponding MD simulations were conducted by Stevan Aleksic (Free Universtiy of Berlin). The findings on side chain dynamics for the carbohydrate binding site were communicated in person and are not included in the cited publication.

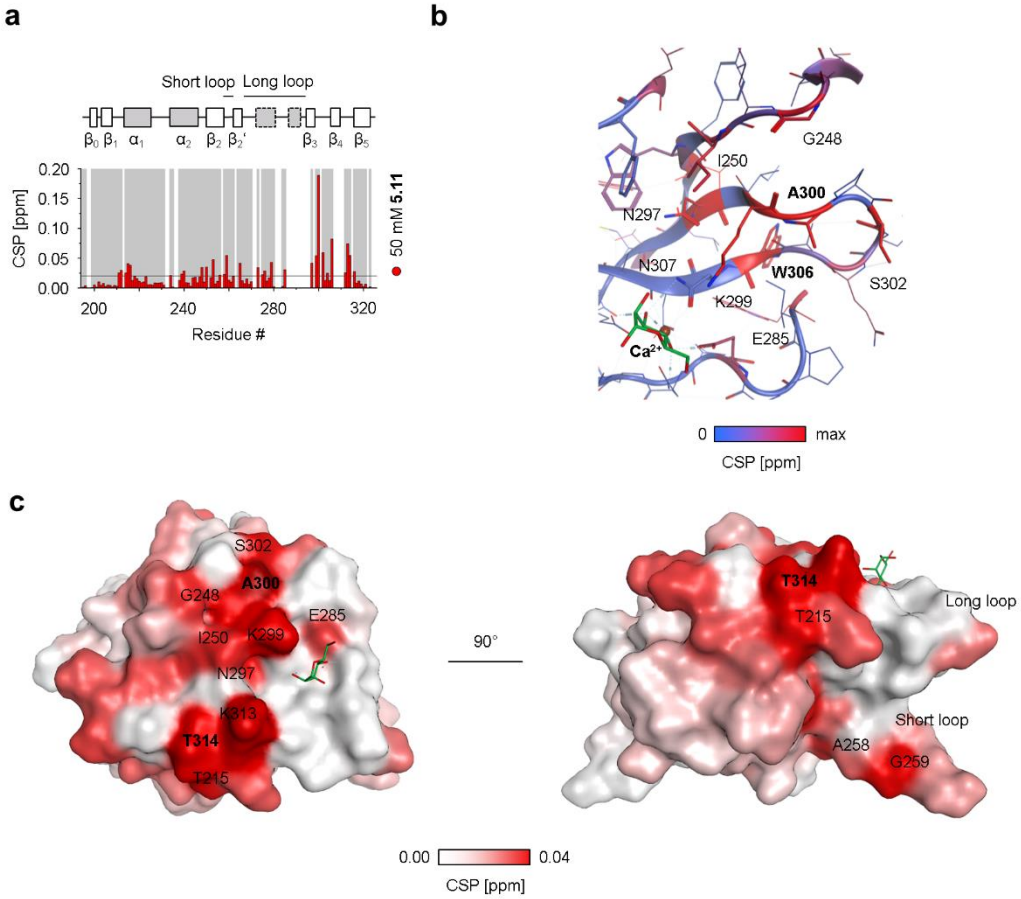


Figure 20. ^{15}N HSQC NMR binding mode anlysis for ManNAc analog 5.11.

a. ^{15}N HSQC NMR experiments with the Langerin CRD revealed the CSP pattern for **5.11**. Assigned resonances detected in the reference spectrum are highlighted (grey). **b.** Mapping the CSPs on the X-ray structure of the Langerin CRD (PDB code: 3P5F) validated a Ca^{2+}-dependent binding mode for **5.11** as indicated by CSPs observed for E285 and K299 (Feinberg *et al.*, 2011). The docking pose for **5.11** predicts the formation of hydrogen bonds between the sulfonate group and N307 as well as K299. While the resonance corresponding to N307 was not assigned, the high CSP values for W306 and K299 are consistent with the formation of this interaction. **c.** Additional CSPs not observed for Man or **5.1** are located two in regions arounds K299 and K313, flanking the sulfonate group (Figure 18b). Here, prominent residues include A300 and T314. As observed for Man and **5.1**, CSPs were also observed in remote regions of the CTLD fold, particularly for G259 and A258 in the short loop region. This indicates a modulation of the previously reported allosteric network upon **5.11** binding (Hanske *et al.*, 2016).

In conclusion, the *in silico* design of ManNAc analogs **5** resulted in the identification of **5.11**, the first glycomimetic ligand reported for Langerin (Wamhoff *et al.*, 2016). The 5.9-fold affinity increase over previously reported naturally occurring monosaccharide ligands along with the introduction of a propargyl group in the anomeric position render **5.11** a valuable chemical probe to investigate DC immunology. The sulfonate group potentially provides improved pharmacokinetic stability compared to sulfated monosaccharides such as Gal-6-OS or Man-6-OS as sulfate groups represent excellent leaving group in nucleophilic substitution reactions. However, the design approach was impeded by the unexpected affinity decrease associated with introduction of the propargyl group and the absence of the hydrogen bond formed between K299 and the hydroxyl group in C2 of Man. Consequently, other monosaccharide scaffolds and linker strategies will be explored in the following chapters to ultimately design more potent glycomimetic Langerin ligands (Chapter 3.3 and 3.4).

The implemented strategy, integrating structure-based *in silico* methods with ^{19}F R_2-filtered and ^{15}N HSQC NMR experiments enabled a detailed and robust characterization of **5.11** recognition by Langerin. While the predicted binding mode was validated, additional CSPs were induced for a relatively large number of residues in proximity of K299 and K313. These effects are less pronounced for Man and indicate conformational flexibility for the carbohydrate binding site. Notably, MD simulations served to validate side chain dynamics for residues associated with the carbohydrate binding site (Hanske *et al.*, 2016).[19] This potentially has important implications for the design of glycomimetic Langerin ligands. The conformational flexibility might be leveraged to explore the formation of additional secondary pockets. On the other hand, molecular docking approaches and the characterization of binding modes *via* ^{15}N HSQC NMR become considerably more complicated under these conditions. In this context, the screening of larger libraries of mono- or oligosaccharides and their synthetic analogs represents a feasible approach to identify glycomimetic Langerin ligands as it will implicitly account for induced fit or conformational selection binding mechanisms.

[19] The corresponding MD simulations were conducted by Stevan Aleksic (Free Universtiy of Berlin). The findings on side chain dynamics for the carbohydrate binding site were communicated in person and are not included in the cited publication.

3.3.2. Modifications in the C1 and the C6 Position of Mannose

While the *in silico* approach presented in the preceding section resulted in the identification of **5.11**, a ManNAc analog displaying a 5.9-fold affinity increase over naturally occurring monosaccharide ligands, more potent glycomimetic Langerin ligands might be required for the efficient delivery of antigens to LCs *via* targeted liposomes (Chapter 3.3.1). The structural characterization of **5.11** recognition by Langerin *via* ^{15}N HSQC NMR and MD simulations suggests conformational flexibility for the carbohydrate binding site. In this context, the screening of larger libraries of carbohydrate analogs represents a suitable approach to efficiently expand the available SAR information and will implicitly account for the formation of additional secondary binding pockets.

Yet, the onerous synthesis of mono- or oligosaccharide analogs as well as other structural glycomimetics critically limits the availability of large and diverse libraries for screening. Despite recent advances in the automated or enzymatic assembly of oligosaccharides and novel diversity-oriented strategies, this limitation represents a major challenge for glycomimetic ligand design (Danby and Withers, 2016; Hahm *et al.*, 2017; Lenci *et al.*, 2016). Here, the utilization of focused glycomimetic libraries originally synthesized to target other GBPs represent an intriguing alternative strategy, particularly if information on the specificity against Langerin is available.

This is typically the case for DC-SIGN inhibitors due to the opposing roles in HIV infection proposed for these CLRs (de Witte *et al.*, 2007; Geijtenbeek *et al.*, 2000). However, none of the reported glycomimetic DC-SIGN ligands displays specificity for Langerin (Andreini *et al.*, 2011; Thépaut *et al.*, 2013; Varga *et al.*, 2013). As DC-SIGN is expressed on macrophages and several DC subsets, this will impede targeted antigen delivery to LCs.

The specificity is of no concern for focused glycomimetic libraries designed to target bacterial adhesins as these receptors will not be present under physiological conditions. *Vice versa*, information on the specificity against human CLRs is typically available to ensure optimal bioavailability and to reduce side effects for anti-infectives. FimH is an adhesin expressed by uropathogenic *E. coli* (Mulvey, 2002). It recognizes Man-type glycans on *e.g.* Uroplakin 1a and thereby promotes the colonization of the urogenital tract. The inhibition of this interaction with potent Man analogs bears considerable therapeutic potential and represents a promising approach to address the emerging resistance against traditional antibiotics (Han *et al.*, 2010; Kleeb *et al.*, 2016). Interestingly, a specificity screening against a set of human CLRs revealed the inhibition of Langerin binding to a Man-type trisaccharide for a Man analog bearing a biphenyl substituents in C1 (Scharenberg *et al.*, 2012). The analogs were screened in a heterogenous assay at concentrations of 1 mM and while no IC_{50} values were reported, the observed inhibition indicated a submillimolar affinity. Moreover, none of the other human CLRs except for the MR displayed comparable affinities for this class of Man analogs. The specificity screening included DC-SIGN, DC-SIGNR, SP-D, MBP-C and BDCA-2. This

renders Man analogs bearing biphenyl substituents in C1 a promising class of glycomimetic Langerin ligands for targeted antigen delivery to LCs. Accordingly, a focused library of Man analogs **12** was composed for screening (Figure 21a). In addition to the biphenyl substituents, these analogs bear phenyl-indolinyl or triazolyl-phenyl substituents in C1. The molecules were previously synthesized in the laboratory of Prof. Beat Ernst (Universität Basel) and kindly provided for affinity characterizations with Langerin (Jiang *et al.*, 2012; Kleeb *et al.*, 2015; Klein *et al.*, 2010; Pang *et al.*, 2012; Schwardt *et al.*, 2011).

Moreover, the K_I determination for a set of monosaccharides presented in the preceding chapters revealed a 6.6-fold affinity increase for Man-6-OS (K_I = 0.8±0.1 mM) over Man. Sulfonamide groups can be utilized as bioisosteres of sulfate groups and this strategy was explored to identify favorable substituents in C6 of the Man scaffold (Meanwell, 2011). To this end, the focused library was complemented with Man analogs **14** previously synthesized in the laboratory of Dr. Alexander Titz (Helmholtz Center for Infection Research) and kindly provided for affinity characterizations with Langerin (Hauck *et al.*, 2013) (Figure 21a). These glycomimetics were originally designed to target the bacterial adhesin LecB which is essential for biofilm formation by *P. aeruginosa*, a multidrug resistant pathogen causing *e.g.* pneumonia. Glycomimetic LecB inhibitors bear considerable therapeutic potential and address the emergence of antibiotic resistance (Johansson *et al.*, 2008).

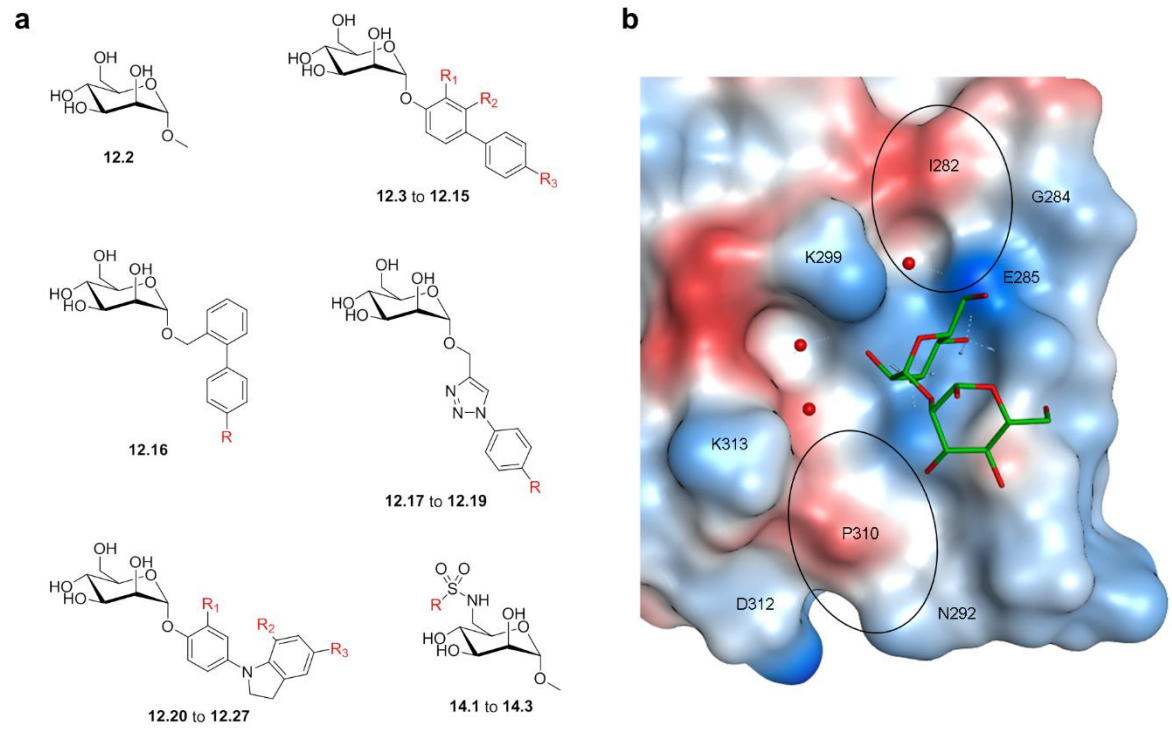


Figure 21. ^{19}F R_2-filtered NMR screening of Man analog libraries 12 and 14.[20]

a. Overall, 29 Man analogs bearing aromatic substituents in in α-configuration of C1 or in C6 were screened by ^{19}F R_2-filtered NMR. A classification according to scaffold structures yields six subsets for the library. The depicted analogs were previously synthesized as anti-infectives targeting the bacterial lectins FimH or LecB (Hauck *et al.*, 2013; Jiang *et al.*, 2012; Kleeb *et al.*, 2015; Klein *et al.*, 2010; Pang *et al.*, 2012; Schwardt *et al.*, 2011). In a specificity screening against human CLRs, members of the subset corresponding to analogs **12.3** to **12.15** were found to interact with Langerin (Scharenberg *et al.*, 2012). Analogs **14.1** to **14.3** were included for their similarity to Man-6-OS. **b.** The binding mode of Man (PDB code: 3P5F) suggests a region of the receptor surface in proximity to K313 potentially targeted by analogs **12** and a region in proximity to K299 potentially targeted by analogs **14** (Feinberg *et al.*, 2011). The receptor surface is colored according to its lipophilicity (lipophilic: red, hydrophilic: blue).

An evaluation of the Man binding mode reveals regions of the Langerin surface potentially targeted by the focused library of Man analogs **12** and **14** (Figure 21b). Assuming a conserved orientation for the Man scaffold, substituents in C1 will potentially form interactions in proximity of K313, P310, D312 and N292, while substituents in C6 will target a region in proximity of K299, E285, G284 and I282 harboring a structural H_2O molecule. In contrast to the receptor surface in the axial direction of C2, these regions are highly solvent exposed and devoid of secondary binding pockets.

The composed library of Man analogs **12** and **14** was screened utilizing the ^{19}F R_2-filtered NMR assay (Table 7, Table 8 and Table 9, Appendix B.4). In total, it comprised 29 analogs of which a considerable number was provided in submilligram quantities. Accordingly, the optimization of the assay for the efficient elucidation of SARs proved essential to implement the screening (Chapter 3.2.2). Estimated K_I values were determined for each analog from experiments at a single concentration in presence of 10% DMSO (Setup 2) (Chapter 3.2.2). Overall, three subsets of the focused library displayed promising SARs. First, the reported affinity increase for biphenyl

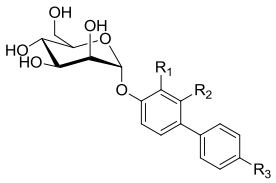
[20] The molecules were provided by Dr. Alexander Titz (Helmholtz Center for Infection Research) and Prof. Dr. Beat Ernst (Universität Basel).

substituents in C1 was validated in the ^{19}F R_2-filtered NMR assay (Table 7) (Scharenberg *et al.*, 2012). Additionally, analogs bearing phenyl-indolinyl substituents in C1 displayed comparable estimated K_I values and the affinity increase observed for Man-6-OS was maintained for the bioisoteric sulfonamide substituents in C6 (Table 8 and Table 9).

The evaluation of the SAR obtained for the biphenyl subset of the Man analog library **12** revealed unexpected trends (Table 7). **12.11** ($K_{I,est}$ = 0.2 mM) displayed the highest affinity within this subset. This analog bears a trifluoromethyl group in ortho position of the first phenyl ring (R_1) and a carboxyl group in para position of the second phenyl ring (R_3). Strikingly, the introduction of the trifluoromethyl group accounted for a 45-fold affinity increase over the reference molecule **12.3** ($K_{I,est}$ = 9 mM). A comparable affinity was obtained for **12.5** ($K_{I,est}$ = 0.3 mM) for which the trifluoromethyl group is substituted by a chloride group. This affinity increase was neither observed for **12.4** ($K_{I,est}$ = 8 mM) which bears the chloride group in meta position of the first phenyl (R_2) nor for other substituents in R_1. Moreover, carboxyl ester or carboxamide formation for the carboxyl group resulted in decreased affinities as exemplified by **12.6** ($K_{I,est}$ = 8 mM) and **12.12** ($K_{I,est}$ = 2 mM). Interestingly, the comparison of estimated K_I values for **12.3**, **12.5**, **12.7** ($K_{I,est}$ = 2 mM) and **12.8** ($K_{I,est}$ = 4 mM) indicated that the contributions of substituents in R_1 and R_3 to the affinity increase are not independent of each other. Substitutions in both positions have been proposed to influence the dihedral angles determining the relative orientation of the two phenyl rings and the Man scaffold, resulting in a non-additive SAR (Klein *et al.*, 2010). Hence, the correct combination of substituents is likely required to optimally orient the two phenyl rings for the formation of favorable interactions with the Langerin surface. Notably, conformational entropy and solvation effects might also contribute to the observed SAR. The latter have been reported to be of particular importance for carbohydrate-protein interactions (Navarra *et al.*, 2017). Yet, both contributions are difficult to quantify experimentally.

As the phenyl-indolinyl subset of the focused library comprised fewer analogs, the obtained SAR information was limited (Table 8). Yet, the evaluation of estimated K_I values revealed an additional class of potent glycomimetic Langerin ligands, complementing the findings for the biphenyl subset. In similarity to the latter, all analogs displaying submillimolar affinities are substituted with a chloride group in the ortho position of the first phenyl ring. Within this subset, **12.20** ($K_{I,est}$ = 0.4 mM), **12.23** ($K_{I,est}$ = 0.3 mM) and **12.25** ($K_{I,est}$ = 0.3 mM) displayed the lowest estimated K_I values. The indolinyl ring of these glycomimetics is substituted with methyl and chloride groups in different positions (R_1 and R_2). Interestingly, polar substituents on the indolinyl ring did not seem to be required as exemplified by the decreased affinities of **12.22** ($K_{I,est}$ = 0.7 mM) and **12.24** ($K_{I,est}$ = 1 mM). This suggests the formation of interactions which differ from those observed for the biphenyl subset. Alternatively, the recognition of both subsets might be dominated by interactions formed by the aromatic substituents *e.g. via* a potential cation-π bond with K313 (Figure 21b).

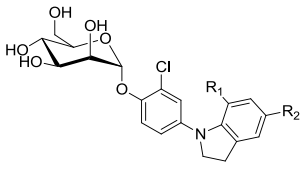
Table 7. Structure-activity relationship of the biphenyl subset of Man analog library 12.[21]

Name	Structure	R_1	R_2	R_3	$K_{I,est}$ [mM][a]	Relative potency[b]
12.3		H	H		9	1.0
12.4		H	Cl		8	1.1
12.5		Cl	H		0.3	30
12.6		Cl	H		6	1.5
12.7		Cl	H		2	4.5
12.8		H	H		4	2.3
12.9		F	H		3	3.0
12.10			H		7	1.3
12.11		CF_3	H		0.2	45
12.12		CF_3	H		2	4.5
12.13			H		1	9.0

[a] Estimated K_I values were determined *via* ^{19}F R_2-filtered NMR experiments at a single competitor concentration.
[b] The relative potency was calculated utilizing the estimated K_I value determined for **12.3**.

[21] In part, these procedures were conducted by Julia Mastouri, Mareike Rentzsch, Lena Kilian and Lennart Schnirch (Freie Universität Berlin) under my supervision.

Table 8. Structure-activity relationship of the phenyl-indolinyl subset of Man analog library 12.[22]

Name	Structure	R_1	R_2	$K_{I,est}$ [mM][a]	Relative potency[b]
12.20				0.4	23
12.21				5	1.8
12.22				0.7	13
12.23				0.3	30
12.24				1	9
12.25				0.3	30

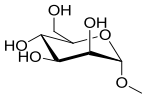
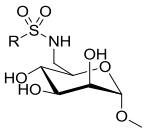
[a] Estimated K_I values were determined *via* ^{19}F R_2-filtered NMR experiments at a single competitor concentration.
[b] The relative potency was calculated utilizing the estimated K_I value determined for **12.3**.

Lastly, the ^{19}F R_2-filtered NMR screening served to estimate K_I values for Man analogs **14** bearing sulfonamide substituents in C6 (Table 9). Here, the SAR is dominated by the affinity decrease associated with α-mannoside formation as exemplified by **12.2** ($K_{I,est}$ = 12 mM). Compared to this reference molecule, all screened analogs maintained the 6.0-fold affinity increase determined for Man-6-OS (K_I = 0.8±0.1 mM). While sulfonamide groups represent suitable bioisosteres of the sulfate group in C6, no affinity increase was observed upon the introduction of phenyl rings for **14.2** ($K_{I,est}$ = 2 mM) and **14.3** ($K_{I,est}$ = 2 mM).

The estimated K_I values for selected Man analogs **12** and **14** were subsequently reproduced in ^{19}F R_2-filtered NMR titration experiments (Setup 1) (Figure 22a and b, Table 10). Overall, the determined affinities were consistent with the screening results and validate analogs **12** bearing either biphenyl or phenyl-indolinyl substituents in C1 as potent glycomimetic Langerin ligands. With a 57-fold affinity increase over reference molecule **12.2** (K_I = 13±3 mM), **12.11** (K_I = 0.23±0.03 mM) represents the most potent member of the focused library. While higher K_I values were determined for **12.5** (K_I = 0.39±0.05 mM) and **12.23** (K_I = 0.33±0.02 mM), both analogs still display a considerable affinity increase over natural occurring carbohydrate ligands. For Man analogs **14** bearing sulfonamide substituents in C6, no affinity increase was observed upon the introduction of a phenyl ring in this position. Consequently, the estimated K_I was only validated for **14.1** (K_I = 3.0±0.2 mM), the least complex member of this subset.

[22] In part, these procedures were conducted by Mareike Rentzsch (Freie Universität Berlin) under my supervision.

Table 9. Structure-activity relationship of Man analog library 14.[23]

Name	Structure	R	$K_{I,est}$ [mM]	Relative potency[a]
12.2			12	1.0
14.1			3	6.0
14.2			2	6.0
14.3			2	6.0

[a] The relative potency was calculated utilizing the K_I value determined for **12.2**.

Additionally, the resonance corresponding to the trifluoromethyl group of **12.11** was monitored in ^{19}F R_2-filtered NMR experiments conducted in presence of 10% DMSO and 0.01% Tween-20. The utilization of Tween-20 served to exclude an aggregative false-positive mechanism for **12.11**. Interestingly, the resonance displayed considerable line broadening $\Delta\nu_{0.5}$ at a negligible CSP (Figure 22c). This observation further validates **12.11** as a potent glycomimetic Langerin ligand. As the experiments were conducted at ligand and receptor concentrations identical to those utilized for **5.1**, the considerable increase of $R_{2,obs}$ ($\Delta R_{2,obs}$ = 5.7 Hz) might be attributed to the 37-fold affinity increase (Figure 22c). Experiments in presence of 10 mM EDTA served to demonstrate the Ca^{2+}-dependency of the interactions.

Furthermore, these findings suggest that **12.11** might represent a suitable reporter molecule for competitive binding experiments utilizing ^{19}F R_2-filtered NMR. The considerably increased $R_{2,obs}$ value in presence of Langerin would translate into an optimized dynamic range, allowing for the utilization of lower receptor concentrations. As discussed above, this would enable the accurate determination of K_I value for high affinity ligands and is therefore of particular interest for the characterization of multivalent glycomimetics (Setup 3) (Chapter 3.2.2). However, the increased affinity of **12.11** might also result in chemical exchange contributions $R_{2,ex}$ to determined $R_{2,obs}$ values and thereby considerably complicate data analysis. Although the utility of **12.11** as a reporter molecule was not investigated in this dissertation, these investigations might prove valuable for the design of next generation glycomimetic Langerin ligands.

[23] In part, these procedures were conducted by Julia Mastouri, Mareike Rentzsch, Lena Kilian and Lennart Schnirch (Freie Universität Berlin) under my supervision.

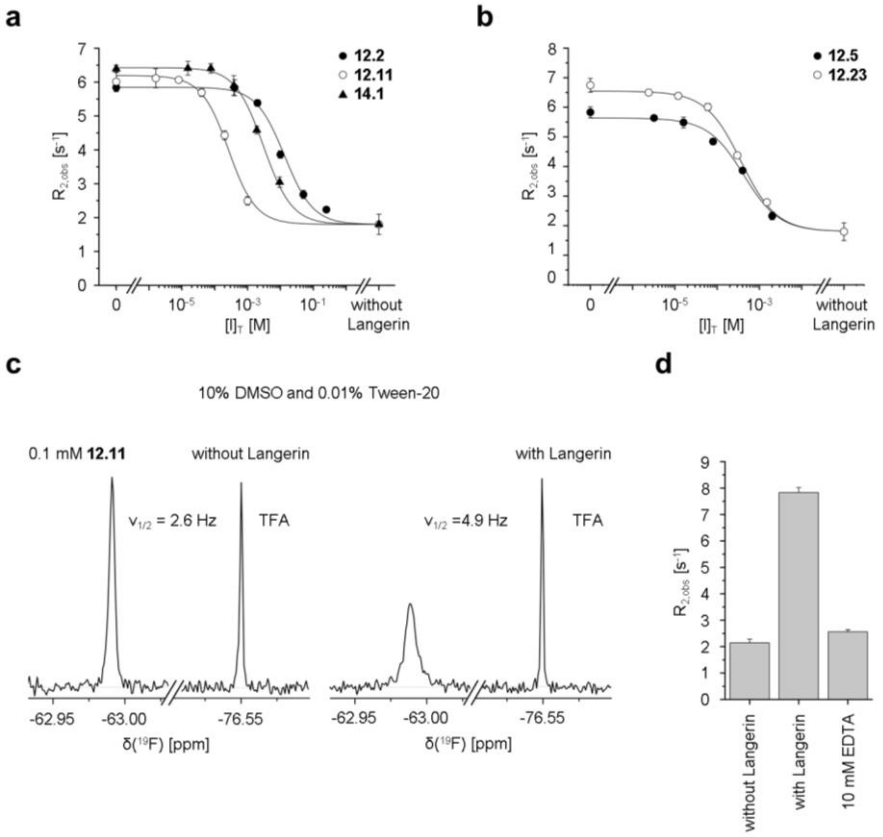


Figure 22. ^{19}F R_2-filtered NMR affinity validation for Man analogs 12.2, 12.5, 12.11, 12.24 and 14.1.[24]

a. and b. Competitive binding experiments in presence of 10% DMSO were conducted to validate the estimated K_I values obtained from the ^{19}F R_2-filtered NMR screening. **12.2** (K_I = 13±3 mM) served as the reference molecule to quantify the affinity increase observed for potent glycomimetic Langerin ligands. **12.11** (K_I = 0.23±0.03 mM) displayed the highest affinity increase, followed by **12.23** (K_I = 0.33±0.02) and **12.5** (K_I = 0.39±0.05 mM). While these analogs bear biphenyl or phenyl-indolinyl substituents in C1 of the Man scaffold, the estimated affinity increase for **14.1** (K_I = 3.0±0.2 mM) bearing a sulfonamide substituent in C6 could be validated as well. **c. and d.** The interaction of **12.11** with Langerin was further validated *via* ^{19}F R_2-filtered NMR experiments conducted in presence of 10% DMSO and 0.01% Tween-20. The trifluoromethyl group of **12.11** displayed line broadening $\Delta\nu_{0.5}$ in presence of the Langerin ECD. Consistent with a 34-fold affinity increase over **5.1**, the observed increase in $R_{2,obs}$ ($\Delta R_{2,obs}$ = 5.7 Hz) was considerably increased. Experiments in presence of 10 mM EDTA served to validate the Ca^{2+}-dependency of the interaction.

^{19}F R_2-filtered NMR-derived K_I values were subsequently validated in ^{15}N HSQC NMR titration experiments (Figure 23, Table 10). Here, **12.11** and **14.1** were selected as they represent the most potent glycomimetics identified that bear substituents in C1 or C6 of the Man scaffold, respectively. The data set was complemented by titration experiments with reference molecule **12.2**. Both the ^{19}F R_2-filtered NMR screening and titration experiments were conducted in presence of 10% DMSO. To ensure the transferability of determined affinities between the assays, ^{15}N HSQC NMR experiments were conducted under the same conditions. Compared to the findings presented in the preceding sections, this resulted in considerably lower maximal CSP values, reducing the accuracy of K_D determination. Moreover, line broadening $\Delta\nu_{0.5}$ or decreased integrals V were observed for a large number of resonances in titration experiments with **12.11.** This observation is indicative of intermediate or slow exchange phenomena and can be attributed to the submillimolar affinity of **12.11**.

[24] In part, these procedures were conducted by Julia Mastouri, Mareike Rentzsch, Lena Kilian and Lennart Schnirch (Freie Universität Berlin) under my supervision.

Thus, the corresponding resonances were not considered for K_D determination from CSPs further reducing the accuracy (Williamson, 2013). Nevertheless, the obtained affinities were generally consistent with ^{19}F R_2-filtered NMR experiments. Albeit less prominent, **12.11** (K_D = 0.5±0.2 mM) still displayed a considerable affinity increase over reference molecule **12.2** (K_D = 7±2 mM). The affinity determined for **14.1** (K_D = 2.0±0.4 mM) is comparable to the corresponding K_I value.

Structural information on the recognition of Man analogs **12** and **14** potentially supports the optimization of this class of glycomimetic Langerin ligands. Here, it is of particular interest to evaluate the potential additivity of the affinity increase observed for **12.11** and **14.1**. To analyze the binding mode of **12.11** and **14.1**, CSPs were mapped on the X-ray structure of Langerin (Figure 24,). In contrast to the results presented in previous sections, this ^{15}N HSQC NMR binding mode analysis is based on the latest version of the resonance assignment (Hanske *et al.*, 2016) (Chapter 3.2.1 and 3.3.1). This limits the transferability of observations to the findings presented titrations with Man, ManNAc, **5.1** and **5.11**. Moreover, the experiments were conducted at an increased ^1H Larmor frequency ω_0 of 700 MHz which potentially alters the chemical exchange regime for indiviual resonances.

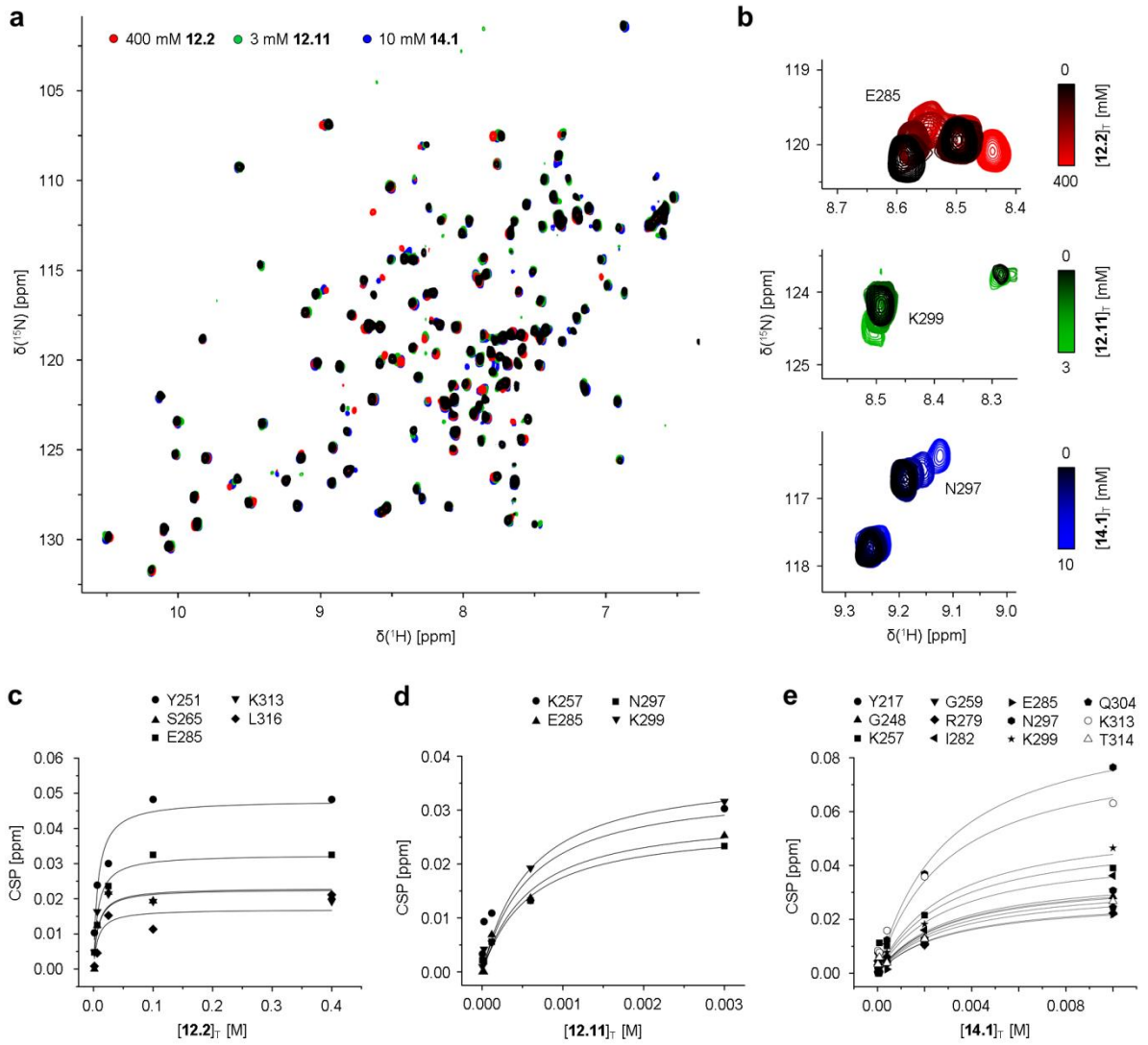
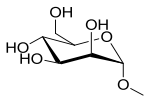
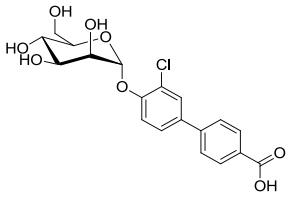
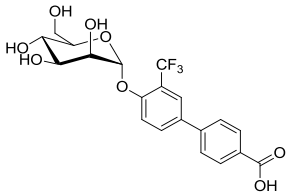
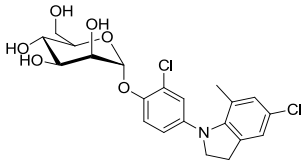
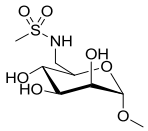


Figure 23. K_D determination for Man analogs 12.2, 12.11 and 14.1. [25]

a. Affinities determined for **12.2**, **12.11** and **14.1** utilizing the ^{19}F R_2-filtered NMR assay were validated *via* ^{15}N HSQC NMR titration experiments with the Langerin CRD in presence of 10% DMSO. **b.** Assigned resonances displaying fast chemical exchange and CSPs larger than 0.02 ppm were selected for the determination of K_D values. Representative CSP trajectories are depicted. **c to e.** Aside from minor dev*ia*tions, the K_D values obtained for **12.2** (K_D = 7±2 mM), **12.11** (K_D = 0.5±0.2 mM) and **14.1** (K_D = 2.9±0.4 mM) are consistent with the results from the ^{19}F R_2-filtered NMR assay and validate **12.11** as a potent glycomimetic Langerin ligand.

[25] In part, these procedures were conducted by Lennart Schnirch (Freie Universität Berlin) under my supervision.

Table 10. Affinity validation for potent Man analogs 12 and 14.[26]

Name	Structure	K_I [mM]	K_D [mM]	Relative potency[a]
12.2		13±3	7±2	1.0
12.5		0.39±0.05		33
12.11		0.23±0.03	0.5±0.2	57
12.24		0.33±0.02		39
14.1		3.0±0.2	2.9±0.4	4.3

[a] The relative potency was calculated utilizing the K_I value determined for **12.2**.

As discussed above, chemical exchange phenomena likely account for the reduced amplitude and number of CSPs observed for **12.11** when compared to **12.2** (Figure 24a). Yet, at a threshold of 0.02 ppm, **12.11** still displayed CSPs for the majority of residues associated with the recognition of the Man scaffold including E285, N297, K299, A300 and K313 (Figure 24b). As observed for all analyzed monosaccharides and their analogs, **12.11** additionally induces CSPs associated with the allosteric network in remote regions of the CTLD fold, prominently K257 and G259 (Hanske *et al.*, 2016). Interestingly, a relative CSP increase is observed for the D312 compared to **12.2**. This residue is found in equatorial direction of C1 of the Man scaffold and might be targeted by the carboxyl group of **12.11** (Figure 21b). It is noteworthy that other residues in the region of the Langerin surface including P310 and N292 are not assigned. In summary, these observations suggest a Ca^{2+}-dependent binding mode for **12.11**. Interactions formed by the Man scaffold seem to be conserved and the biphenyl substituent is likely oriented towards K313 and D312, potentially forming interactions *via* the second phenyl ring or the carboxyl group. The presence of the trifluoromethyl group in ortho position of the first phenyl ring might be required to optimally orient the second phenyl ring for the formation of cation-π

[26] In part, these procedures were conducted by Mareike Rentzsch, Lena Kilian and Lennart Schnirch (Freie Universität Berlin) under my supervision.

interactions or hydrogen bonds. These potential interactions likely contribute critically to the 57-fold affinity increase observed for **12.2**.

The CSP pattern observed for **14.1** resembles that induced by the Man scaffold as exemplified by **12.2 (**Figure 24a). The set of affected residues suggests the conservation of the Ca^{2+}-dependent binding mode and is additionally associated with the allosteric network (Hanske *et al.*, 2016) (Figure 24b). Yet, **14.1** also shows one prominent distinctions compared to **12.2**. The CSP value for I282 and the proximal R279 are increased while the effect for E285 is decreased. This observation might be attributed to interactions formed by the methylsulfonamide group and is consistent with the proposed orientation of C6 towards these residues (Figure 21b).

CSP trajectories, *i.e.* the vector of induced perturbations, provide additional information on the bound state of the receptor and thus the ligand recognition process (Chi *et al.*, 2015; Selvaratnam *et al.*, 2012). Originally, the quantitative analysis of these trajectories was developed to evaluate the effects of mutations on the allosteric activation of EPAC, a cAMP-dependent nucleotide exchange factor. In principal, this approach is transferrable to the comparative analysis of binding modes for glycomimetic ligands. Interestingly, the CSP trajectory of I282 observed for **14.1** is distinct from that for **12.2** and **12.12.** This further validates the suggested interaction with the methylsulfonamide group of **14.1**. Notably, the determination of CSPs for I282 for **12.2** and **12.2** is compromised by partial spectral overlap with the resonance corresponding to Y208. Consequently, an evaluation of CSP trajectories for additional residues served to validate this observation (Figure 25).

Strikingly, trajectories for **12.2** and **12.11** are similar for the majority of residues that display CSPs upon binding of both analogs. For **14.1** this also holds true for residues associated with the allosteric network. However, distinct trajectories are not only induced for I282, but also for a subset of residues associated with the carbohydrate binding site, including N297, K299 and K313. K299 is located in direction of C6 of the Man scaffold and the observed CSP trajectory could result from an interaction with the methylsulfonamide group (Figure 21b). Alternatively, **14.1** might interact with the Ca^{2+}-dependent binding site in an altered orientation for the Man scaffold. Notably, Man has been demonstrated to bind Langerin in two orientations (Feinberg *et al.*, 2001). The alternative orientation involves a 180° rotation of the Man scaffold and would direct C6 towards K313.

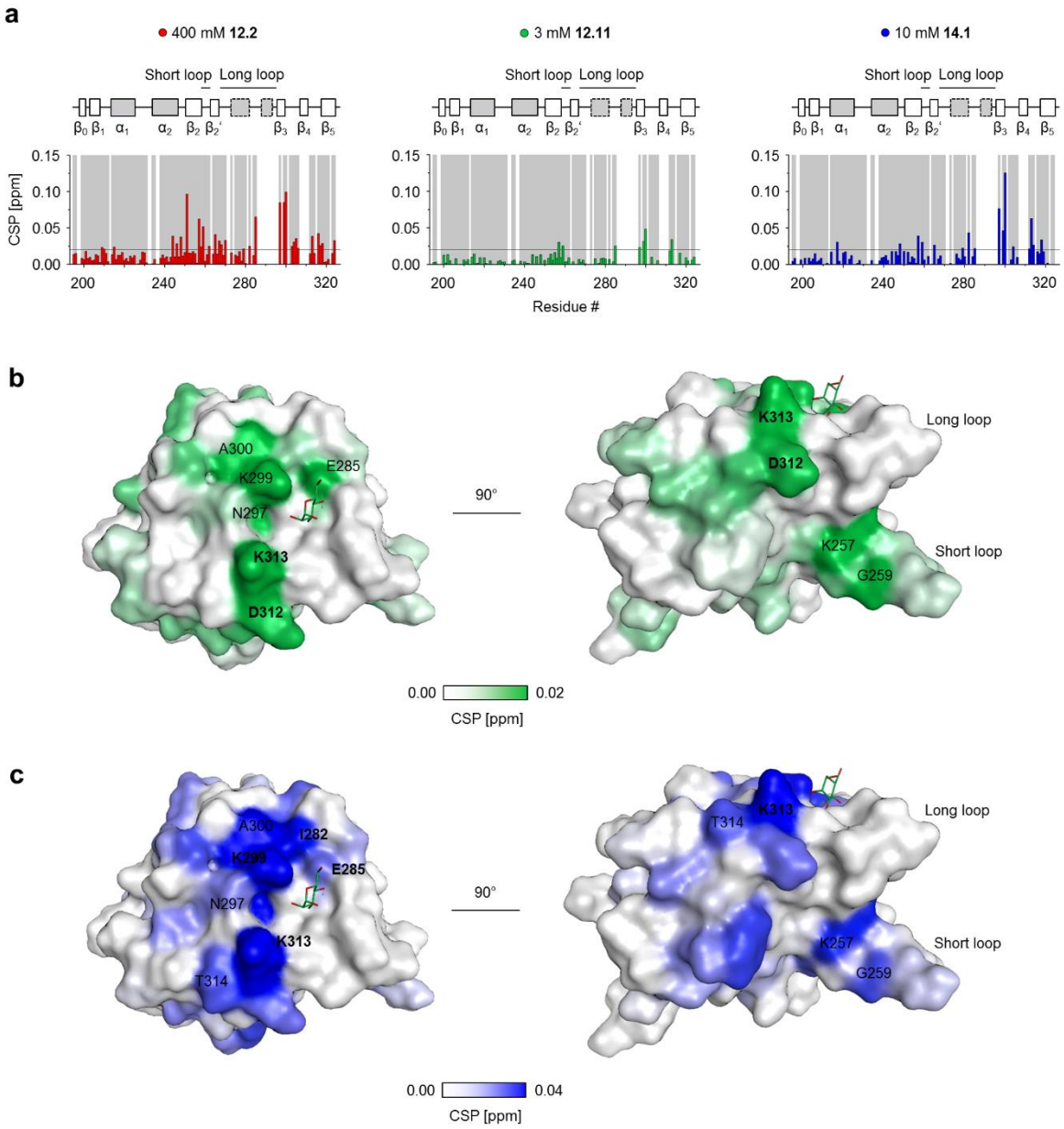


Figure 24. ^{15}N HSQC NMR binding mode analysis for Man analogs 12.2, 12.11 and 14.1.

a. ^{15}N HSQC NMR experiments with the Langerin CRD revealed the CSP patterns for **12.2**, **12.11** and **14.1**. Assigned resonances detected in the reference spectrum are highlighted (grey). Observed CSP values for all three analogs were reduced due to the presence of 10% DMSO. For **12.2**, intermediate and slow exchange phenomena additionally reduced the number of resonances for which CSPs were induced. **b and c.** Mapping the CSPs on the X-ray structure of the Langerin CRD (PDB code: 3P5F) validated a Ca^{2+}-dependent binding mode for both **12.11** (green) and **14.1** (blue) indicated by CSPs observed for E285 and K299 (Feinberg *et al.*, 2011). Additionally, CSPs are observed for N297, K313 and A300, residues also affected upon recognition of Man and ManNAc as well as other Man analogs (Chapter 3.2.1 and 3.3.1). In comparison to reference molecule **12.2**, both **12.11** and **14.1** displayed unique relative CSP patterns consistent with the proposed binding modes for the analogs (Figure 21). For **12.2**, the prominent effects observed for D312 and K313 might be indicative for an interaction with the second phenyl ring or the carboxyl group. For **14.1**, the methylsulfonamide group potentially induces the observed CSPs for I282 and K299. CSPs are also observed in remote regions of the CTLD fold, particularly for K257 and G259 in the short loop region. This indicates a modulation of the previously reported allosteric network (Hanske *et al.*, 2016; Hanske *et al.*, 2017b).

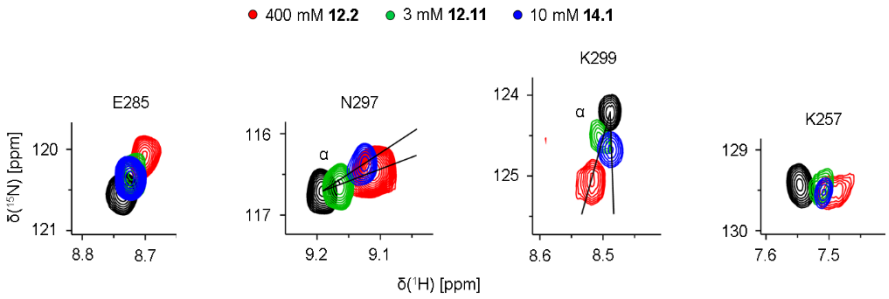


Figure 25. Comparative analysis of CSP trajectories for Man analogs 12.2, 12.11 and 14.1.

12.2, **12.11** and **14.1** induced CSPs for a shared set of residues. The corresponding trajectories yield information on the physicochemical and geometric properties of interactions formed upon the recognition of the respective analog. These vectors were similar for a subset of the residues in proximity of the carbohydrate binding site including E285 and A300. The same trend is observed for residues associated with the allosteric network such as K257. Interestingly, the comparative analysis revealed unique CSP vectors upon titration with **14.1** for a different subset of residues in proximity of the carbohydrate site including I282, N297, K299 and, unexpectedly, K313. This observation can be attributed to interactions formed by the methylsulfonamide group of **14.1** and might alternatively suggest a non-conserved binding mode of the Man scaffold.

To further investigate the binding mode of **12.2**, **12.11** and **14.1**, STD NMR experiments were conducted (Figure 26). All three analogs displayed specific STD effects in presence of Langerin further validating the interactions formed by these glycomimetics (Figure 26a to c). The DPFGSE pulse sequence employed for solvent suppression impeded an analysis for H1 for **12.2** and **14.1** due to spectral overlap with the HDO resonance. Moreover, all Langerin samples contained minor EDTA impurities that could not be removed even after repeated dialysis against H_2O. Yet, the observed coupling patter is indicative of the Ca^{2+}-bound state of the molecule, suggesting that an excess of Ca^{2+} ions is present. Hence, these impurities will likely not interfere with the conducted experiments.

Beyond the qualitative validation of an interaction, a kinetic analysis of the saturation transfer allows for the ligand-observed mapping of binding epitopes (Angulo and Nieto, 2011; Mayer and Meyer, 2001) (Figure 26d, Table 11, Appendix B.5). Based on the NOE and corrected for an R_1 relaxation bias, the parameter STD'_0 is a suitable measure for the distance r between a ligand hydrogen and the receptor surface. Accordingly, the approach is complementary to the ^{15}N HSQC NMR experiments presented above. The epitope obtained for **12.2** reveals similar STD effect for all hydrogens of the Man scaffold with H6 displaying a slightly decreased STD'_0 value. This non-differentiated epitope can be attributed to existence of an alternative orientation previously observed for Man-type glycan binding to Langerin (Feinberg *et al.*, 2011). The methyl group in C1 displays a low relative STD'_0 value indicated high solvent exposure. A similar, yet more differentiated binding epitope has previously been reported for the interaction of a Man-type disaccharide with Langerin (Hanske *et al.*, 2017a).

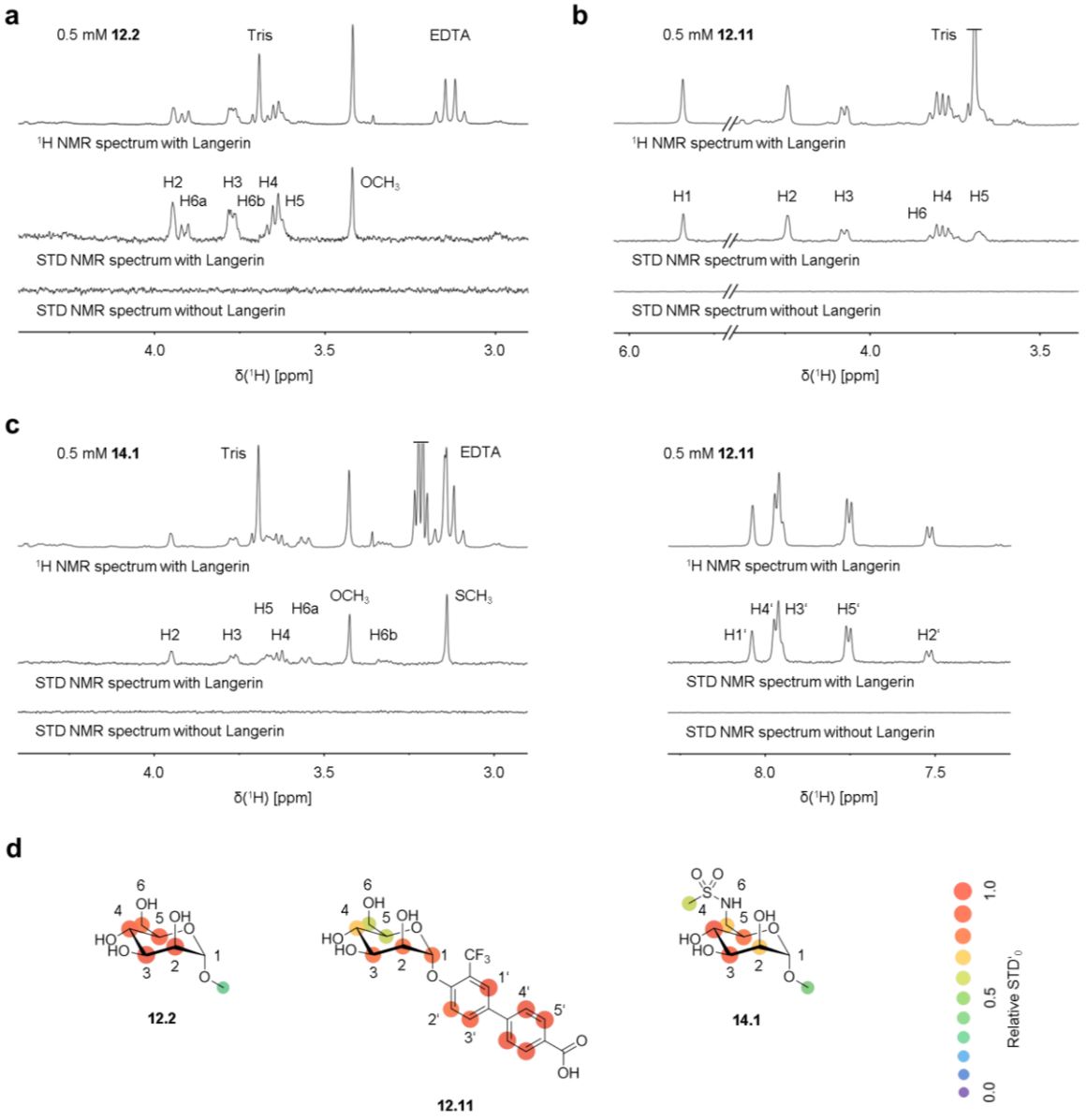


Figure 26. STD NMR epitope mapping for Man analogs 12.2, 12.11 and 14.1. [27]

a. to c. STD NMR experiments served to further validate the interactions of **12.2**, **12.11** and **14.1** with the Langerin ECD. STD NMR spectra were recorded at saturation times t_{sat} of 0.4 s, 0.6 s and 0.4 s and are magnified 32-fold, 4-fold and 4-fold, respectively. **d.** STD NMR epitopes for **12.2**, **12.11** and **14.1** were determined from build-up curves (Appendix B.5). **12.11** displays high STD effects for the biphenyl substituent in C1. While the epitopes determined for the Man scaffold of **12.1** and **12.11** are comparable, **14.1** displays a considerably decreased STD'$_0$ for H2.

[27] In part, these procedures were conducted by Lennart Schnirch (Freie Universität Berlin) under my supervision.

Table 11. Parameters for STD NMR epitope mapping for Man analogs 12.2, 12.11 and 14.1.

	Relative STD'$_0$[a]		
Proton #	**12.1**	**12.11**	**14.1**
1		0.9±0.3	
2	0.98±0.08	1.0±0.3	0.66±0.09
3	1.0±0.2	1.0±0.3	0.8±0.2
4	0.8±0.2	0.8±0.2	1.0±0.1
5	0.9±0.2	0.7±0.2	0.76±0.09
6	0.7±0.2	0.8±0.2	0.64±0.06

[a] The relative STD'$_0$ was calculated based on the highest STD'$_0$ value observed for the Man scaffold. Standard errors were derived directly from the fitting procedure.

For **12.11**, uniformly higher STD effects were observed for the biphenyl substituent in C1, consistent with the proposed interactions in proximity to K313 and D312 and the determined affinity increase (Figure 24b). These effects are more pronounced than for the Man scaffold. The normalization of STD'$_0$ values within the Man scaffold revealed only minor deviations from **12.2,** suggesting a comparable orientation for the Man scaffold for both analogs (Table 11). This is partially contradictory to the suggested 180° rotation for **12.2** as this orientation potentially results in steric clashes for the biphenyl substituent of **12.11** with the Langerin surface. However, obtained standard errors for STD'$_0$ values indicate that subtle differences in the orientation of the Man scaffold might not be detectable by STD NMR experiments. The binding epitope for **14.1**, by contrasts, clearly deviates from that of **12.2**. Particularly, H2 displays a considerably decreased STD effect. This suggests an alternative binding mode for **14.1** and is consistent with the distinct CSP trajectories observed by ^{15}N HSQC NMR (Figure 25). Furthermore, the STD'$_0$ value obtained for the methylsulfonamide group is decreased compared to the Man scaffold. As the introduction of phenyl rings in C6 did not affect the affinities obtained for Man analogs **14**, this observation suggests a solvent exposed orientation for the methylsulfonamide group (Table 9).

In summary, the ^{19}F R_2-filtered NMR screening of a focused library of Man analogs **12** and **14** bearing substituents in C1 and C6 resulted in the identification of **12.11** (K_I = 0.23±0.03 mM) as the most potent glycomimetic ligand reported for Langerin to date (Figure 22a). The analog bears a biphenyl substituents in C1 with a trifluoromethyl group in ortho position of the first phenyl ring and a carboxyl group in para position of the second phenyl ring. ^{19}F R_2-filtered NMR titrations served to determine a 57-fold affinity increase over reference molecule **12.2** (K_I = 13±3 mM), *i.e.* the Man scaffold, and this increase has been reproduced in orthogonal ^{15}N HSQC NMR experiments (Figure 23d). The trifluoromethyl group of **12.11** displayed considerable line broadening $\Delta\nu_{0.5}$ in presence of Langerin and 0.01% Tween-20, further validating the interaction and excluding an aggregative false-positive

mechanism (Figure 22c and c). Additionally, a 4.3-fold affinity increase was observed for **14.1** (K_I = 3.0±0.2 mM) bearing a sulfonamide substituent in C6 (Figure 22a). The substituent represents a bioisoster for the previously identified naturally occurring monosaccharide ligand Man-6-OS. As for **12.11**, the affinity increase was validated in orthogonal ^{15}N HSQC NMR experiments (Figure 23e). Additionally, analogs bearing phenyl-indolinyl substituents in C1 including **12.23** (K_I = 0.33±0.02 mM) were identified as an alternative class of potent glycomimetic Langerin ligands.

The SAR for the Man scaffold suggests that substituents in C1 and C6 might be combined to further increase the affinity of glycomimetic Langerin ligands. Assuming additivity, introducing the substituents of both **12.11** and **14.1** into a single analog would result in a predicted K_I of approximately 0.05 mM corresponding to a 250-fold affinity increase over the reference molecule **12.2**. Yet, additivity of the observed affinity increases critically depends on a conserved binding mode for the Man scaffold. To evaluate this hypothesis the binding modes of **12.2**, **12.11** and **14.1** were analyzed in ^{15}N HSQC and STD NMR experiments. Moreover, this analysis served to further validate the corresponding interactions and the obtained information will generally facilitate the continued optimization of Man analogs as glycomimetic Langerin ligands.

The ^{15}N HSQC NMR experiments for **12.11** indicated additional interactions with the Langerin surface in proximity of K313 and D312 (Figure 24b). In X-ray structures of Langerin in complex with Man-type glycans, these residues are located in equatorial direction of C1 of the Man scaffold (Feinberg *et al.*, 2011). Accordingly, **12.11** potentially forms cation-π interactions or hydrogen bonds with *e.g.* K313 *via* the second phenyl ring or the carboxyl group, respectively. Moreover, observed affinity increases for the biphenyl subset of Man analogs **12** critically depend on the substituent in ortho position of the first phenyl ring (Table 7). As this position has been reported to affect the dihedral angles between the Man scaffold and the two phenyl rings, the trifluoro methyl group is likely required to optimally orient the second phenyl ring to form these interactions (Klein *et al.*, 2010). The determined STD NMR epitope further validates interactions between the biphenyl substituent of **12.12** and the Langerin surface (Figure 26d). Here, aromatic hydrogen atoms displayed uniformly increased STD'$_0$ values compared to the Man scaffold. Furthermore, comparable epitopes are obtained for the Man scaffold of **12.11** and **12.2**. This suggests a conserved binding mode and is consistent with the CSP trajectories observed for these analogs.

Regarding analog **14.1**, ^{15}N HSQC NMR experiments revealed distinct CSP values for I282 and E285 compared to **12.2** (Figure 24b). While this might be attributed to the expected formation of interactions by the sulfonamide substituent in C6 in proximity to K299, additional observations suggest an alternative binding mode for **14.1.** Here, CSP trajectories for a set of residues associated with the recognition of the Man scaffold including N297, K299 and K313 deviate from **12.2** (Figure 25). Moreover, **14.1** displays an STD NMR epitope with a considerably decreased STD'$_0$ value for H2

(Table 11). Overall, this potentially challenges the additivity hypothesis for substituents in C1 and C6 of the Man scaffold.

The employed strategy integrating ^{15}N HSQC and STD NMR experiments provides valuable information on the binding mode of **12.2**, **12.11** and **14.1**. In principal, the experiments account for alternative binding modes as well as differences in binding kinetics and receptor dynamics. However, both NMR techniques display limited spatial resolution, impeding the interpretation of determined CSP patterns and binding epitopes in the context of SARs. Moreover, these interpretations rely on the availability of suitable X-ray structures, that might not account for receptor dynamics, *i.e.* induced fit or conformational selection binding mechanisms. Accordingly, the proposition of an alternative binding mode for **14.1** should be treated cautiously, rendering further evaluation of the additivity hypothesis for substituents in C1 and C6 desirable.

Quantitative approaches to binding mode analysis have been developed for both NMR techniques and provide improved spatial resolution. For STD NMR, the complete relaxation and conformational exchange matrix saturation transfer (CORCEMA-ST) approach enables the validation and optimization of generated docking poses (Jayalakshmi and Rama Krishna, 2004). CORCEMA-ST explicitly accounts for R_1 and cross-relaxation mechanisms as well as the three-dimensional properties of a complex and chemical exchange contributions to STD effects. The high ambiguity driven protein-protein docking (HADDOCK) approach developed for ^{15}N HSQC NMR experiments utilizes CSP values as restraints to generate docking poses (Dominguez *et al.*, 2003). Notably, neither approach accounts for receptor dynamics and the structural flexibility of the Langerin's carbohydrate binding site will complicate their implementation. Here, ensemble docking approaches might be employed to address this limitation (Totrov and Abagyan, 2008). Moreover, the distribution of observed CSPs throughout the CTLD fold due to allosteric effects render the implementation of the HADDOCK approach challenging.

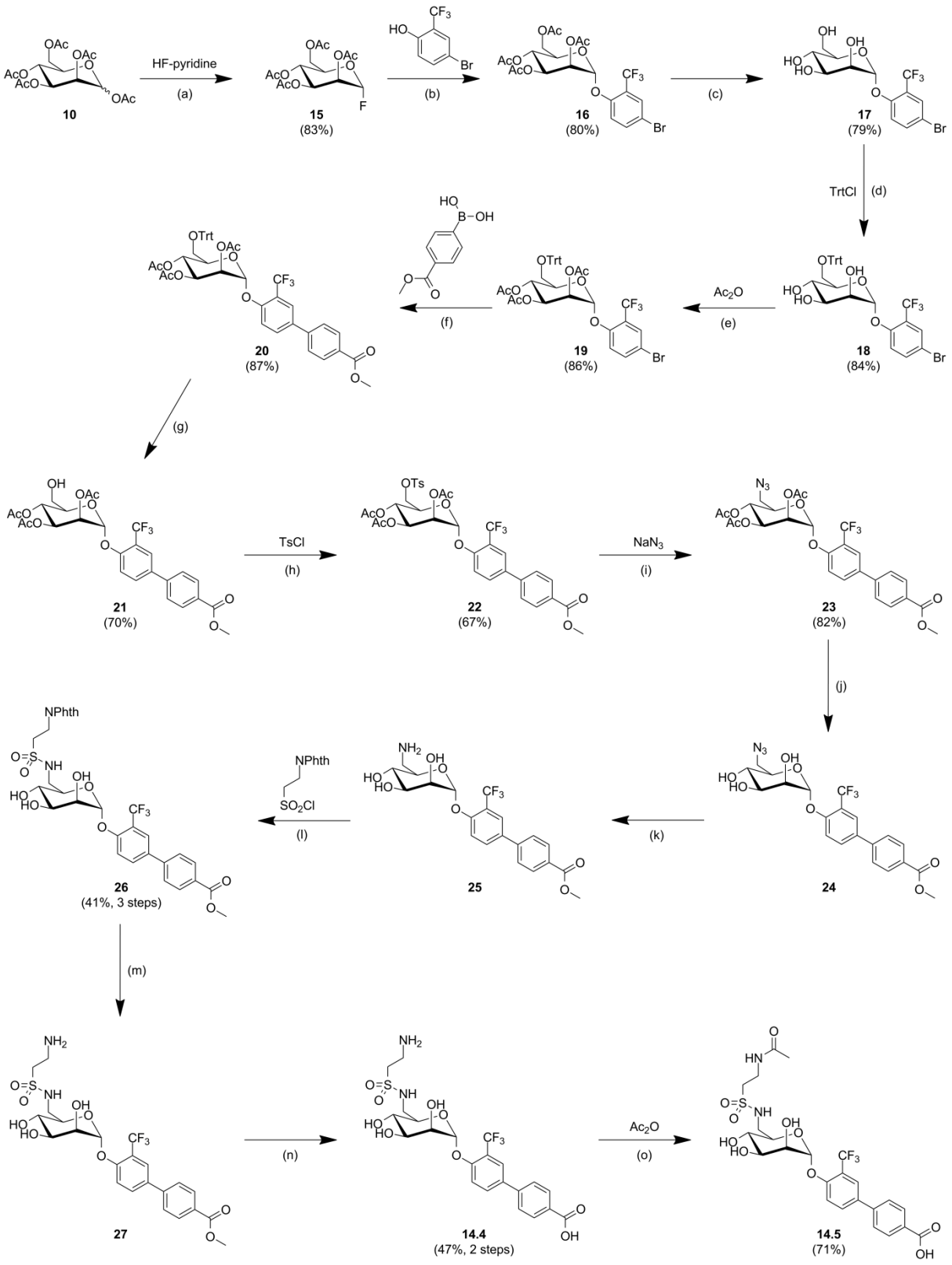
Alternatively, X-ray crystallography generally provides information on the dominant binding mode at atomic resolution and the elucidation of X-ray structures of Langerin complexes with **12.11** and **14.1** will substantially contribute to the interpretation of the obtained SAR. In absence of these X-ray structures, the synthesis and characterization of analogs bearing both biphenyl substituents in C1 and sulfonamide substituents in C6 provides a suitable approach to evaluate the additivity hypothesis. This approach will be presented in the following section and represents an important step towards the utilization of potent Man analogs for the development of targeted delivery application for LCs (Chapter 3.2.3).

3.3.3. Conjugation Strategy for Liposomal Formulations

Man analog **12.11** (K_I = 0.23±0.03 mM) represents the most potent glycomimetic ligand reported for Langerin. Its submillimolar affinity might be sufficient to efficiently target liposomes to LCs and hence, the development of antigen delivery applications. For comparison, the blood group antigen Le^X (K_D = ca. 1 mM) has been reported to bind DC-SIGN with millimolar affinity and its conjugation to liposomes enabled the DC-dependent activation of T cells *in vitro* (Fehres *et al.*, 2015c; Pederson *et al.*, 2014). **12.11** will likely display a favorable specificity profile over other human CLRs as the biphenyl subset of Man analogs **12** has previously been reported to display low affinities for DC-SIGN, DC-SIGNR, SP-D, MBP-C and BDCA-2 (Scharenberg *et al.*, 2012). However, the utilization of **12.11** in targeted delivery applications requires the development of a conjugation strategy and therefore the introduction of suitable functional groups.

As discussed in the preceding sections, modifications in C2, C3, and C4 would considerably decrease the affinity for Langerin (Chapter 3.2.1 and Chapter 3.3.1). An evaluation of the SAR for biphenyl substituents in C1 suggests that conjugation strategies *via* the carboxyl group in para position of the second phenyl ring are not feasible (Table 7). Additionally, the introduction of substituents in other position will potentially alter the relative orientation of the phenyl rings and the Man scaffolds and result decreased affinities. The introduction of a sulfonamide group in C6 of **12.11** by contrast, represents a promising strategy for the synthesis of glycolipids, particularly when considering the steric tolerance for phenyl rings in this position. Assuming additivity for the interactions formed by substituents in C1 and C6, this would additionally result in a substantial affinity increase.

Following this strategy, Man analog **14.4** was synthesized over 14 steps with an overall yield of 2.4% (Scheme 4). Intermediate **16** was prepared as previously published, followed by the orthogonal protection of OH6 with a trityl group (Pang *et al.*, 2012; Scott *et al.*, 1999). Next, the biphenyl substituent was assembled utilizing Suzuki-Miyaura coupling and an amine group was introduced in C6 (Miyaura *et al.*, 1979). Notably, the reduction of the azide group in presence of acetyl groups resulted in their migration to form an acetamide in C6. Accordingly, acetyl groups were removed prior to azide reduction. As the efficient tosylation of OH6 for unprotected Man derivatives has been reported, the utilization of the trityl protecting group might be omitted (Hauck *et al.*, 2013). Thereby the synthetic route would be shortened by 4 steps which would likely provide an optimized overall yield.


HF-pyridine
(a)
(b)
(c)
TrtCl
(d)
Ac2O
(e)
(f)
(g)
TsCl
(h)
NaN3
(i)
(j)
(k)
NPhth
SO2Cl
(l)
(m)
(n)
Ac2O
(o)
10
15
(83%)
16
(80%)
17
(79%)
18
(84%)
19
(86%)
20
(87%)
21
(70%)
22
(67%)
23
(82%)
24
25
26
(41%, 3 steps)
27
14.4
(47%, 2 steps)
14.5
(71%)

Scheme 4. Synthesis of Man analogs 14.4 and 14.5.[28]

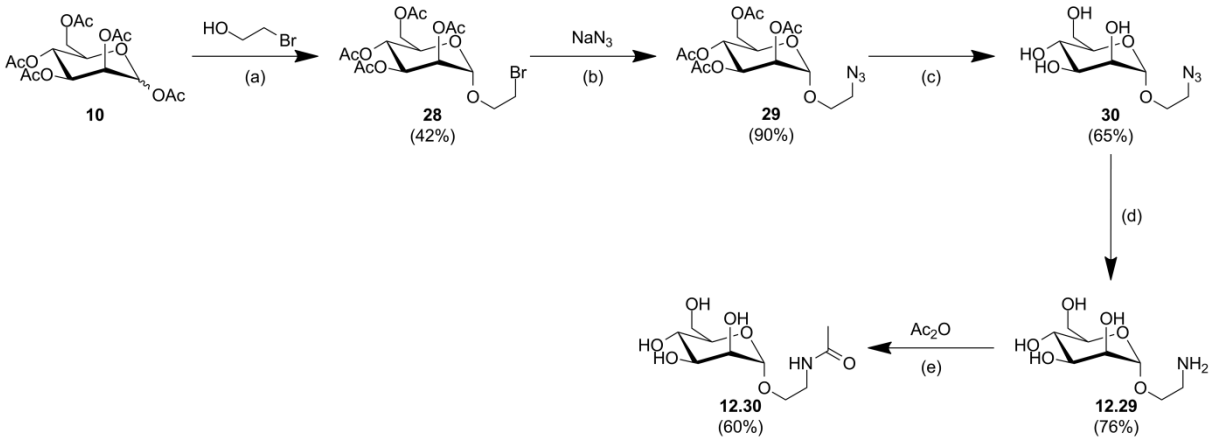
Intermediate **16** was prepared as previously published (Pang *et al.*, 2012; Scott *et al.*, 1999). Reaction conditions for the preparation of Man **14.5**: (a) DCM, 0°C to 40°C; (b) $BF_3{\cdot}OEt_2$, anhydrous DCM, 0°C to room temperature; (c) MeONa, MeOH, room temperature; (d) DMAP, pyridine, room temperature; (e) pyridine, room temperature; (f) $Pd(dppf)Cl_2{\cdot}CH_2Cl_2$, K_3PO_4, anhydrous DMF, 80°C; (g) $FeCl_3$, H_2O, DCM, room temperature; (h) DMAP, anhydrous pyridine, 0°C to room temperature; (i) 1,4,7,10,13,16-hexaoxacyclooctadecane, anhydrous DMF, 0°C to 80°C; (j) MeONa, anhydrous MeOH, room temperature; (k) H_2, Pd/C, 1,4-dioxane, room temperature; (l) Et_3N, anhydrous DMF, 0°C to room temperature; (m) hydrazine monohydrate, MeOH, room temperature; (n) 2 M aqueous NaOH, MeOH, room temperature; (o) MeOH, room temperature.

Next, the amine group in **25** was functionalized with a sulfonyl chloride bearing a phthalimide protected ethylamino linker. Finally, global deprotection yielded **14.4** bearing an amino group in C6 that enables its conjugation to NHS-activated lipids. Importantly, **14.4** was prepared in sufficient quantities for several liposomal formulations. Furthermore, Man analog **14.5** was synthesized as a model molecule to analyze the SAR for the developed conjugation strategy in NMR experiments. The synthesis involved the selective acetylation of the amino group at a yield of 70%.

Man analog **12.29** also bears an ethylamino linker and was selected as a reference molecule to evaluate the conjugation strategy implemented for **12.11** (Scheme 5). **12.29** was synthesized over 4 steps with an overall yield of 19% in the laboratory of Prof. Seitz (Humboldt-Universität zu Berlin) and kindly provided for conjugation to NHS-activated lipids (Hayes *et al.*, 2003). In analogy to **14.4**, the amine was acetylated to afford **12.30** at a yield of 60% for characterization in NMR experiments.

The K_I values of **12.30** and **14.5** were determined by ^{19}F R_2-filtered NMR experiments (Figure 27a). An evaluation of the obtained affinities revealed the SAR for the proposed conjugation strategies for Man analogs (Table 12). The introduction of an ethylamino linker *via* the formation of an α-mannoside results in a 2.2-fold affinity decrease for **12.30** (K_I = 10±1 mM) compared to Man (K_I = 4.5±0.5 mM). This is consistent with decreased affinities observed for other analogs bearing small substituents in C1 such as **12.1** (K_I = 7.6±0.1) and **12.2** (K_I = 13±3 mM). In contrast, **14.5** (K_I = 0.25±0.07 mM) and **12.11** (K_I = 0.23±0.03 mM) displayed comparable affinities. In approximation of the conditions on targeted liposomes, **14.5** provides a 40-fold affinity increase over reference molecule **12.30.**

[28] The procedure was devised by Dr. Oliver Schwardt (Universität Basel) and me. Intermediate **19** was prepared by Dr. Oliver Schwardt.



Scheme 5. Synthesis of Man analogs 12.29 and 12.30.[29]

12.29 was prepared as previously published (Hayes *et al.*, 2003). Reaction conditions for the preparation of **12.30**: (a) $BF_3.OEt_2$, anhydrous DCM, room temperature; (b) anhydrous DMF, 60°C; (c) MeONa, MeOH, room temperature; (d) H_2, Pd/C, anhydrous MeOH, room temperature; (e) MeOH, room temperature.

Determined K_I values for **12.30** (K_D = 12±1 mM) and **14.5** (K_D = 0.46±0.09 mM) were subsequently validated in orthogonal ^{15}N HSQC NMR experiments (Figure 27b to e). Notably, titrations with **12.30** were conducted in absence of additives, while 10% DMSO was present in titrations with **14.5**. As discussed in the preceding section, the presence of DMSO and chemical exchange phenomena likely account for the decreased amplitude and number of CSPs in titrations with **14.5**. Accordingly, the accuracy for K_D determination is compromised resulting in minor deviations compared to ^{19}F R2-filtered NMR experiments. Nevertheless, a 26-fold affinity increase over **12.30** was observed, further validating **14.5** as a potent glycomimetic Langerin ligand.

On the one hand, the obtained SAR characterizes the utilization of an ethylamino linker in C6 as a favorable conjugation strategy for Man analogs. On the other hand, the introduction of the sulfonamide group in this position did not result in the anticipated affinity increase for **14.5** suggesting that the interactions formed by **12.11** and **14.1** are not additive. A binding mode analysis for **12.30** and **14.5** by ^{15}N HSQC and STD NMR experiments served to explain these observations (Figure 28 and Figure 29, Table 13). The presence of DMSO in experiments with **14.5** limits the transferability between the ^{15}N HSQC NMR titrations and the results will therefore be discussed separately.

[29] The procedure was devised by Gunnar Bachem (Humboldt-Universität zu Berlin) and me and conducted by Gunnar Bachem.

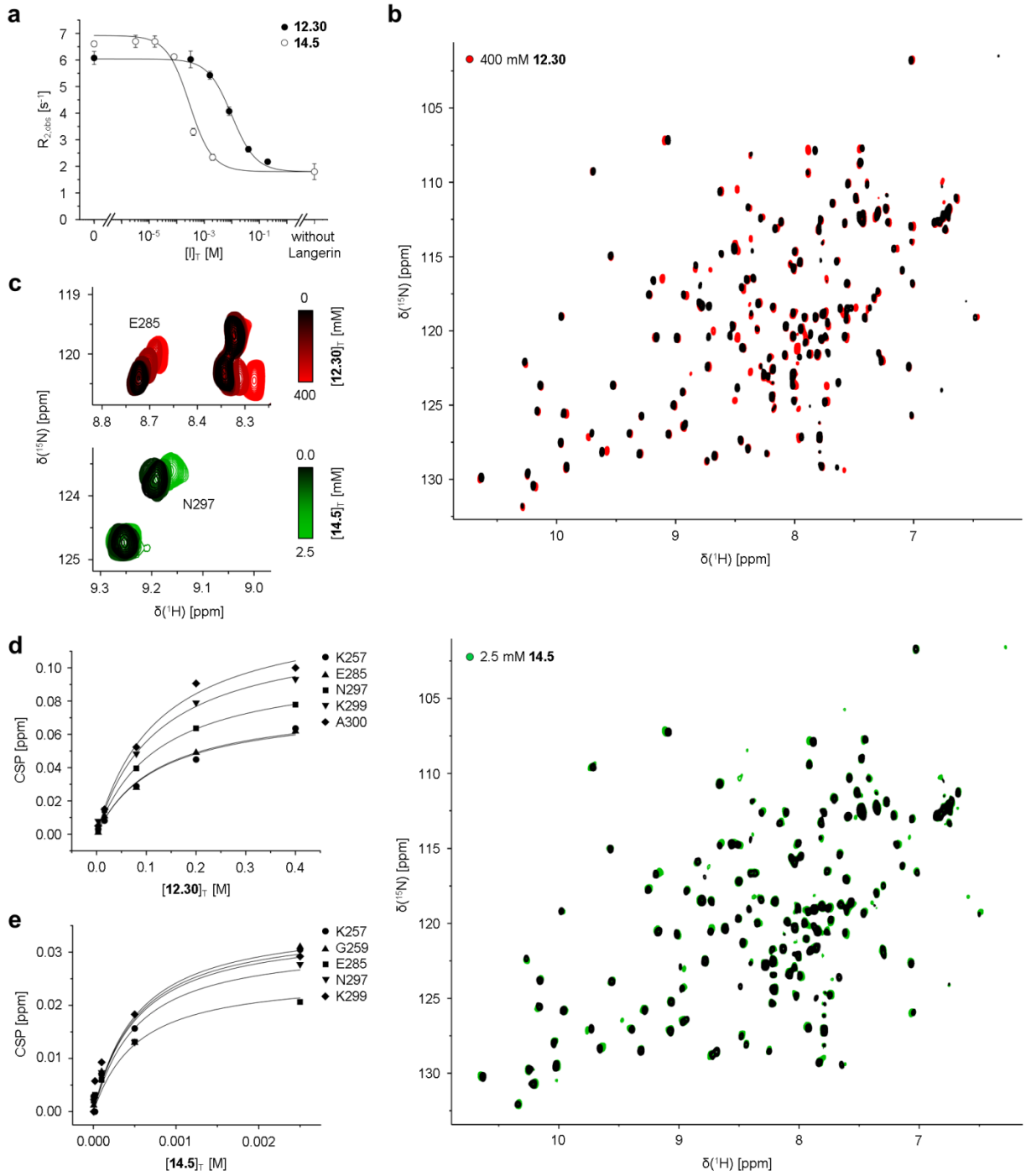


Figure 27. K_I and K_D determination for Man analogs 12.30 and 14.5. [30]

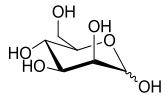
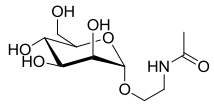
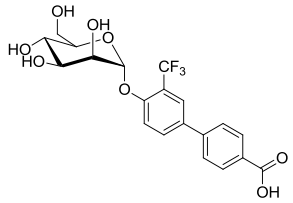
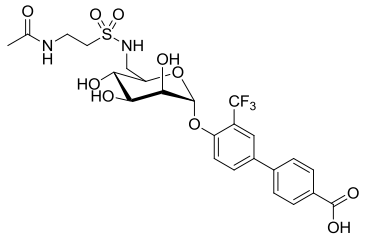
a. ^{19}F R2-filtered NMR experiments served to determine the K_I values of **12.30** (K_I = 10±1 mM) and **14.5** (K_I = 0.25±0.07 mM). These model structures were synthesized to evaluate the conjugation strategy for liposomal formulations. **b.** and **c.** ^{15}N HSQC NMR experiments with the Langerin CRD served to validate obtained K_I. Assigned resonances displaying fast chemical exchange and CSPs larger than 0.06 or 0.02 ppm were selected for the determination of K_D values, respectively. Representative CSP trajectories are depicted. **d** and **e.** Barring minor de*via*tions, determined K_D values for **14.5** (K_D = 0.46±0.09 mM) are **12.30** (K_D = 12±1 mM) are consistent with the ^{19}F R2-filtered NMR assay and validate **14.5** as a potent glycomimetic Langerin ligand. All titrations with **14.5** were conducted in presence of 10% DMSO.

[30] In part, these procedures were conducted by Lennart Schnirch (Freie Universität Berlin) under my supervision.

As expected, the CSP pattern for **12.30** closely resembled that of Man and **12.2** (Figure 28a and b, Figure 24). Prominent residues displaying CSPs upon Langerin binding to **12.30** include E285 and K299. Both residues form interactions with the Man scaffold as determined by X-ray crystallography (Feinberg *et al.*, 2011). Additional CSPs were observed for other residues associated with the carbohydrate binding site including N297, A300 and K313. As detected for other glycomimetics, **12.30** affected the allosteric network involved in Ca^{2+} recognition, particularly K257 and G259 located in the short loop region. Overall, these findings imply that the Ca^{2+}-dependent binding mode of the Man scaffold is maintained and closely resembles the corresponding X-ray structure. Notably, no contributions from the acetylated ethylamino linker are observed, indicating a solvent exposed orientation for this substituent in C1.

For **14.5**, observed CSP values were uniformly decreased, likely due to the presence of 10% DMSO as well as chemical exchange phenomena (Figure 28a). Accordingly, the induced CSP pattern was visualized at a reduced threshold of 0.02 ppm (Figure 28c). This value corresponds to the standard deviation σ previously reported for CSP quantifications with Langerin (Hanske *et al.*, 2016). Utilizing this threshold, most of the residues associated with the recognition of the Man scaffold including E285, N297, K299, A300 and K313 were affected upon **14.5** binding. Likewise, **14.5** also induced CSPs for K257 and G259 in the short loop region, which is part of the previously identified allosteric network (Hanske *et al.*, 2016). A comparison with the titration for **12.11** further validates the proposed orientation of the biphenyl substituent in equatorial direction of C1. While D312 displayed a CSP value (CSP = 0.015 ppm) that is smaller than the threshold, the CSP pattern in this region of the Langerin surface was reproducibly observed for both **12.11** and **14.5** (Figure 24).

Table 12. Structure-activity relationship for the conjugation of Man analogs 12.29 and 14.4.[31]

Name	Structure	K_I [mM]	K_D [mM]	Relative potency[a]
Man		4.5±0.5[b]	5.8±0.3	2.2
12.30		10±1	12±1	1.0
12.11		0.23±0.03	0.5±0.02	44
14.5		0.25±0.07	0.46±0.09	40

[a] The relative potency was calculated utilizing the K_I value determined for **12.30**.
[b] This value was determined from three independent titration experiments.

In distinction to the patterns obtained for **12.2** and **12.11**, **14.5** binding additionally affected G284. This residue is located in direction of C6 and the observed CSP might be induced by the sulfonamide group or the acetylated ethylamino linker. Lastly, the comparative analysis of CSP trajectories, *i.e.* the vector of induced perturbations, revealed comparable trajectories for **12.2**, **12.11** and **14.5** for residues associated with the recognition of the Man scaffold including E285, N297, K299, A300 and K313, providing no indication for an alternative binding mode. In summary, these findings suggest that the Ca^{2+}-dependent binding mode of the Man scaffold is maintained and that **14.5** shares this binding mode with **12.11**. Under this assumption, the biphenyl substituent in C1 is oriented towards K313 and D312 and potentially forms cation-π interactions or hydrogen bonds with in this region of the Langerin surface while the acetylated ethylamino linker in C6 is oriented towards G248.

[31] In part, these procedures were conducted by Lennart Schnirch (Freie Universität Berlin) under my supervision.

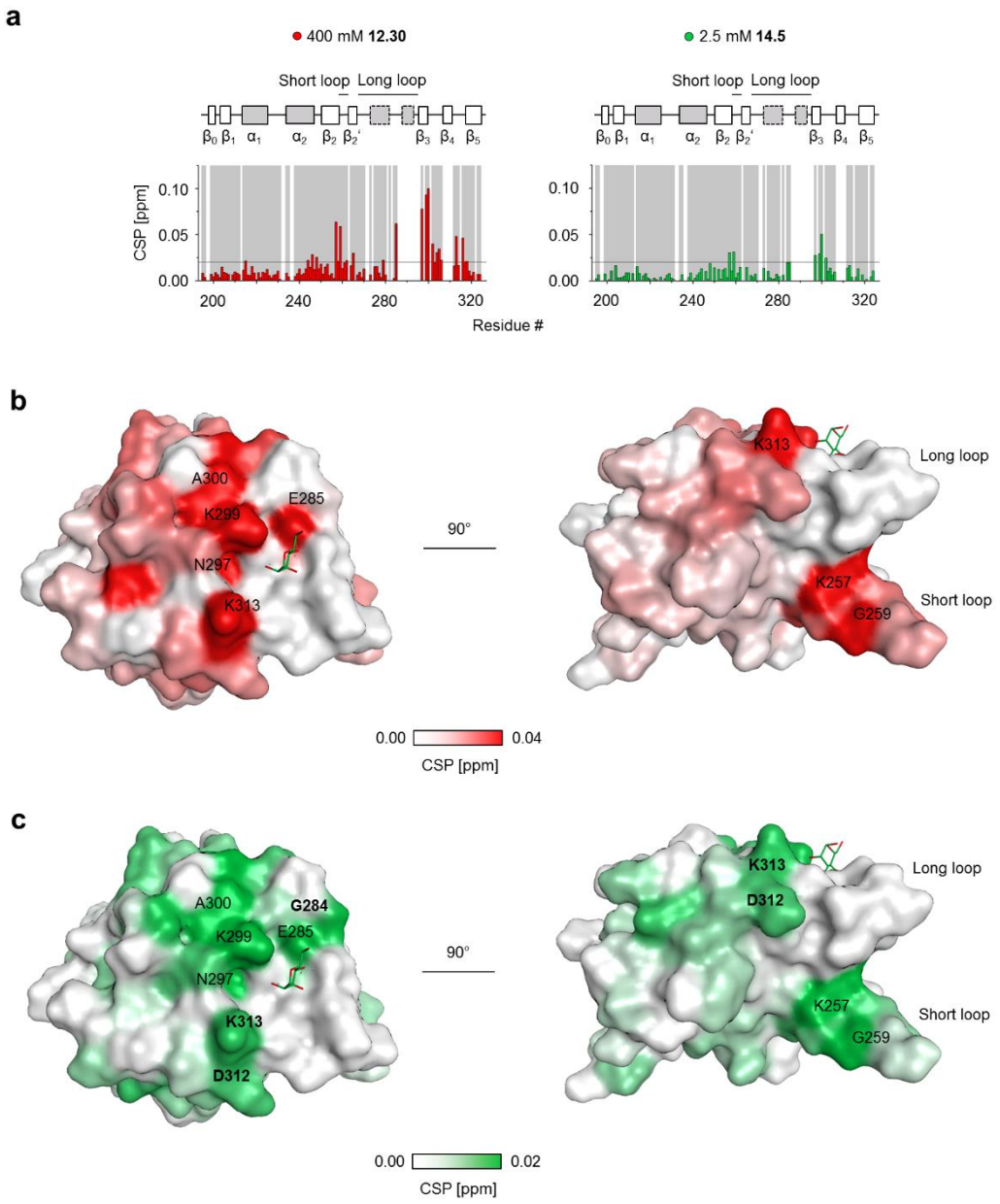


Figure 28. ^{15}N HSQC NMR binding mode analysis for Man analogs 12.30 and 14.5.

a. ^{15}N HSQC NMR experiments with the Langerin CRD reveal the CSP fingerprints for **12.30** and **14.5**. Assigned resonances detected in the reference spectrum are highlighted (grey). Observed CSP values for **14.5** are likely reduced due to the presence of 10% DMSO and intermediate and slow exchange phenomena. **b and c.** Mapping the CSPs on the X-ray structure of the Langerin CRD (PDB code: 3P5F) validates a Ca^{2+}-dependent binding mode for both **12.30** (red) and **14.5** (green) as indicated by CSPs observed for E285 and K299 (Feinberg *et al.*, 2011). Additionally, CSPs are observed for N297, K313 and A300, residues also affected upon recognition of Man and ManNAc as well as other Man analogs (Chapter 3.2.1, 3.3.1 and 3.3.2). While the CSP pattern for reference molecule **12.30** is comparable to Man, two distinct features are observed for **14.5**. Here, prominent effects observed for D312 might be indicative for an interaction with the second phenyl ring or the carboxyl group while the CSP for G284 might be inducted by either the sulfonamide group or the acetylated ethylamino linker. These findings suggest that the binding mode of the Man scaffold is maintained and comparable to **12.11**. CSPs are also observed in remote regions of the CTLD fold, particularly for K257 and G259 in the short loop region. This indicates a modulation of the previously reported allosteric network (Hanske *et al.*, 2016).

STD NMR epitope mapping with **12.30** and **14.5** served to evaluate the orientation of the acetylated ethylamino linker and to further validate the findings from the ^{15}N HSQC NMR experiments (Figure 29, Table 13, Appendix B.5). Both analogs displayed specific STD effects in presence of Langerin confirming the interactions formed by these glycomimetics (Figure 29a and b). Notably, spectral overlap with the HDO resonance impeded an analysis for H1 of **12.30**. Furthermore, no STD'$_0$ value could be determined for H5 of **14.5** due to spectral overlap and strong J coupling effect with H6a. However, the corresponding multiplet displayed an STD effect. As discussed in the previous sections, all Langerin samples contained minor EDTA impurities which could not be removed after repeated dialysis against H_2O. Yet, these impurities will likely not interfere with the conducted experiments.

The STD NMR epitope obtained for **12.30** revealed similar STD effects for all hydrogens of the Man scaffold with H6 displaying a slightly decreased STD'$_0$ value (Table 13). This non-differentiated epitope is comparable to that of **12.2** and might be attributed to the alternative orientation observed for Man-type glycan binding to Langerin (Feinberg *et al.*, 2011). A similar, yet more differentiated binding epitope has previously been reported for the interactions of a Man-type disaccharide with Langerin (Hanske *et al.*, 2017a). Importantly, low STD'$_0$ values were observed for the acetylated ethylamino linker suggesting a solvent exposed orientation (Figure 29c).

This observation also holds true for **14.5**, displaying comparably low relative STD'$_0$ values for the linker (Figure 29c). Overall, the binding epitope for **14.5** is dominated by uniformly strong STD effects for the biphenyl substituent in C1, consistent with the relative STD'$_0$ values determined for **12.11**. This not only validates the proposed formation of interactions with the Langerin surface in proximity of D312 and K313, it is also consistent with the observed affinity increase. The normalization of STD'$_0$ values within the Man scaffold revealed a similar yet more differentiated epitope compared to **12.30** as well as **12.11** with decreased STD effects for both H5 and H6 (Table 13). The findings are in accordance with the epitope previously reported for the Langerin interactions with a Man-type disaccharide (Hanske *et al.*, 2017a). Hence, STD NMR experiments indicate a conservation of the binding mode of the Man scaffold determined *via* X-ray crystallography for **14.5** (Feinberg *et al.*, 2001).

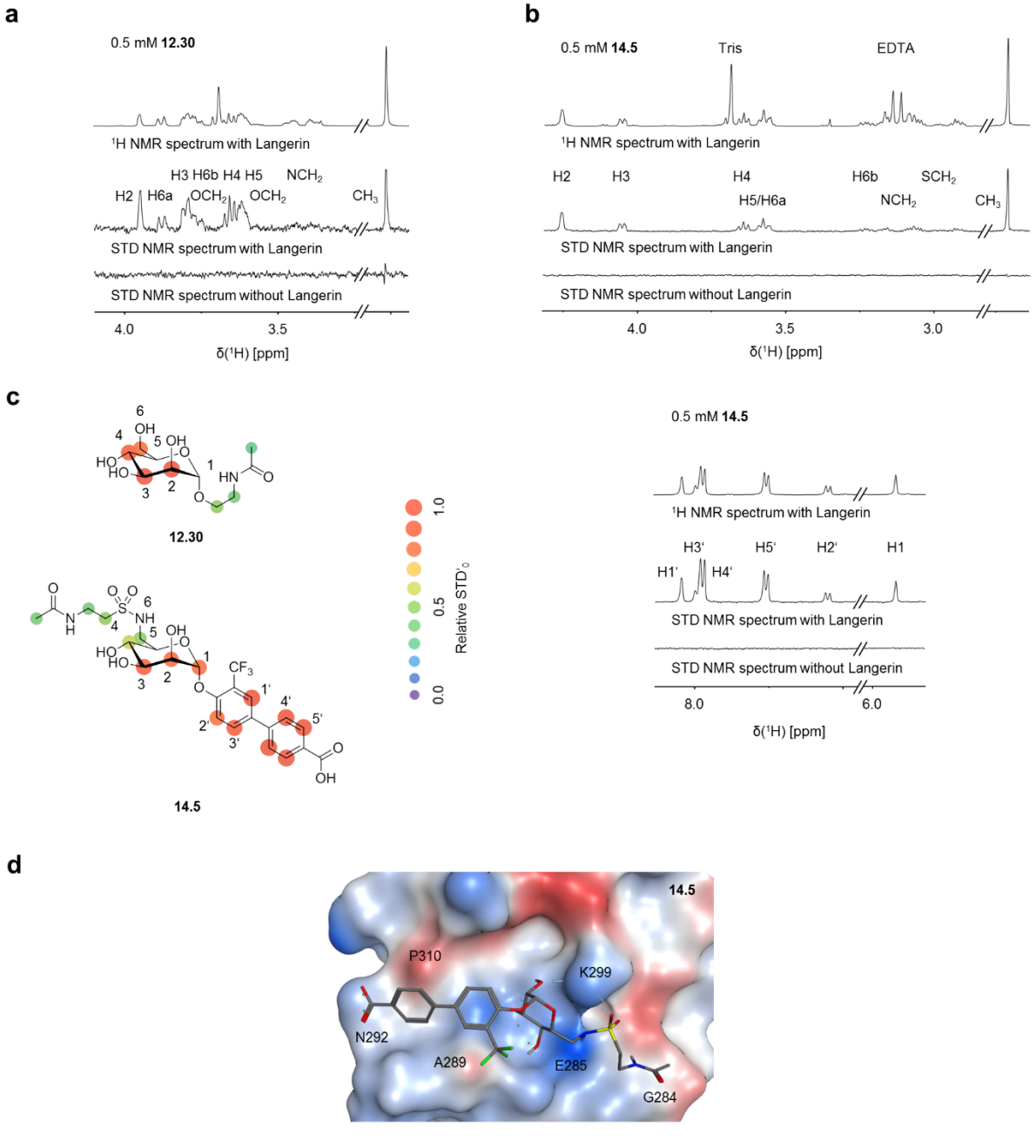


Figure 29. STD NMR epitope mapping and docking pose for Man analogs 12.30 and 14.5. [32]

a. and b. STD NMR experiments served to further validate the interactions formed by **12.30** and **14.5** with the Langerin ECD. STD NMR spectra were recorded at saturation times t_{sat} of 0.4 s and are magnified 8-fold and 64-fold, respectively. **c.** STD NMR epitopes for **12.30** and **14.5** were determined from build-up curves (Appendix B.5). Similar to **12.11**, **14.5** displayed pronounced STD effects for the biphenyl substituent in C1. The epitopes determined for the Man scaffold of **12.30** and **14.5** were comparable. Both analogs exhibited low relative STD'$_0$ values for the acetylated ethylamino linker, a strong indication for a solvent exposed orientation. **d. 14.5** was docked into the carbohydrate binding site to rationalize the observations from ^{15}N HSQC and STD NMR experiments. The selected docking pose for **14.5** predicts the formation of a hydrogen bond between the carboxyl group and N292. Additionally, the second phenyl ring is located proximal to P310 and the trifluoromethyl group is forms van der Waals interactions with A289. The linker is oriented towards G284 and displaying high solvent exposure. The receptor surface is colored according to its lipophilicity (lipophilic: red, hydrophilic: blue).

[32] In part, these procedures were conducted by Lennart Schnirch (Freie Universität Berlin) under my supervision.

Table 13. Parameters for STD NMR epitope mapping for Man analogs 12.30 and 14.5.

	Relative STD'$_0$	
Proton #	**12.30**	**14.5**
1		0.8±0.2
2	1.0±0.2	0.9±0.2
3	0.9±0.2	1.0±0.3
4	0.9±0.1	0.8±0.2
5		0.6±0.2
6	0.7±0.1	0.6±0.1

[a] The relative STD'$_0$ was calculated utilizing the highest STD'$_0$ observed for the Man scaffold. Standard errors were derived directly from the fitting procedure.

To rationalize the findings from the ^{15}N HSQC and STD NMR experiments and to visualize potential interactions formed by **14.5**, molecular docking studies were conducted (Figure 29d). In analogy to the procedure implemented for the *in silico* design of ManNAc analogs **5**, two alternative K313 were considered and a conformation database for **14.5** was generated with MOE's Conformation Import (Chemical Computing Group, 2016). Conformers were placed in the carbohydrate binding site utilizing the pharmacophore model previously developed to constrain the binding mode of the Man scaffold (Figure 17). Generated docking poses were scored using the proprietary London ΔG function (Chemical Computing Group, 2016). Highly scored poses were refined utilizing MM simulations, rescored *via* the GBVI/WSA ΔG function and filtered by the pharmacophore model (Corbeil *et al.*, 2012; Massova and Kollman, 2000). Overall, ten docking poses were generated and evaluated visually in the context of the hypotheses generated from ^{15}N HSQC and STD NMR experiments. Seven of these docking poses predicted the orientation of the biphenyl substituent in C1 in direction of K313 and D312 and a representative pose was selected for discussion (Figure 29d). However, interactions with these residues were not observed. Instead, all docking poses displayed the formation of a hydrogen bond between the carboxyl group and N292. The second phenyl ring is located in proximity of P310 while the trifluoromethyl group forms van-der-Waals interactions with A289. Notably, the resonances corresponding to these residues were not assigned and ^{15}N HSQC NMR experiments do not provide direct evidence of the predicted interactions. Nevertheless, the generated docking pose might be envisioned to induce the increased CSP values for D312 and K313 and represents a valuable three-dimensional model to rationalize to affinity increase observed for **14.5**. In contrast to the biphenyl substituent, various conformations were generated for the sulfonamide group in C6 and no conserved interactions were identified. Yet, for all seven docking poses the acetylated ethylamino linker was oriented in approximate direction of I282 or G284 and displayed considerable solvent exposure. These findings are consistent with the low STD'$_0$ values of the linker and the non-additive SAR.

While the implemented approach implicitly accounted for minor conformational changes of the carbohydrate binding site and the Man scaffold, generated docking poses were considerably constrained. Accordingly, the presented model should be treated cautiously, particularly as an alternative orientation of the Man scaffold would allow for the formation of interactions with *e.g.* K313 or D312 and as this orientation might not be detectable *via* ^{15}N HSQC and STD NMR experiments. More elaborate techniques including the aforementioned CORCEMA-ST and HADDOCK approaches potentially provide improved spatial resolution (Dominguez *et al.*, 2003; Jayalakshmi and Rama Krishna, 2004). Additionally, trNOE NMR experiments enable the determination of conformations corresponding to the bound state of **14.5** and might be utilized to validate proposed orientation of the biphenyl substituent in the future (Fiege *et al.*, 2012; Munoz-Garcia *et al.*, 2015). While ensemble docking approaches would account for the conformational flexibility of the carbohydrate binding site, the elucidation of an X-ray structure of the Langerin complex with **14.5** is highly desirable for the continued optimization of the class glycomimetic ligands (Totrov and Abagyan, 2008).

To summarize, the obtained SAR and the comparative binding mode analysis for model molecules **12.30** and **14.5** in ^{15}N HSQC and STD NMR experiments as well as docking studies yield two central findings for the proposed conjugation strategy. First, the introduction of a linker in C6 or C1 of the Man scaffold does not result in steric hindrance with the Langerin surface. Obtained STD NMR epitopes suggest a solvent exposed orientation for both positions and obtained affinities are comparable to **12.2** and **12.11**, respectively (Figure 27 and Figure 29, Table 12). Secondly, the introduction of the sulfonamide group in C6 did not result in the anticipated affinity increase for **14.5 (**Figure 27, Table 12). ^{15}N HSQC and STD NMR experiments suggest that **12.11** and **14.5** share a similar binding mode distinct from that of **14.1** (Figure 28 and Figure 29). This binding mode is likely comparable to that of Man-type glycans observed *via* X-ray crystallography (Feinberg *et al.*, 2001). Consequently, the geometric requirements for interactions formed by **14.1** *via* the sulfonamide group are not met for **14.5**, resulting in a non-additive SAR. As the sulfonamide group seems to be irrelevant to maintain the affinity of **14.5**, a simplified conjugation strategy for liposomal conjugation might be envisioned. After global deprotection of intermediate **25**, the corresponding product would be applicable for conjugation to NHS-activated lipids. This approach results in a substantially shortened synthetic route comprising only 8 steps when considering the proposed omission of the trityl protecting group discussed above (Hauck *et al.*, 2013).

In conclusion, the presented findings render glycomimetic **14.4** a promising targeting ligand for antigen delivery to LCs due to its reactive ethylamino linker. Hitherto, **14.4** represents the most potent molecular probe reported for Langerin, displaying a submillimolar affinity (K_I = 0.25±0.07 mM) that exceeds that of naturally occurring mono- or oligosaccharides. The availability of an additional targeting ligand, **12.29**, with a substantially decreased K_I value will enable investigations on the

influence of monovalent affinities on the endocytosis of liposomes and the processing of antigens by LCs. Importantly, both analogs were prepared in sufficient quantities for their conjugation to NHS-activated lipids and liposomal formulation utilizing **14.4** are currently being established in our laboratory. Aside from the required affinity to enable the efficient endocytosis of liposomes, specificity for the targeted CLR represents an essential requirement for targeting ligands in antigen delivery applications. A favorable specificity profile against other human CLRs including DC-SIGN, DC-SIGNR, SP-D, MBP-C and BDCA-2 has been reported for Man analogs bearing biphenyl substituents in C1. However, the introduction of the sulfonamide group and the ethylamino linker in C6 might alter this profile (Scharenberg *et al.*, 2012).

While **14.4** has not been utilized for liposomal formulations within the scope of this dissertation and its specificity against other human CLRs remains to be elucidated, these aspects are currently under investigation in our laboratory. In parallel to the work presented in this chapter, an alternative class of glycomimetic Langerin ligands, GlcNS analogs, was explored for their potential as targeting ligands for antigen delivery applications. The design of these analogs as well as the evaluation of their specificity and their utilization with liposomes in cell-based *in vitro* and *ex vivo* experiments will be presented in the following chapter (Chapter 3.4).

3.4. Heparin-Inspired Glycomimetic Ligands for Langerhans Cell Targeting

3.4.1. Structure-Activity Relationship of *N*-Acetylglucosamine Derivatives and their Synthetic Analogs

Aside from its function as a PRR, Langerin has been reported to interact with sulfated and non-sulfated GAG self-antigens (Chabrol *et al.*, 2012; Hanske *et al.*, 2017b; Munoz-Garcia *et al.*, 2015; van den Berg *et al.*, 2015; Zhao *et al.*, 2016). GAGs are linear polysaccharides whose disaccharide repeating units consist of Gal or uronic acids (UrA) and differentially sulfated GlcNAc derivatives. They are involved in various processes including cell-cell interactions, signaling and structural scaffolding of the ECM. Except for hyaluronic acid and heprain, GAGs are expressed as proteoglycans and typically secreted to the ECM or presented on the cell membrane. Langerin-GAG interactions promote LC clustering with DCs and have been proposed to mediate the formation of Birbeck granules (Thepaut *et al.*, 2009; van den Berg *et al.*, 2015).

Recently, a Ca^{2+}-independent binding mode has been proposed for heparin (Munoz-Garcia *et al.*, 2015). Interactions at the trimerization interface of the Langerin ECD were demonstrated to emerge for heparin-derived hexasaccharides while a Ca^{+}-dependent affinity was determined for a trisaccharide (K_D = 0.49±0.05 mM) *via* STD NMR experiments. Investigations conducted in our laboratory further revealed that heparin-derived hexasaccharides induce Ca^{2+}-independent CSPs in ^{15}N HSQC NMR experiments with the CRD (Hanske *et al.*, 2017b). The observed CSP pattern depends on the sulfation pattern and is associated with an allosteric network involved in Ca^{2+} recognition by Langerin (Hanske *et al.*, 2016). ^{19}F R_2-filtered NMR experiments additionally demonstrated partial competition with reporter molecule **5.1** binding ($K_{I,est}$ = 1 to 20 mM) and validated the influence of the sulfation pattern on the binding mode. These experiments were complemented by the determination of K_I values for set of differentially sulfated GlcNAc derivatives present in heparin (Figure 30a, Table 14). Interestingly, obtained affinities for GlcNS (K_I = 1.4±0.2 mM), GlcNAc-6-OS (K_I = 0.6±0.1 mM) and GlcNS-6-OS (K_I = 0.28±0.06 mM) were comparable or higher than those observed for heparin-derived hexasaccharides and other monosaccharides including Glc (K_I = 21±4 mM), GlcNAc (K_I = 4.1±0.7 mM), Man (K_I = 4.5±0.5 mM) and Man-6-OS (K_I = 0.8±0.1 mM) (Table 3 and Table 14).

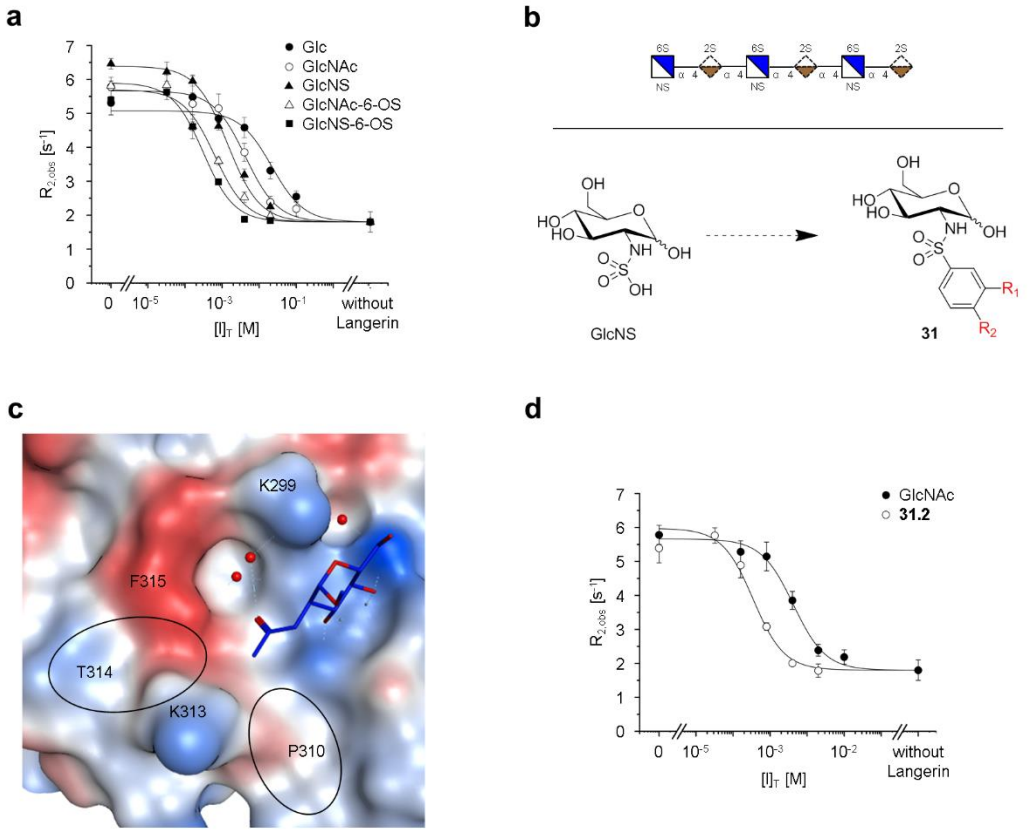
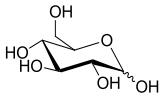
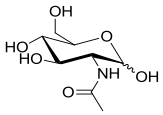
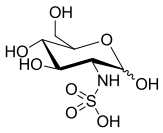
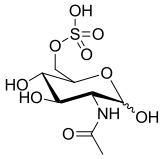
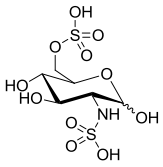


Figure 30. Heparin-inspired design of GlcNS analogs 31.

a. The K_I determination for heparin-derived GlcNAc derivatives *via* ^{19}F R_2-filtered NMR revealed the impact of sulfation patterns on monosaccharide affinity. **b.** GlcNS was identified as a feasible monosaccharide scaffold for glycomimetic ligand design. Analogous to a fragment growing approach, differentially substituted phenyl systems were attached to the Glc scaffold. **c.** The binding mode of GlcNAc (PDB code: 4N32) suggested two regions of the receptor surface in proximity to K313 potentially targeted by GlcNS analogs **31** (Feinberg *et al.*, 2013). **d.** With an affinity increase of up to 13-fold over GlcNAc, the synthesized GlcNS analogs **31** represent an intriguing class of glycomimetic Langerin ligands. The receptor surface is colored according to its lipophilicity (lipophilic: red, hydrophilic: blue).

These findings render sulfated GlcNAc derivatives favorable scaffolds for the design of glycomimetic Langerin ligands (Figure 30b). GlcNS-6-OS represents the most potent monosaccharide identified as sulfation in C2 and C6 results in an additive SAR and a 14-fold affinity increase over GlcNAc. An evaluation of the X-ray structure for the Langerin complex with GlcNAc suggested that the sulfate groups form salt bridges with K299 and K313 (Figure 30c) (Feinberg *et al.*, 2013). Furthermore, the 5-fold affinity increase over Glc observed for GlcNAc is likely the result of an H_2O-mediated hydrogen bond with K299 that might also be formed by GlcNS and GlcNS-6-OS. Importantly, these interactions can be leveraged for the design glycomimetic ligands *via* the bioisoteric substitution of the sulfate groups with a sulfonamide linker (Meanwell, 2011). In this context, the synthesis of GlcNS analogs represents a feasible fragment growing approach to explore the carbohydrate binding site for favorable secondary interactions. Small aromatic substituents in C2 were envisioned to form cation-π interactions with K313 or π-π and H-π interactions with F315 and P310, respectively.

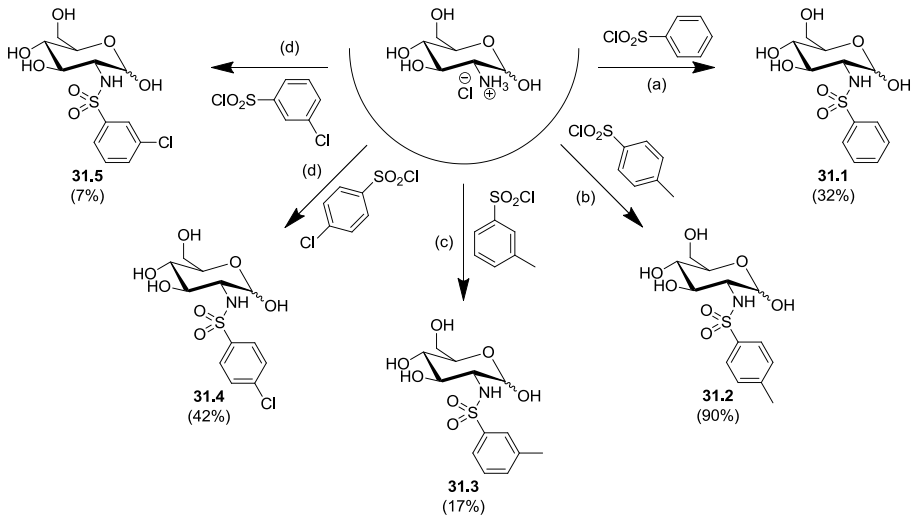
Table 14. Structure-activity relationship of GlcNAc derivatives.[33]

Name	Structure	K_I [mM]	Relative potency[a]
Glc		21±4	0.20
GlcNAc		4.1±0.7	1.0
GlcNS		1.4±0.2	2.9
GlcNAc-6-OS		0.6±0.1	6.8
GlcNS-6-OS		0.28±0.06	14

[a] The relative potency was calculated utilizing the K_I value determined for GlcNAc.

Accordingly, a focused library of GlcNS analogs **31** bearing differentially substituted phenyl rings was prepared for the characterization *via* the ^{19}F R_2-filtered NMR assay (Scheme 6). The synthesis involved the functionalization of glucosamine (GlcN) with commercially available sulfonyl chlorides at yields ranging from 7% to 90%. The varying yields were predominantly the result of reduced reactivities for some of the sulfonyl chlorides and the formation of side products *via* the hydroxyl groups of GlcN. The simplicity of implemented strategy considerably reduced the synthetic effort for library preparation. Moreover, as the formation of β-glucosides was avoided, GlcNS analogs will not display unfavorable contributions from substituents in C1 as observed for the focused library of ManNAc analogs **5** (Chapter 3.3.1).

[33] In part, these procedures were conducted by Julia Mastouri (Freie Universität Berlin) under my supervision.



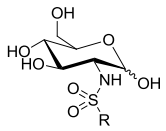
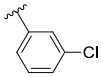
Scheme 6. Synthesis of GlcNS analog library 31.[34]

Reaction conditions for the preparation of library **31**: (a) Et_3N, anhydrous MeOH, 0°C to room temperature; (b) acetone, room temperature; (c) DIPEA, DMSO, room temperature; (d) DIPEA, anhydrous MeOH, room temperature.

The determination of K_I values revealed increased affinities over GlcNAc for all prepared analogs **31** (Figure 30d, Table 15, Appendix B.6). With a 13-fold affinity increase, **31.2** (K_I = 0.32±0.05 mM) represents the most potent member of the focused library. The analog bears a methyl group in para position of the phenyl ring, that does not form favorable interactions as exemplified by the affinity obtained for **31.1** (K_I = 0.37±0.04 mM) without substituents on the phenyl ring. The substitution of the methyl group with a chloride group for **31.4** (K_I = 0.6±0.2 mM) resulted in a decreased affinity presumably due to steric clashes, solvation effects or an alteration of the electronic structure of the phenyl ring. Similarly, increased K_I values were observed for the introduction of substituents in the meta position of the phenyl ring for **31.3** (K_I = 0.56±0.09 mM) and **31.5** (K_I = 0.59±0.06 mM).

[34] In part, the procedures were conducted by David Hartmann (University of Cambridge) under my supervision or by these Gunnar Bachem (Humboldt-Universität zu Berlin).

Table 15. Structure-activity relationship of GlcNS analog library 31.[35]

Name	Structure	R	K_I [mM]	Relative potency[a]
31.1			0.37±0.04	11
31.2			0.32±0.05	13
31.3			0.56±0.09	7.3
31.4			0.60±0.02	6.8
31.5			0.59±0.06	7.0

[a] The relative potency was calculated utilizing the K_I value determined for GlcNAc.

A detailed analysis of the binding mode of these GlcNS analogs is presented in the following section (Chapter 3.4.2). Yet, the obtained SAR indicates that the introduction of the phenyl ring results in the proposed aromatic interactions with K299, P310 or F315. Further design approaches might involve modifications in C6 *via* an additional sulfonamide linker to explore the carbohydrate binding site for favorable interactions. Although the introduction of a sulfate group in C6 for **31.2** would potentially results in a 5-fold additional affinity increase, the elaborate protecting group strategies required for the corresponding synthesis render this approach unfeasible (Noti *et al.*, 2006). Moreover, sulfate groups display unfavorable pharmacokinetic stability compared to sulfonamide groups. Alternatively, modifications in C1 of the Glc scaffold enable the integration of the SAR obtained for Man analogs **12** and **14**. Notably, the synthesis of α-glucosides for GlcNS is challenging and will considerably impede the preparation of focused libraries (Orgueira *et al.*, 2002). However, these suggestions critically depend on a conserved binding mode for the Glc scaffold and will be evaluated experimentally in the future.

In summary, the dissection of previously reported interactions between heparin and Langerin led to the identification of sulfated GlcNAc derivatives as favorable scaffolds for the design of glycomimetic ligands (Figure 30, Table 14). Based on the available X-ray structure of the Langerin complex with GlcNAc, a focused library of GlcNS analogs **31** was designed, synthesized and characterized in ^{19}F R_2-filtered NMR experiments (Figure 30, Table 15, Scheme 6). The introduction of phenyl substituents in C2 resulted in increased affinities for all library members suggesting the formation of

[35] In part, these procedures were conducted by David Hartmann (University of Cambridge) under my supervision.

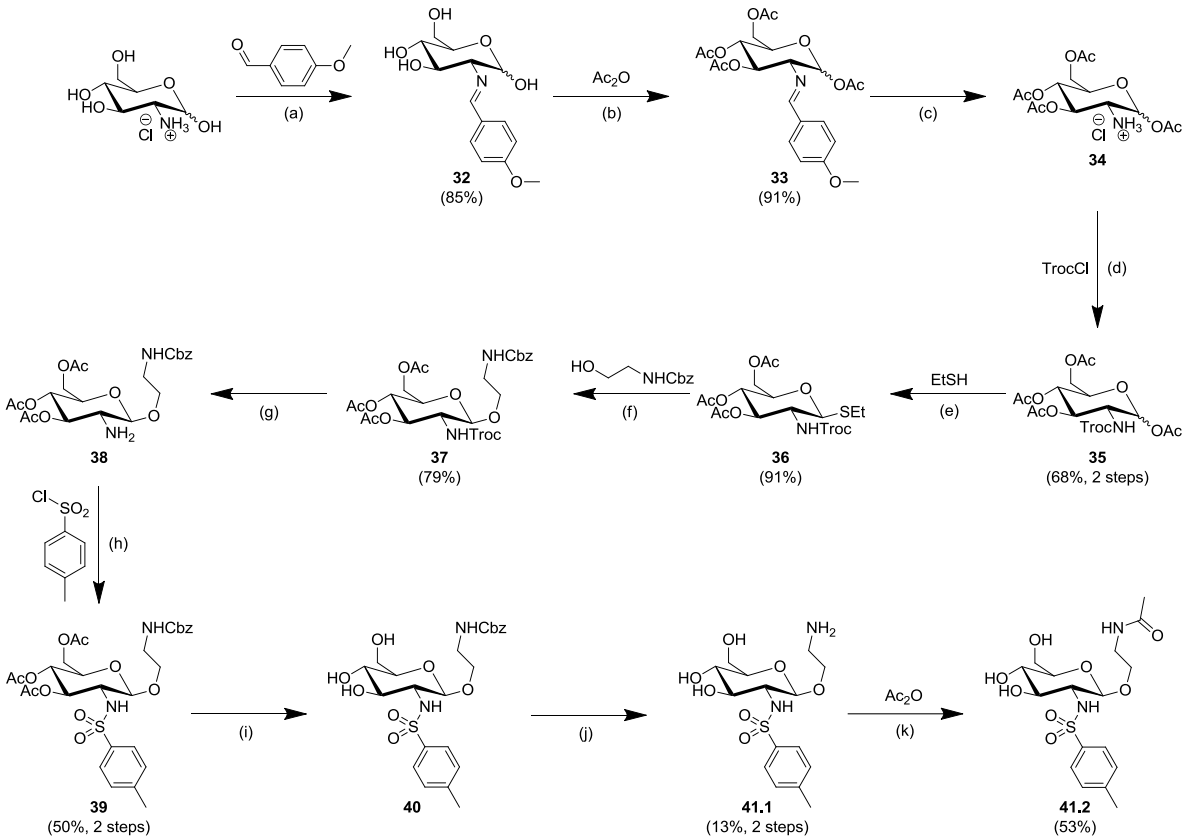
favorable aromatic interactions with *e.g.* K299, P310 or F315. With a 13-fold affinity increase over GlcNAc, **31.2** (K_I = 0.32±0.05 mM) represents another promising glycomimetic Langerin ligand for the development of targeted delivery applications for LCs. Importantly, the findings presented in this chapter are complementary to the design of analog **14.5**, rather than redundant (Chapter 3.3). For instance, different monosaccharide scaffolds will likely display differential specificity profiles and the availability of several ligand classes thus increases the probability to discover specific targeting ligands for Langerin.

3.4.2. Conjugation Strategy for Liposomal Formulations

Despite its low complexity, **31.2** displays an affinity comparable to that of the previously designed targeting ligand **14.5** which might be sufficient to efficiently target liposomes to LCs. This assumption is based on the observation that the blood group antigen Le^X (K_D = ca. 1 mM) has been successfully utilized to target liposomes to DCs to activate T cells *in vitro* (Fehres *et al.*, 2015c; Pederson *et al.*, 2014). Accordingly, advancing **31.2** towards targeted delivery applications was prioritized over further affinity optimization and a conjugation strategy for liposomal formulations was developed. In analogy to Man analogs **12.29** and **14.4**., this strategy involved the introduction of a ethylamino linker in C1 of the Glc scaffold to yield targeting ligand **41.1** *via* the formation of a β-glucoside.

41.1 was prepared over 10 steps at an overall yield of 2.5% in the laboratory of Prof. Oliver Seitz (Humboldt-Universität zu Berlin) and kindly provided for conjugation to NHS-activated lipids. The synthesis involved the orthogonal protection of the amino group in C2 utilizing a Troc group, the subsequent formation of a β-thioglucoside and the introduction of the Cbz-protected ethylamino linker in C1. The selective Zn-mediated removal of the Troc group was followed by the introduction of the phenyl substituent in C2 *via* the formation of a sulfonamide. Finally, global deprotection yielded **41.1** bearing an amino group in C1 that enables its conjugation to NHS-activated lipids. Importantly, **41.1** was prepared in sufficient quantities for several liposomal formulations. Furthermore, GlcNS analog **41.2** was synthesized as a model molecule to analyze the SAR for the developed conjugation strategy *via* NMR experiments. The synthesis involved the selective acetylation of the amino group at a yield of 53%.

The K_I value of **41.2** was determined in ^{19}F R_2-filtered NMR experiments (Figure 31a). The comparison with previously obtained affinities for Man, Man analog **12.30** and **31.2** revealed the SAR for the proposed conjugation strategy of GlcNS analogs (Table 16). Strikingly, the introduction of an ethylamino linker *via* the formation of a β-glucoside results in a minor affinity increase for **41.2** (K_I = 0.24±0.03) compared to **31.2** (K_I = 0.32±0.05). This contrasts the affinity decrease observed for the formation of α-mannosides as exemplified by Man analog **12.30** (K_I = 10±1 mM)**.** Thus, **41.2** displays a 42-fold affinity increase over reference molecule **12.30** with both molecules serving as a model for liposomal formulations.



Scheme 7. Synthesis of GlcNS analogs 41.1 and 41.2.[36]

Intermediate **36** was prepared as previously published.(Ellervik and Magnusson, 1996) Reaction conditions for the preparation of **41.2**: (a) 1 M aqueous NaOH, 0°C; (b) pyridine, 0°C to room temperature; (c) 5 M aqueous HCl, acetone, reflux; (d) pyridine, 0°C; (e) $BF_3.OEt_2$, anhydrous DCM, 0°C to room temperature; (f) DMTST, anhydrous DCM, room temperature; (g) Zn, AcOH, room temperature; (h) pyridine, room temperature; (i) MeONa, anhydrous MeOH, room temperature; (j) H_2, Pd/C, EtOH, room temperature; (k) MeOH, room temperature.

[36] The procedure was devised by Gunnar Bachem (Humboldt-Universität zu Berlin) and me and conducted by Gunnar Bachem.

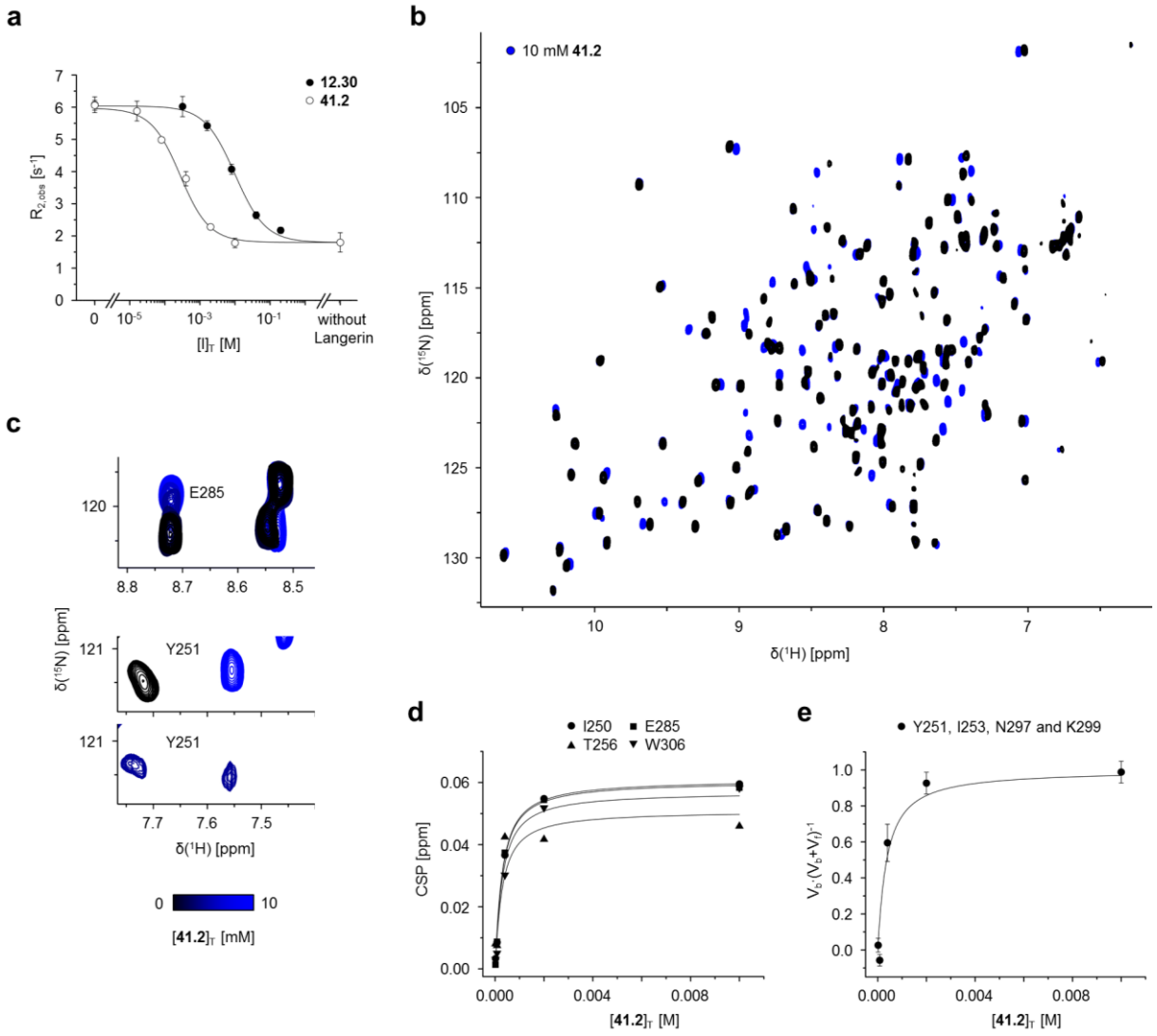
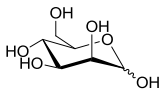
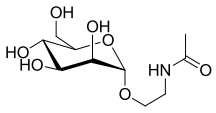
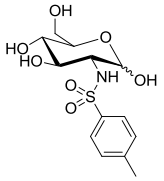
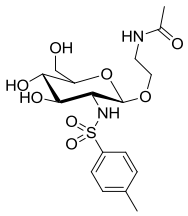


Figure 31. K_I and K_D determination for GlcNS analog 41.2.

a. ^{19}F R_2-filtered NMR experiments with the Langerin CRD served to determine the K_I value of **41.2** (K_I = 0.24±0.03 mM). The analog represents a model structures synthesized to evaluate the conjugation strategy for liposomal formulations. The binding isotherm for titrations with Man analog **12.30** is depicted for comparison. **b. and c.** ^{15}N HSQC NMR experiments served to validate the obtained K_I value for **41.2**. Assigned resonances displaying fast chemical exchange and CSPs larger than 0.04 ppm were selected for the determination of K_D values. Additionally, a set of residues including Y251, I253, N297 and K299 displayed slow exchange phenomena. For these residues, integrals V_f and V_b of resonances corresponding to the free and the bound state of the Langerin were utilized to determine K_D values. Representative CSP trajectories are depicted for both approaches. **d** and **e.** Both, K_D values determined *via* the fast exchange approach (K_D = 0.23±0.07 mM) and the slow exchange approach ($K_{D,slow}$ = 0.3±0.1 mM) are consistent with ^{19}F R_2-filtered NMR experiments and validate **41.2** as a potent glycomimetic Langerin ligand.

Table 16. Structure-activity relationship for the conjugation of GlcNS analog 41.1.

Name	Structure	K_I [mM]	K_D [mM]	Relative potency[a]
Man		4.5±0.5[b]	5.8±0.3	2.2
12.30		10±1	12±1	1.0
31.2		0.32±0.05	0.46±0.04 0.5±0.2[c]	31
41.2		0.24±0.03	0.23±0.07 0.3±0.1[c]	42

[a] The relative potency was calculated utilizing the K_I value determined for **12.30**.
[b] This value was determined from three independent titration experiments.
[c] This value was determined from integrals V of resonances observed to be in slow exchange.

Next, the K_I values for **31.2** (K_D = 0.46±0.04 mM) and **41.2** (K_D = 0.23±03 mM) mM were confirmed in orthogonal ^{15}N HSQC NMR experiments (Figure 31b to e, Table 16, Appendix B.7). As the experiments were conducted in absence of DMSO, a higher amplitude and number of CSPs were observed compared to titrations with *e.g.* **14.5**. However, a considerable fraction of the resonances displaying CSPs additionally displayed line broadening $\Delta\nu_{0.5}$ of more than 10 Hz, indicative of intermediate exchange phenomena. Accordingly, these resonances were not considered for K_D determination. Simultaneously, slow exchange phenomena were observed for a set of resonances corresponding to Y251, I253, N297 and K299 (Figure 31c and e, Appendix B.7). For these residues, the integral V_f of the free state of the Langerin CRD was reduced upon binding of **31.2** or **41.2** and a proximal resonance was assigned to the bound state. Integrals V_f and V_b served to quantify the bound fraction of the receptor p_b and to determine K_D values for both **31.2** ($K_{D,slow}$ = 0.5±0.2 mM) and **41.2** ($K_{D,slow}$ = 0.3±0.1 mM). Typically, additional experiments such as EXSY NMR are required to provide rigorous evidence for chemical exchange between receptor resonances (Demers and Mittermaier, 2009). However, as the determined affinities were consistent with both ^{19}F R_2-filtered NMR experiments and K_D values obtained from CSP values, the assignment of bound state resonances was assumed to be correct. Overall, the ^{15}N HSQC NMR experiments with **41.2** validate the 42-fold affinity increase over reference molecule **12.30**.

The obtained SAR indicated that the introduction of an ethylamino linker in C1 represents a favorable conjugation strategy for GlcNS analogs. In contrast to the unfavorable contributions for α-mannosides, no affinity decrease was observed for the formation β-glucosides for GlcNS analogs. To further investigate the binding mode of targeting ligand **41.2**, ^{15}N HSQC and STD NMR experiments were combined with molecular docking studies (Figure 32 and Figure 33). Here, the orientation of the linker was of particular interest to evaluate the compatibility of the binding mode with the presentation of **41.2** on liposomes.

^{15}N HSQC NMR experiments with GlcNS analog **41.2** revealed globally increased CSP values compared to titrations with Man and Man analog **12.30** (Figure 32a). This might be attributed to stronger changes in chemical environment of backbone resonances induced either directly *via* the formation protein-ligand interactions or indirectly *via* conformational dynamics. Yet, the evaluation of the CSP patterns identified common motives for the recognition of the GlcNS analog and the Man scaffold (Figure 32b and c). **41.2** induced CSPs for E285 and K299 providing evidence for a Ca^{2+}-dependent binding mode. Moreover, high CSP values were observed for other residues associated with the recognition of the Man scaffold including N297, A300 and S302. By contrast, Y251 and I250 displayed considerably increased CSP values compared to Man analog **12.30**, while a relative decrease was observed for K313. Simultaneously, the proximal T314 displayed an increased CSP value. Additional relative increases in CSP values were observed for W252 and W306. These distinctions were accompanied by considerable changes in the CSP trajectories, *i.e.* the vector of induced perturbations, of residues associated with the carbohydrate binding site including E285 and W252 (Figure 32d). As observed for other glycomimetic Langerin ligands, **41.2** additionally induced CSPs associated with the allosteric network in remote regions of the CTLD fold, prominently K257 and G259 (Hanske *et al.*, 2016). For these residues CSP trajectories were conserved compared to titrations with Man analog **12.30**.

The ^{15}N HSQC NMR binding mode analysis revealed unique features for the recognition of the Glc scaffold. The differences observed in comparison to Man analog **12.30** might be attributed to the altered stereochemistry in C2 as well as potential interactions formed by the sulfonamide group and the phenyl ring of **41.2**. Interestingly, residues displaying increased CSP values are predominantly located in proximity to F315, which was originally envisioned to form interactions with the phenyl ring. As the CSP value for K313 is decreased compared to titrations with Man analog **12.30**, this observation indicates the phenyl ring might indeed be oriented towards F315 and K299 rather than K313 or P310.

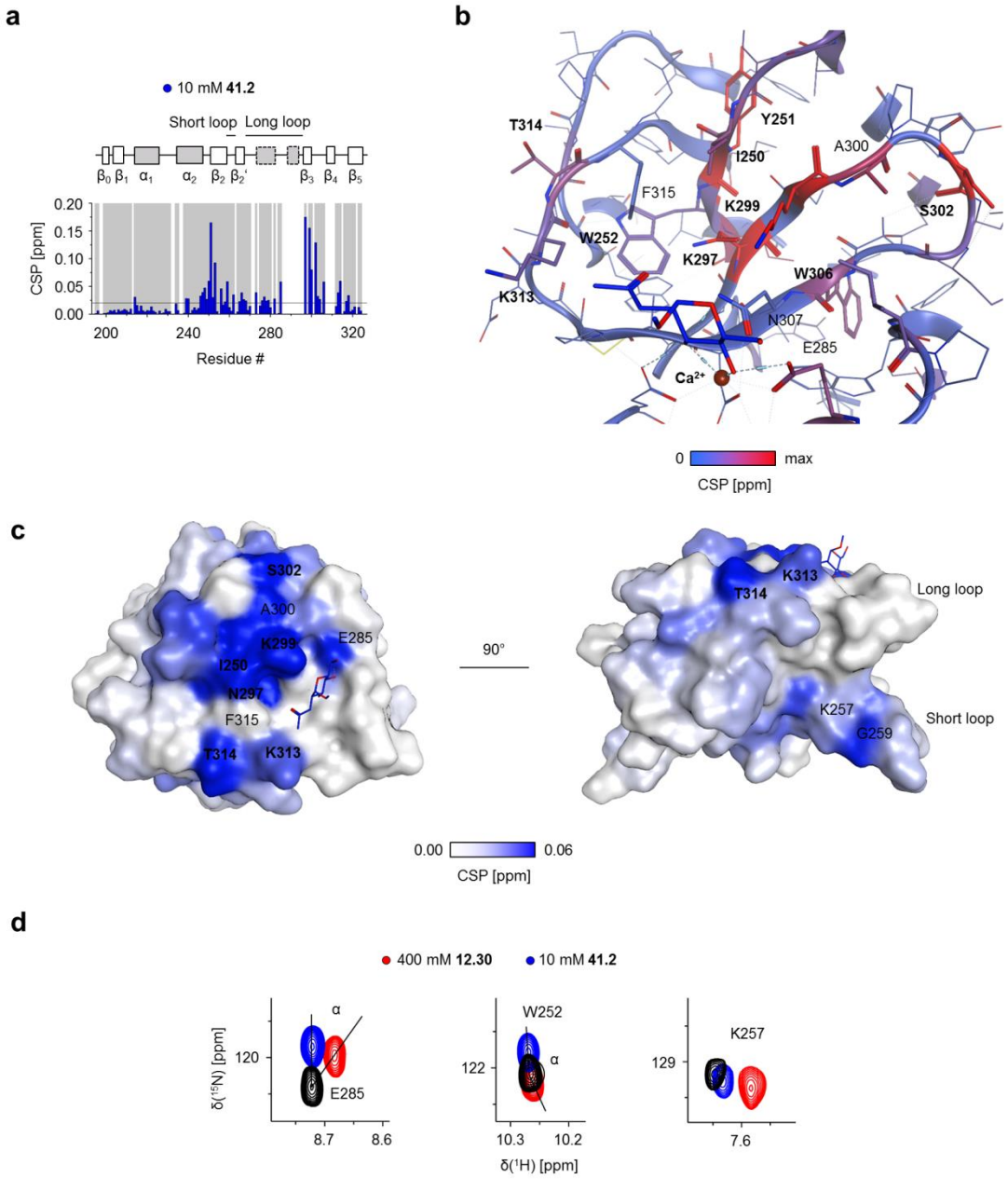


Figure 32. ^{15}N HSQC NMR binding mode analysis for GlcNS analog 41.2.

a. ^{15}N HSQC NMR experiments with the Langerin CRD revealed the CSP pattern for **41.2**. Assigned resonances detected in the reference spectrum are highlighted (grey). Observed CSP values for **42.1** displayed globally increased amplitudes when compared to titrations with Man and Man analogs. **b and c.** Mapping the CSPs on the X-ray structure of the Langerin CRD (PDB code: 4N32) validated a Ca^{2+}-dependent binding mode as indicated by CSPs observed for E285 and K299 (Feinberg *et al.*, 2013). Additionally, CSPs were observed for N297, A300 and S302 residues also affected upon recognition of Man or **12.30** (Chapter 3.2.1 and 3.3.3). By contrast, Y251 and I250 displayed considerably increased CSP values compared to **12.30**, while a relative decrease was observed for K313. This decrease was accompanied by a relative increase for the proximal T314. Notably, residues that display considerably increased CSP values can predominantly associated with F315 which was not assigned. This also hold true for W252 and W306 that displayed smaller relative increases. Accordingly, the observed CSP pattern might be induced by interactions formed between **41.2** and F313 rather than K313. Similar to Man and Man analog **12.30**, CSPs were also observed in remote regions of the CTLD fold, particularly for K257 and G259 in the short loop region. This might indicate a modulation of the previously reported allosteric network (Hanske *et al.*, 2016). **d.** A comparison of titrations with Man analog **12.30** and **41.2** revealed distinct CSP trajectories for residues associated with the carbohydrate binding site such as E285 or W252 while trajectories of residues located in remote regions of the CTLD fold were conserved.

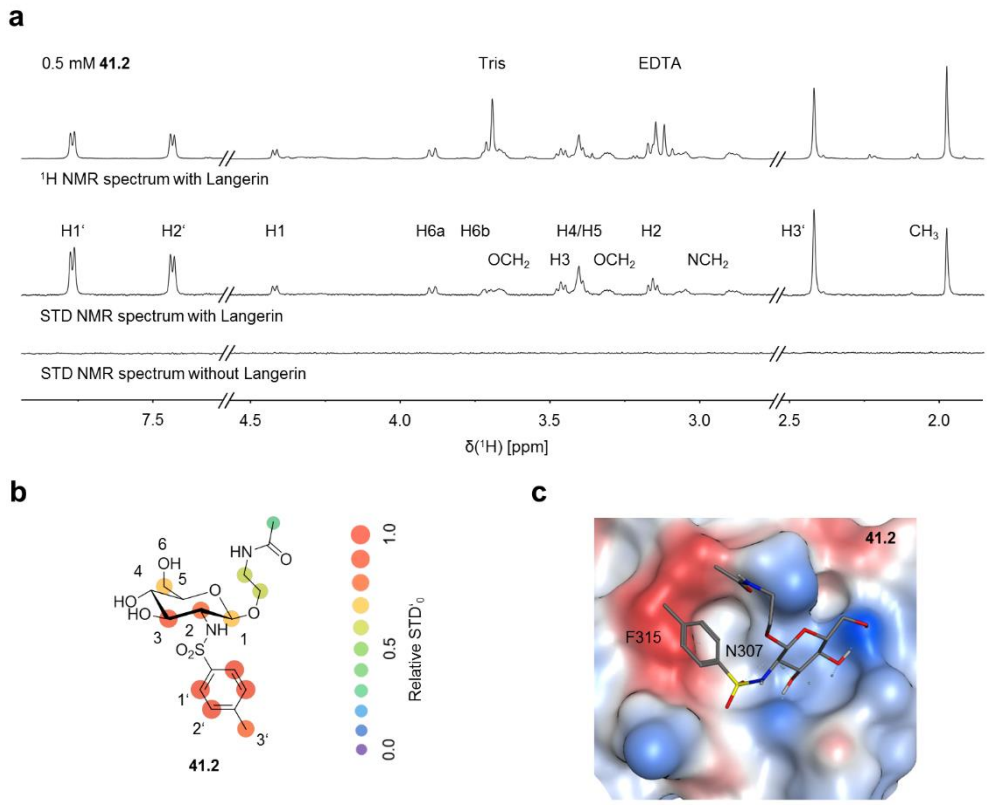


Figure 33. STD NMR epitope mapping and docking pose for GlcNS analog 41.2.

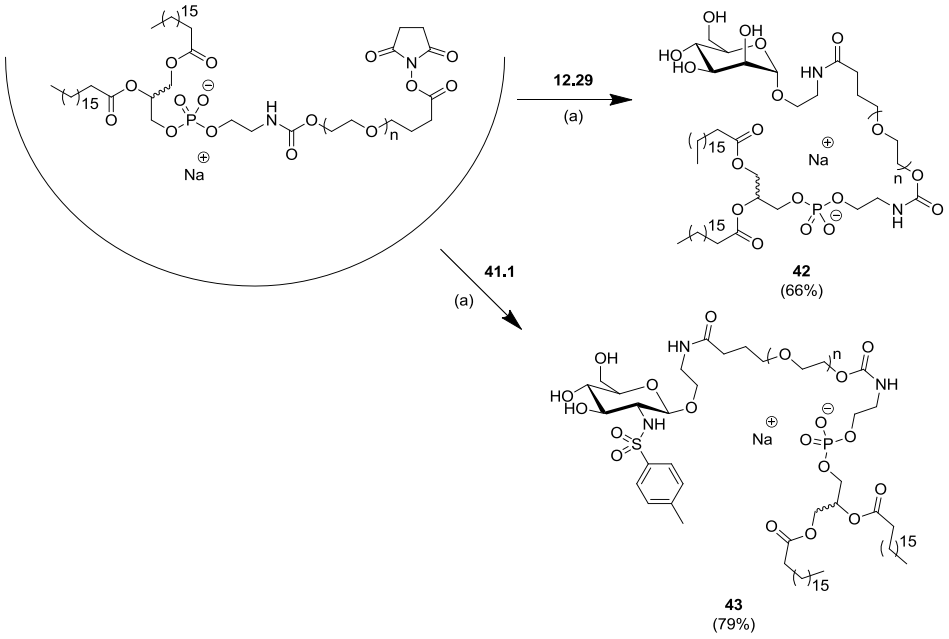
a. and b. STD NMR experiments served to further validate the interactions formed by **41.2** with the Langerin ECD. STD NMR spectra were recorded at saturation times t_{sat} of 0.4 s and are magnified 8-fold. **c.** The STD NMR epitopes for **41.2** was determined from build-up curves (Appendix B.8). The highest STD effects are observed for the phenyl substituent in C2 including both aromatic hydrogens and the methyl group in para position, suggesting the formation of interactions with the Langerin surface. While a quantitative analysis for H4 and H5 was impeded by spectral overlap and strong J coupling, other hydrogens of the Glc scaffolds also display high relative STD'$_0$ values. In contrast, low relative STD'$_0$ for the acetylated ethylamino linker indicate a solvent exposed orientation. **c. 41.2** was docked into the carbohydrate binding site to rationalize the observations from ^{15}N HSQC and STD NMR experiments. The selected docking pose for **41.2** predicts the formation of π-π interactions between the phenyl ring and F315 as well as the formation of a hydrogen bond between the sulfonamide group and N307. The linker displays high solvent exposure. The receptor surface is colored according to its lipophilicity (lipophilic: red, hydrophilic: blue)

To complement the observations from ^{15}N HSQC NMR experiments and to analyze the orientation of the acetylated ethylamino linker, STD NMR epitope mapping with **41.2** was conducted (Figure 33). As discussed in the previous sections, all Langerin samples contained minor EDTA impurities that could not be removed after repeated dialysis against H_2O. These impurities were assumed to not interfere with the implemented experiments. **41.2** displays specific STD effects in presence of Langerin, further confirming the interactions formed by this GlcNS analog (Figure 33a). Notably, spectral overlap and strong J coupling impeded the determination of STD'$_0$ values for H4 and H5 while the corresponding multiplet displays a substantial STD effect. Accordingly, an analysis of the GlcNS binding mode was compromised, particularly considering the similar STD'$_0$ values observed for H1, H2, H3 and H6 (Figure 33b).

Overall, the binding epitope of **41.2** is dominated by uniformly high STD effects for the phenyl substituent in C2 of the Glc scaffold (Figure 33b). Comparable STD'$_0$ values were determined for the

aromatic hydrogens and the methyl group in the para position. This observation corroborated the suggested formation of favorable interactions by the phenyl substituent with the Langerin surface. The acetylated ethylamino linker did, by contrast, display uniformly low STD effects indicating a solvent exposed orientation and validating the developed conjugation strategy for GlcNS analogs.

Molecular docking studies served to visualize potential interactions formed by **41.2** and to interpret the findings from ^{15}N HSQC and STD NMR experiments (Figure 33c). The procedure was conducted in analogy to Man analog **14.5**, utilizing the X-ray structure of the Langerin complex with GlcNAc (PDB code: 4N32) (Feinberg *et al.*, 2013). Overall, ten docking poses were generated and evaluated visually in the context of previously proposed binding hypotheses. Five of these docking poses predicted the orientation of the phenyl substituent in C2 in direction of F315 and K299 while an orientation towards K313 was observed for two poses. Calculated GBVI/WSA ΔG scores for these conformations were comparable and a representative docking pose was selected to visualize the formation of potential interactions (Figure 33c). Indeed, orienting the phenyl ring towards F315 and K299 resulted in the formation of π-π interactions. Interestingly, this orientation also coincided with the formation of a hydrogen bond between the sulfonamide group and N307. These observations are consistent with ^{15}N HSQC NMR experiments for **41.2**. First, the most pronounced CSP values are observed for residues that are associated with F315 including I250, Y251, N297 and K299. Simultaneously, the relative amplitude for K313 is decreased, when compared to experiments with Man analog **12.30**. Additionally, relative increases for CSP values in this region of the CRD are observed for T315 and W252, both residues are associated with F315. Notably, the generated docking pose for ManNAc analog **5.11** also predicted the formation a hydrogen bond with N307 (Chapter 3.3.1). While the resonance corresponding to N307 was not assigned, ^{15}N HSQC NMR experiments revealed a considerably increased CSP value for the adjacent W306. While less pronounced, this increase is also observed for titrations with **41.2**. Furthermore, for nine of the ten generated docking poses the acetylated ethylamino linker displayed high solvent exposure and no conserved interactions was observed. This orientation is consistent with the low STD'$_0$ values determined in STD NMR experiments.



Scheme 8. Synthesis of glycolipids 42 and 43.[37]

Reaction conditions for the preparation of **42** and **43**: (a) DMF:1M aqueous $NaHCO_3$ (1:10), pH 8.4, room temperature.

Whereas the predicted docking pose was corroborated by several experimental observations, the model should be treated cautiously. As discussed for Man analog **14.5**, the conformational flexibility of the carbohydrate binding site and a non-conserved orientation of the Glc scaffold might result in an alternative binding mode for **41.2**. These changes might not be detectable by ^{15}N HSQC or STD NMR experiments and are not explicitly accounted for by the implemented *in silico* approach. Moreover, structural H_2O molecules play an essential role for the recognition of GlcNAc (Feinberg *et al.*, 2013). Yet, H_2O molecules were removed prior to molecular docking. CORCEMA-ST and HADDOCK approaches potentially provide improved spatial resolution for the utilized NMR techniques (Dominguez *et al.*, 2003; Jayalakshmi and Rama Krishna, 2004). Additionally, trNOE NMR experiments enable the evaluation of the orientation of the phenyl substituent in the bound state of **41.2** (Fiege *et al.*, 2012; Munoz-Garcia *et al.*, 2015). Finally, ensemble docking represents a powerful approach to account for the conformational flexibility of the carbohydrate binding site (Totrov and Abagyan, 2008). While these techniques potentially provide valuable structural information for the continued optimization of GlcNS analogs, the elucidation of an X-ray structure of the Langerin complex with **41.2** will ultimately be required.

[37] These procedures were devised by Jessica Schulze (Max Planck Institute of Colloids and Interfaces) and me and conducted by Jessica Schulze.

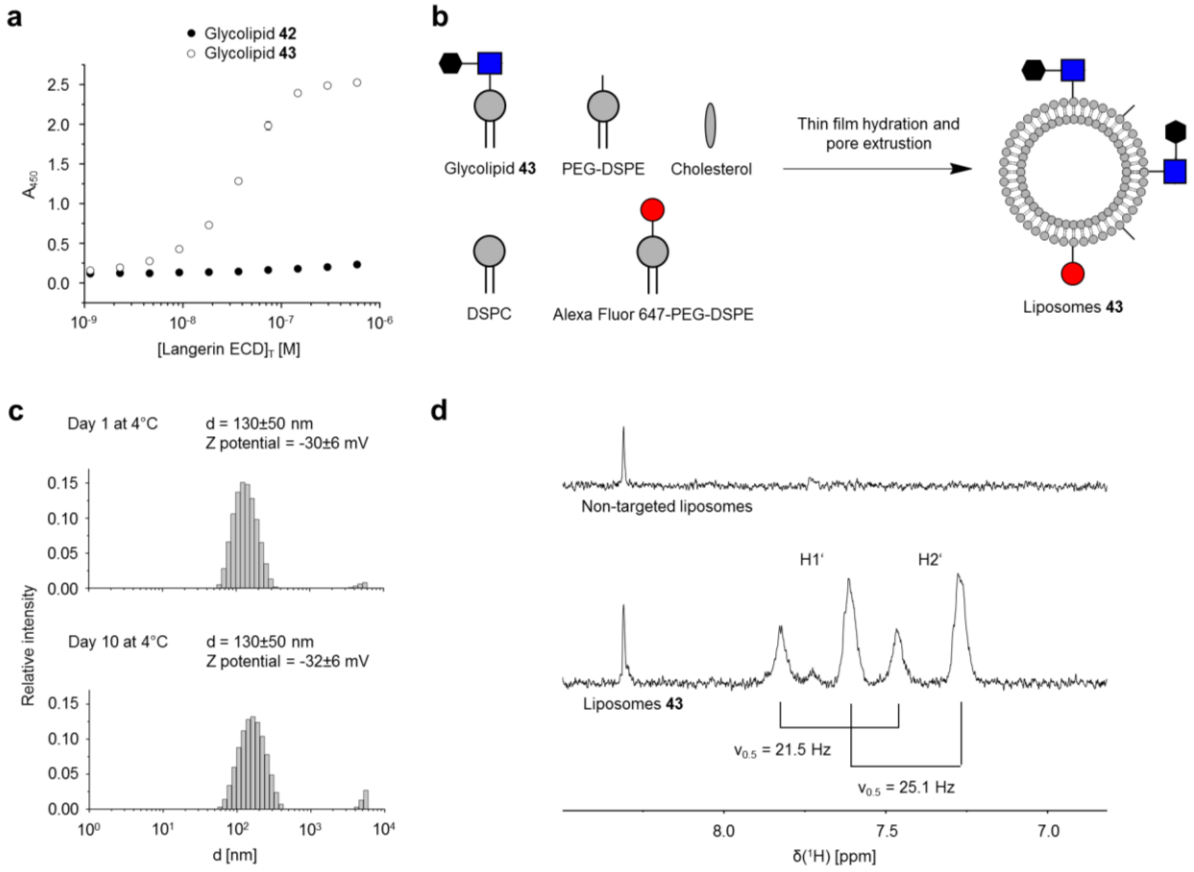


Figure 34. Characterization of glycolipids 42 and 43 and liposomal formulations.[38]

a. ELLA experiments served to evaluate the affinity of glycolipids **42** and **43** for the Langerin ECD. No binding was detected for glycolipid **42**, bearing Man analog **12.29**. In contrast, glycolipid **43**, bearing GlcNS analog **41.1** display a dose-dependent response validating the expected affinity increase over reference molecule **42**. **b.** Targeted liposomes were prepared *via* thin film hydration and subsequent pore extrusion as previously reported (Rillahan *et al.*, 2012). Liposomal formulation contained 57% DSPC, 38% Cholesterol, 0.25% Alexa Fluro 647-PEG-DSPE and up to 4.75% of glycolipid **43**. Liposomes for glycolipid **42** were prepared analogously. **c.** Dispersity and stability of the prepared liposomes were analyzed *via* DLS and electrophoresis experiments. The liposomes are monodisperse (d = 130±50 nm) and stable (Z potential = -32±6 mV) for at least ten days when stored at 4°C in PBS. **d.** ^{1}H NMR experiments with liposomes **43** revealed two states for the GlcNS analogs **41.1** as observed for resonances corresponding to H1' and H2' of the phenyl ring. Both states display a linewidth $\nu_{0.5}$ smaller than 30 Hz, indicating accessibly and residual flexibility due to their presentation on extended PEG chains. The alternative state potentially corresponds to targeting ligands oriented towards the lumen of the liposomes.

The obtained SAR, conducted NMR experiments and molecular docking studies characterized the introduction of the ethylamino linker in C1 of the Glc scaffolds as a suitable conjugation strategy for liposomal formulations. Following this strategy, Man analog **12.29** and **41.1** were conjugated to NHS-activated lipids to synthesize glycolipids **42** and **43** (Scheme 8). This procedure as well as the formulation of liposomes and their characterization *via* DLS and electrophoresis experiments were conducted by Jessica Schulze (Max Planck Institute of Colloids and Interfaces). The affinity of these glycolipids for Langerin was evaluated by Dr. Jonas Aretz (Max Planck Institute of Colloids and Interfaces) in a previously reported plate-based ELLA (Figure 34a) (Aretz, 2017; Fehres *et al.*, 2015c).

[38] The ELLA was conducted by Felix Fuchsberger (University of Utrecht) under supervision by Dr. Jonas Aretz (Max Planck Institute of Colloids and Interfaces). The formulation of liposomes as well as their characterization *via* DLS and electrophoresis experiments was conducted by Jessica Schulze (Max Planck Institute of Colloids and Interfaces).

The assay utilizes the immobilization of glycolipids to achieve multivalent presentation to the trimeric Langerin ECD. Subsequently, Langerin binding is detected *via* an HRP-dependent colorimetric reaction. While a dose-dependent interaction with Langerin was observed for glycolipid **43**, bearing GlcNS analog **41.1**, no interaction was detected for the immobilization of glycolipid **42**. On the one hand, this validates the expected affinity increase for **43** over reference molecule **42**. On the other hand, the absence of binding for **42** might be attributed to the limited sensitivity of the conducted ELLA.

Next, glycolipids **42** and **43** were utilized to prepare targeted liposomes *via* thin film hydration and subsequent pore extrusion as previously reported (Chen *et al.*, 2010b) (Figure 34a). The dispersity and stability of the prepared liposomes were evaluated by DLS and electrophoresis experiments (Figure 34b). In these experiments, liposomes **43** were monodisperse (d = 130±50 nm) and stable (Z potential = -32±6 mV) for at least ten days when stored at 4°C in PBS. Finally, ^{1}H NMR experiments served to characterize the properties of GlcNS analog **41.1** on the surface of the liposomes. Interestingly, two states were observed for the resonances corresponding to H1' and H2' of the phenyl ring. Both states displayed linewidths $\nu_{0.5}$ smaller than 30 Hz suggesting residual flexibility due the presentation of the targeting ligand on extended PEG chains. Hence, GlcNS analog **41.1** is likely accessible for interactions with Langerin in the context of targeted delivery applications. The alternative state might correspond to targeting ligands oriented towards the lumen of the liposomes, which would reduce the effective concentration of accessible GlcNS analog **41.1**. Notably, liposomal formulations with Man analog **14.4** potentially enable more elaborate ^{19}F NMR experiments *via* its trifluoromethyl group as background signals from *e.g.* PEG or Langerin will not interfere with data analysis.

In summary, the presented findings demonstrate that the introduction of an ethylamino linker in C1 of the Glc scaffolds represents a suitable conjugation strategy for liposomal formulations. In contrast to Man analog **12.30**, the formation of the corresponding β-glucoside resulted in an additional, yet minor, affinity increase for model molecule **41.2** (Figure 31, Table 16, Scheme 7). ^{15}N HSQC and STD NMR experiments were combined with molecular docking studies to investigate the binding mode of **14.5** and to rationalize the observed 42-fold affinity increase (Figure 32 and Figure 33). These investigations suggest the formation of π-π interactions with F315 and a hydrogen bond between the sulfonamide group and N307. Moreover, the acetylated ethylamino linker likely displays a solvent exposed orientation and will thus enable a favorable presentation of targeting ligand **41.1** on liposomes. Next, reference molecule **12.29** and **41.1** were conjugated to NHS-activated lipids to synthesize glycolipids **42** and **43**, respectively (Scheme 8). The affinity of **42**, bearing GlcNS analog **41.1**, for Langerin was validated in ELLA experiments and both glycolipids were subsequently utilized for liposomal formulations (Figure 34a). Prepared liposomes were demonstrated to be monodisperse and stable for at least ten days (Figure 34b and c). Finally, ^{1}H NMR experiments served to validate the accessibility of **41.1** on the liposomal surface (Figure 34d).

Overall, **41.1** represents a promising targeting ligand for LCs. It displays submillimolar affinity (K_I = 0.24±0.03 mM) exceeding that of naturally occurring carbohydrates. Hence, **41.1** likely meets the affinity requirements for efficient endocytosis by LCs. This assumption is corroborated by the capacity of the Le^X trisaccharide (K_D = ca. 1 mM) to target liposomes to DCs and activate T cells *in vitro* (Fehres *et al.*, 2015c; Pederson *et al.*, 2014). Importantly, a favorable specificity profile against other GBPs is essential for the applicability of **41.1** to targeted antigen delivery. With the prepared targeted liposomes at hand, this aspect was investigated in a cellular context both *in vitro* as well as *ex vivo* utilizing human epidermal and whole skin cell suspensions (Chapter 3.4.3).

3.4.3. Specificity against DC-SIGN and *ex vivo* Targeting of Langerhans Cells

The specificity of glycomimetic targeting ligands for CLRs has important implications for both therapeutic applications and basic research. When utilizing these ligands to *e.g.* deliver antigens to an individual subset of DCs, off-target affinity might reduce the efficiency of the approach and induce adverse effects. The endocytosis or adsorption of targeted liposomes mediated by other GBPs will reduce the effective concentration of delivered antigens. For example, ASGPR, a CLR expressed on hepatocytes, is of particular relevance for systemic administrations as it will promote the clearance of targeted liposomes (D'Souza and Devarajan, 2015). Similarly, the soluble CLR MBP-C present in human blood has been implicated in the Man-type glycan-dependent clearance of pathogens and self-antigens (Takahashi *et al.*, 2006). Moreover, off-target DC subsets potentially induce a tolerogenic response to the delivered antigen (Clausen and Stoitzner, 2015). Generally, the underlying mechanisms governing the endocytosis, processing and cross-presentation of antigens by DCs to efficiently induce immunogenic T cell responses remain largely elusive (Fehres *et al.*, 2014b). In this context, specific molecular probes are required to dissect contributions from individual CLRs or DC subsets.

The human epidermis consists of melanocytes and keratinocytes as well as a limited population of immune cells (Pasparakis *et al.*, 2014). This immune cell population is dominated by LCs and additionally comprises resident memory T cells and cytotoxic T cells. Upon epidermal administration, targeted liposomes will thus encounter a limited set of potential off-target GBPs and no APCs other than LCs. Notably, keratinocytes have been reported to expresses the MR, a CLR recognizing Man-type glycans (Szolnoky *et al.*, 2001). By contrast, several types of APCs including dermal DCs, macrophages and monocytes are present in the human dermis (Pasparakis *et al.*, 2014). Consequently, specificity against a considerable number of GBPs including CLRs such as DC-SIGN, Dectin-1 or potentially Mincle will be required for targeted liposomes administered to the dermis to avoid off-target effects (Kostarnoy *et al.*, 2017; Stoitzner *et al.*, 2014).

Prior to the evaluation of the targeted liposomes in a cellular context, the specificity of both **12.30** and **41.2** against DC-SIGN was quantified in ^{19}F R_2-filtered NMR experiments (Figure 35, Table 17, Appendix B.9). Strikingly, **41.2** (K_I = 15±3 mM) displayed a considerably decreased K_I compared to Langerin corresponding to 63-fold specificity. For both targeting ligands, the introduction of the acetylated ethylamino linker in C1 results in a minor affinity increase. This results in 3.7-fold specificity for DC-SIGN for Man analog **12.30** (K_I = 2.7±0.3 mM).

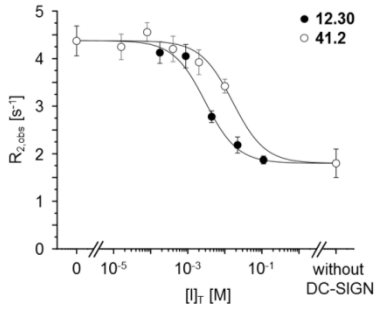
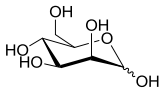
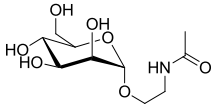
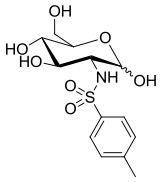
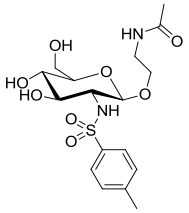


Figure 35. K_I determination for Man analog 12.30 and GlcNS analog 41.2 with DC-SIGN.

^{19}F R_2-filtered NMR experiments reveal the specificity of **12.30** (K_I = 2.7±0.3 mM) and **41.2** (K_I = 15±3 mM) against DC-SIGN.

Encouraged by these findings, targeted liposomes **42** and **43** were evaluated in *in vitro* flow cytometry experiments with human Raji cells (Figure 36a). The cells were stably transfected to express Langerin, DC-SIGN or Dectin-1 and kindly provided by Prof. Nina van Sorge (University of Utrecht). The experiments were conducted by Jessica Schulze (Max Planck Institute of Colloids and Interfaces). Briefly, cells were cultured in multi-well plates, incubated with liposomes for 1 h at 4°C and resuspended for analysis *via* flow cytometry. Initial titration experiments revealed dose-dependent binding of liposome**s 43**, bearing GlcNS analog **41.1**, to Langerin$^+$ cells (Figure 36b). Unspecific effects were negligible and observed MFI values depended on the fraction of glycolipid **43** utilized for liposomal formulations (Figure 36c). Liposome binding *via* the Ca^{2+}-dependent carbohydrate binding site of Langerin was further confirmed by the addition of EDTA and mannan (Figure 36d). Analogously, liposomes **42**, bearing Man analog **12.29**, were observed to interact with DC-SIGN$^+$ Raji cells (Figure 36e). Strikingly, binding of these liposomes was not detected for Langerin$^+$ or Dectin-1$^+$ cells, suggesting that the determined 3.7-fold affinity increase is sufficient to convey specificity in a cellular context (Figure 36f). An increased expression level and the formation of tetramers by DC-SIGN are potentially causative for these findings as both parameters amplify avidity effects. Additionally, DC-SIGN has been reported to form membrane-associated nanoclusters for DCs further promoting multivalent ligand recognition (Cambi *et al.*, 2004). Previously. the blood group antigen Le^Y was explored as a targeting ligand for LCs and dermal DCs (Fehres *et al.*, 2015c). While DC-SIGN and Langerin displayed comparable affinities for Le^Y in plate-based assays, the corresponding liposomes failed to target LCs. Overall these findings suggest that increased affinities are required to target liposomes to Langerin$^+$ cells. Importantly, liposomes **43** were demonstrated to be specific for Langerin$^+$ Raji cells (Figure 36f). While specificity was expected for DC-SIGN based on ^{19}F R_2-filtered NMR experiments, Dectin-1 has been reported to recognize Glc-type glycans (Drummond and Brown, 2011). As both CLRs are expressed in the dermis of the human skin, these findings further corroborate the suitability of GlcNS analog **41.1** for specific LC targeting.

Table 17. Structure-activity relationship for the conjugation of Man analog 12.29 and GlcNS analog 41.1 with DC-SIGN.

Name	Structure	K_I [mM]	Relative potency[a]	Specificity for Langerin[b]
Man		3.0±0.3	0.90	0.67
12.30		2.7±0.3	1.0	0.27
31.2		17±1	0.16	53
41.2		15±3	0.18	63

[a] The relative potency was calculated utilizing the K_I value determined for **12.30**.
[b] The specificity for Langerin was calculated from the corresponding K_I values listed in Table 16.

To advance liposomes **43**, bearing GlcNS analog **41.1**, towards targeted delivery applications, the prepared liposomes were transferred to the laboratory of Assoc. Prof. Patrizia Stoitzner (Medizinische Universität Innsbruck). The specificity for and endocytosis by human LCs were investigated in *ex vivo* flow cytometry experiments with epidermal and whole skin cell suspensions (Figure 37a). Cell suspensions were prepared as previously published (Ortner *et al.*, 2017). Briefly, skin samples were collected from healthy humans and subcutaneous fat was removed mechanically. For epidermal cell suspensions, the samples were subsequently incubated with dispase II and trypsin, followed by the mechanical removal and homogenization of the epidermis. Whole skin cell suspensions were obtained after incubation with collagenase IV and mechanical homogenization of the samples. Unless indicated otherwise, the cell suspensions were incubated with liposomes **42** and **43** as well as non-targeted liposomes for 1 h at 37°C and resuspended for flow cytometry experiments. Gating strategies were implemented as previously described and served to discriminate APCs such as LCs, dermal DCs, macrophages and monocytes from other cell populations found in the human skin (Ober-Blobaum *et al.*, 2017; Stoitzner *et al.*, 2014). LCs and DCs isolated by the described procedure are generally considered to be immature.

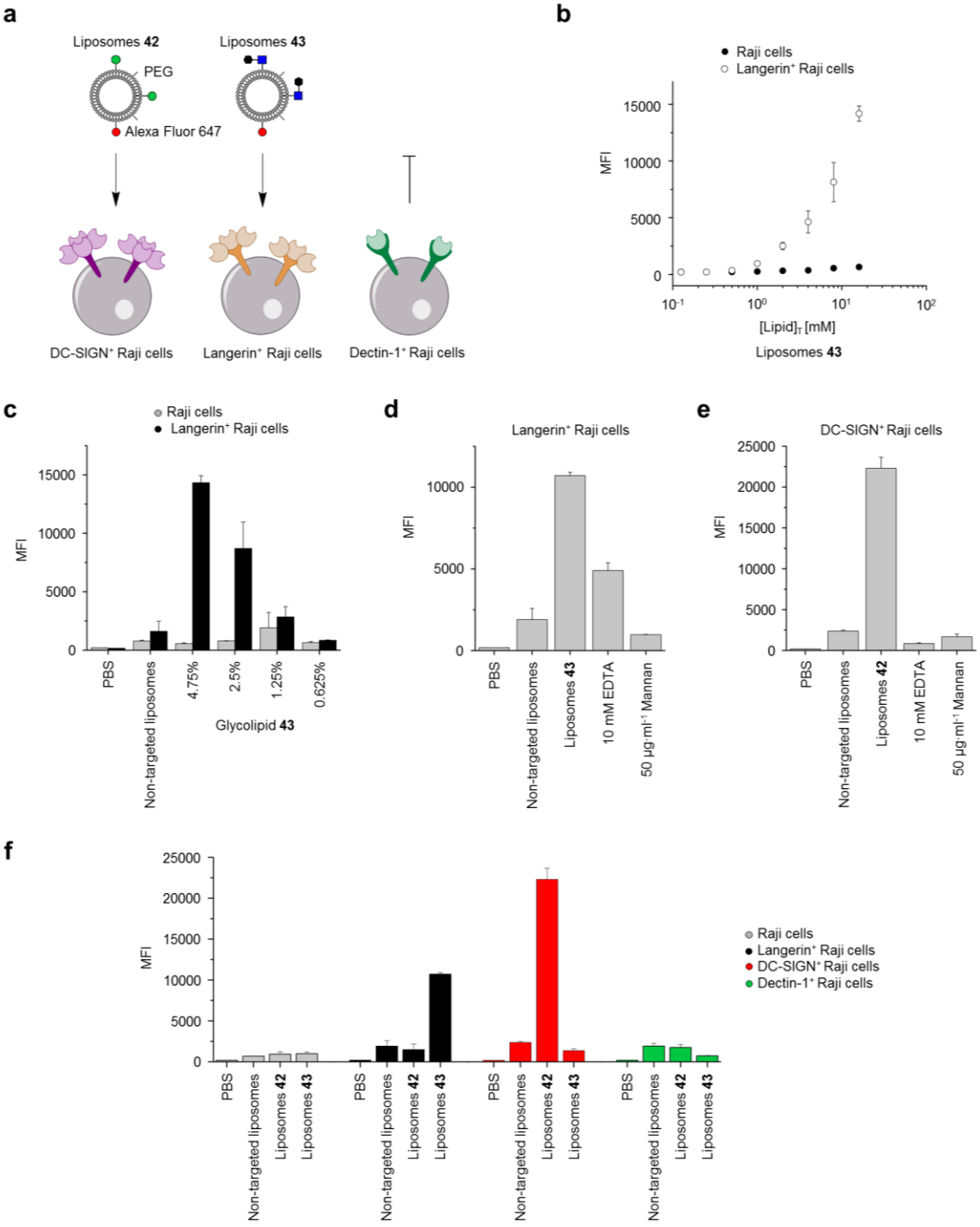

a
Liposomes 42
Liposomes 43
PEG
Alexa Fluor 647
DC-SIGN+ Raji cells
Langerin+ Raji cells
Dectin-1+ Raji cells
b
Raji cells
Langerin+ Raji cells
MFI
[Lipid]T [mM]
Liposomes 43
c
Raji cells
Langerin+ Raji cells
MFI
PBS
Non-targeted liposomes
4.75%
2.5%
1.25%
0.625%
Glycolipid 43
d
Langerin+ Raji cells
MFI
PBS
Non-targeted liposomes
Liposomes 43
10 mM EDTA
50 µg·ml-1 Mannan
e
DC-SIGN+ Raji cells
MFI
PBS
Non-targeted liposomes
Liposomes 42
10 mM EDTA
50 µg·ml-1 Mannan
f
MFI
PBS
Non-targeted liposomes
Liposomes 42
Liposomes 43
Raji cells
Langerin+ Raji cells
DC-SIGN+ Raji cells
Dectin-1+ Raji cells

Figure 36. *In vitro* targeting of Langerin$^+$ human Raji cells.[39]

a. The binding of liposomes **42** and **43** to human Raji cells expressing Langerin, DC-SIGN and Dectin-1 was investigated in flow cytometry experiments. For experiments with Dectin-1 the murine ortholog was utilized. **b.** Prior to detection of liposome binding *via* the fluorescence of Alexa Fluor 647, cells were gated for *via*bility. Dose-dependent binding of liposomes **43**, bearing GlcNS analog **41.1**, to Langerin$^+$ cells was observed. The number of liposomes utilized is expressed as the total concentration of lipids $[Lipid]_T$. All subsequent experiments were conducted at a concentration $[Lipid]_T$ of 16 µM. **c.** The binding of liposomes **43** to Langerin$^+$ cells was demonstrated to depend on GlcNS analog **41.1** in titration experiments. Only negligible unspecific binding of non-targeted liposomes to Raji cells was observed. All subsequent experiments were conducted utilizing 4.75% of glycolipids **42** or **43** for liposomal formulations **d.** Competition experiments with EDTA and mannan served to demonstrate that Langerin$^+$ Raji cells bind liposomes **43** *via* the Ca^{2+}-dependent carbohydrate binding site. **e.** Analogously, liposomes **42**, bearing Man analog **12.29**, were observed to bind DC-SIGN$^+$ Raji cells *via* the Ca^{2+}-dependent carbohydrate binding site. **f.** Among the set of analyzed CLRs, liposomes **43** are specific for Langerin$^+$ cells, while liposomes **42** are specific for DC-SIGN$^+$ cells. Mean MFI values were determined from 3 independent experiments.

Upon incubation of epidermal suspensions, binding to LCs was exclusively observed for liposomes **43**, bearing GlcNS analog **41.1** (Figure 37b). In these experiments, LCs were identified as viable HLA-DR$^+$-CD45$^+$-CD1a^{high} cells and constituted approximately 0.3% of the total population of viable cells. This value is decreased compared to LC populations of approximately 1.9% reported for the human epidermis (Bauer *et al.*, 2001). The deviation can likely be attributed to the more sophisticated quantification methods utilized in the cited study. While non-targeted liposomes neither bound to keratinocytes, melanocytes, T cells nor LCs, liposomes **43** displayed remarkable specificity for LCs (Figure 37b). The kinetics of endocytosis by LCs were evaluated by addition of EDTA at different time points during the incubation (Figure 37c). The simultaneous incubation with liposomes **43** and EDTA resulted in complete inhibition of endocytosis. By contrast, the addition of EDTA 10 min after incubation at 37°C did not affect the fraction of Alexa Fluor 647$^+$ LCs, demonstrating the efficient internalization of targeted liposomes by more than 97% of the cells on this timescale. Expectedly, endocytosis was considerably delayed at 4°C. Finally, the internalization of liposomes **43** by LCs was additionally validated *via* confocal microscopy (Figure 37d). Notably, the utilization of an anti-Langerin antibody in preliminary experiments did not result in a complete abrogation of binding and endocytosis by LCs (data not shown). While these experiments would provide rigorous evidence for Langerin-dependency, it is not clear whether the utilized antibody is a potent inhibitor of Ca^{2+}-dependent glycomimetic-Langerin interactions.

[39] The experiments were devised and conducted by Jessica Schulze (Max Planck Institute of Colloids and Interfaces).

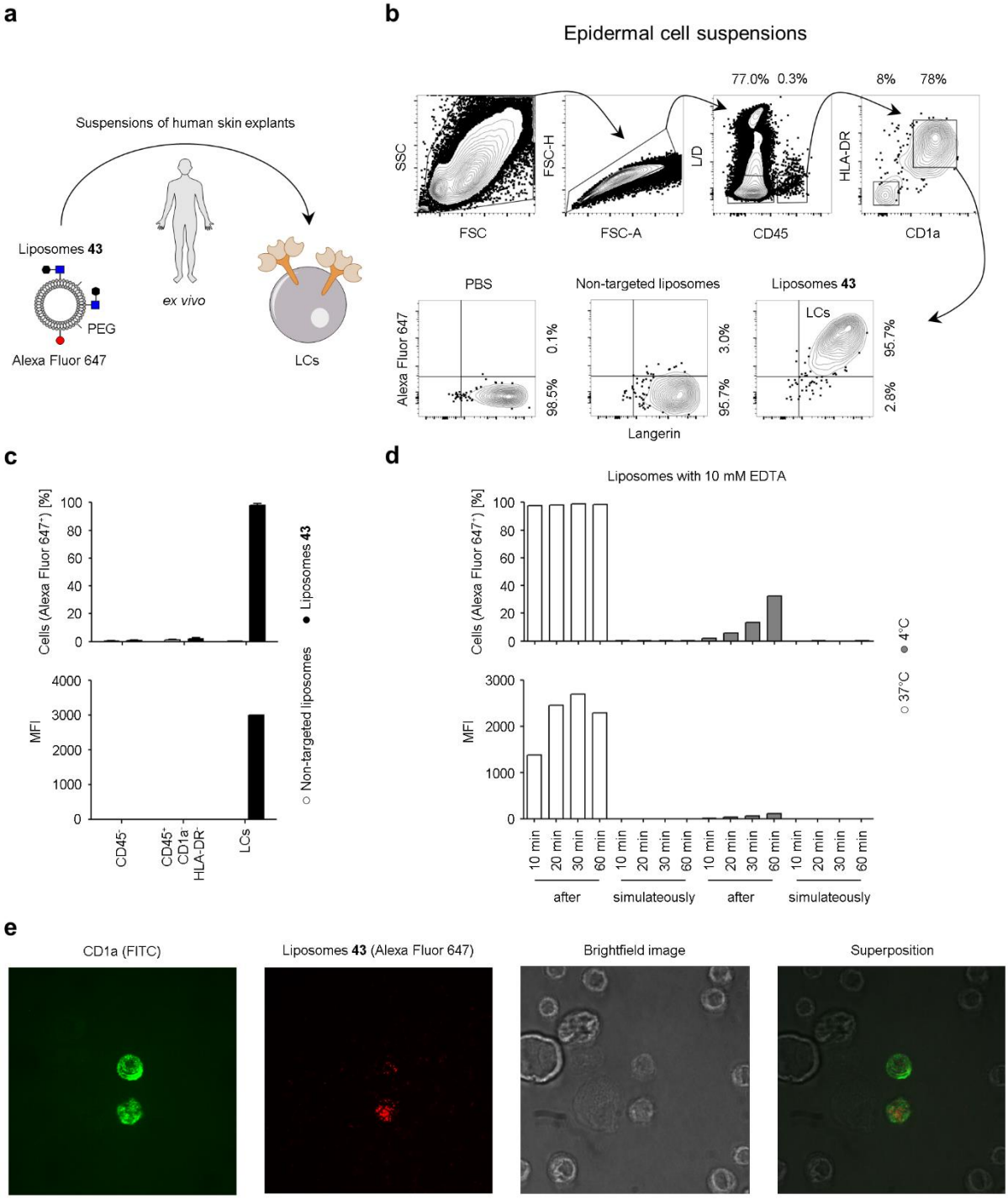

a
Suspensions of human skin explants
Liposomes 43
PEG
Alexa Fluor 647
ex vivo
LCs
b
Epidermal cell suspensions
SSC
FSC
FSC-H
FSC-A
77.0%
0.3%
L/D
CD45
8%
78%
HLA-DR
CD1a
PBS
Non-targeted liposomes
Liposomes 43
Alexa Fluor 647
0.1%
98.5%
3.0%
95.7%
LCs
95.7%
2.8%
Langerin
c
Cells (Alexa Fluor 647+) [%]
Liposomes 43
Non-targeted liposomes
MFI
CD45-
CD45+ CD1a- HLA-DR-
LCs
d
Liposomes with 10 mM EDTA
Cells (Alexa Fluor 647+) [%]
4°C
37°C
MFI
10 min
20 min
30 min
60 min
after
simulateously
e
CD1a (FITC)
Liposomes 43 (Alexa Fluor 647)
Brightfield image
Superposition

Figure 37. ***Ex vivo*** **targeting of human LCs in epidermal cell suspensions.**[40]

a. LC targeting by liposomes **43** was investigated utilizing human skin explants in flow cytometry experiments. To this end, epidermal cell suspensions were prepared as previously described (Ortner *et al.*, 2017). **b.** LCs were identified as *via*ble HLA-DR^{+}-CD45^{+}-CD1a^{high} cells (Stoitzner *et al.*, 2014). The binding of liposomes **43** to human LCs was detected *via* the fluorescence signal of Alexa Fluor 647. **c.** Selectivity for LCs was reproducibly demonstrated in 5 independent experiments and quantified *via* the fraction of Alexa Fluor 647^{+} cells. Additionally, the corresponding MFI for a representative experiment is depicted. **d.** The kinetics of liposome **43** endocytosis by LCs were analyzed at different temperatures in 3 independent experiments. Simultaneous incubation with liposomes **43** and EDTA resulted in complete inhibition of endocytosis. By contrast, the addition of EDTA 10 min after incubation with **43**-liposmoes at 37°C did not affect the fraction of Alexa Fluor 647^{+} LCs, indicating efficient endocytosis. As expected, endocytosis was considerably delayed at 4°C. **e.** LCs in epidermal cell suspensions were identified by addition of a fluorescently labeled anti-CD1a antibody and incorporation of liposomes **43** was visualized by confocal microscopy. All experiments were conducted at a concentration $[Lipid]_T$ of 16 µM

As discussed above the human dermis comprises additional APCs including dermal DCs, macrophages and monocytes (Pasparakis *et al.*, 2014). These APCs express a variety of GBPs such as the DC-SIGN and Dectin-1 which represent potential off-target receptors for glycomimetics (Stoitzner *et al.*, 2014). To evaluate whether the specificity of liposomes **43** for LCs is maintained under more physiological conditions, *ex vivo* flow cytometry experiments with whole skin cells suspensions were conducted (Figure 38). As observed for epidermal cell suspensions, targeted liposomes were efficiently internalized by LCs. The corresponding gate displayed an increased Alexa Fluor 647 fluorescence for 94% of Langerin^{+} cells. Additionally, minor populations of CD1a^{intermediate}-Langerin^{+} dermal DCs also capable of internalizing liposomes **43** were identified. Remarkably, neither CD1a^{intermediate}-Langerin^{-} dermal DCs nor CD14^{+} macrophages and monocytes were targeted by liposomes **43** (McGovern *et al.*, 2014). While minor Alexa Fluor 647^{+} populations below 2% were observed, the effects were comparable to experiments with non-targeted liposomes and might be attributed to receptor-independent, unspecific endocytosis.

[40] The experiments were devised by Assoc. Prof. Patrizia Stoitzner and conducted by Lydia Bellmann and Dr. Martin Hermann (Medizinische Universität Innsbruck).

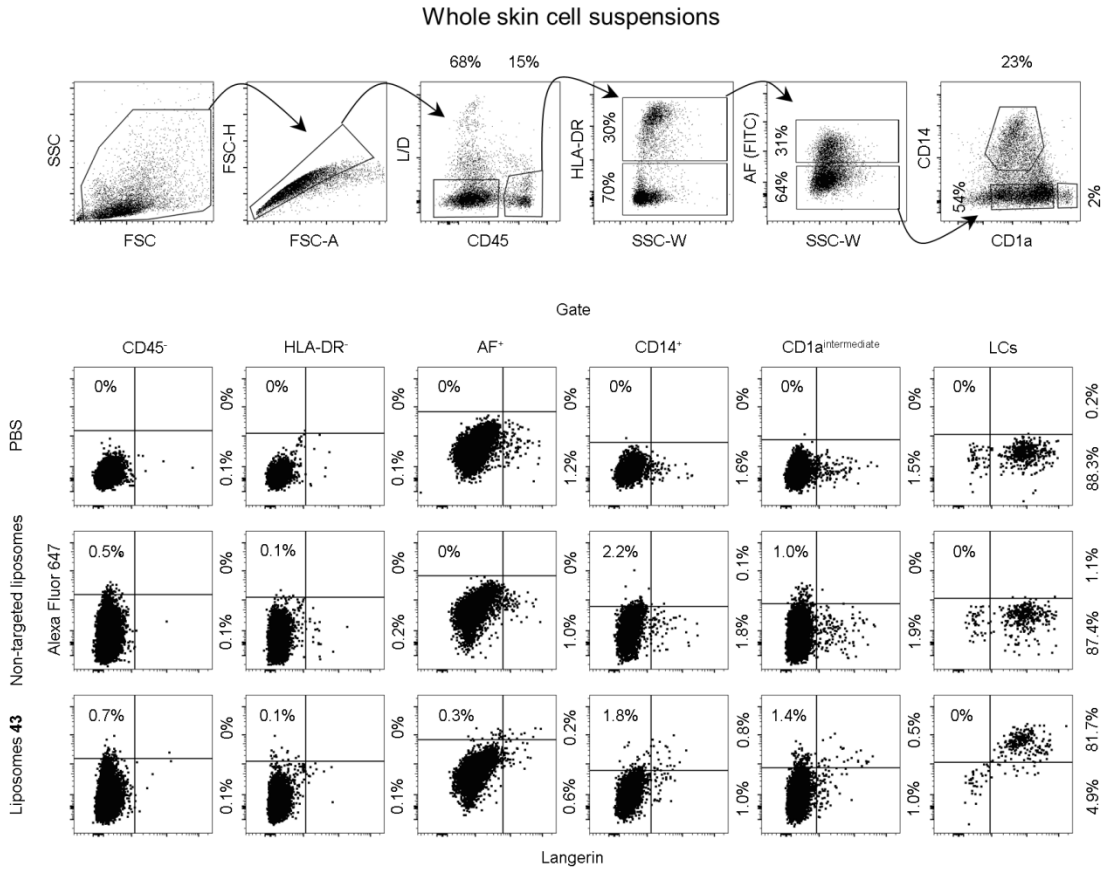


Figure 38. ***Ex vivo*** **targeting of human LCs in whole skin cell suspensions.** [41]

The specificity of liposomes **43** for LCs in the human skin was evaluated in flow cytometry experiments. To this end, whole skin suspensions were prepared as previously published (Ortner *et al.*, 2017). LCs were identified as *via*ble HLA-DR$^+$-CD45$^+$-CD1a^{high} cells while dermal DCs were identified as *via*ble HLA-DR$^+$-CD45$^+$-CD1a^{intermediate} cells (Ober-Blobaum *et al.*, 2017; Stoitzner *et al.*, 2014). Monocytes and macrophages were discriminated by the expression of CD14. The autofluorescence (AF) gate was utilized to enable the detection smaller populations of immune cells. Strikingly, binding of liposomes **43** was exclusively observed for LCs. The depicted plots are representative of 3 independent experiments. All experiments were conducted at a concentration $[Lipid]_T$ of 16 µM

To summarize, the specificity of targeting ligand **41.1** was evaluated in both *in vitro* and *ex vivo* experiments. Prior to cell-based experiments, the affinities of GlcNS analog **41.2** and Man analog **12.30** for Langerin and DC-SIGN were quantified by ^{19}F R_2-filtered NMR (Figure 35, Table 17). A comparison with flow cytometry studies utilizing model cells expressing Langerin, DC-SIGN and Dectin-1 revealed two central observations (Figure 36). First, liposomes **43**, bearing targeting ligand **41.1** were found to exclusively bind Langerin$^+$ cells, consistent with the obtained K_I values. Second, liposomes **42**, bearing targeting ligand **12.29**, were specific for DC-SIGN$^+$ cells. As Man analog **12.30** was shown to interact with Langerin in both ^{19}F R_2-filtered and ^{15}N HSQC NMR experiments, this suggests the requirement of increased affinities to target liposomes to Langerin$^+$ cells. Importantly, the submillimolar affinity determined for GlcNS analog **41.2** was sufficient to induce the efficient endocytosis of liposomes **43** by LCs in experiments with epidermal cell suspensions (Figure 37). Additionally, the liposomes were found to be specific for LCs and a minor population of

[41] The procedures were devised by Assoc. Prof. Patrizia Stoitzner and conducted by Lydia Bellmann (Medizinische Universität Innsbruck).

CD1a^{intermediate}-Langerin^{+} dermal DCs in the context of the human skin in *ex vivo* experiments (Figure 38).

A hallmark study by Steinman *et al.* first demonstrated that the antibody-mediated delivery of antigens to Langerin^{+} DCs induces efficient cross-priming of cytotoxic T cells *in vivo* (Idoyaga *et al.*, 2011). However, these experiments were conducted in mice and the elicited T cell response was dependent on CD8α^{+} DCs, a subset that does not exist in humans. Murine LCs are capable of cross-presentation but they predominantly induce tolerance (Flacher *et al.*, 2014). Human LCs, by contrast, have been shown to efficiently cross-prime cytotoxic T cells (Klechevsky *et al.*, 2008). These findings are consistent with the conservation of cross-priming-associated expression modules between murine CD8α^{+} DCs and human LCs and their absence in murine LCs (Artyomov *et al.*, 2015). While this characterizes human LCs as viable therapeutic targets for antigen delivery and the development of tumor vaccination strategies, available experimental evidence is exclusively derived from *ex vivo* or in *vitro* studies. Moreover, murine Langerin has been shown to display a specificity profile for oligosaccharides recognition that is distinct from its human ortholog and potentially translates into a decreased affinity for targeting ligand **41.1** (Hanske *et al.*, 2017a).

These considerations have important implications for the development of antigen delivery strategies for tumor vaccination utilizing liposomes **43**. While the targeted liposomes potentially display clinical relevance and might be utilized in human trials in the future, GlcNS analog **41.1** might not be applicable to animal studies in mice. Moreover, findings from these studies are potentially non-transferrable to the human system because of the distinct immunological characteristic of Langerin^{+} DC subsets (Clausen and Stoitzner, 2015). Trials in non-human primates represent a long-term alternative. Yet, these studies imply ethical concerns as well as considerable administrative and experimental efforts, particularly as the required adoptive T cell transfer and tumor challenge models are well-established for mice (Bouwer *et al.*, 2014; Idoyaga *et al.*, 2011). To justify these efforts, more elaborate *ex vivo* studies with isolated human LCs will enable cross-priming experiments with autologous T cells. Over 90% of the human population are infected with Epstein-Barr virus (EBV) and the virus typically persist throughout an individual's entire life (Cohen, 2000). Latency is associated with the expression of a limited set of antigens including EBNA-1 and LMP-2 that have been reported to efficiently induce both helper and cytotoxic T cell responses in infected humans (Petersen *et al.*, 1989; Taylor *et al.*, 2004). Consequently, the co-formulation of these antigens with liposomes **43** will allow for the suggested *ex vivo* cross-priming experiments with LCs isolated from healthy humans (Stoitzner *et al.*, 2010).

The intracellular routing of internalized liposomes **43** and the efficient release of targeting ligands in the endosome will likely determine the efficiency of antigen processing and presentation to T cells. For example, the modification of antigens with an anti-Langerin antibody resulted in altered intracellular routing and an increased cross-presentation by human LCs compared to receptor-

independent, unspecific endocytosis (Fehres *et al.*, 2015b). An attempt to utilize anti-Langerin antibodies to target liposomes to human LCs, on the other hand, failed to activate cytotoxic T cells (Fehres *et al.*, 2015c). Additionally, the capacity of LCs to cross-prime T cells depends on the co-administration of adjuvants. Here, TLR-3 and -7 agonists were observed to activated human LCs and dermal DCs and promoted cross-priming of cytotoxic T cells (Fehres *et al.*, 2014a; Fehres *et al.*, 2015b). Finally, efficient anti-tumor T cell responses will also require the DC-mediated activation of NK cells (Bouwer *et al.*, 2014). These aspects remain to be investigated for antigen delivery *via* the targeted liposomes and will critically influence their suitability for the development of cancer immunotherapies.

In conclusion, the heparin-inspired design strategy integrating carbohydrate chemistry and NMR spectroscopy led to the discovery of a potent and specific targeting ligand for Langerin. Prominently, the developed ^{19}F R_2-filtered NMR assay enabled the robust and efficient elucidation of the SAR for a focused library of GlcNS analogs **31** (Figure 30, Table 15, Scheme 6). The subsequent identification of a suitable conjugation strategy was guided by structural information obtained *via* molecular docking studies as well as ^{15}N HSQC and STD NMR experiments (Figure 31, Figure 32 and Figure 33, Table 16). Targeting ligand **41.2** displayed a 42-fold affinity increase over naturally occurring mono- and oligosaccharides and its specificity against DC-SIGN was demonstrated *via* the ^{19}F R_2-filtered NMR assay (Figure 35, Table 17). Strikingly, findings obtained *via* NMR spectroscopy were readily transferable to cell-based experiments both *in vitro* and *ex vivo*, highlighting the potential of the implemented design strategy for other CLRs. Following the synthesis of glycolipid **43**, flow cytometry experiments served to optimize liposomal formulations and to validate the binding of targeted liposomes to Langerin$^+$ model cells (Figure 36). Finally, findings from *ex vivo* experiments demonstrate the capacity of the targeted liposomes **43** to specifically target LCs in the human skin. The liposomes were efficiently endocytosed and thus represent a promising antigen delivery platform.

The presented findings constitute an important milestone for the research field of DC immunology and the development of DC-based immunotherapies for the treatment of cancer and autoimmune diseases. The designed glycomimetics address current limitations encountered for the utilization of naturally occurring oligosaccharides or antibodies as targeting ligands for CLRs. In particular, they provide the required potency and specificity while maintaining Ca^{2+}-dependent interactions with the carbohydrate binding site of Langerin. Thereby they are likely to mimic physiological and pathophysiological process including potential signaling profiles or the pH-dependent release of targeting ligands in endosomes. Future investigations in our laboratory will focus on the utilization of the designed liposomes to target antigens to LCs. To this end, the affinity of targeting ligand **41.1** for murine Langerin will be evaluated to explore its applicability to animal studies in mice. Additionally, fragment-based ligands were developed to target the murine ortholog (Aretz, 2017). Regarding the human system, the proposed *ex vivo* studies with EBV-derived antigens and isolated human LCs will

be prioritized to evaluate the capacity of liposomes **43** to induce the cross-priming of naïve $CD8^+$ T cells. In parallel, confocal microscopy will serve to investigate the intracellular routing and processing of both the liposomes and delivered antigens. Notably, the availability of targeting ligands **12.29**, **14.4** and **41.1** enables the systematic evaluation of the influence of *e.g.* the affinity or the monosaccharide scaffold on these processes. As the designed delivery platform enables facile co-formulation of adjuvants and other immunomodulatory molecules, potential applications extend beyond the development of cancer immunotherapies and include the treatment of autoimmune diseases.

3.5. Multivalent Organization of Glycomimetic Ligands

The multivalent organization of mono- and oligosaccharides or glycomimetic ligands on scaffold structures represents a powerful strategy to address the challenges encountered for the design of potent CLR ligands outlined in the previous chapters (Chapter 1.4, 3.1 and 3.3). Chelate cooperativity effects potentially induce avidity increases that are specific for the geometry of oligomeric CLRs bearing multiple CRDs. Over the last two decades, the development of synthetic strategies for glycoclusters, glycodendrimers, glycopolymers and glyconanoparticles has enabled the discovery of potent multivalent glycomimetic ligands (Cecioni *et al.*, 2015).

As highlighted above, nucleic acids represent an attractive class of scaffold structures as they provide excellent control over the valency and spatial orientation of organized ligands (Chapter 1.4). The unique fidelity of Watson-Crick base pairing enables the programmed self-assembly of monodisperse PNA-DNA duplexes displaying optimized thermodynamic and enzymatic stability. To design potent divalent glycomimetic Langerin ligands, a library of PNA-DNA duplexes organizing GlcNS analog **41.1** was prepared in the laboratory of Prof. Oliver Seitz (Humboldt-Universität zu Berlin). Langerin forms homotrimers in the plasma membrane of LCs and the structure of the trimeric ECD has been elucidated *via* X-ray crystallography (Feinberg *et al.*, 2010). The CRDs are organized symmetrically by the neck region with the Ca^{2+}-dependent carbohydrate binding sites located at the top of the trimer. The distance between the Ca^{2+} ions equals approximately 42 Å and the programmed nucleic acid scaffolds were designed to match this geometry to leverage chelate cooperativity effects (Figure 39a). Notably, the concave receptor surface between the binding sites does not require bending of the PNA-DNA duplexes.

PNA-DNA duplexes were prepared in analogy to a previously published procedure (Scheibe *et al.*, 2011). Briefly, DNA templates and PNA fragments were assembled *via* automated solid-phase synthesis. An extended linker structure bearing a maleimide group was introduced to GlcNS analog **41.1** and the modified targeting ligand was conjugated to thiol-modified PNA fragments by thioester formation. Overall, five **41.1**-modified PNA fragments displaying unique anticodons and modifications in different positions were prepared. Combinations of one or two of these fragments and one unmodified PNA fragment were hybridized with a DNA template resulting in the self-assembly of divalent glycomimetics. The length of the DNA template was kept constant and the sequence was designed to display three codons complementary to the anticodons of the PNA fragments. The variation of these codons enabled the differential organization of GlcNS analog **41.1** at defined distances d. Second generation PNA-DNA duplexes were optimized *via* the introduction of β-lysine residues to the PNA fragments to address potential problems with solubility.

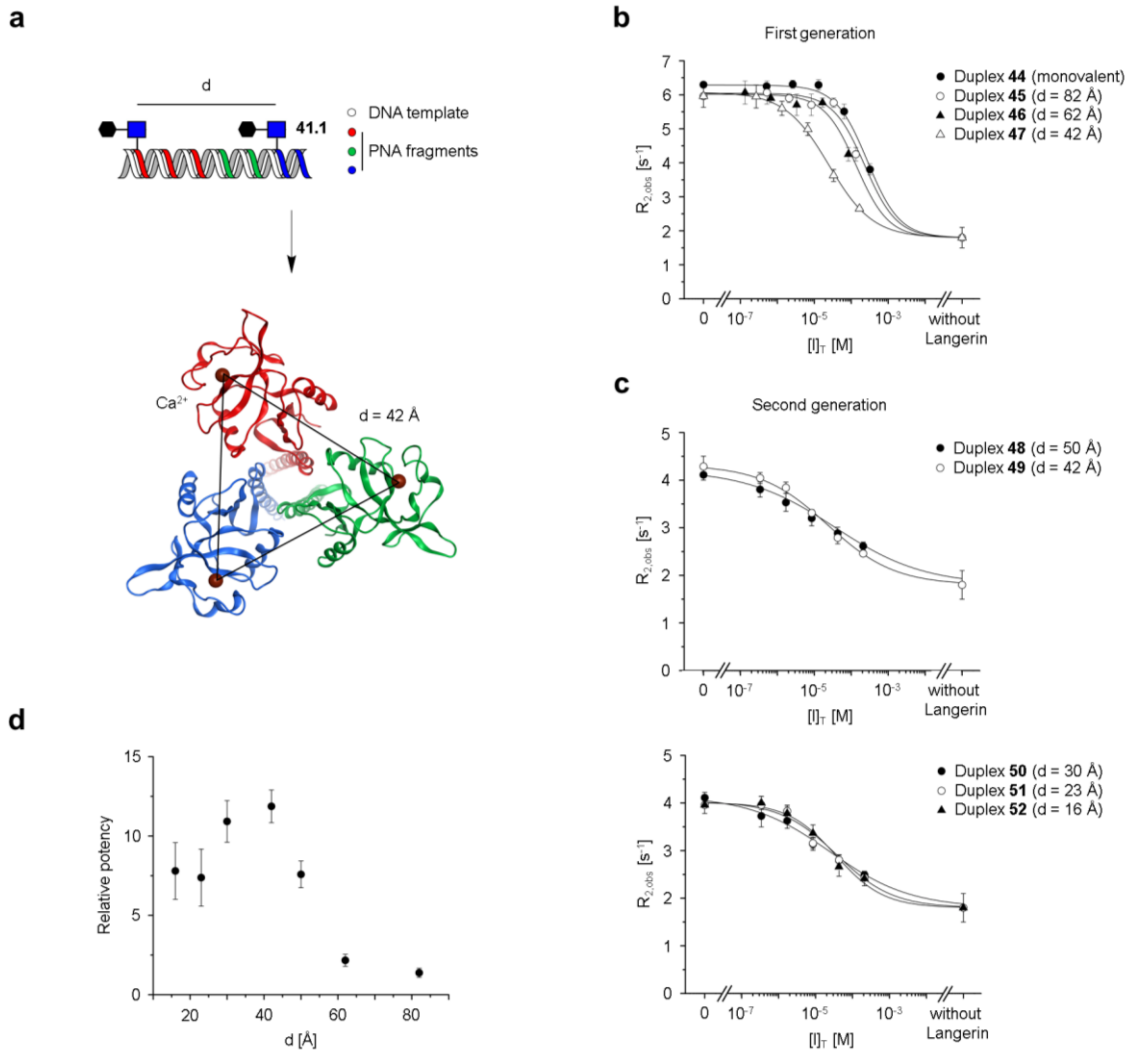


Figure 39. Divalent organization of GlcNS analog 41.1 on programmed nucleic acid scaffolds.[42]

a. A library of DNA-PNA duplexes was prepared to target the trimeric Langerin ECD as previously described (Scheibe *et al.*, 2011). PNA fragments were organized on a DNA template to program divalent glycomimetics displaying defined distances d between GlcNS analogs **41.1**. The X-ray structure of the trimeric Langerin ECD revealed the distance (d = 42 Å) between Ca^{2+}-dependent carbohydrate binding sites (PDB code: 3KQG) (Feinberg *et al.*, 2010). **b.** The evaluation of the first generation of divalent glycomimetics *via* ^{19}F R_2-filtered NMR revealed a distance-dependent 11-fold avidity increase for duplex **47** (IC_{50} = 25±1 µM) over the monovalent reference duplex **44** (IC_{50} = 273±32 µM). **c.** A second generation of divalent glycomimetics was prepared to address potential problems with solubility. These duplexes were evaluated at a reduced receptor concentration ($[P]_T$ = 25 µM) to address potential aggregative false-positive mechanisms and the assays upper limit of accurate affinity determination. At reduced distances between GlcNS analogs **41.1** (d = 23 and 16 Å), duplexes **51 (**IC_{50} = 37±9 µM**)** and **52** (IC_{50} = 35±9 µM) displayed decreased affinities. **d.** The distance dependency of observed avidity increases was visualized and suggests an optimal organization of GlcNS analogs **41.1** at a distance of approximately 42 Å, matching the geometry of the trimeric Langerin ECD.

[42] The DNA-PNA duplexes were designed and prepared by Gunnar Bachem (Humboldt-Universität zu Berlin).

Table 18. Structure-activity relationship for programmed divalent nucleic acid glycomimetics. [43]

Name	d [Å]	IC_{50} [µM]	p	Relative potency[a]
		First generation		
44	Monovalent	273±32	1[b]	1.0
45	82	198±42	1[b]	1.4
46	62	126±22	1[b]	2.2
47	42	25±1	0.73±0.02	11
		Second generation		
48	50	36±4	0.36±0.02	7.6
49	42	23±2	0.48±0.03	12
50	30	25±3	0.40±0.03	11
51	23	37±9	0.57±0.09	7.4
52	16	35±8	0.67±0.10	7.8

[a] The relative potency was calculated utilizing the IC_{50} value determined for monovalent duplex **44**.
[b] The Hill factor p was fixed during the fitting procedure.

Next, the divalent glycomimetics were characterized in ^{19}F R_2-filtered NMR experiments (Figure 39b to d, Table 18). Due to the complex multi-state equilibria encountered for interactions between divalent glycomimetics and the trimeric Langerin ECD, avidities were quantified by the determination of IC_{50} values (Setup 3) (Chapter 3.2). Strikingly, duplex **47** (d = 42 Å, IC_{50} = 25 ±1 µM) displayed an 11-fold avidity increase of over monovalent reference duplex **44** (IC_{50} = 273±32 µM) (Figure 39b). The avidity of PNA-DNA duplexes was found to be distance-dependent with larger distances corresponding to higher IC_{50} values. Notably, complete competition was observed for neither duplex **44**, **45** nor **46**. As the utilization of higher concentrations was impeded by potential solubility limitations for the PNA fragments and the considerable synthetic efforts required to prepared the PNA-DNA duplexes, IC_{50} values were estimated by assuming a Hill factor p of 1 during fitting procedures.

Several important parameters potentially influence the determination of IC_{50} values and Hill factors *via* the conducted ^{19}F R_2-filtered NMR experiments. First, positive chelate cooperativity effects likely contribute to the observed avidity increase. This assumption is supported by the fact that the distance for duplex **47** resembles the distance between carbohydrate binding site in the Langerin trimer. However, the corresponding binding isotherm and the decreased Hill factor (p = 0.73±0.02) indicate that these cooperativity effects are non-optimal. For comparison, increased Hill factors (p = 1.1 to 1.5) were predicted for a one-site binding model in the relevant affinity range (Chapter 3.2.2). Alternatively, duplex **47** might bind the trimeric Langerin ECD in an aggregative false-positive

[43] The DNA-PNA duplexes were designed and prepared by Gunnar Bachem (Humboldt-Universität zu Berlin).

mechanism as reported for other GBPs (Sacchettini *et al.*, 2001; Sutkeviciute *et al.*, 2014). This would result artificially increased $R_{2,obs}$ values for reporter molecule **5.1** at higher concentrations of duplex **47** and thereby contribute to the determination of decreased avidities and Hill factors. Lastly, titrations were conducted at a receptor concentration $[P]_T$ of 50 µM and the observed IC_{50} value might be artificially increased due to the upper limit of accurate affinity determination (Chapter 3.2.2)

Accordingly, additional ^{19}F R_2-filtered NMR experiments were conducted at a reduced receptor concentration of 25 µM to account for potential aggregative false-positive mechanisms and to accurately determine IC_{50} values in the low micromolar range (Figure 39c). In these experiments, a second generation of PNA-DNA duplexes was characterized. The duplexes were designed to sample additional distances and to account for potential solubility limitations. Importantly, the avidity increase for duplex **47** was validated under these experimental conditions as duplexes **49** (d = 42 Å, IC_{50} = 23±2 mM) and **50** (d = 30 Å, IC_{50} = 25±3 mM) displayed comparable IC_{50} values. By contrast, decreased affinities were observed at distances shorter than 30 Å. It is noteworthy that the utilization of a reduced receptor concentration resulted in a reduced dynamic range for the assay as well as increased standard errors for IC_{50} and p values. The general decrease observed for p values is consistent with values predicted for a one-site binding model (p = 1.0 to 1.2) under the changed experimental conditions (Chapter 3.2.2).

To summarize, a distance-dependent SAR was observed for the divalent organization of GlcNS analog **41.1** on PNA-DNA duplexes (Figure 39d). With a 11-fold avidity increase, duplexes **47** (IC_{50} = 25 ±1 µM) and **49** (IC_{50} = 23±2 mM) displayed optimal spacing at approximately 42 Å. The same distance was measured between the Ca^{2+} ions in the X-ray structure of the trimeric Langerin ECD, suggesting positive chelate cooperativity effects for the divalent interaction. Importantly, a potential aggregative false-positive mechanism was excluded in titrations at a reduced receptor concentration. The maximal avidity effects for a multivalent system depend on the ΔG value of the corresponding monovalent interactions. Assuming perfect additivity and no favorable contributions from the scaffold structure, the divalent organization of GlcNS analog **41.1** might potentially result in a more than 3000-fold avidity increase. While this optimal scenario is unlikely, divalent PNA-DNA duplexes displaying increases of more than 60-fold have been reported for comparable monovalent interactions (Scheibe *et al.*, 2011; Scheibe *et al.*, 2013). Hence, further optimization of duplexes **47** and **49** will likely result in more potent divalent glycomimetic Langerin ligands and two orthogonal design approaches might be envisioned. First, the rigidification of the linker structure potentially avoids unfavorable entropic contributions due to an increased conformational flexibility for GlcNS analog **41.1** in the free state. Alternatively, the rigidity of the RNA-DNA duplex might hinder the optimal orientation of GlcNS analog **41.1** in Langerin's carbohydrate binding site. Here, the introduction short unpaired regions for the DNA template represents a well-established approach to gradually increase the flexibility of a duplex (Scheibe *et al.*, 2011).

Furthermore, it will be necessary to validate the observed avidity effects in orthogonal assays. For example, SPR experiments have successfully been utilized to determine apparent K_D values for multivalent systems. However, the robustness of this parameter is compromised as observed binding isotherms critically depend on the immobilization density (Scheibe *et al.*, 2011). This observation is also relevant for other heterogeneous assays including ELLAs. Alternatively, ITC experiments enable the determination of apparent thermodynamic parameters in homogeneous systems. Lastly, additional ^{19}F R_2-filtered NMR experiments utilizing the monomeric Langerin CRD will be conducted to assess the potential contributions from the nucleic acid scaffold.

The distance-dependent avidity increase for duplexes **47** and **49** suggest that the divalent organization of GlcNS analog **41.1** will convey additional specificity against other CLRs: On the one hand, avidity effects for monomeric CLRs displaying a single CRD such as L-Selectin or E-Selectin will exclusively depend on their spacing on the plasma membrane. On the other hand, many oligomeric CLRs display distinct geometries thereby reducing potential chelate cooperativity effects. The ECD of DC-SIGNR for instance, forms tetramers and the distances between the carbohydrate binding sites range from 40 to 80 Å (Feinberg *et al.*, 2005). Importantly, and in contrast to Langerin, the receptor surface between the carbohydrate binding sites is convex as the bindings sites are oriented in opposite directions. As X-ray structures are only available in selected cases, other approaches such as SAXS or mapping by multivalent glycomimetics served to estimate the geometry for DC-SIGN tetramers and ASGPR trimers or tetramers, respectively (Huang *et al.*, 2017; Tabarani *et al.*, 2009). The distances for the tetrameric DC-SGIN were proposed to be similar to DC-SIGNR, yet the orientation of the binding sites is likely comparable to Langerin. For ASGPR, shorter distances of approximately 20 Å were estimated. Hence, the utilization of the designed divalent glycomimetics represents an attractive strategy to further improve the specificity profile and targeting efficiency of liposomes (Chapter 3.4). However, this would require the development of a suitable conjugation strategy as the prepared DNA-PNA duplexes bear several amino groups.

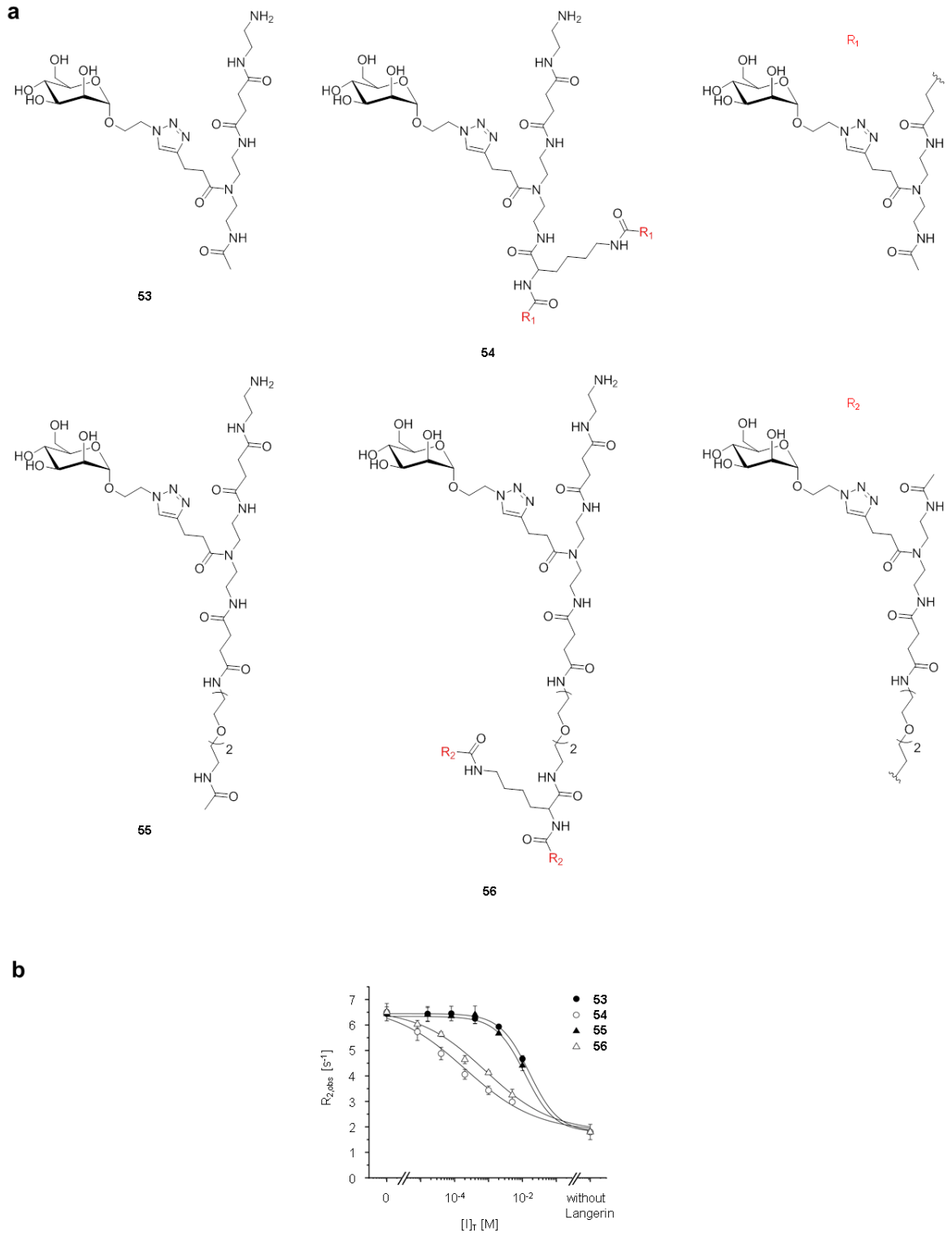


Figure 40. Trivalent organization of Man on glycoclusters.[44]

a. A library of glycoclusters was prepared to target the trimeric Langerin ECD as previously described (Ponader *et al.*, 2014). Modifications of the published procedure include the introduction of a β-lysine residue to obtain branched scaffolds. **b**. The evaluation of the trivalent glycomimetics *via* ^{19}F R_2-filtered NMR revealed an 80-fold avidity increase for **54** (IC_{50} = 0.20±0.08 mM) over the monovalent reference molecule **53** (IC_{50} = 16±4 mM). Avidity effects for the extended scaffold of **56** (IC_{50} = 0.8±0.1 mM) are less pronounced.

[44] The glycoclusters were designed and prepared by Kira Neuhaus (Heinrich Heine Universtiy of Düsseldorf)

In an alternative approach, trivalent glycoclusters were designed to target Langerin (Figure 40a). The glycomimetics bear a polymeric amidoamino scaffold which can be extended by the introduction of PEG groups. In contrast to typical glycopolymers, the scaffolds are monodisperse and enabled the defined trivalent organization of Man. The approach has not only been successfully utilized to prepare heteromultivalent glycomimetics to target oligomeric plant lectins but also led to the discovery of shielding effects for multivalent interactions (Ponader *et al.*, 2014). The scaffold displays substantial flexibility and thus provides less control over the spatial organization of ligands than PNA-DNA duplexes. However, the glycomimetics can be prepared in larger quantities and are compatible with the developed conjugation strategy for liposomal formulations. The glycoclusters were prepared in the laboratory of Prof. Laura Hartmann (Heinrich-Heine Universität Düsseldorf) in analogy to a previously described procedure (Ponader *et al.*, 2014). Briefly, the scaffolds were assembled *via* automated solid-phase synthesis. The introduction of an ethylazido group *via* α-mannoside formation and enabled the *in situ* conjugation of Man to alkyne-modified β-alanine-like building blocks by Cu(I)-catalyzed Huisgen 1,3-dipolar cycloaddition (Rostovtsev *et al.*, 2002).

Subsequently, the trimeric glycoclusters were characterized in ^{19}F R_2-filtered NMR competitive binding experiments (Figure 40b, Table 19). Due to the complexity of multivalent equilibria, avidities were quantified by the determination of IC_{50} values (Setup 3) (Chapter 3.2). The evaluation of the obtained SAR reveals an 80-fold avidity increase for **54** (IC_{50} = 0.20±0.08 mM) over the monovalent reference molecule **53** (IC_{50} = 16±4 mM). In comparison, the avidity effects for the extended scaffold of **56** (IC_{50} = 0.8±0.1 mM) were less pronounced. This observation might be attributed to an increased flexibility and the corresponding unfavorable entropic contributions. Considering the low affinity for the monovalent interactions, the observed avidity for **54** is remarkable and indicates considerable positive chelate cooperativity effects. However, both trivalent glycomimetics display decreased Hill factors (p = 0.41±0.08 and 0.44±0.09) suggesting that these effects are non-optimal. A comparison with the PNA-DNA duplexes requires the normalization of avidity increases with respect to the valency. This normalization reveals that avidity effects for **54** are 5.4-fold more pronounced than for duplex **47**, despite weaker monovalent interactions. Hence, it can be argued that the branched architecture of the trimeric glycoclusters efficiently accounts for the spatial organization of CRDs in the Langerin trimer and thereby potentially conveys specificity against other CLRs. Furthermore, the flexibility provided by the amidoamino scaffold might be required to optimally orient ligands in the Ca^{2+}-dependent carbohydrate binding site.

Table 19. Structure-activity relationship for trivalent glycoclusters. [45]

Name	IC_{50} [mM]	p	Relative potency
53	16±4	1[a]	1.0[b]
54	0.20±0.08	0.41±0.08	80[b]
55	12±3	1[a]	1.0[c]
56	0.8±0.1	0.44±0.09	15[c]

[a] The Hill factor p was fixed during the fitting procedure.
[b] The relative potency was calculated utilizing the IC_{50} value determined for monovalent reference molecule **53**.
[c] The relative potency was calculated utilizing the IC_{50} value determined for monovalent reference molecule **55**.

While these initial findings are encouraging, the observed avidity effects should be validated in additional experiments to exclude *e.g.* aggregative false-positive mechanisms. As discussed for the PNA-DNA duplexes, the implementation of ^{19}F R_2-filtered NMR experiments at reduced receptor concentrations and orthogonal biophysical techniques such as SPR and ITC represent a suitable strategy. Alternatively, the trivalent glycoclusters can be readily conjugated to NHS-activated lipids and the properties of the targeted liposomes in cell-based *in vitro* experiments are currently under investigation in our laboratory. Notably, this approach enables the utilization of an increased concentration of targeting ligands while maintaining the required stability for formulated liposomes. In the future, the ethylamino linker of GlcNS analog **41.1** might be substituted with an azido group to leverage the scaffold of **54** for trivalent organization to further optimize the specificity profile and targeting efficiency of liposomes (Chapter 3.4).

[45] The glycoclusters were designed and prepared by Kira Neuhaus (Heinrich Heine University of Düsseldorf)

4. Conclusion and Perspectives

Cancer Immunotherapy and Langerhans Cells

Immune evasion represents an important hallmark of cancer progression (Hanahan and Weinberg, 2011). In this context, the induction of tumor-specific cytotoxic T cell immunity to overcome immunological tolerance has become a focal point of cancer immunotherapy (Palucka and Banchereau, 2013). DCs play a pivotal role in immune regulation and have been recognized for their potential to endocytose and cross-present exogenous antigens *via* MHC-I to efficiently prime naïve $CD8^+$ T cells. This has rendered DCs attractive targets for the development of novel vaccination strategies.

Adoptive DC therapy has been approved for the treatment of prostate cancer and provides the proof-of-principle for DC-based vaccines (Kantoff *et al.*, 2010). However, *ex vivo* therapies display limited therapeutic efficacy and are considered laborious as well as expensive (Tacken *et al.*, 2007). The targeted delivery of TAAs to DCs *in vivo* has been extensively explored to address these limitations (Palucka and Banchereau, 2013). CLRs constitute a family of endocytic receptors *i.a.* expressed by DCs and other immune cells. CLR-dependent endocytosis promotes cross-presentation and Steinmann *et al.* have established the utility of anti-CLR antibody-antigen conjugates to elicit cytotoxic T cell immunity (Bonifaz *et al.*, 2002). Prominently, targeting TAAs to DEC-205 has translated into clinical trials for the treatment of *e.g.* melanoma (Dhodapkar *et al.*, 2014).

However, the underlying mechanisms determining the efficiency of DC-mediated cross-presentation to prime naïve $CD8^+$ T cells remain largely elusive (Fehres *et al.*, 2014b). Consequently, elucidating the contributions from individual CLRs and DC subsets is of particular importance for basic research and potentially facilitates the development of novel vaccination strategies (Palucka and Banchereau, 2013).

Compared to other CLRs such as DEC-205, Langerin displays a highly restricted expression profile and is predominantly present on human LCs (Clausen and Stoitzner, 2015). This DC subset, residing in the epidermis of the human skin, has been demonstrated to efficiently activate cytotoxic T cells (Klechevsky *et al.*, 2008). Moreover, the induction of cytotoxic T cell immunity can be facilitated by the Langerin-dependent endocytosis of exogenous antigens to promote cross-presentation (Fehres *et al.*, 2015b; Fehres *et al.*, 2015c). Hence, Langerin and LCs represent attractive therapeutic targets and are ideally suited to investigate DC immunology.

The central aim of the work presented in this dissertation was the design of glycomimetic Langerin ligands to target liposomes to LCs (Figure 41). Over the last decades, liposomes have emerged as

versatile delivery platforms displaying several advantages over antibody-TAA conjugates (Sercombe *et al.*, 2015). Prominently, they enable the co-formulation of adjuvants such as TLR-3 agonists that are required to induce the maturation of LCs and other DC subsets (Boks *et al.*, 2015; Fehres *et al.*, 2015a). Notably, immature LCs have been observed to induce peripheral tolerance rather than TAA-specific T cell immunity (Seneschal *et al.*, 2012).

Langerin exerts its physiological functions as an endocytic PRR *via* the Ca^{2+}-dependent recognition of pathogen- or self-associated glycans (de Witte *et al.*, 2007; Feinberg *et al.*, 2011). These interactions are abrogated at decreased Ca^{2+} concentration and pH values, routing antigens to the early or late endosomal compartment to promote cross-presentation (Fehres *et al.*, 2015b; Hanske *et al.*, 2016). As Langerin is a recycling CLR, antigen release in the endosome additionally enhances the endocytic capacity of LCs (Gidon *et al.*, 2012; McDermott *et al.*, 2002). Structural glycomimetics, particularly carbohydrate analogs, will share a subset of the interactions formed by glycans. Accordingly, liposomal targeting by glycomimetic Langerin ligands potentially facilitates endocytosis, endosomal antigen release and cross-presentation by LCs compared to utilization of antibody TAA conjugates (O'Reilly *et al.*, 2011).

Discovery of Glycomimetic Langerin Ligands

However, the design of specific and potent glycomimetics is challenging (Ernst and Magnani, 2009). GPIs are typically weak and promiscuous as Ca^{2+}-dependent CLR carbohydrate binding sites display high solvent exposure and hydrophilicity (Weis and Drickamer, 1996). Moreover, the synthesis of carbohydrate analogs and other structural glycomimetics is considered onerous due to diverse stereo- and regioisomerism as well as the high density of hydroxyl groups (Hahm *et al.*, 2017; Seeberger and Werz, 2007). Modern strategies for the design of structural glycomimetics either involve the replacement of the carbohydrate scaffold with non-carbohydrate ligands, the reduction of the scaffold or its derivatization to explore the CLR surface for favorable secondary interactions (Barra *et al.*, 2016; Binder *et al.*, 2012; Garber *et al.*, 2010). Alternatively, GPIs can be modulated by functional glycomimetics interacting with secondary binding pockets leveraging allosteric mechanisms or steric hindrance (Aretz *et al.*, 2017; Liu and Finzel, 2014). Concepts from FBLD are ideally suited to guide glycomimetic ligand design: Fragment screening facilitates the identification of non-carbohydrate CLR ligands and fragment growing or linking can be employed to discover potent carbohydrate analogs (Aretz *et al.*, 2017; Egger *et al.*, 2013; Liu and Finzel, 2014).

The implementation of FBLD-inspired approaches requires the existence of druggable binding pockets. In this context, the structure-based *in silico* analysis of 21 X-ray structures served to evaluate the general feasibility of glycomimetic ligand design for human CLRs (Chapter 3.1). This analysis, utilizing DoGSiteScorer, corroborated the classification of Ca^{2+}-dependent carbohydrate binding sites as undruggable or challenging targets (Volkamer *et al.*, 2012b). Moreover, druggable secondary

binding pockets in proximity to Ca^{2+} ions were exclusively identified for CLRs of limited therapeutic relevance. These findings are contrasted by experimental hit rates obtained from fragment screenings against Langerin, DC-SIGN and MCL conducted in our laboratory (Aretz *et al.*, 2014). While only few GPI inhibitors were identified, hit rates for Ca^{2+}-dependent interactions were considerably increased (Aretz *et al.*, 2016; Wamhoff *et al.*, 2016). Notably, structural changes upon Ca^{2+} recognition, particularly in the long loop region have been reported for several CLRs and can be leveraged for the design of functional glycomimetics (Aretz *et al.*, 2017; Furukawa *et al.*, 2013; Hanske *et al.*, 2016). The grid-based algorithm utilized by DoGSiteScorer does not account for structural flexibility which impedes the identification of secondary binding pockets potentially targeted *via* conformational selection or induced fit mechanisms (Volkamer *et al.*, 2010). This limitation represents a major challenge for *in silico* druggability predictions and has recently addressed by Sal *et al.* (Cimermancic *et al.*, 2016).

As no druggable secondary binding pockets were identified *via* DoGSiteScorer, the glycomimetic ligand design for Langerin focused on fragment growing-inspired approaches to discover potent carbohydrate analogs (Chapter 3.3 and 3.4). Several strategies were employed to reduce the synthetic efforts throughout the design process and to efficiently explore the carbohydrate binding site for favorable secondary interactions. Langerin contains an EPN motif and recognizes Fuc- and Man-type glycans (Drickamer, 1992; Feinberg *et al.*, 2011). As Man-type oligosaccharides display no affinity increase for the monomeric CRD, the corresponding monosaccharide was selected as an initial scaffold for glycomimetic ligand design (Hanske, 2016; Holla and Skerra, 2011).

The visual assessment of X-ray structures of Langerin in complex with different Man-type oligosaccharides revealed two potential binding pockets adjacent to the carbohydrate binding site (Feinberg *et al.*, 2011). These pockets were targeted *via* the structure-based *in silico* screening of substituents in C2 of the Man scaffold to design a focused library of ManNAc analogs **5** (Chapter 3.3.1). Overall, twelve analogs were synthesized utilizing commercially available carboxylic acids and subsequently characterized in titration experiments. **5.11** (K_I = 1.3±0.1 mM) was identified as the first glycomimetic ligand reported for Langerin displaying a 5.9-fold affinity increase over Man (Figure 41) (Wamhoff *et al.*, 2016). This affinity increase can be attributed to the formation of hydrogen bonds between the sulfonate group and K299 as well as N307. Notably, **5.11** additionally bears a propargyl group in axial orientation of C1 that enables the conjugation to liposomes.

The SAR of the Man scaffold was further elucidated by screening an existing focused library of Man analogs **12** and **14** bearing substituents in C1 and C6 (Chapter 3.3.2 and 3.3.3). One subset of this library was originally synthesized to target the bacterial adhesions FimH and selected based on the previously reported specificity profile against several human CLRs (Scharenberg *et al.*, 2012). The other subset comprised bioisosteres of Man-6-OS (K_I = 0.8±0.1 mM) and was originally designed as inhibitors of the bacterial adhesin LecB (Hauck *et al.*, 2013). The implemented screening led to the

discovery of **12.11** (K_I = 0.23±0.03 mM) displaying a 44-fold affinity increase over Man. **12.11** bears a biphenyl substituent in C1 of the Man scaffold and the binding mode analysis for the analog suggested the formation of hydrogen bonds with N292 as well as hydrophobic interactions with P310 and A289. Moreover, the introduction of a methylsulfonamide substituent in C6 for **14.1** (K_I = 3.0±0.2 mM) resulted in a 4.3-fold affinity increase over naturally occurring carbohydrates. However, the SAR of the Man scaffold was observed to be non-additive as **14.5** (K_I = 0.25±0.07 mM), bearing substituents in both C1 and C6, did not display an additional affinity increase (Figure 41). Importantly, this approach nevertheless enabled the introduction of an ethylamino linker in C6 of **14.4** to develop a conjugation strategy for liposomal formulations.

The discovery of heparin-derived sulfated GlcNAc derivatives as potent monosaccharide ligands for Langerin led to the implementation of an alternative design strategy (Chapter 3.4.1 and 3.4.2) (Hanske *et al.*, 2017b). Based on the X-ray structure of Langerin in complex with GlcNAc, aromatic substituents in C2 were envisioned to form *e.g.* cation-π or π-π interactions with either P310, K313 or F315. Leveraging the favorable contributions observed for sulfonamide groups, five GlcNS analogs **31** were synthesized. The subsequent determination of K_I values for the focused library revealed **31.2** (K_I = 0.32±0.05 mM) as a potent glycomimetic displaying a 31-fold affinity increase over Man. The introduction of an acetylated ethylamino linker in equatorial orientation of C1 to enable liposomal formulations resulted in additional affinity increase for **41.2** (K_I = 0.24±0.03 mM) (Figure 41). **41.2** bears a para-methyl phenyl substituent in C2. The binding mode analysis for the analog validated the formation of π-π interactions with F315 and suggested the formation of a hydrogen bond between the sulfonamide linker and N307.

In summary, three potent glycomimetic Langerin ligands, *i.e.* **5.11**, **14.4** and **41.1**, bearing propargyl or ethylamino linkers were discovered (Figure 41). Consequently, these ligands represent valuable molecular probes to investigate DC immunology and to develop novel liposome-based cancer immunotherapies. The designed carbohydrate analogs differ with respect to their affinity for Langerin, their carbohydrate scaffold and the utilized conjugation strategy. Additionally, they were demonstrated to form unique interactions with the Ca^{2+}-dependent carbohydrate binding site. These differences are of particular value as they potentially enable the elucidation of the underlying mechanisms determining the efficiency of endocytosis, routing and cross-presentation of TAAs by LCs (Fehres *et al.*, 2014b).

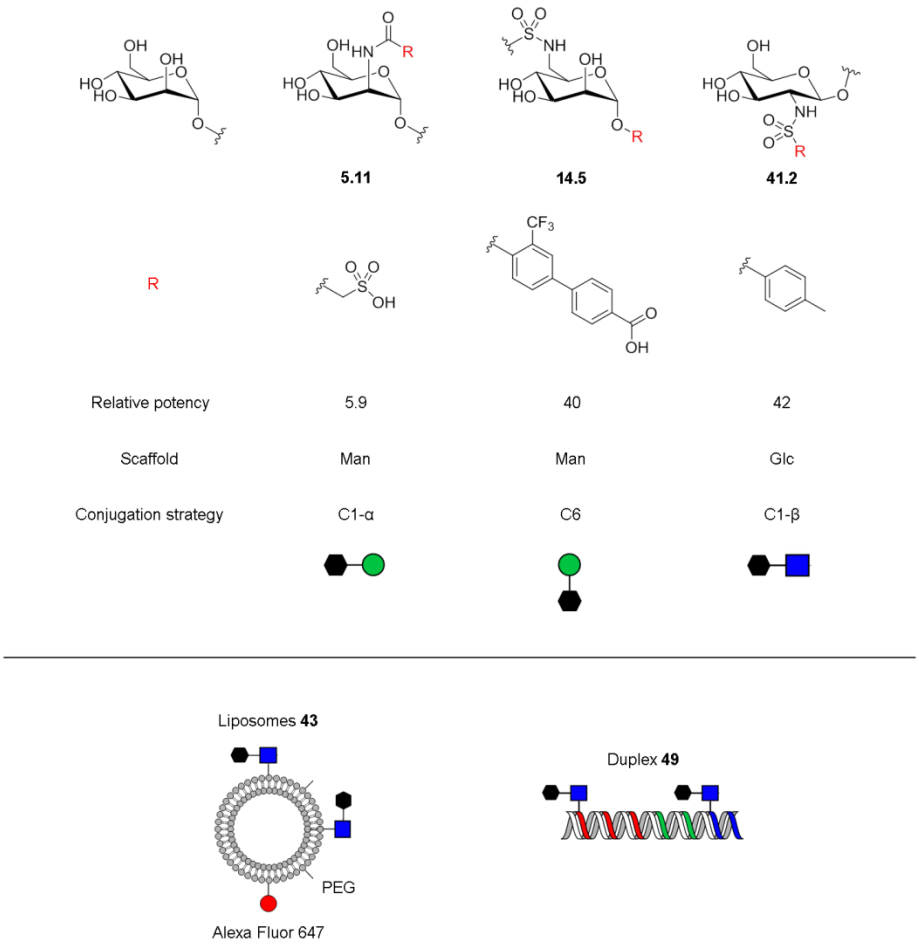


Figure 41. Overview of the designed glycomimetic Langerin ligands.

Three carbohydrate analogs, *i.e.* **5.11** (K_I = 1.3±0.1 mM), **14.5** (K_I = 0.25±0.07 mM) and **41.2** (K_I = 0.24±0.03) were identified as potent Langerin ligands. These analogs display differential affinities and are based on either Man or Glc scaffolds. Moreover, the corresponding reactive intermediates, *i.e.* **5.11**, **14.4** and **41.1**, bear propargyl or ethylamino groups that enable the conjugation to liposomes or other scaffold structures. The linker structures were introduced in C1 or C6 of the respective monosaccharide scaffold with varying stereochemistry. Furthermore, the glycomimetics form distinct interactions with the Ca^{2+}-dependent carbohydrate binding site of Langerin and are likely to display differential specificity profiles for other GBPs. GlcNS analog **41.1** was utilized to develop liposomes specifically targeting LCs and to design potent divalent glycomimetics (IC_{50} = 23±2 µM).

The observed affinity increases for Man analog **14.5** and GlcNS analog **41.2** are comparable to first generation glycomimetics discovered for other CLRs such as DC-SIGN (Garber *et al.*, 2010; Timpano *et al.*, 2008). The design of next generation glycomimetics for Langerin will be guided by ITC experiments to reveal the thermodynamic parameters of the interactions formed by these analogs. Furthermore, X-ray crystallography provides optimized resolution for the validation of the proposed binding modes and will thus enable the development of more sophisticated structure-based *in silico* design approaches. Finally, both **14.4** and **41.1** represent suitable molecular probes for the implementation of integrated solid-phase synthesis and *in situ* screening strategies to efficiently generate next generation carbohydrate analog libraries (Dasgupta *et al.*, 2014; Rillahan *et al.*, 2012; Rillahan *et al.*, 2013)

The multivalent organization of carbohydrates or their synthetic analogs to match the geometry of the Langerin trimer represents an alternative strategy to design potent and specific glycomimetics (Cecioni *et al.*, 2015). Nucleic acid scaffolds provide excellent control over the valency as well as the spatial orientation of organized ligands and are thus ideally suited to leverage cooperativity effects (Fasting *et al.*, 2012; Scheibe *et al.*, 2011). Strikingly, the conjugation of **41.1** to PNA-DNA duplexes to design divalent glycomimetics (IC_{50} = 23±2 µM) resulted in an additional distance-dependent 12-fold avidity increase (Figure 41). In an alternative approach, trivalent Man-bearing glycoclusters (IC_{50} = 0.20±0.08 mM) displayed an 80-fold avidity increase over the corresponding monosaccharide. While these findings are encouraging, observed avidity effects should be validated in orthogonal assays such as SPR or ITC experiments to exclude *e.g.* aggregative false-positive mechanisms. The optimization of the flexibility of the scaffold and linker structures will guide the design of next generation multivalent glycomimetics. Additionally, the trivalent organization of **41.1** on the glycoclusters can be envisioned. Intriguingly, both the PNA-DNA duplexes and the glycoclusters are compatible with liposomal formulations, potentially facilitating the targeted delivery of TAAs to LCs or other DC subsets in the future.

Characterizing Langerin-Ligand Interactions by NMR

The challenges encountered for glycomimetic ligand design extend to the development of suitable assays, particularly during the initial stages of the design process. The low affinities observed for GPIs translate into sensitivity problems and a lack of reporter molecules. This has been compensated for *via* the multivalent organization of glycans and CLRs in heterogeneous assays (Cecioni *et al.*, 2015). For Langerin, heterogeneous assays were utilized to quantify its specificity for mono- and oligosaccharides (Chabrol *et al.*, 2012; Stambach and Taylor, 2003). However, these methods introduce complex multi-state equilibria and surface phenomena, complicating data analysis as well as limiting the transferability of determined affinities between assay formats (Kitov and Bundle, 2003). Alternatively, sensitive biophysical techniques enable the development of homogeneous and monovalent assays. ITC and STD NMR experiments enabled the determination of K_D values in the micro to millimolar range for Langerin-oligosaccharide interactions (Holla and Skerra, 2011; Munoz-Garcia *et al.*, 2015). Notably, both techniques provide non-optimal throughput and material consumption in the context of glycomimetic ligand design.

The limitations outlined above were addressed *via* the development of a monovalent and homogeneous ^{19}F R_2-filtered NMR assay providing excellent sensitivity for the determination of K_I and IC_{50} values (Chapter 3.1). Importantly, the optimization of the assay setup with respect to throughput and material consumption proved instrumental for the discovery of potent carbohydrate analogs and multivalent glycomimetics. Moreover, the optimized setup enabled the implementation of an explorative ^{19}F R_2-filtered NMR fragment screening and the identification of the first non-carbohydrate inhibitor reported for Langerin (Wamhoff *et al.*, 2016). The ^{19}F R_2-filtered NMR assay is

in principle transferrable to other CLRs and was successfully applied to DC-SIGN to evaluate the specificity of designed carbohydrate analogs. Notably, the simulation of binding isotherms at different K_I values indicated that satisfactory sensitivity is maintained for affinities in the micro- to millimolar range. Consequently, alternative assays will be required for the characterization of next generation glycomimetics potentially displaying nanomolar affinities. Here, FP and ITC experiments represent attractive techniques that have been successfully utilized for other GBPs (Han *et al.*, 2010; Pang *et al.*, 2012).

While ^{19}F R_2-filtered NMR experiments served as the primary screening and characterization assay, the affinities of potent carbohydrate analogs were validated by STD and ^{15}N HSQC NMR. This integrated strategy comprises both protein- and ligand-observed experiments to detect GPIs, either directly or *via* competitive binding experiments and is thus ideally suited to exclude false-positive mechanisms such as the aggregation of glycomimetics or the denaturation of CLRs (Gossert and Jahnke, 2016). Furthermore, STD and ^{15}N HSQC NMR experiments complemented by molecular docking studies enabled the analysis of Ca^{2+}-dependent binding modes for selected carbohydrate analogs. Importantly, these investigations guided the structure-based design of glycomimetic Langerin ligands and enabled the development of suitable conjugation strategies for liposomal formulations. As discussed above, the elucidation of X-ray structures for Langerin in complex with **14.5** and **41.2** will be required to validate the proposed binding modes, particularly considering the conformational flexibility of side chains in the carbohydrate binding site (Hanske *et al.*, 2016)[46].

***Ex vivo* Targeting of Liposomes to Langerhans Cells**

As highlighted above, the specificity of glycomimetic targeting ligands for Langerin has important implications for basic research as well as the development of novel cancer immunotherapies. The location of LCs in the epidermis of the human skin enables the topical or dermal administration of vaccines, potentially improving patient care and facilitating TAA delivery (Flacher *et al.*, 2009; Flacher *et al.*, 2010). Under these conditions, targeted liposomes will encounter other GBPs including several CLRs. For instance, keratinocytes express MR while dermal DCs express Dectin-1 (Stoitzner *et al.*, 2014; Szolnoky *et al.*, 2001). Another prominent APC population of the human dermis are DC-$SIGN^+$-$CD14^+$ macrophages (McGovern *et al.*, 2014).

GlcNS analog **41.2** displayed remarkable specificity against DC-SIGN in ^{19}F R_2-filtered NMR experiments. Encouraged by these findings, the reactive intermediate **41.1** was utilized to prepare targeted liposomes **43** (Figure 41) (Chapter 3.4.2 and 3.4.3). Flow cytometry experiments were employed to optimize liposomal formulations and to validate the binding of the liposomes to

[46] The corresponding MD simulations were conducted by Stevan Aleksic (Free Universtiy of Berlin). The findings on side chain dynamics for the carbohydrate binding site were communicated in person and are not included in the cited publication.

Langerin$^+$ model cells *in vitro*. By contrast, no binding was observed for DC-SIGN$^+$ and Dectin-1$^+$ model cells. Finally, the utility of the targeted liposomes for TAA delivery strategies to LCs was evaluated in *ex vivo* flow cytometry experiments with epidermal and whole skin cell suspensions. The suspensions were obtained from healthy humans and comprise physiologically relevant cell populations (Ortner *et al.*, 2017). Liposomal targeting was demonstrated to be specific for LCs and efficiently promoted endocytosis.

These findings render the targeted liposomes **43** an attractive delivery platform enabling the co-formulation of adjuvants such as TLR-3 agonists to promote the activation of cytotoxic T cells (Boks *et al.*, 2015; Fehres *et al.*, 2015a). Consequently, this dissertation constitutes an important advancement towards the development of novel cancer immunotherapies. Moreover, the liposomes might be utilized to induce antigen-specific LC-mediated tolerance for the treatment autoimmune diseases (Seneschal *et al.*, 2012). Future research in our laboratory will focus on the liposomal delivery of antigens to elicit cytotoxic T cell immunity. In this context, the *ex vivo* targeting of EBV-derived antigens to isolated human LCs represents an intriguing strategy to evaluate the cross-priming efficiency for autologous CD8$^+$ T cells (Stoitzner *et al.*, 2010; Taylor *et al.*, 2004). In parallel, confocal microscopy will serve to investigate the endosomal routing and processing of both the targeted liposomes and delivered antigens.

Notably, the transferability of the developed delivery platform to *in vivo* studies is compromised as murine DC subsets display distinct immunological characteristics compared to their human orthologs (Artyomov *et al.*, 2015; Flacher *et al.*, 2014; Idoyaga *et al.*, 2013). Moreover, the glycan specificity of murine and human Langerin is not conserved, potentially translating into a decreased affinity for targeting ligand **41.1** (Hanske *et al.*, 2017a). Hence, the exploration of alternative animal models such as non-human primates might be required to advance the targeted liposomes **43** towards clinical trials. Yet, these studies imply ethical concerns as well as considerable administrative and experimental efforts, particularly as adoptive T cell transfer and tumor challenge models are well-established for mice (Bouwer *et al.*, 2014; Idoyaga *et al.*, 2011).

References

Abergel, C., Chenivesse, S., Stinnakre, M. G., Guasco, S., Brechot, C., Claverie, J. M., Devinoy, E. and Christa, L. Crystallization and preliminary crystallographic study of HIP/PAP, a human C-lectin overexpressed in primary liver cancers. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **1999,** *55* (8), 1487-9.

Accapezzato, D., Visco, V., Francavilla, V., Molette, C., Donato, T., Paroli, M., Mondelli, M. U., Doria, M., Torrisi, M. R. and Barnaba, V. Chloroquine enhances human CD8+ T cell responses against soluble antigens in vivo. *J. Exp. Med.* **2005,** *202* (6), 817-28.

Albert, M. L., Darnell, J. C., Bender, A., Francisco, L. M., Bhardwaj, N. and Darnell, R. B. Tumor-specific killer cells in paraneoplastic cerebellar degeneration. *Nat. Med.* **1998,** *4* (11), 1321-4.

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. and Bernardi, A. Second generation of fucose-based DC-SIGN ligands: affinity improvement and specificity versus Langerin. *Org. Biomol. Chem.* **2011,** *9* (16), 5778-86.

Angulo, J., Enriquez-Navas, P. M. and Nieto, P. M. Ligand-receptor binding affinities from saturation transfer difference (STD) NMR spectroscopy: the binding isotherm of STD initial growth rates. *Chem. Eur. J.* **2010,** *16* (26), 7803-12.

Angulo, J. and Nieto, P. M. STD-NMR: Application to transient interactions between biomolecules - a quantitative approach. *Eur. Biophys. J.* **2011,** *40* (12), 1357-69.

Apostolopoulos, V., Pietersz, G. A., Tsibanis, A., Tsikkinis, A., Drakaki, H., Loveland, B. E., Piddlesden, S. J., Plebanski, M., Pouniotis, D. S., Alexis, M. N., McKenzie, I. F. and Vassilaros, S. Pilot phase III immunotherapy study in early-stage breast cancer patients using oxidized mannan-MUC1. *Breast. Cancer Res.* **2006,** *8* (3), R27.

Appay, V., Douek, D. C. and Price, D. A. CD8+ T cell efficacy in vaccination and disease. *Nat. Med.* **2008,** *14* (6), 623-8.

Aretz, J. *Fragment-based Design of Novel Lectin Ligands.* Freie Universität Berlin, **2017**.

Aretz, J., Baukmann, H., Shanina, E., Hanske, J., Wawrzinek, R., Zapol'skii, V. A., Seeberger, P. H., Kaufmann, D. E. and Rademacher, C. Fragment screening against DC-SIGN identifies multiple druggable secondary sites. *Angew. Chem. Int. Ed.* **2017**, accepted.

Aretz, J., Kondoh, Y., Honda, K., Anumala, U. R., Nazare, M., Watanabe, N., Osada, H. and Rademacher, C. Chemical fragment arrays for rapid druggability assessment. *Chem. Commun.* **2016,** *52* (58), 9067-70.

Aretz, J., Wamhoff, E.-C., Hanske, J., Heymann, D. and Rademacher, C. Computational and experimental prediction of human C-type lectin receptor druggability. *Front. Immunol.* **2014,** *5* (323).

Armon, A., Graur, D. and Ben-Tal, N. ConSurf: an algorithmic tool for the identification of functional regions in proteins by surface mapping of phylogenetic information. *J. Mol. Biol.* **2001,** *307* (1), 447-63.

Artyomov, M. N., Munk, A., Gorvel, L., Korenfeld, D., Cella, M., Tung, T. and Klechevsky, E. Modular expression analysis reveals functional conservation between human Langerhans cells and mouse cross-priming dendritic cells. *J. Exp. Med.* **2015**.

Bakan, A., Nevins, N., Lakdawala, A. S. and Bahar, I. Druggability assessment of allosteric proteins by dynamics simulations in the presence of probe molecules. *J. Chem. Theory Comput.* **2012,** *8* (7), 2435-2447.

Banchereau, J. and Steinman, R. M. Dendritic cells and the control of immunity. *Nature* **1998,** *392* (6673), 245-52.

Barenholz, Y. Doxil(R) - the first FDA-approved nano-drug: lessons learned. *J. Control. Release* **2012,** *160* (2), 117-34.

Barra, P. A., Jimenez, V. A., Gavin, J. A., Daranas, A. H. and Alderete, J. B. Discovery of new E-Selectin inhibitors by virtual screening, fluorescence binding assays and STD NMR experiments. *ChemMedChem* **2016,** *11* (9), 1008-14.

Bauer, J., Bahmer, F. A., Worl, J., Neuhuber, W., Schuler, G. and Fartasch, M. A strikingly constant ratio exists between Langerhans cells and other epidermal cells in human skin. A stereologic study using the optical disector method and the confocal laser scanning microscope. *J. Invest. Dermatol.* **2001,** *116* (2), 313-8.

Belizaire, R. and Unanue, E. R. Targeting proteins to distinct subcellular compartments reveals unique requirements for MHC class I and II presentation. *Proc. Natl. Acad. Sci. U. S. A.* **2009,** *106* (41), 17463-8.

Berthold, M. R., Cebron, N., Dill, F., Gabriel, T. R., Koetter, T., Meinl, T., Ohl, P., Sieb, C., Thiel, K. and Wiswedel, B. KNIME: the Konstanz information miner. In *Studies in Classification, Data Analysis, and Knowledge Organization*, Springer: **2007**.

Bider, M. D., Wahlberg, J. M., Kammerer, R. A. and Spiess, M. The oligomerization domain of the asialoglycoprotein receptor preferentially forms 2:2 heterotetramers in vitro. *J. Biol. Chem.* **1996,** *271* (50), 31996-2001.

Bigley, V., McGovern, N., Milne, P., Dickinson, R., Pagan, S., Cookson, S., Haniffa, M. and Collin, M. Langerin-expressing dendritic cells in human tissues are related to CD1c+ dendritic cells and distinct from Langerhans cells and CD141high XCR1+ dendritic cells. *J. Leukoc. Biol.* **2015,** *97* (4), 627-34.

Binder, F. P., Lemme, K., Preston, R. C. and Ernst, B. Sialyl Lewis(x): a "pre-organized water oligomer"? *Angew. Chem. Int. Ed.* **2012,** *51* (29), 7327-31.

Blomhoff, R., Tolleshaug, H. and Berg, T. Binding of calcium ions to the isolated asialo-glycoprotein receptor - implications for receptor function in suspended hepatocytes. *J. Biol. Chem.* **1982,** *257* (13), 7456-9.

Bodenhausen, G. and Ruben, D. J. Natural abundance nitrogen-15 NMR by enhanced heteronuclear spectroscopy. *Chem. Phys. Lett.* **1980,** *69* (1), 185-189.

Boks, M. A., Ambrosini, M., Bruijns, S. C., Kalay, H., van Bloois, L., Storm, G., Garcia-Vallejo, J. J. and van Kooyk, Y. MPLA incorporation into DC-targeting glycoliposomes favours anti-tumour T cell responses. *J. Control. Release* **2015,** *216*, 37-46.

Bonifaz, L., Bonnyay, D., Mahnke, K., Rivera, M., Nussenzweig, M. C. and Steinman, R. M. Efficient targeting of protein antigen to the dendritic cell receptor DEC-205 in the steady state leads to antigen presentation on major histocompatibility complex class I products and peripheral CD8+ T cell tolerance. *J. Exp. Med.* **2002,** *196* (12), 1627-38.

Borrok, M. J. and Kiessling, L. L. Non-carbohydrate inhibitors of the lectin DC-SIGN. *J. Am. Chem. Soc.* **2007,** *129* (42), 12780-5.

Bouwer, A. L., Saunderson, S. C., Caldwell, F. J., Damani, T. T., Pelham, S. J., Dunn, A. C., Jack, R. W., Stoitzner, P. and McLellan, A. D. NK cells are required for dendritic cell–based immunotherapy at the time of tumor challenge. *J. Immunol.* **2014,** *192* (5), 2514-2521.

Broecker, F. and Seeberger, P. H. Synthetic glycan microarrays. *Methods Mol. Biol.* **2017,** *1518*, 227-240.

Brown, D. and Superti-Furga, G. Rediscovering the sweet spot in drug discovery. *Drug. Discov. Today* **2003,** *8* (23), 1067-77.

Brown, J., O'Callaghan, C. A., Marshall, A. S., Gilbert, R. J., Siebold, C., Gordon, S., Brown, G. D. and Jones, E. Y. Structure of the fungal beta-glucan-binding immune receptor dectin-1: implications for function. *Protein Sci.* **2007,** *16* (6), 1042-52.

Burgdorf, S., Kautz, A., Bohnert, V., Knolle, P. A. and Kurts, C. Distinct pathways of antigen uptake and intracellular routing in CD4 and CD8 T cell activation. *Science* **2007,** *316* (5824), 612-6.

Burgdorf, S., Scholz, C., Kautz, A., Tampe, R. and Kurts, C. Spatial and mechanistic separation of cross-presentation and endogenous antigen presentation. *Nat. Immunol.* **2008,** *9* (5), 558-66.

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., Van Hulst, N. F. and Figdor, C. G. Microdomains of the C-type lectin DC-SIGN are portals for virus entry into dendritic cells. *J. Cell Biol.* **2004,** *164* (1), 145-155.

Carr H. Y. and M., P. E. Effects of diffusion on free precession in nuclear magnetic resonance experiments. *Phys. Rev.* **1954,** *94* (630).

Case, D. A., Darden, T. A., Cheatham, T. E., Simmerling, C. L., Wang, J., Duke, R. E., Luo, R., Crowley, M., R.C.Walker, Zhang, W., Merz, K. M., B.Wang, Hayik, S., Roitberg, A., Seabra, G., Kolossváry, I., K.F.Wong, Paesani, F., Vanicek, J., X.Wu, Brozell, S. R., Steinbrecher, T., Gohlke, H., Yang, L., Tan, C., Mongan, J., Hornak, V., Cui, G., Mathews, D. H., Seetin, M. G., Sagui, C., Babin, V. and P.A. Kollman. *AMBER*, 10. San Francisco, U. S. A., **2008**.

Cecioni, S., Imberty, A. and Vidal, S. Glycomimetics versus multivalent glycoconjugates for the design of high affinity lectin ligands. *Chem. Rev.* **2015,** *115* (1), 525-61.

Chabrol, E., Nurisso, A., Daina, A., Vassal-Stermann, E., Thepaut, M., Girard, E., Vives, R. R. and Fieschi, F. Glycosaminoglycans are interactants of Langerin: comparison with gp120 highlights an unexpected calcium-independent binding mode. *PLoS One* **2012,** *7* (11), 50722.

Chang, J., Patton, J. T., Sarkar, A., Ernst, B., Magnani, J. L. and Frenette, P. S. GMI-1070, a novel pan-selectin antagonist, reverses acute vascular occlusions in sickle cell mice. *Blood* **2010,** *116* (10), 1779-86.

Chappell, C. P., Giltiay, N. V., Draves, K. E., Chen, C., Hayden-Ledbetter, M. S., Shlomchik, M. J., Kaplan, D. H. and Clark, E. A. Targeting antigens through blood dendritic cell antigen 2 (BDCA2) on plasmacytoid dendritic cells promotes immunologic tolerance. *J. Immunol.* **2014,** *192* (12), 5789-5801.

Chatterjee, B., Smed-Sorensen, A., Cohn, L., Chalouni, C., Vandlen, R., Lee, B. C., Widger, J., Keler, T., Delamarre, L. and Mellman, I. Internalization and endosomal degradation of receptor-bound

antigens regulate the efficiency of cross presentation by human dendritic cells. *Blood* **2012,** *120* (10), 2011-20.

ChemAxon. *Marvin*, 14.7.7.0. Budapest, Hungary, **2014**.

Chemical Computing Group. *Molecular Operating Environment*, 2016.08. Montreal, Canada, **2016**.

Chen, I. J. and Hubbard, R. E. Lessons for fragment library design: analysis of output from multiple screening campaigns. *J. Comput. Aided Mol. Des.* **2009,** *23* (8), 603-20.

Chen, V. B., Arendall, W. B., 3rd, Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S. and Richardson, D. C. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. Sect. D-Biol. Crystallogr.* **2010a,** *66* (Pt 1), 12-21.

Chen, W. C., Completo, G. C., Sigal, D. S., Crocker, P. R., Saven, A. and Paulson, J. C. In vivo targeting of B-cell lymphoma with glycan ligands of CD22. *Blood* **2010b,** *115* (23), 4778-86.

Cheng, A. C., Coleman, R. G., Smyth, K. T., Cao, Q., Soulard, P., Caffrey, D. R., Salzberg, A. C. and Huang, E. S. Structure-based maximal affinity model predicts small-molecule druggability. *Nat. Biotechnol.* **2007,** *25* (1), 71-5.

Chi, C. N., Vogeli, B., Bibow, S., Strotz, D., Orts, J., Guntert, P. and Riek, R. A structural ensemble for the enzyme cyclophilin reveals an orchestrated mode of action at atomic resolution. *Angew. Chem. Int. Ed.* **2015,** *54* (40), 11657-61.

Chien, Y. Y., Jan, M. D., Adak, A. K., Tzeng, H. C., Lin, Y. P., Chen, Y. J., Wang, K. T., Chen, C. T., Chen, C. C. and Lin, C. C. Globotriose-functionalized gold nanoparticles as multivalent probes for Shiga-like toxin. *ChemBioChem* **2008,** *9* (7), 1100-9.

Cimermancic, P., Weinkam, P., Rettenmaier, T. J., Bichmann, L., Keedy, D. A., Woldeyes, R. A., Schneidman-Duhovny, D., Demerdash, O. N., Mitchell, J. C., Wells, J. A., Fraser, J. S. and Sali, A. CryptoSite: expanding the druggable proteome by characterization and prediction of cryptic binding sites. *J. Mol. Biol.* **2016,** *428* (4), 709-19.

Ciobanu, M., Huang, K. T., Daguer, J. P., Barluenga, S., Chaloin, O., Schaeffer, E., Mueller, C. G., Mitchell, D. A. and Winssinger, N. Selection of a synthetic glycan oligomer from a library of DNA-templated fragments against DC-SIGN and inhibition of HIV gp120 binding to dendritic cells. *Chem. Commun.* **2011,** *47* (33), 9321-3.

Clausen, B. E. and Stoitzner, P. Functional specialization of skin dendritic cell subsets in regulating T cell responses. *Front. Immunol.* **2015,** *6*, 534.

Cohen, J. I. Epstein-Barr virus infection. *N. Engl. J. Med.* **2000,** *343* (7), 481-92.

Collins, B. E., Blixt, O., Han, S., Duong, B., Li, H., Nathan, J. K., Bovin, N. and Paulson, J. C. High-affinity ligand probes of CD22 overcome the threshold set by cis ligands to allow for binding, endocytosis, and killing of B cells. *J. Immunol.* **2006,** *177* (5), 2994-3003.

Congreve, M., Langmead, C. J., Mason, J. S. and Marshall, F. H. Progress in structure based drug design for G protein-coupled receptors. *J. Med. Chem.* **2011,** *54* (13), 4283-311.

Connolly, M. L. The molecular surface package. *J. Mol. Graph.* **1993,** *11* (2), 139-141.

Coombs, P. J., Graham, S. A., Drickamer, K. and Taylor, M. E. Selective binding of the scavenger receptor C-type lectin to Lewisx trisaccharide and related glycan ligands. *J. Biol. Chem.* **2005,** *280* (24), 22993-9.

Corbeil, C. R., Williams, C. I. and Labute, P. Variability in docking success rates due to dataset preparation. *J. Comput. Aided Mol. Des.* **2012,** *26* (6), 775-786.

Coste, J., Le-Nguyen, D. and Castro, B. PyBOP®: a new peptide coupling reagent devoid of toxic by-product. *Tetrahedron Lett.* **1990,** *31* (2), 205-208.

D'Souza, A. A. and Devarajan, P. V. Asialoglycoprotein receptor mediated hepatocyte targeting - strategies and applications. *J. Control. Release* **2015,** *203*, 126-39.

Dalvit, C. Ligand- and substrate-based 19F NMR screening: Principles and applications to drug discovery. *Prog. Nucl. Mag. Res. Sp.* **2007,** *51* (4), 243-271.

Dalvit, C., Mongelli, N., Papeo, G., Giordano, P., Veronesi, M., Moskau, D. and Kummerle, R. Sensitivity improvement in 19F NMR-based screening experiments: theoretical considerations and experimental applications. *J. Am. Chem. Soc.* **2005,** *127* (38), 13380-5.

Danby, P. M. and Withers, S. G. Advances in enzymatic glycoside synthesis. *ACS Chem. Biol.* **2016,** *11* (7), 1784-94.

Dasgupta, S., Kitov, P. I., Sadowska, J. M. and Bundle, D. R. Discovery of inhibitors of Shiga Toxin Type 2 by on-plate generation and screening of a focused compound library. *Angew. Chem. Int. Ed.* **2014,** *53* (6), 1510-5.

Davis, B. J. and Erlanson, D. A. Learning from our mistakes: the 'unknown knowns' in fragment screening. *Bioorg. Med. Chem. Lett.* **2013,** *23* (10), 2844-52.

de Jong, M. A., de Witte, L., Taylor, M. E. and Geijtenbeek, T. B. Herpes simplex virus type 2 enhances HIV-1 susceptibility by affecting Langerhans cell function. *J. Immunol.* **2010a,** *185* (3), 1633-41.

de Jong, M. A., Vriend, L. E., Theelen, B., Taylor, M. E., Fluitsma, D., Boekhout, T. and Geijtenbeek, T. B. C-type lectin Langerin is a beta-glucan receptor on human Langerhans cells that recognizes opportunistic and pathogenic fungi. *Mol. Immunol.* **2010b,** *47* (6), 1216-25.

de Witte, L., Nabatov, A., Pion, M., Fluitsma, D., de Jong, M. A., de Gruijl, T., Piguet, V., van Kooyk, Y. and Geijtenbeek, T. B. Langerin is a natural barrier to HIV-1 transmission by Langerhans cells. *Nat. Med.* **2007,** *13* (3), 367-71.

Decout, A., Silva-Gomes, S., Drocourt, D., Barbe, S., Andre, I., Cueto, F. J., Lioux, T., Sancho, D., Perouzel, E., Vercellone, A., Prandi, J., Gilleron, M., Tiraby, G. and Nigou, J. Rational design of adjuvants targeting the C-type lectin Mincle. *Proc. Natl. Acad. Sci. U. S. A.* **2017,** *114* (10), 2675-2680.

Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J. and Bax, A. NMRPipe: A multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* **1995,** *6* (3), 277-293.

Delamarre, L., Pack, M., Chang, H., Mellman, I. and Trombetta, E. S. Differential lysosomal proteolysis in antigen-presenting cells determines antigen fate. *Science* **2005,** *307* (5715), 1630-4.

DeMarco, M. L. and Woods, R. J. Structural glycobiology: a game of snakes and ladders. *Glycobiology* **2008,** *18* (6), 426-40.

Demers, J.-P. and Mittermaier, A. Binding mechanism of an SH3 domain studied by NMR and ITC. *J. Am. Chem. Soc.* **2009,** *131* (12), 4355-4367.

Dhodapkar, M. V., Sznol, M., Zhao, B., Wang, D., Carvajal, R. D., Keohan, M. L., Chuang, E., Sanborn, R. E., Lutzky, J., Powderly, J., Kluger, H., Tejwani, S., Green, J., Ramakrishna, V., Crocker, A., Vitale, L., Yellin, M., Davis, T. and Keler, T. Induction of antigen-specific immunity with a vaccine targeting NY-ESO-1 to the dendritic cell receptor DEC-205. *Sci. Transl. Med.* **2014,** *6* (232), 232-51.

Dominguez, C., Boelens, R. and Bonvin, A. M. HADDOCK: a protein-protein docking approach based on biochemical or biophysical information. *J. Am. Chem. Soc.* **2003,** *125* (7), 1731-7.

Dong, H., Strome, S. E., Salomao, D. R., Tamura, H., Hirano, F., Flies, D. B., Roche, P. C., Lu, J., Zhu, G., Tamada, K., Lennon, V. A., Celis, E. and Chen, L. Tumor-associated B7-H1 promotes T-cell apoptosis: a potential mechanism of immune evasion. *Nat. Med.* **2002,** *8* (8), 793-800.

Drickamer, K. Engineering galactose-binding activity into a C-type mannose-binding protein. *Nature* **1992,** *360* (6400), 183-6.

Drickamer, K. and Taylor, M. E. Recent insights into structures and functions of C-type lectins in the immune system. *Curr. Opin. Struct. Biol.* **2015,** *34*, 26-34.

Drummond, R. A. and Brown, G. D. The role of Dectin-1 in the host defence against fungal infections. *Curr. Opin. Microbiol.* **2011,** *14* (4), 392-399.

Dumitrescu, L., Eppe, G., Tikad, A., Pan, W., Bkassiny, S. E., Gurcha, S. S., Ardá, A., Jiménez-Barbero, J., Besra, G. S. and Vincent, S. P. Selectfluor and NFSI exo-glycal fluorination strategies applied to the enhancement of the binding affinity of galactofuranosyltransferase GlfT2 inhibitors. *Chem. Eur. J.* **2014,** *20* (46), 15208-15215.

Duus, J., Gotfredsen, C. H. and Bock, K. Carbohydrate structural determination by NMR spectroscopy: modern methods and limitations. *Chem. Rev.* **2000,** *100* (12), 4589-614.

Dzionek, A., Sohma, Y., Nagafune, J., Cella, M., Colonna, M., Facchetti, F., Gunther, G., Johnston, I., Lanzavecchia, A., Nagasaka, T., Okada, T., Vermi, W., Winkels, G., Yamamoto, T., Zysk, M., Yamaguchi, Y. and Schmitz, J. BDCA-2, a novel plasmacytoid dendritic cell-specific type II C-type lectin, mediates antigen capture and is a potent inhibitor of interferon alpha/beta induction. *J. Exp. Med.* **2001,** *194* (12), 1823-34.

Ebner, S., Ehammer, Z., Holzmann, S., Schwingshackl, P., Forstner, M., Stoitzner, P., Huemer, G. M., Fritsch, P. and Romani, N. Expression of C-type lectin receptors by subsets of dendritic cells in human skin. *Int. Immunol.* **2004,** *16* (6), 877-87.

Edfeldt, F. N., Folmer, R. H. and Breeze, A. L. Fragment screening to predict druggability (ligandability) and lead discovery success. *Drug Discov. Today* **2011,** *16* (7-8), 284-7.

Egger, J., Weckerle, C., Cutting, B., Schwardt, O., Rabbani, S., Lemme, K. and Ernst, B. Nanomolar E-selectin antagonists with prolonged half-lives by a fragment-based approach. *J. Am. Chem. Soc.* **2013,** *135* (26), 9820-8.

Egner, U. and Hillig, R. C. A structural biology view of target drugability. *Expert. Opin. Drug. Discov.* **2008,** *3* (4), 391-401.

Eid, S., Saleh, N., Zalewski, A. and Vedani, A. Exploring the free-energy landscape of carbohydrate–protein complexes: development and validation of scoring functions considering the binding-site topology. *J. Comput. Aided Mol. Des.* **2014,** *28* (12), 1191-1204.

Eldridge, M. D., Murray, C. W., Auton, T. R., Paolini, G. V. and Mee, R. P. Empirical scoring functions: I. the development of a fast empirical scoring function to estimate the binding affinity of ligands in receptor complexes. *J. Comput. Aided Mol. Des.* **1997,** *11* (5), 425-45.

Ellervik, U. and Magnusson, G. Glycosylation with N-Troc-protected glycosyl donors. *Carbohydr. Res.* **1996,** *280* (2), 251-60.

Elliott, S., Lorenzini, T., Asher, S., Aoki, K., Brankow, D., Buck, L., Busse, L., Chang, D., Fuller, J., Grant, J., Hernday, N., Hokum, M., Hu, S., Knudten, A., Levin, N., Komorowski, R., Martin, F., Navarro, R., Osslund, T., Rogers, G., Rogers, N., Trail, G. and Egrie, J. Enhancement of therapeutic protein in vivo activities through glycoengineering. *Nat. Biotechnol.* **2003,** *21* (4), 414-21.

Engering, A., Geijtenbeek, T. B., van Vliet, S. J., Wijers, M., van Liempt, E., Demaurex, N., Lanzavecchia, A., Fransen, J., Figdor, C. G., Piguet, V. and van Kooyk, Y. The dendritic cell-specific adhesion receptor DC-SIGN internalizes antigen for presentation to T cells. *J. Immunol.* **2002,** *168* (5), 2118-26.

Engering, A. J., Cella, M., Fluitsma, D., Brockhaus, M., Hoefsmit, E. C., Lanzavecchia, A. and Pieters, J. The mannose receptor functions as a high capacity and broad specificity antigen receptor in human dendritic cells. *Eur. J. Immunol.* **1997,** *27* (9), 2417-25.

Eriksson, M. and Nielsen, P. E. Solution structure of a peptide nucleic acid-DNA duplex. *Nat. Struct. Biol.* **1996,** *3* (5), 410-3.

Erlanson, D. A., Fesik, S. W., Hubbard, R. E., Jahnke, W. and Jhoti, H. Twenty years on: the impact of fragments on drug discovery. *Nat. Rev. Drug Discov.* **2016,** *15* (9), 605-619.

Ernst, B. and Magnani, J. L. From carbohydrate leads to glycomimetic drugs. *Nat. Rev. Drug. Discov.* **2009,** *8* (8), 661-77.

Fasting, C., Schalley, C. A., Weber, M., Seitz, O., Hecht, S., Koksch, B., Dernedde, J., Graf, C., Knapp, E.-W. and Haag, R. Multivalency as a chemical organization and action principle. *Angew. Chem. Int. Ed.* **2012,** *51* (42), 10472-10498.

Feasley, C. L., Johnson, J. M., West, C. M. and Chia, C. P. Glycopeptidome of a heavily N-glycosylated cell surface glycoprotein of Dictyostelium implicated in cell adhesion. *J. Proteome Res.* **2010,** *9* (7), 3495-3510.

Fehres, C. M., Bruijns, S. C., Sotthewes, B. N., Kalay, H., Schaffer, L., Head, S. R., de Gruijl, T. D., Garcia-Vallejo, J. J. and van Kooyk, Y. Phenotypic and functional properties of human steady state CD14+ and CD1a+ antigen presenting cells and epidermal Langerhans cells. *PLoS One* **2015a,** *10* (11), e0143519.

Fehres, C. M., Bruijns, S. C., van Beelen, A. J., Kalay, H., Ambrosini, M., Hooijberg, E., Unger, W. W., de Gruijl, T. D. and van Kooyk, Y. Topical rather than intradermal application of the TLR7 ligand imiquimod leads to human dermal dendritic cell maturation and CD8+ T-cell cross-priming. *Eur. J. Immunol.* **2014a,** *44* (8), 2415-24.

Fehres, C. M., Duinkerken, S., Bruijns, S. C. M., Kalay, H., van Vliet, S. J., Ambrosini, M., de Gruijl, T. D., Unger, W. W. J., Garcia-Vallejo, J. J. and van Kooyk, Y. Langerin-mediated internalization of a modified peptide routes antigens to early endosomes and enhances cross-presentation by human Langerhans cells. *Cell. Mol. Immunol.* **2015b**.

Fehres, C. M., Kalay, H., Bruijns, S. C., Musaafir, S. A., Ambrosini, M., van Bloois, L., van Vliet, S. J., Storm, G., Garcia-Vallejo, J. J. and van Kooyk, Y. Cross-presentation through Langerin and DC-SIGN targeting requires different formulations of glycan-modified antigens. *J. Control. Release* **2015c,** *203*, 67-76.

Fehres, C. M., Unger, W. W., Garcia-Vallejo, J. J. and van Kooyk, Y. Understanding the biology of antigen cross-presentation for the design of vaccines against cancer. *Front. Immunol.* **2014b,** *5*, 149.

Fehres, C. M., van Beelen, A. J., Bruijns, S. C., Ambrosini, M., Kalay, H., van Bloois, L., Unger, W. W., Garcia-Vallejo, J. J., Storm, G., de Gruijl, T. D. and van Kooyk, Y. In situ delivery of antigen to DC-SIGN(+)CD14(+) dermal dendritic cells results in enhanced CD8(+) T cell responses. *J. Invest. Dermatol.* **2015d,** *135* (9), 2228-36.

Feinberg, H., Guo, Y., Mitchell, D. A., Drickamer, K. and Weis, W. I. Extended neck regions stabilize tetramers of the receptors DC-SIGN and DC-SIGNR. *J. Biol. Chem.* **2005,** *280* (2), 1327-35.

Feinberg, H., Mitchell, D. A., Drickamer, K. and Weis, W. I. Structural basis for selective recognition of oligosaccharides by DC-SIGN and DC-SIGNR. *Science* **2001,** *294* (5549), 2163-6.

Feinberg, H., Park-Snyder, S., Kolatkar, A. R., Heise, C. T., Taylor, M. E. and Weis, W. I. Structure of a C-type carbohydrate recognition domain from the macrophage mannose receptor. *J. Biol. Chem.* **2000,** *275* (28), 21539-48.

Feinberg, H., Powlesland, A. S., Taylor, M. E. and Weis, W. I. Trimeric structure of Langerin. *J. Biol. Chem.* **2010,** *285* (17), 13285-93.

Feinberg, H., Rowntree, T. J., Tan, S. L., Drickamer, K., Weis, W. I. and Taylor, M. E. Common polymorphisms in human Langerin change specificity for glycan ligands. *J. Biol. Chem.* **2013,** *288* (52), 36762-71.

Feinberg, H., Taylor, M. E., Razi, N., McBride, R., Knirel, Y. A., Graham, S. A., Drickamer, K. and Weis, W. I. Structural basis for Langerin recognition of diverse pathogen and mammalian glycans through a single binding site. *J. Mol. Biol.* **2011,** *405* (4), 1027-39.

Ferguson, D. M. and Raber, D. J. A new approach to probing conformational space with molecular mechanics: random incremental pulse search. *J. Am. Chem. Soc.* **1989,** *111* (12), 4371-4378.

Fiege, B., Rabbani, S., Preston, R. C., Jakob, R. P., Zihlmann, P., Schwardt, O., Jiang, X., Maier, T. and Ernst, B. The tyrosine gate of the bacterial lectin FimH: a conformational analysis by NMR spectroscopy and X-ray crystallography. *ChemBioChem* **2015,** *16* (8), 1235-46.

Fiege, B., Rademacher, C., Cartmell, J., Kitov, P. I., Parra, F. and Peters, T. Molecular details of the recognition of blood group antigens by a human norovirus as determined by STD NMR spectroscopy. *Angew. Chem. Int. Ed.* **2012,** *51* (4), 928-932.

Fielding, L. NMR methods for the determination of protein–ligand dissociation constants. *Prog. Nucl. Mag. Res. Sp.* **2007,** *51* (4), 219-242.

Flacher, V., Sparber, F., Tripp, C. H., Romani, N. and Stoitzner, P. Targeting of epidermal Langerhans cells with antigenic proteins: attempts to harness their properties for immunotherapy. *Cancer Immunol. Immunother.* **2009,** *58* (7), 1137-47.

Flacher, V., Tripp, C. H., Mairhofer, D. G., Steinman, R. M., Stoitzner, P., Idoyaga, J. and Romani, N. Murine Langerin+ dermal dendritic cells prime CD8+ T cells while Langerhans cells induce cross-tolerance. *EMBO Mol. Med.* **2014,** *6* (9), 1191-204.

Flacher, V., Tripp, C. H., Stoitzner, P., Haid, B., Ebner, S., Del Frari, B., Koch, F., Park, C. G., Steinman, R. M., Idoyaga, J. and Romani, N. Epidermal Langerhans cells rapidly capture and present antigens from C-type lectin-targeting antibodies deposited in the dermis. *J. Invest. Dermatol.* **2010,** *130* (3), 755-62.

Flynn, B. J., Kastenmuller, K., Wille-Reece, U., Tomaras, G. D., Alam, M., Lindsay, R. W., Salazar, A. M., Perdiguero, B., Gomez, C. E., Wagner, R., Esteban, M., Park, C. G., Trumpfheller, C., Keler, T., Pantaleo, G., Steinman, R. M. and Seder, R. Immunization with HIV Gag targeted to dendritic cells followed by recombinant New York vaccinia virus induces robust T-cell immunity in nonhuman primates. *Proc. Natl. Acad. Sci. U. S. A.* **2011,** *108* (17), 7131-6.

Frank, M. Computational docking as a tool for the rational design of carbohydrate-based drugs. In *Carbohydrates as Drugs*, Springer: **2014**, Vol. 12, 53-72.

Freeman, G. J., Wherry, E. J., Ahmed, R. and Sharpe, A. H. Reinvigorating exhausted HIV-specific T cells via PD-1-PD-1 ligand blockade. *J. Exp. Med.* **2006,** *203* (10), 2223-7.

Freeze, H. H. and Wolgast, D. Structural analysis of N-linked oligosaccharides from glycoproteins secreted by Dictyostelium discoideum. Identification of mannose 6-sulfate. *J. Biol. Chem.* **1986,** *261* (1), 127-34.

Friesner, R. A., Banks, J. L., Murphy, R. B., Halgren, T. A., Klicic, J. J., Mainz, D. T., Repasky, M. P., Knoll, E. H., Shelley, M., Perry, J. K., Shaw, D. E., Francis, P. and Shenkin, P. S. Glide: a new approach for rapid, accurate docking and scoring: 1. method and assessment of docking accuracy. *J. Med. Chem.* **2004,** *47* (7), 1739-49.

Furukawa, A., Kamishikiryo, J., Mori, D., Toyonaga, K., Okabe, Y., Toji, A., Kanda, R., Miyake, Y., Ose, T., Yamasaki, S. and Maenaka, K. Structural analysis for glycolipid recognition by the C-type lectins Mincle and MCL. *Proc. Natl. Acad. Sci. U. S. A.* **2013,** *110* (43), 17438-43.

Garber, K. C., Wangkanont, K., Carlson, E. E. and Kiessling, L. L. A general glycomimetic strategy yields non-carbohydrate inhibitors of DC-SIGN. *Chem. Commun.* **2010,** *46* (36), 6747-9.

Gardner, A. and Ruffell, B. Dendritic cells and cancer immunity. *Trends Immunol.* **2016,** *37* (12), 855-865.

Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M. R., Appel, R. D. and Bairoch, A. Protein identification and analysis tools on the ExPASy server. In *The Proteomics Protocols Handbook*, Walker, J. M., Ed. Humana Press: **2005**, 571-607.

Geijtenbeek, T. B. and Gringhuis, S. I. Signalling through C-type lectin receptors: shaping immune responses. *Nat. Rev. Immunol.* **2009,** *9* (7), 465-79.

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. DC-SIGN, a dendritic cell–specific HIV-1-binding protein that enhances trans-infection of T cells. *Cell* **2000,** *100* (5), 587-597.

Gerbaud, V., Pignol, D., Loret, E., Bertrand, J. A., Berland, Y., Fontecilla-Camps, J. C., Canselier, J. P., Gabas, N. and Verdier, J. M. Mechanism of calcite crystal growth inhibition by the N-terminal undecapeptide of lithostathine. *J. Biol. Chem.* **2000,** *275* (2), 1057-64.

Gerber, P. R. and Müller, K. MAB, a generally applicable molecular force field for structure modelling in medicinal chemistry. *J. Comput. Aided Mol. Des.* **1995,** *9* (3), 251-68.

Gerner, M. Y., Casey, K. A. and Mescher, M. F. Defective MHC class II presentation by dendritic cells limits CD4 T cell help for antitumor CD8 T cell responses. *J. Immunol.* **2008,** *181* (1), 155-64.

Gidon, A., Bardin, S., Cinquin, B., Boulanger, J., Waharte, F., Heliot, L., de la Salle, H., Hanau, D., Kervrann, C., Goud, B. and Salamero, J. A Rab11A/Myosin Vb/Rab11-FIP2 complex frames two late recycling steps of Langerin from the ERC to the plasma membrane. *Traffic* **2012,** *13* (6), 815-833.

Goodford, P. J. A computational procedure for determining energetically favorable binding sites on biologically important macromolecules. *J. Med. Chem.* **1985,** *28* (7), 849-857.

Gossert, A. D. and Jahnke, W. NMR in drug discovery: a practical guide to identification and validation of ligands interacting with biological macromolecules. *Prog. Nucl. Magn. Reson. Spectrosc.* **2016,** *97*, 82-125.

Grant, O. C., Smith, H. M., Firsova, D., Fadda, E. and Woods, R. J. Presentation, presentation, presentation! Molecular-level insight into linker effects on glycan array screening data. *Glycobiology* **2014,** *24* (1), 17-25.

Graves, B. J., Crowther, R. L., Chandran, C., Rumberger, J. M., Li, S., Huang, K. S., Presky, D. H., Familletti, P. C., Wolitzky, B. A. and Burns, D. K. Insight into E-selectin/ligand interaction from the crystal structure and mutagenesis of the lec/EGF domains. *Nature* **1994,** *367* (6463), 532-8.

Grim, J. C., Garber, K. C. A. and Kiessling, L. L. Glycomimetic building blocks: a divergent synthesis of epimers of shikimic acid. *Org. Lett.* **2011,** *13* (15), 3790-3793.

Gringhuis, S. I., den Dunnen, J., Litjens, M., van der Vlist, M. and Geijtenbeek, T. B. Carbohydrate-specific signaling through the DC-SIGN signalosome tailors immunity to Mycobacterium tuberculosis, HIV-1 and Helicobacter pylori. *Nat. Immunol.* **2009a,** *10* (10), 1081-8.

Gringhuis, S. I., den Dunnen, J., Litjens, M., van der Vlist, M., Wevers, B., Bruijns, S. C. and Geijtenbeek, T. B. Dectin-1 directs T helper cell differentiation by controlling noncanonical NF-kappaB activation through Raf-1 and Syk. *Nat. Immunol.* **2009b,** *10* (2), 203-13.

Gringhuis, S. I., Kaptein, T. M., Wevers, B. A., Mesman, A. W. and Geijtenbeek, T. B. Fucose-specific DC-SIGN signalling directs T helper cell type-2 responses via IKKepsilon- and CYLD-dependent Bcl3 activation. *Nat. Commun.* **2014,** *5*, 3898.

Gromme, M., Uytdehaag, F. G., Janssen, H., Calafat, J., van Binnendijk, R. S., Kenter, M. J., Tulp, A., Verwoerd, D. and Neefjes, J. Recycling MHC class I molecules and endosomal peptide loading. *Proc. Natl. Acad. Sci. U. S. A.* **1999,** *96* (18), 10326-31.

Guermonprez, P., Saveanu, L., Kleijmeer, M., Davoust, J., Van Endert, P. and Amigorena, S. ER-phagosome fusion defines an MHC class I cross-presentation compartment in dendritic cells. *Nature* **2003,** *425* (6956), 397-402.

Guo, Y., Feinberg, H., Conroy, E., Mitchell, D. A., Alvarez, R., Blixt, O., Taylor, M. E., Weis, W. I. and Drickamer, K. Structural basis for distinct ligand-binding and targeting properties of the receptors DC-SIGN and DC-SIGNR. *Nat. Struct. Mol. Biol.* **2004,** *11* (7), 591-8.

Hahm, H. S., Schlegel, M. K., Hurevich, M., Eller, S., Schuhmacher, F., Hofmann, J., Pagel, K. and Seeberger, P. H. Automated glycan assembly using the Glyconeer 2.1 synthesizer. *Proc. Natl. Acad. Sci. U. S. A.* **2017**.

Hajduk, P. J., Huth, J. R. and Fesik, S. W. Druggability indices for protein targets derived from NMR-based screening data. *J. Med. Chem.* **2005a,** *48* (7), 2518-25.

Hajduk, P. J., Huth, J. R. and Tse, C. Predicting protein druggability. *Drug Discov. Today* **2005b,** *10* (23-24), 1675-1682.

Hajduk, P. J., Olejniczak, E. T. and Fesik, S. W. One-dimensional relaxation- and diffusion-edited NMR methods for screening compounds that bind to macromolecules. *J. Am. Chem. Soc.* **1997,** *119* (50), 12257-12261.

Halgren, T. A. Merck molecular force field. I. basis, form, scope, parameterization, and performance of MMFF94. *J. Comput. Chem.* **1996,** *17* (5-6), 490-519.

Halgren, T. A. Identifying and characterizing binding sites and assessing druggability. *J. Chem. Inf. Model.* **2009,** *49* (2), 377-89.

Han, Z., Pinkner, J. S., Ford, B., Obermann, R., Nolan, W., Wildman, S. A., Hobbs, D., Ellenberger, T., Cusumano, C. K., Hultgren, S. J. and Janetka, J. W. Structure-based drug design and optimization of mannoside bacterial FimH antagonists. *J. Med. Chem.* **2010,** *53* (12), 4779-4792.

Hanahan, D. and Weinberg, R. A. Hallmarks of cancer: the next generation. *Cell* **2011,** *144* (5), 646-74.

Hann, M. M. Molecular obesity, potency and other addictions in drug discovery. *Med. Chem. Commun.* **2011,** *2* (5), 349-355.

Hann, M. M., Leach, A. R. and Harper, G. Molecular complexity and its impact on the probability of finding leads for drug discovery. *J. Chem. Inf. Comput. Sci.* **2001,** *41* (3), 856-64.

Hanske, J. *Investigation of the structural basis of ligand recognition for the C-type lectin receptor Langerin.* Freie Universität Berlin, **2016**.

Hanske, J., Aleksic, S., Ballaschk, M., Jurk, M., Shanina, E., Beerbaum, M., Schmieder, P., Keller, B. G. and Rademacher, C. Intradomain allosteric network modulates calcium affinity of the C-type lectin receptor Langerin. *J. Am. Chem. Soc.* **2016,** *138* (37), 12176-86.

Hanske, J., Schulze, J., Aretz, J., McBride, R., Loll, B., Schmidt, H., Knirel, Y., Rabsch, W., Wahl, M. C., Paulson, J. C. and Rademacher, C. Bacterial polysaccharide specificity of the pattern recognition receptor Langerin is highly species dependent. *J. Biol. Chem.* **2017a,** *292* (3), 862-871.

Hanske, J., Wawrzinek, R., Geissner, A., Wamhoff, E.-C., Sellrie, K., Schmidt, H., Seeberger, P. H. and Rademacher, C. Calcium-independent activation of an allosteric network in Langerin by heparin oligosaccharides. *ChemBioChem* **2017b**, accepted.

Harding, F. A., Stickler, M. M., Razo, J. and DuBridge, R. B. The immunogenicity of humanized and fully human antibodies: residual immunogenicity resides in the CDR regions. *MAbs* **2010,** *2* (3), 256-65.

Harner, M. J., Frank, A. O. and Fesik, S. W. Fragment-based drug discovery using NMR spectroscopy. *J. Biomol. NMR* **2013,** *56* (2), 65-75.

Harris, R., Kiddle, G. R., Field, R. A., Milton, M. J., Ernst, B., Magnani, J. L. and Homans, S. W. Stable-isotope-assisted NMR studies on 13C-enriched sialyl LewisX in solution and bound to E-Selectin. *J. Am. Chem. Soc.* **1999,** *121* (11), 2546-2551.

Hasegawa, T., Numata, M., Okumura, S., Kimura, T., Sakurai, K. and Shinkai, S. Carbohydrate-appended curdlans as a new family of glycoclusters with binding properties both for a polynucleotide and lectins. *Org. Biomol. Chem.* **2007,** *5* (15), 2404-12.

Hauck, D., Joachim, I., Frommeyer, B., Varrot, A., Philipp, B., Moller, H. M., Imberty, A., Exner, T. E. and Titz, A. Discovery of two classes of potent glycomimetic inhibitors of Pseudomonas aeruginosa LecB with distinct binding modes. *ACS Chem. Biol.* **2013,** *8* (8), 1775-84.

Hayes, W., Osborn, H. M. I., Osborne, S. D., Rastall, R. A. and Romagnoli, B. One-pot synthesis of multivalent arrays of mannose mono- and disaccharides. *Tetrahedron* **2003,** *59* (40), 7983-7996.

Holla, A. and Skerra, A. Comparative analysis reveals selective recognition of glycans by the dendritic cell receptors DC-SIGN and Langerin. *Protein Eng. Des. Sel.* **2011,** *24* (9), 659-69.

Hopkins, A. L. and Groom, C. R. The druggable genome. *Nat. Rev. Drug. Discov.* **2002,** *1* (9), 727-30.

Hopkins, A. L., Keseru, G. M., Leeson, P. D., Rees, D. C. and Reynolds, C. H. The role of ligand efficiency metrics in drug discovery. *Nat. Rev. Drug. Discov.* **2014,** *13* (2), 105-121.

Huang, N. and Shoichet, B. K. Exploiting ordered waters in molecular docking. *J. Med. Chem.* **2008,** *51* (16), 4862-5.

Huang, X., Leroux, J. C. and Castagner, B. Well-defined multivalent ligands for hepatocytes targeting via asialoglycoprotein receptor. *Bioconjug. Chem.* **2017,** *28* (2), 283-295.

Hudson, K. L., Bartlett, G. J., Diehl, R. C., Agirre, J., Gallagher, T., Kiessling, L. L. and Woolfson, D. N. Carbohydrate-aromatic interactions in proteins. *J. Am. Chem. Soc.* **2015,** *137* (48), 15152-60.

Hudson, S. A., McLean, K. J., Surade, S., Yang, Y. Q., Leys, D., Ciulli, A., Munro, A. W. and Abell, C. Application of fragment screening and merging to the discovery of inhibitors of the Mycobacterium tuberculosis cytochrome P450 CYP121. *Angew. Chem. Int. Ed.* **2012,** *51* (37), 9311-6.

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. and Modlin, R. L. Langerhans cells utilize CD1a and Langerin to efficiently present nonpeptide antigens to T cells. *J. Clin. Invest.* **2004,** *113* (5), 701-708.

Hunter, C. A. and Anderson, H. L. What is cooperativity? *Angew. Chem. Int. Ed.* **2009,** *48* (41), 7488-7499.

Huysamen, C., Willment, J. A., Dennehy, K. M. and Brown, G. D. CLEC9A is a novel activation C-type lectin-like receptor expressed on BDCA3+ dendritic cells and a subset of monocytes. *J. Biol. Chem.* **2008,** *283* (24), 16693-701.

Hwang, T. L. and Shaka, A. J. Water suppression that works - excitation sculpting using arbitrary wave-forms and pulsed-field gradients. *J. Magn. Reson., Ser. A* **1995,** *112* (2), 275-279.

Idoyaga, J., Fiorese, C., Zbytnuik, L., Lubkin, A., Miller, J., Malissen, B., Mucida, D., Merad, M. and Steinman, R. M. Specialized role of migratory dendritic cells in peripheral tolerance induction. *J. Clin. Invest.* **2013,** *123* (2), 844-54.

Idoyaga, J., Lubkin, A., Fiorese, C., Lahoud, M. H., Caminschi, I., Huang, Y., Rodriguez, A., Clausen, B. E., Park, C. G., Trumpfheller, C. and Steinman, R. M. Comparable T helper 1 (Th1) and CD8 T-cell immunity by targeting HIV gag p24 to CD8 dendritic cells within antibodies to Langerin, DEC205, and Clec9A. *Proc. Natl. Acad. Sci. U. S. A.* **2011,** *108* (6), 2384-2389.

Idoyaga, J., Suda, N., Suda, K., Park, C. G. and Steinman, R. M. Antibody to Langerin/CD207 localizes large numbers of CD8alpha+ dendritic cells to the marginal zone of mouse spleen. *Proc. Natl. Acad. Sci. U. S. A.* **2009,** *106* (5), 1524-9.

Jahn, P. S., Zanker, K. S., Schmitz, J. and Dzionek, A. BDCA-2 signaling inhibits TLR-9-agonist-induced plasmacytoid dendritic cell activation and antigen presentation. *Cell. Immunol.* **2010,** *265* (1), 15-22.

Jahnke, W., Kolb, H. C., Blommers, M. J. J., Magnani, J. L. and Ernst, B. Comparison of the bioactive conformations of sialyl LewisX and a potent sialyl LewisX mimic. *Angew. Chem. Int. Ed.* **1997,** *36* (23), 2603-2607.

Jayalakshmi, V. and Rama Krishna, N. CORCEMA refinement of the bound ligand conformation within the protein binding pocket in reversibly forming weak complexes using STD-NMR intensities. *J. Magn. Reson.* **2004,** *168* (1), 36-45.

Jegouzo, S. A., Feinberg, H., Dungarwalla, T., Drickamer, K., Weis, W. I. and Taylor, M. E. A Novel Mechanism for Binding of Galactose-terminated Glycans by the C-type Carbohydrate Recognition Domain in Blood Dendritic Cell Antigen 2. *J. Biol. Chem.* **2015,** *290* (27), 16759-71.

Jencks, W. P. On the attribution and additivity of binding energies. *Proc. Natl. Acad. Sci. U. S. A.* **1981,** *78* (7), 4046-50.

Jiang, X., Abgottspon, D., Kleeb, S., Rabbani, S., Scharenberg, M., Wittwer, M., Haug, M., Schwardt, O. and Ernst, B. Antiadhesion therapy for urinary tract infections—a balanced PK/PD profile proved to be key for success. *J. Med. Chem.* **2012,** *55* (10), 4700-4713.

Johansson, E. M. V., Crusz, S. A., Kolomiets, E., Buts, L., Kadam, R. U., Cacciarini, M., Bartels, K.-M., Diggle, S. P., Cámara, M., Williams, P., Loris, R., Nativi, C., Rosenau, F., Jaeger, K.-E., Darbre, T. and Reymond, J.-L. Inhibition and dispersion of Pseudomonas aeruginosa biofilms by glycopeptide dendrimers targeting the fucose-specific lectin LecB. *Cell Chem. Biol.* **2008,** *15* (12), 1249-1257.

Jongbloed, S. L., Kassianos, A. J., McDonald, K. J., Clark, G. J., Ju, X., Angel, C. E., Chen, C. J., Dunbar, P. R., Wadley, R. B., Jeet, V., Vulink, A. J., Hart, D. N. and Radford, K. J. Human CD141+ (BDCA-3)+ dendritic cells (DCs) represent a unique myeloid DC subset that cross-presents necrotic cell antigens. *J. Exp. Med.* **2010,** *207* (6), 1247-60.

Jordan, J. B., Poppe, L., Xia, X., Cheng, A. C., Sun, Y., Michelsen, K., Eastwood, H., Schnier, P. D., Nixey, T. and Zhong, W. Fragment based drug discovery: practical implementation based on 19F NMR spectroscopy. *J. Med. Chem.* **2012,** *55* (2), 678-87.

Jug, G., Anderluh, M. and Tomasic, T. Comparative evaluation of several docking tools for docking small molecule ligands to DC-SIGN. *J. Mol. Model.* **2015,** *21* (6), 164.

Kantoff, P. W., Higano, C. S., Shore, N. D., Berger, E. R., Small, E. J., Penson, D. F., Redfern, C. H., Ferrari, A. C., Dreicer, R., Sims, R. B., Xu, Y., Frohlich, M. W., Schellhammer, P. F. and Investigators, I. S. Sipuleucel-T immunotherapy for castration-resistant prostate cancer. *N. Engl. J. Med.* **2010,** *363* (5), 411-22.

Karanikas, V., Hwang, L. A., Pearson, J., Ong, C. S., Apostolopoulos, V., Vaughan, H., Xing, P. X., Jamieson, G., Pietersz, G., Tait, B., Broadbent, R., Thynne, G. and McKenzie, I. F. Antibody and T cell responses of patients with adenocarcinoma immunized with mannan-MUC1 fusion protein. *J. Clin. Invest.* **1997,** *100* (11), 2783-92.

Kastenmuller, W., Kastenmuller, K., Kurts, C. and Seder, R. A. Dendritic cell-targeted vaccines - hope or hype? *Nat. Rev. Immunol.* **2014,** *14* (10), 705-11.

Kastrup, J. S., Nielsen, B. B., Rasmussen, H., Holtet, T. L., Graversen, J. H., Etzerodt, M., Thogersen, H. C. and Larsen, I. K. Structure of the C-type lectin carbohydrate recognition domain of human tetranectin. *Acta Crystallogr., Sect D: Biol. Crystallogr.* **1998,** *54* (Pt 5), 757-66.

Kawasaki, N., Vela, J. L., Nycholat, C. M., Rademacher, C., Khurana, A., van Rooijen, N., Crocker, P. R., Kronenberg, M. and Paulson, J. C. Targeted delivery of lipid antigen to macrophages via the CD169/sialoadhesin endocytic pathway induces robust invariant natural killer T cell activation. *Proc. Natl. Acad. Sci. U. S. A.* **2013,** *110* (19), 7826-31.

Keeler, J. *Understanding NMR Spectroscopy*. 2nd ed., Wiley: **2011**.

Keseru, G. M., Erlanson, D. A., Ferenczy, G. G., Hann, M. M., Murray, C. W. and Pickett, S. D. Design principles for fragment libraries: maximizing the value of learnings from pharma fragment-based drug discovery (FBDD) programs for use in academia. *J. Med. Chem.* **2016,** *Online*.

Kim, H. J., Brennan, P. J., Heaslip, D., Udey, M. C., Modlin, R. L. and Belisle, J. T. Carbohydrate-dependent binding of Langerin to SodC, a cell wall glycoprotein of Mycobacterium leprae. *J. Bacteriol.* **2015,** *197* (3), 615-625.

Kirschner, K. N., Yongye, A. B., Tschampel, S. M., Gonzalez-Outeirino, J., Daniels, C. R., Foley, B. L. and Woods, R. J. GLYCAM06: a generalizable biomolecular force field. carbohydrates. *J. Comput. Chem.* **2008,** *29* (4), 622-55.

Kitov, P. I. and Bundle, D. R. On the nature of the multivalency effect: a thermodynamic model. *J. Am. Chem. Soc.* **2003,** *125* (52), 16271-84.

Kitov, P. I., Sadowska, J. M., Mulvey, G., Armstrong, G. D., Ling, H., Pannu, N. S., Read, R. J. and Bundle, D. R. Shiga-like toxins are neutralized by tailored multivalent carbohydrate ligands. *Nature* **2000,** *403* (6770), 669-72.

Klechevsky, E., Morita, R., Liu, M., Cao, Y., Coquery, S., Thompson-Snipes, L., Briere, F., Chaussabel, D., Zurawski, G., Palucka, A. K., Reiter, Y., Banchereau, J. and Ueno, H. Functional specializations of human epidermal Langerhans cells and CD14+ dermal dendritic cells. *Immunity* **2008,** *29* (3), 497-510.

Kleeb, S., Jiang, X., Frei, P., Sigl, A., Bezencon, J., Bamberger, K., Schwardt, O. and Ernst, B. FimH antagonists: phosphate prodrugs improve oral bioavailability. *J. Med. Chem.* **2016,** *59* (7), 3163-82.

Kleeb, S., Pang, L., Mayer, K., Eris, D., Sigl, A., Preston, R. C., Zihlmann, P., Sharpe, T., Jakob, R. P., Abgottspon, D., Hutter, A. S., Scharenberg, M., Jiang, X., Navarra, G., Rabbani, S., Smiesko, M., Ludin, N., Bezencon, J., Schwardt, O., Maier, T. and Ernst, B. FimH antagonists: bioisosteres to improve the in vitro and in vivo PK/PD profile. *J. Med. Chem.* **2015,** *58* (5), 2221-39.

Klein, T., Abgottspon, D., Wittwer, M., Rabbani, S., Herold, J., Jiang, X., Kleeb, S., Luthi, C., Scharenberg, M., Bezencon, J., Gubler, E., Pang, L., Smiesko, M., Cutting, B., Schwardt, O. and Ernst, B. FimH antagonists for the oral treatment of urinary tract infections: from design and synthesis to in vitro and in vivo evaluation. *J. Med. Chem.* **2010,** *53* (24), 8627-41.

Kolb, H. C. and Ernst, B. Development of tools for the design of selectin antagonists. *Chem. Eur. J.* **1997,** *3* (10), 1571-1578.

Kostarnoy, A. V., Gancheva, P. G., Lepenies, B., Tukhvatulin, A. I., Dzharullaeva, A. S., Polyakov, N. B., Grumov, D. A., Egorova, D. A., Kulibin, A. Y., Bobrov, M. A., Malolina, E. A., Zykin, P. A., Soloviev, A. I., Riabenko, E., Maltseva, D. V., Sakharov, D. A., Tonevitsky, A. G., Verkhovskaya, L. V., Logunov, D. Y., Naroditsky, B. S. and Gintsburg, A. L. Receptor Mincle promotes skin allergies and is capable of recognizing cholesterol sulfate. *Proc. Natl. Acad. Sci. U. S. A.* **2017,** *114* (13), E2758-E2765.

Kovacsovics-Bankowski, M. and Rock, K. L. A phagosome-to-cytosol pathway for exogenous antigens presented on MHC class I molecules. *Science* **1995,** *267* (5195), 243-6.

Kuntz, I. D., Chen, K., Sharp, K. A. and Kollman, P. A. The maximal affinity of ligands. *Proc. Natl. Acad. Sci. U. S. A.* **1999,** *96* (18), 9997-10002.

Labute, P. LowModeMD - implicit low-mode velocity filtering applied to conformational search of macrocycles and protein loops. *J. Chem. Inf. Model.* **2010,** *50* (5), 792-800.

Lang, R., Schoenen, H. and Desel, C. Targeting Syk-Card9-activating C-type lectin receptors by vaccine adjuvants: findings, implications and open questions. *Immunobiology* **2011,** *216* (11), 1184-91.

Le Guilloux, V., Schmidtke, P. and Tuffery, P. Fpocket: An open source platform for ligand pocket detection. *BMC Bioinform.* **2009,** *10*, 11.

Lee, Y. C., Townsend, R. R., Hardy, M. R., Lonngren, J., Arnarp, J., Haraldsson, M. and Lonn, H. Binding of synthetic oligosaccharides to the hepatic Gal/GalNAc lectin. Dependence on fine structural features. *J. Biol. Chem.* **1983,** *258* (1), 199-202.

Lehotzky, R. E., Partch, C. L., Mukherjee, S., Cash, H. L., Goldman, W. E., Gardner, K. H. and Hooper, L. V. Molecular basis for peptidoglycan recognition by a bactericidal lectin. *Proc. Natl. Acad. Sci. U. S. A.* **2010,** *107* (17), 7722-7.

Lenci, E., Menchi, G. and Trabocchi, A. Carbohydrates in diversity-oriented synthesis: challenges and opportunities. *Org. Biomol. Chem.* **2016,** *14* (3), 808-825.

Lenter, M., Levinovitz, A., Isenmann, S. and Vestweber, D. Monospecific and common glycoprotein ligands for E- and P-selectin on myeloid cells. *J. Cell Biol.* **1994,** *125* (2), 471-81.

Leon, B. and Ardavin, C. Monocyte-derived dendritic cells in innate and adaptive immunity. *Immunol. Cell. Biol.* **2008,** *86* (4), 320-4.

Ley, K. The role of selectins in inflammation and disease. *Trends Mol. Med.* **2003,** *9* (6), 263-8.

Li, S. S. Specificity and versatility of SH3 and other proline-recognition domains: structural basis and implications for cellular signal transduction. *Biochem. J.* **2005,** *390* (Pt 3), 641-53.

Lipinski, C. A., Lombardo, F., Dominy, B. W. and Feeney, P. J. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv. Drug Deliv. Rev.* **2001,** *46* (1-3), 3-26.

Liu, L. K. and Finzel, B. C. Fragment-based identification of an inducible binding site on cell surface receptor CD44 for the design of protein-carbohydrate interaction inhibitors. *J. Med. Chem.* **2014,** *57* (6), 2714-25.

Liu, M., Mao, X.-a., Ye, C., Huang, H., Nicholson, J. K. and Lindon, J. C. Improved WATERGATE pulse sequences for solvent suppression in NMR spectroscopy. *J. Magn. Reson.* **1998,** *132* (1), 125-129.

Luczkowiak, J., Sattin, S., Sutkeviciute, I., Reina, J. J., Sanchez-Navarro, M., Thepaut, M., Martinez-Prats, L., Daghetti, A., Fieschi, F., Delgado, R., Bernardi, A. and Rojo, J. Pseudosaccharide functionalized dendrimers as potent inhibitors of DC-SIGN dependent Ebola pseudotyped viral infection. *Bioconjug. Chem.* **2011,** *22* (7), 1354-65.

Lutz, M. B. Buy one, get one free: additional functions of GM-CSF in DC maturation. *Eur. J. Immunol.* **2012,** *42* (1), 35-8.

Macauley, M. S., Pfrengle, F., Rademacher, C., Nycholat, C. M., Gale, A. J., von Drygalski, A. and Paulson, J. C. Antigenic liposomes displaying CD22 ligands induce antigen-specific B cell apoptosis. *J. Clin. Invest.* **2013,** *123* (7), 3074-3083.

Mammen, M., Choi, S.-K. and Whitesides, G. M. Polyvalent interactions in biological systems: implications for design and use of multivalent ligands and inhibitors. *Angew. Chem. Int. Ed.* **1998,** *37* (20), 2754-2794.

Mandal, S., Eksteen-Akeroyd, Z. H., Jacobs, M. J., Hammink, R., Koepf, M., Lambeck, A. J. A., van Hest, J. C. M., Wilson, C. J., Blank, K., Figdor, C. G. and Rowan, A. E. Therapeutic nanoworms: towards novel synthetic dendritic cells for immunotherapy. *Chem. Sci.* **2013,** *4* (11), 4168-4174.

Mandal, S., Hammink, R., Tel, J., Eksteen-Akeroyd, Z. H., Rowan, A. E., Blank, K. and Figdor, C. G. Polymer-based synthetic dendritic cells for tailoring robust and multifunctional T cell responses. *ACS Chem. Biol.* **2015,** *10* (2), 485-92.

Mangold, S. L., Prost, L. R. and Kiessling, L. L. Quinoxalinone inhibitors of the lectin DC-SIGN. *Chem. Sci.* **2012,** *3* (3), 772-777.

Mannhold, R., Poda, G. I., Ostermann, C. and Tetko, I. V. Calculation of molecular lipophilicity: state-of-the-art and comparison of log P methods on more than 96,000 compounds. *J. Pharm. Sci.* **2009,** *98* (3), 861-93.

Manzo, C., Torreno-Pina, J. A., Joosten, B., Reinieren-Beeren, I., Gualda, E. J., Loza-Alvarez, P., Figdor, C. G., Garcia-Parajo, M. F. and Cambi, A. The neck region of the C-type lectin DC-SIGN regulates its surface spatiotemporal organization and virus-binding capacity on antigen-presenting cells. *J. Biol. Chem.* **2012,** *287* (46), 38946-55.

Marcelo, F., Garcia-Martin, F., Matsushita, T., Sardinha, J., Coelho, H., Oude-Vrielink, A., Koller, C., Andre, S., Cabrita, E. J., Gabius, H. J., Nishimura, S., Jimenez-Barbero, J. and Canada, F. J. Delineating binding modes of Gal/GalNAc and structural elements of the molecular recognition of tumor-associated mucin glycopeptides by the human macrophage galactose-type lectin. *Chem. Eur. J.* **2014,** *20* (49), 16147-55.

Marr, D. and Hildreth, E. Theory of edge detection. *Proc. R. Soc. Lond. B Biol. Sci.* **1980,** *207* (1167), 187-217.

Martinez-Lostao, L., Anel, A. and Pardo, J. How do cytotoxic lymphocytes kill cancer cells? *Clin. Cancer Res.* **2015,** *21* (22), 5047-56.

Mashalidis, E. H., Sledz, P., Lang, S. and Abell, C. A three-stage biophysical screening cascade for fragment-based drug discovery. *Nat. Protoc.* **2013,** *8* (11), 2309-24.

Massova, I. and Kollman, P. A. Combined molecular mechanical and continuum solvent approach (MM-PBSA/GBSA) to predict ligand binding. *Perspect. Dug Discov.* **2000,** *18* (1), 113-135.

Matei, E., Andre, S., Glinschert, A., Infantino, A. S., Oscarson, S., Gabius, H. J. and Gronenborn, A. M. Fluorinated carbohydrates as lectin ligands: dissecting glycan-cyanovirin interactions by using 19F NMR spectroscopy. *Chem. Eur. J.* **2013,** *19* (17), 5364-74.

MathWorks. *Bioinformatics Toolbox*, 4.7. Natick, U. S. A., **2016a**.

MathWorks. *MatLab*, 9.0. Natick, U. S. A., **2016b**.

Mattos, C. and Ringe, D. Locating and characterizing binding sites on proteins. *Nat. Biotech.* **1996,** *14* (5), 595-599.

Mayer, M. and Meyer, B. Characterization of ligand binding by saturation transfer difference NMR spectroscopy. *Angew. Chem. Int. Ed.* **1999,** *38* (12), 1784-1788.

Mayer, M. and Meyer, B. Group epitope mapping by saturation transfer difference NMR to identify segments of a ligand in direct contact with a protein receptor. *J. Am. Chem. Soc.* **2001,** *123* (25), 6108-17.

McDermott, R., Ziylan, U., Spehner, D., Bausinger, H., Lipsker, D., Mommaas, M., Cazenave, J. P., Raposo, G., Goud, B., de la Salle, H., Salamero, J. and Hanau, D. Birbeck granules are subdomains of endosomal recycling compartment in human epidermal Langerhans cells, which form where Langerin accumulates. *Mol. Biol. Cell.* **2002,** *13* (1), 317-35.

McGovern, N., Schlitzer, A., Gunawan, M., Jardine, L., Shin, A., Poyner, E., Green, K., Dickinson, R., Wang, X. N., Low, D., Best, K., Covins, S., Milne, P., Pagan, S., Aljefri, K., Windebank, M., Miranda-Saavedra, D., Larbi, A., Wasan, P. S., Duan, K., Poidinger, M., Bigley, V., Ginhoux, F., Collin, M. and Haniffa, M. Human dermal CD14(+) cells are a transient population of monocyte-derived macrophages. *Immunity* **2014,** *41* (3), 465-77.

Meanwell, N. A. Synopsis of some recent tactical application of bioisosteres in drug design. *J. Med. Chem.* **2011,** *54* (8), 2529-91.

Meiboom, S. and Gill, D. Modified spin- echo method for measuring nuclear relaxation times. *Rev. Sci. Intrum.* **1958,** *29* (8), 688-691.

Meier, M., Bider, M. D., Malashkevich, V. N., Spiess, M. and Burkhard, P. Crystal structure of the carbohydrate recognition domain of the H1 subunit of the asialoglycoprotein receptor. *J. Mol. Biol.* **2000,** *300* (4), 857-65.

Mesch, S., Moser, D., Strasser, D. S., Kelm, A., Cutting, B., Rossato, G., Vedani, A., Koliwer-Brandl, H., Wittwer, M., Rabbani, S., Schwardt, O., Kelm, S. and Ernst, B. Low molecular weight antagonists of the myelin-associated glycoprotein: synthesis, docking, and biological evaluation. *J. Med. Chem.* **2010,** *53* (4), 1597-1615.

Mestrelab Research. *MestreNova*, 11.0.2. Santiago de Compostela, Spain, **2016**.

Meyer-Wentrup, F., Cambi, A., Joosten, B., Looman, M. W., de Vries, I. J., Figdor, C. G. and Adema, G. J. DCIR is endocytosed into human dendritic cells and inhibits TLR8-mediated cytokine production. *J. Leukoc. Biol.* **2009,** *85* (3), 518-25.

Meyer, B. and Peters, T. NMR spectroscopy techniques for screening and identifying ligand binding to protein receptors. *Angew. Chem. Int. Ed.* **2003,** *42* (8), 864-90.

Mitchell, D. A., Fadden, A. J. and Drickamer, K. A novel mechanism of carbohydrate recognition by the C-type lectins DC-SIGN and DC-SIGNR: subunit organization and binding to multivalent ligands. *J. Biol. Chem.* **2001,** *276* (31), 28939-28945.

Miyaura, N., Yamada, K. and Suzuki, A. A new stereospecific cross-coupling by the palladium-catalyzed reaction of 1-alkenylboranes with 1-alkenyl or 1-alkynyl halides. *Tetrahedron Lett.* **1979,** *20* (36), 3437-3440.

Morris, G. M., Huey, R., Lindstrom, W., Sanner, M. F., Belew, R. K., Goodsell, D. S. and Olson, A. J. AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *J. Comput. Chem.* **2009,** *30* (16), 2785-91.

Morse, M. A., Chapman, R., Powderly, J., Blackwell, K., Keler, T., Green, J., Riggs, R., He, L. Z., Ramakrishna, V., Vitale, L., Zhao, B., Butler, S. A., Hobeika, A., Osada, T., Davis, T., Clay, T. and Lyerly, H. K. Phase I study utilizing a novel antigen-presenting cell-targeted vaccine with Toll-like receptor stimulation to induce immunity to self-antigens in cancer patients. *Clin. Cancer Res.* **2011,** *17* (14), 4844-53.

Moschen, T., Grutsch, S., Juen, M. A., Wunderlich, C. H., Kreutz, C. and Tollinger, M. Measurement of ligand–target residence times by (1)H relaxation dispersion NMR spectroscopy. *J. Med. Chem.* **2016,** *59* (23), 10788-10793.

Mullin, N. P., Hall, K. T. and Taylor, M. E. Characterization of ligand binding to a carbohydrate-recognition domain of the macrophage Mannose Receptor. *J. Biol. Chem.* **1994,** *269* (45), 28405-13.

Mulvey, M. A. Adhesion and entry of uropathogenic Escherichia coli. *Cell. Microbiol.* **2002,** *4* (5), 257-71.

Munoz-Garcia, J. C., Chabrol, E., Vives, R. R., Thomas, A., de Paz, J. L., Rojo, J., Imberty, A., Fieschi, F., Nieto, P. M. and Angulo, J. Langerin-heparin interaction: two binding sites for small and large ligands as revealed by a combination of NMR spectroscopy and cross-linking mapping experiments. *J. Am. Chem. Soc.* **2015,** *137* (12), 4100-10.

Murray, C. W. and Verdonk, M. L. The consequences of translational and rotational entropy lost by small molecules on binding to proteins. *J. Comput. Aided Mol. Des.* **2002,** *16* (10), 741-53.

Mydock-McGrane, L., Cusumano, Z., Han, Z., Binkley, J., Kostakioti, M., Hannan, T., Pinkner, J. S., Klein, R., Kalas, V., Crowley, J., Rath, N. P., Hultgren, S. J. and Janetka, J. W. Antivirulence C-Mannosides as antibiotic-sparing, oral therapeutics for urinary tract infections. *J. Med. Chem.* **2016,** *59* (20), 9390-9408.

N'Go, I., Golten, S., Ardá, A., Cañada, J., Jiménez-Barbero, J., Linclau, B. and Vincent, S. P. Tetrafluorination of sugars as a strategy for enhancing protein–carbohydrate affinity: application to UDP-Galp Mutase inhibition. *Chem. Eur. J.* **2014,** *20* (1), 106-112.

Nagae, M., Ikeda, A., Kitago, Y., Matsumoto, N., Yamamoto, K. and Yamaguchi, Y. Crystal structures of carbohydrate recognition domain of blood dendritic cell antigen-2 (BDCA2) reveal a common domain-swapped dimer. *Proteins: Struct., Funct., Bioinf.* **2014,** *82* (7), 1512-8.

Navarra, G., Zihlmann, P., Jakob, R. P., Stangier, K., Preston, R. C., Rabbani, S., Smiesko, M., Wagner, B., Maier, T. and Ernst, B. Carbohydrate–lectin interactions: an unexpected contribution to affinity. *ChemBioChem* **2017,** *18* (6), 539-544.

Nelson, R. M., Dolich, S., Aruffo, A., Cecconi, O. and Bevilacqua, M. P. Higher-affinity oligosaccharide ligands for E-selectin. *J. Clin. Invest.* **1993,** *91* (3), 1157-66.

Ng, K. K., Park-Snyder, S. and Weis, W. I. Ca2+-dependent structural changes in C-type mannose-binding proteins. *Biochemistry* **1998,** *37* (51), 17965-76.

Ng, W. C., Londrigan, S. L., Nasr, N., Cunningham, A. L., Turville, S., Brooks, A. G. and Reading, P. C. The C-type lectin Langerin functions as a receptor for attachment and infectious entry of Influenza A virus. *J. Virol.* **2016,** *90* (1), 206-221.

Nielbo, S., Thomsen, J. K., Graversen, J. H., Jensen, P. H., Etzerodt, M., Poulsen, F. M. and Thogersen, H. C. Structure of the plasminogen kringle 4 binding calcium-free form of the C-type lectin-like domain of tetranectin. *Biochemistry* **2004,** *43* (27), 8636-43.

Nielsen, B. B., Kastrup, J. S., Rasmussen, H., Holtet, T. L., Graversen, J. H., Etzerodt, M., Thogersen, H. C. and Larsen, I. K. Crystal structure of tetranectin, a trimeric plasminogen-binding protein with an alpha-helical coiled coil. *FEBS Lett.* **1997,** *412* (2), 388-96.

Noti, C., de Paz, J. L., Polito, L. and Seeberger, P. H. Preparation and use of microarrays containing synthetic heparin oligosaccharides for the rapid analysis of heparin-protein interactions. *Chem. Eur. J.* **2006,** *12* (34), 8664-86.

Nycholat, C. M., Rademacher, C., Kawasaki, N. and Paulson, J. C. In silico-aided design of a glycan ligand of Sialoadhesin for in vivo targeting of macrophages. *J. Am. Chem. Soc.* **2012,** *134* (38), 15696-9.

O'Reilly, M. K., Tian, H. and Paulson, J. C. CD22 is a recycling receptor that can shuttle cargo between the cell surface and endosomal compartments of B cells. *J. Immunol.* **2011,** *186* (3), 1554-63.

Ober-Blobaum, J. L., Ortner, D., Haid, B., Brand, A., Tripp, C., Clausen, B. E. and Stoitzner, P. Monitoring skin dendritic cells in steady state and inflammation by immunofluorescence microscopy and flow cytometry. *Methods Mol. Biol.* **2017,** *1559*, 37-52.

Onizuka, T., Shimizu, H., Moriwaki, Y., Nakano, T., Kanai, S., Shimada, I. and Takahashi, H. NMR study of ligand release from asialoglycoprotein receptor under solution conditions in early endosomes. *FEBS J.* **2012,** *279* (15), 2645-56.

Oprea, T. I. Property distribution of drug-related chemical databases. *J. Comput. Aided Mol. Des.* **2000,** *14* (3), 251-64.

Ordanini, S., Varga, N., Porkolab, V., Thepaut, M., Belvisi, L., Bertaglia, A., Palmioli, A., Berzi, A., Trabattoni, D., Clerici, M., Fieschi, F. and Bernardi, A. Designing nanomolar antagonists of DC-SIGN-mediated HIV infection: ligand presentation using molecular rods. *Chem. Commun.* **2015,** *51* (18), 3816-9.

Orgueira, H. A., Bartolozzi, A., Schell, P. and Seeberger, P. H. Conformational locking of the glycosyl acceptor for stereocontrol in the key step in the synthesis of heparin. *Angew. Chem. Int. Ed.* **2002,** *41* (12), 2128-31.

OriginLab. *OriginPro*, 9.1. Northhampton, U. S. A., **2015**.

Ortner, D., Tripp, C. H., Komenda, K., Dubrac, S., Zelger, B., Hermann, M., Doppler, W., Tymoszuk, P. Z., Boon, L., Clausen, B. E. and Stoitzner, P. Langerhans cells and NK cells cooperate in the inhibition of chemical skin carcinogenesis. *Oncoimmunology* **2017,** *6* (2), e1260215.

Ostrand-Rosenberg, S. and Sinha, P. Myeloid-derived suppressor cells: linking inflammation and cancer. *J. Immunol.* **2009,** *182* (8), 4499-506.

Palucka, K. and Banchereau, J. Cancer immunotherapy via dendritic cells. *Nat. Rev. Cancer* **2012,** *12* (4), 265-277.

Palucka, K. and Banchereau, J. Dendritic-cell-based therapeutic cancer vaccines. *Immunity* **2013,** *39* (1), 38-48.

Pang, L., Kleeb, S., Lemme, K., Rabbani, S., Scharenberg, M., Zalewski, A., Schadler, F., Schwardt, O. and Ernst, B. FimH antagonists: structure-activity and structure-property relationships for biphenyl alpha-D-mannopyranosides. *ChemMedChem* **2012,** *7* (8), 1404-22.

Paradis, E., Claude, J. and Strimmer, K. APE: analyses of phylogenetics and evolution in R language. *Bioinformatics* **2004,** *20* (2), 289-90.

Park, H., Adsit, F. G. and Boyington, J. C. The 1.4 angstrom crystal structure of the human oxidized low density lipoprotein receptor Lox-1. *J. Biol. Chem.* **2005,** *280* (14), 13593-9.

Pasparakis, M., Haase, I. and Nestle, F. O. Mechanisms regulating skin immunity and inflammation. *Nat. Rev. Immunol.* **2014,** *14* (5), 289-301.

Pederson, K., Mitchell, D. A. and Prestegard, J. H. Structural characterization of the DC-SIGN–LeX complex. *Biochemistry* **2014,** *53* (35), 5700-5709.

Peer, D., Karp, J. M., Hong, S., Farokhzad, O. C., Margalit, R. and Langer, R. Nanocarriers as an emerging platform for cancer therapy. *Nat. Nanotechnol.* **2007,** *2* (12), 751-60.

Pervushin, K., Riek, R., Wider, G. and Wuthrich, K. Attenuated T2 relaxation by mutual cancellation of dipole-dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution. *Proc. Natl. Acad. Sci. U. S. A.* **1997,** *94* (23), 12366-71.

Peters, T., Scheffler, K., Ernst, B., Katopodis, A., Magnani, J. L., Wang, W. T. and Weisemann, R. Determination of the bioactive conformation of the carbohydrate ligand in the E-Selectin/sialyl LewisX complex. *Angew. Chem. Int. Ed.* **1995,** *34* (17), 1841-1844.

Petersen, B. O., Vinogradov, E., Kay, W., Wurtz, P., Nyberg, N. T., Duus, J. O. and Sorensen, O. W. H2BC: a new technique for NMR analysis of complex carbohydrates. *Carbohydr. Res.* **2006,** *341* (4), 550-6.

Petersen, J., Rhodes, G., Patrick, K., Roudier, J. and Vaughan, J. H. Human T cell responses to the Epstein-Barr nuclear antigen-1 (EBNA-1) as evaluated by synthetic peptides. *Cell. Immunol.* **1989,** *123* (2), 325-33.

Plosker, G. L. and Figgitt, D. P. Rituximab: a review of its use in non-Hodgkin's lymphoma and chronic lymphocytic leukaemia. *Drugs* **2003,** *63* (8), 803-43.

Ponader, D., Maffre, P., Aretz, J., Pussak, D., Ninnemann, N. M., Schmidt, S., Seeberger, P. H., Rademacher, C., Nienhaus, G. U. and Hartmann, L. Carbohydrate-lectin recognition of sequence-defined heteromultivalent glycooligomers. *J. Am. Chem. Soc.* **2014,** *136* (5), 2008-2016.

Preston, R. C., Jakob, R. P., Binder, F. P., Sager, C. P., Ernst, B. and Maier, T. E-selectin ligand complexes adopt an extended high-affinity conformation. *J. Mol. Cell. Biol.* **2016,** *8* (1), 62-72.

Probert, F., Whittaker, S. B.-M., Crispin, M., Mitchell, D. A. and Dixon, A. M. Solution NMR analyses of the C-type carbohydrate recognition domain of DC-SIGNR protein reveal different binding modes for HIV-derived oligosaccharides and smaller glycan fragments. *J. Biol. Chem.* **2013,** *288* (31), 22745-22757.

Prost, L. R., Grim, J. C., Tonelli, M. and Kiessling, L. L. Noncarbohydrate glycomimetics and glycoprotein surrogates as DC-SIGN antagonists and agonists. *ACS Chem. Biol.* **2012,** *7* (9), 1603-8.

Punta, M., Coggill, P. C., Eberhardt, R. Y., Mistry, J., Tate, J., Boursnell, C., Pang, N., Forslund, K., Ceric, G., Clements, J., Heger, A., Holm, L., Sonnhammer, E. L. L., Eddy, S. R., Bateman, A. and Finn, R. D. The Pfam protein families database. *Nucleic Acids Res.* **2012,** *40* (D1), D290-D301.

Quakkelaar, E. D. and Melief, C. J. Experience with synthetic vaccines for cancer and persistent virus infections in nonhuman primates and patients. *Adv. Immunol.* **2012,** *114*, 77-106.

R Core Team. *R: A Language and Environment for Statistical Computing*, 3.2.2. Vienna, Austria, **2014**.

Rademacher, C., Landstrom, J., Sindhuwinata, N., Palcic, M. M., Widmalm, G. and Peters, T. NMR-based exploration of the acceptor binding site of human blood group B galactosyltransferase with molecular fragments. *Glycoconj. J.* **2010,** *27* (3), 349-58.

Ramphal, J. Y., Zheng, Z. L., Perez, C., Walker, L. E., DeFrees, S. A. and Gaeta, F. C. Structure-activity relationships of sialyl Lewis x-containing oligosaccharides. Effect of modifications of the fucose moiety. *J. Med. Chem.* **1994,** *37* (21), 3459-63.

Regnault, A., Lankar, D., Lacabanne, V., Rodriguez, A., Thery, C., Rescigno, M., Saito, T., Verbeek, S., Bonnerot, C., Ricciardi-Castagnoli, P. and Amigorena, S. Fcgamma receptor-mediated induction of dendritic cell maturation and major histocompatibility complex class I-restricted antigen presentation after immune complex internalization. *J. Exp. Med.* **1999,** *189* (2), 371-80.

Restifo, N. P., Dudley, M. E. and Rosenberg, S. A. Adoptive immunotherapy for cancer: harnessing the T cell response. *Nat. Rev. Immunol.* **2012,** *12* (4), 269-81.

Ribeiro-Viana, R., Sanchez-Navarro, M., Luczkowiak, J., Koeppe, J. R., Delgado, R., Rojo, J. and Davis, B. G. Virus-like glycodendrinanoparticles displaying quasi-equivalent nested polyvalency upon glycoprotein platforms potently block viral infection. *Nat. Commun.* **2012,** *3*, 1303.

Ribeiro, J. P., Diercks, T., Jimenez-Barbero, J., Andre, S., Gabius, H. J. and Canada, F. J. Fluorinated carbohydrates as lectin ligands: (19)F-based direct STD monitoring for the detection of anomeric selectivity. *Biomolecules* **2015,** *5* (4), 3177-92.

Richardson, M. B. and Williams, S. J. MCL and Mincle: C-Type Lectin Receptors That Sense Damaged Self and Pathogen-Associated Molecular Patterns. *Front. Immunol.* **2014,** *5*, 288.

Rillahan, C. D., Schwartz, E., McBride, R., Fokin, V. V. and Paulson, J. C. Click and pick: identification of sialoside analogues for Siglec-based cell targeting. *Angew. Chem. Int. Ed.* **2012,** *51* (44), 11014-11018.

Rillahan, C. D., Schwartz, E., Rademacher, C., McBride, R., Rangarajan, J., Fokin, V. V. and Paulson, J. C. On-chip synthesis and screening of a sialoside library yields a high affinity ligand for Siglec-7. *ACS Chem. Biol.* **2013,** *8* (7), 1417-22.

Rinnbauer, M., Ernst, B., Wagner, B., Magnani, J., Benie, A. J. and Peters, T. Epitope mapping of sialyl Lewis(x) bound to E-selectin using saturation transfer difference NMR experiments. *Glycobiology* **2003,** *13* (6), 435-43.

Rock, K. L., Reits, E. and Neefjes, J. Present yourself! By MHC class I and MHC class II molecules. *Trends Immunol.* **2016,** *37* (11), 724-737.

Rodriguez, A., Regnault, A., Kleijmeer, M., Ricciardi-Castagnoli, P. and Amigorena, S. Selective transport of internalized antigens to the cytosol for MHC class I presentation in dendritic cells. *Nat. Cell Biol.* **1999,** *1* (6), 362-8.

Roehrl, M. H., Wang, J. Y. and Wagner, G. A general framework for development and data analysis of competitive high-throughput screens for small-molecule inhibitors of protein-protein interactions by fluorescence polarization. *Biochemistry* **2004,** *43* (51), 16056-66.

Romani, N., Clausen, B. E. and Stoitzner, P. Langerhans cells and more: Langerin-expressing dendritic cell subsets in the skin. *Immunol. Rev.* **2010,** *234* (1), 120-41.

Rostovtsev, V. V., Green, L. G., Fokin, V. V. and Sharpless, K. B. A stepwise huisgen cycloaddition process: copper(I)-catalyzed regioselective "ligation" of azides and terminal alkynes. *Angew. Chem. Int. Ed.* **2002,** *41* (14), 2596-9.

Sacchettini, J. C., Baum, L. G. and Brewer, C. F. Multivalent protein-carbohydrate interactions: a new paradigm for supermolecular assembly and signal transduction. *Biochemistry* **2001,** *40* (10), 3009-15.

Sancho, D., Mourao-Sa, D., Joffre, O. P., Schulz, O., Rogers, N. C., Pennington, D. J., Carlyle, J. R. and Reis e Sousa, C. Tumor therapy in mice via antigen targeting to a novel, DC-restricted C-type lectin. *J. Clin. Invest.* **2008,** *118* (6), 2098-110.

Scharenberg, M., Schwardt, O., Rabbani, S. and Ernst, B. Target Selectivity of FimH Antagonists. *J. Med. Chem.* **2012,** *55* (22), 9810-6.

Scheffler, K., Brisson, J. R., Weisemann, R., Magnani, J. L., Wong, W. T., Ernst, B. and Peters, T. Application of homonuclear 3D NMR experiments and 1D analogs to study the conformation of sialyl Lewis(x) bound to E-selectin. *J. Biomol. NMR.* **1997,** *9* (4), 423-36.

Scheibe, C., Bujotzek, A., Dernedde, J., Weber, M. and Seitz, O. DNA-programmed spatial screening of carbohydrate-lectin interactions. *Chem. Sci.* **2011,** *2* (4), 770-775.

Scheibe, C., Wedepohl, S., Riese, S. B., Dernedde, J. and Seitz, O. Carbohydrate–PNA and aptamer–PNA conjugates for the spatial screening of lectins and lectin assemblies. *ChemBioChem* **2013,** *14* (2), 236-250.

Schietinger, A., Philip, M., Liu, R. B., Schreiber, K. and Schreiber, H. Bystander killing of cancer requires the cooperation of CD4(+) and CD8(+) T cells during the effector phase. *J. Exp. Med.* **2010,** *207* (11), 2469-77.

Schmidtke, P. and Barril, X. Understanding and predicting druggability. A high-throughput method for detection of drug binding sites. *J. Med. Chem.* **2010,** *53* (15), 5858-67.

Schwardt, O., Rabbani, S., Hartmann, M., Abgottspon, D., Wittwer, M., Kleeb, S., Zalewski, A., Smieško, M., Cutting, B. and Ernst, B. Design, synthesis and biological evaluation of mannosyl triazoles as FimH antagonists. *Bioorg. Med. Chem.* **2011,** *19* (21), 6454-6473.

Schwarz, U. S. and Alon, R. L-selectin-mediated leukocyte tethering in shear flow is controlled by multiple contacts and cytoskeletal anchorage facilitating fast rebinding events. *Proc Natl Acad Sci U S A* **2004,** *101* (18), 6940-5.

Schwizer, D., Patton, J. T., Cutting, B., Smiesko, M., Wagner, B., Kato, A., Weckerle, C., Binder, F. P., Rabbani, S., Schwardt, O., Magnani, J. L. and Ernst, B. Pre-organization of the core structure of E-selectin antagonists. *Chem. Eur. J.* **2012,** *18* (5), 1342-51.

Scott, I. L., Market, R. V., DeOrazio, R. J., Meckler, H. and Kogan, T. P. Stereospecific α-d-mannosylation. *Carbohydr. Res.* **1999,** *317* (1–4), 210-216.

Seeberger, P. H. and Werz, D. B. Synthesis and medical applications of oligosaccharides. *Nature* **2007,** *446* (7139), 1046-1051.

Segawa, T. F. and Bodenhausen, G. Determination of transverse relaxation rates in systems with scalar-coupled spins: the role of antiphase coherences. *J. Magn. Reson.* **2013,** *237*, 139-46.

Selvaratnam, R., VanSchouwen, B., Fogolari, F., Mazhab-Jafari, Mohammad T., Das, R. and Melacini, G. The projection analysis of NMR chemical shifts reveals extended EPAC autoinhibition determinants. *Biophys. J.* **2012,** *102* (3), 630-639.

Seneschal, J., Clark, R. A., Gehad, A., Baecher-Allan, C. M. and Kupper, T. S. Human epidermal Langerhans cells maintain immune homeostasis in skin by activating skin resident regulatory T cells. *Immunity* **2012,** *36* (5), 873-84.

Sercombe, L., Veerati, T., Moheimani, F., Wu, S. Y., Sood, A. K. and Hua, S. Advances and challenges of liposome-assisted drug delivery. *Front. Pharmacol.* **2015,** *6*, 286.

Sharma, P. and Allison, J. P. The future of immune checkpoint therapy. *Science* **2015,** *348* (6230), 56-61.

Shelke, S. V., Cutting, B., Jiang, X., Koliwer-Brandl, H., Strasser, D. S., Schwardt, O., Kelm, S. and Ernst, B. A fragment-based in situ combinatorial approach to identify high-affinity ligands for unknown binding sites. *Angew. Chem. Int. Ed.* **2010,** *49* (33), 5721-5.

Sheridan, R. P., Maiorov, V. N., Holloway, M. K., Cornell, W. D. and Gao, Y. D. Drug-like density: a method of quantifying the "bindability" of a protein target based on a very large set of pockets and drug-like ligands from the Protein Data Bank. *J. Chem. Inf. Model.* **2010,** *50* (11), 2029-40.

Sheriff, S., Chang, C. Y. and Ezekowitz, R. A. Human mannose-binding protein carbohydrate recognition domain trimerizes through a triple alpha-helical coiled-coil. *Nat. Struct. Biol.* **1994,** *1* (11), 789-94.

Shoichet, B. K. Screening in a spirit haunted world. *Drug Discov.Today* **2006,** *11* (13–14), 607-615.

Shrive, A. K., Martin, C., Burns, I., Paterson, J. M., Martin, J. D., Townsend, J. P., Waters, P., Clark, H. W., Kishore, U., Reid, K. B. and Greenhough, T. J. Structural characterisation of ligand-binding determinants in human lung surfactant protein D: influence of Asp325. *J. Mol. Biol.* **2009,** *394* (4), 776-88.

Shuker, S. B., Hajduk, P. J., Meadows, R. P. and Fesik, S. W. Discovering high-affinity ligands for proteins: SAR by NMR. *Science* **1996,** *274* (5292), 1531-4.

Smith, A. G., Tasker, P. A. and White, D. J. The structures of phenolic oximes and their complexes. *Coord. Chemi. Rev.* **2003,** *241* (1–2), 61-85.

Somers, W. S., Tang, J., Shaw, G. D. and Camphausen, R. T. Insights into the molecular basis of leukocyte tethering and rolling revealed by structures of P- and E-selectin bound to SLe(X) and PSGL-1. *Cell* **2000,** *103* (3), 467-79.

Sommer, R., Exner, T. E. and Titz, A. A biophysical study with carbohydrate derivatives explains the molecular basis of monosaccharide selectivity of the Pseudomonas aeruginosa lectin LecB. *PLoS One* **2014,** *9* (11), e112822.

Spiliotopoulos, D., Zhu, J., Wamhoff, E. C., Deerain, N., Marchand, J. R., Aretz, J., Rademacher, C. and Caflisch, A. Virtual screen to NMR (VS2NMR): discovery of fragment hits for the CBP bromodomain. *Bioorg. Med. Chem. Lett.* **2017,** *27* (11), 2472-2478.

Sprokholt, J. K., Hertoghs, N. and Geijtenbeek, T. B. Flow Cytometry-Based Bead-Binding Assay for Measuring Receptor Ligand Specificity. *Methods Mol. Biol.* **2016,** *1390*, 121-9.

Stahl, P., Schlesinger, P. H., Sigardson, E., Rodman, J. S. and Lee, Y. C. Receptor-mediated pinocytosis of mannose glycoconjugates by macrophages: characterization and evidence for receptor recycling. *Cell* **1980,** *19* (1), 207-15.

Stambach, N. S. and Taylor, M. E. Characterization of carbohydrate recognition by Langerin, a C-type lectin of Langerhans cells. *Glycobiology* **2003,** *13* (5), 401-10.

Steegmaier, M., Levinovitz, A., Isenmann, S., Borges, E., Lenter, M., Kocher, H. P., Kleuser, B. and Vestweber, D. The E-selectin-ligand ESL-1 is a variant of a receptor for fibroblast growth factor. *Nature* **1995,** *373* (6515), 615-20.

Stefan, M. I. and Le Novère, N. Cooperative binding. *PLOS Comput. Biol.* **2013,** *9* (6), 1003106.

Steinman, R. M. Decisions about dendritic cells: past, present, and future. *Annu. Rev. Immunol.* **2012,** *30*, 1-22.

Stoitzner, P., Green, L. K., Jung, J. Y., Price, K. M., Tripp, C. H., Malissen, B., Kissenpfennig, A., Hermans, I. F. and Ronchese, F. Tumor immunotherapy by epicutaneous immunization requires Langerhans cells. *J. Immunol.* **2008,** *180* (3), 1991-8.

Stoitzner, P., Romani, N., McLellan, A. D., Tripp, C. H. and Ebner, S. Isolation of skin dendritic cells from mouse and man. *Methods Mol. Biol.* **2010,** *595*, 235-48.

Stoitzner, P., Schaffenrath, S., Tripp, C. H., Reider, D., Komenda, K., Del Frari, B., Djedovic, G., Ebner, S. and Romani, N. Human skin dendritic cells can be targeted in situ by intradermal injection of antibodies against lectin receptors. *Exp. Dermatol.* **2014,** *23* (12), 909-15.

Sukumar, S., Conrad, D. H., Szakal, A. K. and Tew, J. G. Differential T cell-mediated regulation of CD23 (FcεRII) in B cells and follicular dendritic cells. *J. Immunol.* **2006,** *176* (8), 4811-4817.

Sutkeviciute, I., Thepaut, M., Sattin, S., Berzi, A., McGeagh, J., Grudinin, S., Weiser, J., Le Roy, A., Reina, J. J., Rojo, J., Clerici, M., Bernardi, A., Ebel, C. and Fieschi, F. Unique DC-SIGN clustering activity of a small glycomimetic: A lesson for ligand design. *ACS Chem. Biol.* **2014,** *9* (6), 1377-85.

Swaminathan, G. J., Weaver, A. J., Loegering, D. A., Checkel, J. L., Leonidas, D. D., Gleich, G. J. and Acharya, K. R. Crystal structure of the eosinophil major basic protein at 1.8 A. An atypical lectin with a paradigm shift in specificity. *J. Biol. Chem.* **2001,** *276* (28), 26197-203.

Szolnoky, G., Bata-Csörgö, Z., Kenderessy, A. S., Kiss, M., Pivarcsi, A., Novák, Z., Nagy Newman, K., Michel, G., Ruzicka, T., Maródi, L., Dobozy, A. and Kemény, L. A Mannose-Binding Receptor is expressed on human keratinocytes and mediates killing of Candida albicans. *J. Invest. Dermatol.* **2001,** *117* (2), 205-213.

Tabarani, G., Thepaut, M., Stroebel, D., Ebel, C., Vives, C., Vachette, P., Durand, D. and Fieschi, F. DC-SIGN neck domain is a pH-sensor controlling oligomerization: SAXS and hydrodynamic studies of extracellular domain. *J. Biol. Chem.* **2009,** *284* (32), 21229-40.

Tacken, P. J., de Vries, I. J., Torensma, R. and Figdor, C. G. Dendritic-cell immunotherapy: from ex vivo loading to in vivo targeting. *Nat. Rev. Immunol.* **2007,** *7* (10), 790-802.

Tacken, P. J. and Figdor, C. G. Targeted antigen delivery and activation of dendritic cells in vivo: steps towards cost effective vaccines. *Semin. Immunol.* **2011,** *23* (1), 12-20.

Tacken, P. J., Ginter, W., Berod, L., Cruz, L. J., Joosten, B., Sparwasser, T., Figdor, C. G. and Cambi, A. Targeting DC-SIGN via its neck region leads to prolonged antigen residence in early endosomes, delayed lysosomal degradation, and cross-presentation. *Blood* **2011,** *118* (15), 4111-9.

Takahashi, K., Ip, W. E., Michelow, I. C. and Ezekowitz, R. A. The mannose-binding lectin: a prototypic pattern recognition molecule. *Curr. Opin. Immunol.* **2006,** *18* (1), 16-23.

Talpaz, M., Hehlmann, R., Quintas-Cardama, A., Mercer, J. and Cortes, J. Re-emergence of interferon-alpha in the treatment of chronic myeloid leukemia. *Leukemia* **2013,** *27* (4), 803-12.

Tateno, H., Ohnishi, K., Yabe, R., Hayatsu, N., Sato, T., Takeya, M., Narimatsu, H. and Hirabayashi, J. Dual specificity of Langerin to sulfated and mannosylated glycans via a single C-type carbohydrate recognition domain. *J. Biol. Chem.* **2010,** *285* (9), 6390-400.

Taylor, G. S., Haigh, T. A., Gudgeon, N. H., Phelps, R. J., Lee, S. P., Steven, N. M. and Rickinson, A. B. Dual stimulation of Epstein-Barr virus (EBV)-specific CD4+- and CD8+-T-cell responses by a chimeric antigen construct: potential therapeutic vaccine for EBV-positive nasopharyngeal carcinoma. *J. Virol.* **2004,** *78* (2), 768-78.

Taylor, M. E., Bezouska, K. and Drickamer, K. Contribution to ligand binding by multiple carbohydrate-recognition domains in the macrophage mannose receptor. *J. Biol. Chem.* **1992,** *267* (3), 1719-26.

Taylor, M. E. and Drickamer, K. Structural insights into what glycan arrays tell us about how glycan-binding proteins interact with their ligands. *Glycobiology* **2009,** *19* (11), 1155-1162.

Thépaut, M., Guzzi, C., Sutkeviciute, I., Sattin, S., Ribeiro-Viana, R., Varga, N., Chabrol, E., Rojo, J., Bernardi, A., Angulo, J., Nieto, P. M. and Fieschi, F. Structure of a glycomimetic ligand in the carbohydrate recognition domain of C-type lectin DC-SIGN: structural requirements for selectivity and ligand design. *J. Am. Chem. Soc.* **2013,** *135* (7), 2518-2529.

Thepaut, M., Valladeau, J., Nurisso, A., Kahn, R., Arnou, B., Vives, C., Saeland, S., Ebel, C., Monnier, C., Dezutter-Dambuyant, C., Imberty, A. and Fieschi, F. Structural studies of Langerin and Birbeck granules: a macromolecular organization model. *Biochemistry* **2009,** *48* (12), 2684-98.

Thoma, G., Banteli, R., Jahnke, W., Magnani, J. L. and Patton, J. T. A readily available, highly potent E-Selectin antagonist. *Angew. Chem. Int. Ed.* **2001a,** *40* (19), 3644-3647.

Thoma, G., Kinzy, W., Bruns, C., Patton, J. T., Magnani, J. L. and Banteli, R. Synthesis and biological evaluation of a potent E-selectin antagonist. *J. Med. Chem.* **1999,** *42* (23), 4909-13.

Thoma, G., Magnani, J. L. and Patton, J. T. Synthesis and biological evaluation of a sialyl Lewis X mimic with significantly improved E-selectin inhibition. *Bioorg. Med. Chem. Lett.* **2001b,** *11* (7), 923-5.

Timpano, G., Tabarani, G., Anderluh, M., Invernizzi, D., Vasile, F., Potenza, D., Nieto, P. M., Rojo, J., Fieschi, F. and Bernardi, A. Synthesis of novel DC-SIGN ligands with an alpha-fucosylamide anchor. *ChemBioChem* **2008,** *9* (12), 1921-30.

Totrov, M. and Abagyan, R. Flexible ligand docking to multiple receptor conformations: a practical alternative. *Curr. Opin. Struct. Biol.* **2008,** *18* (2), 178-184.

Tullett, K. M., Leal Rojas, I. M., Minoda, Y., Tan, P. S., Zhang, J. G., Smith, C., Khanna, R., Shortman, K., Caminschi, I., Lahoud, M. H. and Radford, K. J. Targeting CLEC9A delivers antigen to human CD141+ DC for CD4+ and CD8+T cell recognition. *JCI Insight* **2016,** *1* (7), e87102.

Unione, L., Alcala, M., Echeverria, B., Serna, S., Arda, A., Franconetti, A., Canada, J., Diercks, T., Reichardt, N. and Jimenez-Barbero, J. Fluoroacetamide moieties as NMR probes for molecular recognition of GlcNAc-containing sugars: modulation of the CH-pi stacking interactions by different fluorination patterns. *Chem. Eur. J.* **2017**.

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. and Saeland, S. Langerin, a novel C-type lectin specific to Langerhans cells, is an endocytic receptor that induces the formation of Birbeck granules. *Immunity* **2000,** *12* (1), 71-81.

van den Berg, L. M., Cardinaud, S., van der Aar, A. M., Sprokholt, J. K., de Jong, M. A., Zijlstra-Willems, E. M., Moris, A. and Geijtenbeek, T. B. Langerhans cell-dendritic cell cross-talk via Langerin and hyaluronic acid mediates antigen transfer and cross-presentation of HIV-1. *J. Immunol.* **2015,** *195* (4), 1763-73.

van den Berg, L. M., Ribeiro, C. M., Zijlstra-Willems, E. M., de Witte, L., Fluitsma, D., Tigchelaar, W., Everts, V. and Geijtenbeek, T. B. Caveolin-1 mediated uptake via langerin restricts HIV-1 infection in human Langerhans cells. *Retrovirology* **2014,** *11*, 123.

van der Vlist, M., de Witte, L., de Vries, R. D., Litjens, M., de Jong, M. A., Fluitsma, D., de Swart, R. L. and Geijtenbeek, T. B. Human Langerhans cells capture measles virus through Langerin and present viral antigens to CD4(+) T cells but are incapable of cross-presentation. *Eur. J. Immunol.* **2011,** *41* (9), 2619-31.

van Montfoort, N., Camps, M. G., Khan, S., Filippov, D. V., Weterings, J. J., Griffith, J. M., Geuze, H. J., van Hall, T., Verbeek, J. S., Melief, C. J. and Ossendorp, F. Antigen storage compartments in

mature dendritic cells facilitate prolonged cytotoxic T lymphocyte cross-priming capacity. *Proc. Natl. Acad. Sci. U. S. A.* **2009,** *106* (16), 6730-5.

van Vliet, S. J., Aarnoudse, C. A., Broks-van den Berg, V. C., Boks, M., Geijtenbeek, T. B. and van Kooyk, Y. MGL-mediated internalization and antigen presentation by dendritic cells: a role for tyrosine-5. *Eur. J. Immunol.* **2007,** *37* (8), 2075-81.

Varga, N., Sutkeviciute, I., Guzzi, C., McGeagh, J., Petit-Haertlein, I., Gugliotta, S., Weiser, J., Angulo, J., Fieschi, F. and Bernardi, A. Selective targeting of Dendritic Cell-Specific Intercellular Adhesion Molecule-3-Grabbing Nonintegrin (DC-SIGN) with mannose-based glycomimetics: synthesis and interaction studies of bis(benzylamide) derivatives of a pseudomannobioside. *Chem. Eur. J.* **2013,** *19* (15), 4786-4797.

Varki, A. *Essentials of Glycobiology*. 2nd ed., Cold Spring Harbor Laboratory Press: **2009**.

Verdier, L., Gharbi-Benarous, J., Bertho, G., Evrard-Todeschi, N., Mauvais, P. and Girault, J.-P. Dissociation-equilibrium constant and bound conformation for weak antibiotic binding interaction with different bacterial ribosomes. *Perkin. Trans. 2* **2000,** (12), 2363-2371.

Verdonk, M. L. and Rees, D. C. Group efficiency: a guideline for hits-to-leads chemistry. *ChemMedChem* **2008,** *3* (8), 1179-80.

Verkhivker, G. M., Bouzida, D., Gehlhaar, D. K., Rejto, P. A., Arthurs, S., Colson, A. B., Freer, S. T., Larson, V., Luty, B. A., Marrone, T. and Rose, P. W. Deciphering common failures in molecular docking of ligand-protein complexes. *J. Comput. Aided Mol. Des.* **2000,** *14* (8), 731-51.

Veselý, J., Ledvina, M., Jindrich, J., Šaman, D. and Trnka, T. Improved synthesis of 1,2-trans-acetates and 1,2-trans ethyl 1-thioglycosides derived from 3,4,6-tri-O-acetyl-2-deoxy-2-phthalimido-d-hexopyranosides. *Collect. Czech. Chem. Comm.* **2003,** *68* (7), 1264-1274.

Volkamer, A., Griewel, A., Grombacher, T. and Rarey, M. Analyzing the topology of active sites: on the prediction of pockets and subpockets. *J. Chem. Inf. Model.* **2010,** *50* (11), 2041-52.

Volkamer, A., Kuhn, D., Grombacher, T., Rippmann, F. and Rarey, M. Combining global and local measures for structure-based druggability predictions. *J. Chem. Inf. Model.* **2012a,** *52* (2), 360-72.

Volkamer, A., Kuhn, D., Rippmann, F. and Rarey, M. DoGSiteScorer: a web server for automatic binding site prediction, analysis and druggability assessment. *Bioinformatics* **2012b,** *28* (15), 2074-2075.

Vranken, W. F., Boucher, W., Stevens, T. J., Fogh, R. H., Pajon, A., Llinas, M., Ulrich, E. L., Markley, J. L., Ionides, J. and Laue, E. D. The CCPN data model for NMR spectroscopy: development of a software pipeline. *Proteins* **2005,** *59* (4), 687-96.

Wamhoff, E.-C., Hanske, J., Schnirch, L., Aretz, J., Grube, M., Varon Silva, D. and Rademacher, C. 19F NMR-guided design of glycomimetic Langerin ligands. *ACS Chem. Biol.* **2016,** *11* (9), 2407–2413.

Wang, L., Wu, Y., Deng, Y., Kim, B., Pierce, L., Krilov, G., Lupyan, D., Robinson, S., Dahlgren, M. K., Greenwood, J., Romero, D. L., Masse, C., Knight, J. L., Steinbrecher, T., Beuming, T., Damm, W., Harder, E., Sherman, W., Brewer, M., Wester, R., Murcko, M., Frye, L., Farid, R., Lin, T., Mobley, D. L., Jorgensen, W. L., Berne, B. J., Friesner, R. A. and Abel, R. Accurate and reliable prediction of relative ligand binding potency in prospective drug discovery by way of a modern free-energy calculation protocol and force field. *J. Am .Chem. Soc.* **2015,** *137* (7), 2695-703.

Watson, A. A., Brown, J., Harlos, K., Eble, J. A., Walter, T. S. and O'Callaghan, C. A. The crystal structure and mutational binding analysis of the extracellular domain of the platelet-activating receptor CLEC-2. *J. Biol. Chem.* **2007,** *282* (5), 3165-72.

Watson, A. A., Lebedev, A. A., Hall, B. A., Fenton-May, A. E., Vagin, A. A., Dejnirattisai, W., Felce, J., Mongkolsapaya, J., Palma, A. S., Liu, Y., Feizi, T., Screaton, G. R., Murshudov, G. N. and O'Callaghan, C. A. Structural flexibility of the macrophage dengue virus receptor CLEC5A: implications for ligand binding and signaling. *J. Biol. Chem.* **2011,** *286* (27), 24208-18.

Weis, W. I. and Drickamer, K. Structural basis of lectin-carbohydrate recognition. *Annu. Rev. Biochem.* **1996,** *65*, 441-73.

Weis, W. I., Drickamer, K. and Hendrickson, W. A. Structure of a C-type mannose-binding protein complexed with an oligosaccharide. *Nature* **1992,** *360* (6400), 127-34.

Williamson, M. P. Using chemical shift perturbation to characterise ligand binding. *Prog. Nucl. Magn. Reson. Spectrosc.* **2013,** *73*, 1-16.

Wragg, S. and Drickamer, K. Identification of amino acid residues that determine pH dependence of ligand binding to the Asialoglycoprotein Receptor during endocytosis. *J. Biol. Chem.* **1999,** *274* (50), 35400-35406.

Wurzburg, B. A., Tarchevskaya, S. S. and Jardetzky, T. S. Structural changes in the lectin domain of CD23, the low-affinity IgE receptor, upon calcium binding. *Structure* **2006,** *14* (6), 1049-58.

Wyatt, P. G., Woodhead, A. J., Berdini, V., Boulstridge, J. A., Carr, M. G., Cross, D. M., Davis, D. J., Devine, L. A., Early, T. R., Feltell, R. E., Lewis, E. J., McMenamin, R. L., Navarro, E. F., O'Brien, M. A., O'Reilly, M., Reule, M., Saxty, G., Seavers, L. C., Smith, D. M., Squires, M. S., Trewartha, G., Walker, M. T. and Woolford, A. J. Identification of N-(4-piperidinyl)-4-(2,6-dichlorobenzoylamino)-1H-pyrazole-3-carboxamide (AT7519), a novel cyclin dependent kinase inhibitor using fragment-based X-ray crystallography and structure based drug design. *J. Med. Chem.* **2008,** *51* (16), 4986-99.

Yamasaki, S., Ishikawa, E., Sakuma, M., Hara, H., Ogata, K. and Saito, T. Mincle is an ITAM-coupled activating receptor that senses damaged cells. *Nat. Immunol.* **2008,** *9* (10), 1179-88.

Yang, K., Park, C. G., Cheong, C., Bulgheresi, S., Zhang, S. S., Zhang, P., He, Y. X., Jiang, L. Y., Huang, H. P., Ding, H. H., Wu, Y. P., Wang, S. G., Zhang, L., Li, A. Y., Xia, L. X., Bartra, S. S., Plano, G. V., Skurnik, M., Klena, J. D. and Chen, T. Host Langerin (CD207) is a receptor for Yersinia pestis phagocytosis and promotes dissemination. *Immunol. Cell Biol.* **2015,** *93* (9), 815-824.

Yang, L., Pang, Y. and Moses, H. L. TGF-beta and immune cells: an important regulatory axis in the tumor microenvironment and progression. *Trends Immunol.* **2010,** *31* (6), 220-7.

Zaccai, N. R., Maenaka, K., Maenaka, T., Crocker, P. R., Brossmer, R., Kelm, S. and Jones, E. Y. Structure-guided design of sialic acid-based Siglec inhibitors and crystallographic analysis in complex with sialoadhesin. *Structure* **2003,** *11* (5), 557-67.

Zelensky, A. N. and Gready, J. E. The C-type lectin-like domain superfamily. *FEBS J.* **2005,** *272* (24), 6179-217.

Zhang, J. G., Czabotar, P. E., Policheni, A. N., Caminschi, I., Wan, S. S., Kitsoulis, S., Tullett, K. M., Robin, A. Y., Brammananth, R., van Delft, M. F., Lu, J., O'Reilly, L. A., Josefsson, E. C., Kile, B. T., Chin, W. J., Mintern, J. D., Olshina, M. A., Wong, W., Baum, J., Wright, M. D., Huang, D. C., Mohandas, N., Coppel, R. L., Colman, P. M., Nicola, N. A., Shortman, K. and Lahoud, M. H. The dendritic cell receptor Clec9A binds damaged cells via exposed actin filaments. *Immunity* **2012,** *36* (4), 646-57.

Zhang, J. H., Chung, T. D. and Oldenburg, K. R. A simple statistical parameter for use in evaluation and validation of high throughput screening assays. *J. Biomol. Screen.* **1999,** *4* (2), 67-73.

Zhang, S., Xiao, Q., Sherman, S. E., Muncan, A., Ramos Vicente, A. D., Wang, Z., Hammer, D. A., Williams, D., Chen, Y., Pochan, D. J., Vertesy, S., Andre, S., Klein, M. L., Gabius, H. J. and Percec, V. Glycodendrimersomes from sequence-defined Janus glycodendrimers reveal high activity and sensor capacity for the agglutination by natural variants of human lectins. *J. Am. Chem. Soc.* **2015,** *137* (41), 13334-44.

Zhao, H. Y., Yang, Y. D., von Itzstein, M. and Zhou, Y. Q. Carbohydrate-binding protein identification by coupling structural similarity searching with binding affinity prediction. *J. Comput. Chem.* **2014,** *35* (30), 2177-2183.

Zhao, J., Liu, X., Kao, C., Zhang, E., Li, Q., Zhang, F. and Linhardt, R. J. Kinetic and structural studies of interactions between glycosaminoglycans and Langerin. *Biochemistry* **2016**.

Zierke, M., Smiesko, M., Rabbani, S., Aeschbacher, T., Cutting, B., Allain, F. H., Schubert, M. and Ernst, B. Stabilization of branched oligosaccharides: Lewis(x) benefits from a nonconventional C-H...O hydrogen bond. *J. Am. Chem. Soc.* **2013,** *135* (36), 13464-72.

Zou, W. Immunosuppressive networks in the tumour environment and their therapeutic relevance. *Nat. Rev. Cancer.* **2005,** *5* (4), 263-74.

Zusammenfassung

Immunevasion ist ein wichtiges Kennzeichen von fortgeschrittenen Krebserkrankungen. Die Aktivierung von tumorspezifischen zytotoxischen T-Zellen kann genutzt werden um die entwickelte immunologische Toleranz zu überwinden und ist in den letzten Jahren in den Fokus für die Entwicklung neuer Krebstherapien gerückt. Langerhans-Zellen gehören zu den dendritischen Zellen und sind in der Epidermis der menschlichen Haut zu finden. Diese Zellen zeichnen sich durch eine effiziente Endozytose und MHC-I-abhängigen Cross-Presentation von exogenen Antigenen aus, ein Merkmal, welches das *Priming* von naiven $CD8^+$ T-Zellen gewährleistet. Langerin ist ein C-Typ Lektin mit endozytotischer Aktivität, das in der Ca^{2+}-abhängigen Erkennung von Glykanen auf Pathogenen oder körpereigenen Strukturen beteiligt und überwiegend auf Langerhans-Zellen zu finden ist. Folglich stellt das Langerin-vermittelte *Targeted Delivery* von tumorassoziierten Antigenen einen vielversprechenden Ansatz für die Entwicklung von neue Impfstrategien dar.

Liposomen sind vielseitige Plattformen für solche *Targeted Delivery* Ansätze und die Konjugation von spezifischen Liganden für Langerin ermöglicht ihren Einsatz um Langerhans-Zellen zu adressieren. Da die Interaktionen zwischen C-Typ Lektinen und Glykanen typischerweise schwach und unspezifisch sind, ist hierfür das Design von synthetischen Glykomimetika erforderlich. Der Designprozess erweist sich jedoch im Allgemeinen als eine Herausforderung: Zum einen ist die Synthese solcher Liganden anspruchsvoll und arbeitsintensiv, zum anderen sind Kohlenhydrat-Bindungstaschen sehr hydrophil und in hohem Maße zugänglich für Lösungsmittelmoleküle. Die strukturbasierte *in silico*-Analyse von 21 Kristallstrukturen, die im Rahmen dieser Dissertation durchgeführt wurde, unterstützt diese Klassifikation von C-Typ Lektinen als *undruggable* oder herausfordernde Proteine für das Design von Liganden. *Druggable* sekundäre Bindungstaschen in der Nähe der Kohlenhydrat-Bindungstasche wurden lediglich für C-Typ Lektine mit eingeschränkter therapeutischer Relevanz identifiziert.

Mehrere Strategien wurden angewendet um die genannten Herausforderungen für die Identifikation von glykomimetischen Liganden für Langerin zu überwinden. Das strukturbasierte *in silico Screening* von Substituenten in C2 des Mannose-*Scaffolds* diente dem Design einer initialen fokussierten Bibliothek. Die Struktur-Wirkungsbeziehung dieses *Scaffolds* wurde des Weiteren mittels der Bestimmung von Affinitäten für eine bereits existierende Bibliothek von Mannoseanaloga aufgeklärt. Diese Analoga waren in C1 und C6 derivatisiert. Eine alternative Strategie verfolgte die strukturbasierte Identifizierung von geeigneten Substituenten in C2 von Glucosamine-2-Sulfat. Die Untersuchungen führten letztendlich zum Design von affinen Mannose- (K_I = 0.25±0.07 mM) und Glucosamine-2-Sulfat-Analoga (K_I = 0.24±0.03 mM), die im Vergleich zu natürlichen Kohlenhydratliganden einen 40- bis 42-fachen Affinitätsgewinn aufweisen.

Die multivalente Organisation von Kohlenhydraten oder ihren synthetischen Analoga in einer Geometrie, die der des Langerin-Trimers entspricht, ermöglicht die zusätzliche Optimierung von Spezifität und Avidität. Das identifizierte Glucosamine-2-Sulfat Analogon wurde an ein Nukleinsäure-*Scaffold* konjugiert um divalente Glykomimetika zu designen (IC_{50} = 23±2 µM). Dieser Ansatz resultierte in einem zusätzlichen, geometrie-abhängigen 12-fachen Aviditätsgewinn. Des Weiteren wurden trivalente Mannose-präsentierende *Glycocluster* (IC_{50} = 0.20±0.08 mM) entwickelt, die einen 80-fachen Aviditätsgewinn im Vergleich zur natürlich vorkommenden Kohlenhydratliganden aufwiesen.

Die Entwicklung einer sensitiven ^{19}F R_2-gefilterten Kernspinresonanzspektroskopie-basierten Nachweismethode ermöglichte die Bestimmung von K_I- und IC_{50}-Werten für Langerin. Hierbei war die Optimierung des Durchsatzes und des Materialverbrauchs der Nachweismethode entscheidend für die effiziente Identifizierung von affinen Kohlenhydratanaloga sowie multivalenten Glykomimetika. Zusätzlich ermöglichte diese Optimierung die Implementierung eines explorativen ^{19}F R_2-gefilterten Kernspinresonanzspektroskopie-basierten *Fragment Screenings*, welches zur Identifizierung des ersten, für Langerin berichteten, nicht-Kohlenhydrat Inhibitors führte. Die erfolgreiche Übertragung der Nachweismethode auf das C-Typ Lektin DC-SIGN konnte für die Untersuchung der Spezifität der designten Kohlenhydratanaloga genutzt werden.

Während ^{19}F R_2-gefilterte Kernspinresonanzspektroskopie als die primäre *Screening*- und Nachweismethode diente, wurden *Saturation Transfer Difference* und ^{15}N *Heteronuclear Single Quantum Coherence* Experimente genutzt um bestimmte Affinitäten zu validieren. Darüber hinaus, wurden diese Techniken der Kernspinresonanzspektroskopie mit *Molecular Docking* kombiniert um die strukturelle Aspekte des Ca^{2+}-abhängigen Bindung der designten Mannose- und Glucsoamine-2-Sulfat-Analoga zu analysieren. Die Untersuchungen ermöglichten die Entwicklung einer geeigneten Strategie für die Konjugation der Liganden an Liposomen.

Per Kernspinresonanzspektroskopie konnte eine erstaunliche Spezifität des designten Glucsoamine-2-Sulfat-Analogons gegen DC-SIGN gezeigt werden. Daraufhin wurde dieser Ligand für die Herstellung von *targeted* Liposomen verwendet. Experimente, die auf Durchflusszytometrie basierten, dienten der Optimierung der Formulierung dieser Liposomen, sowie der Validierung von deren Bindung an Langerin$^+$ Modellzellen *in vitro*. Abschließend konnte in *ex vivo*-Experimenten gezeigt werden, dass die *targeted* Liposomen im Kontext der menschlichen Haut spezifisch Langerhans-Zellen adressieren und zudem effizient von diesen endozytiert werden. Somit stellen diese eine vielversprechende *Targeted Delivery* Plattform für tumorassoziierte Antigene dar.

Zusammenfassend kann die integrierte Designstrategie, welche Kohlenhydratchemie, strukturbasierte *in silico*-Methoden und Kernspinresonanzspektroskopie vereint, als wesentliche Voraussetzung für die erfolgreiche Identifizierung von affinen Kohlenhydratanaloga und multivalenten Glykomimetika für

das C-Typ Lektin Langerin herausgestellt werden. Die identifizierten Liganden sind geeignet für die Herstellung von *targeted* Liposomen, die das spezifische Adressieren von Langerhans-Zellen im Kontext der menschlichen Haut ermöglichen und zudem effizient endozytiert werden. Folglich, stellen die in dieser Dissertation präsentierten Erkenntnisse einen wichtigen Fortschritt für die Erforschung der Immunologie der dendritischen Zellen sowie für die Entwicklung von neuen Krebstherapien dar.

Appendix

A. Supplementary Data – Materials and Methods

A.1. Quality Controls for Receptor Samples

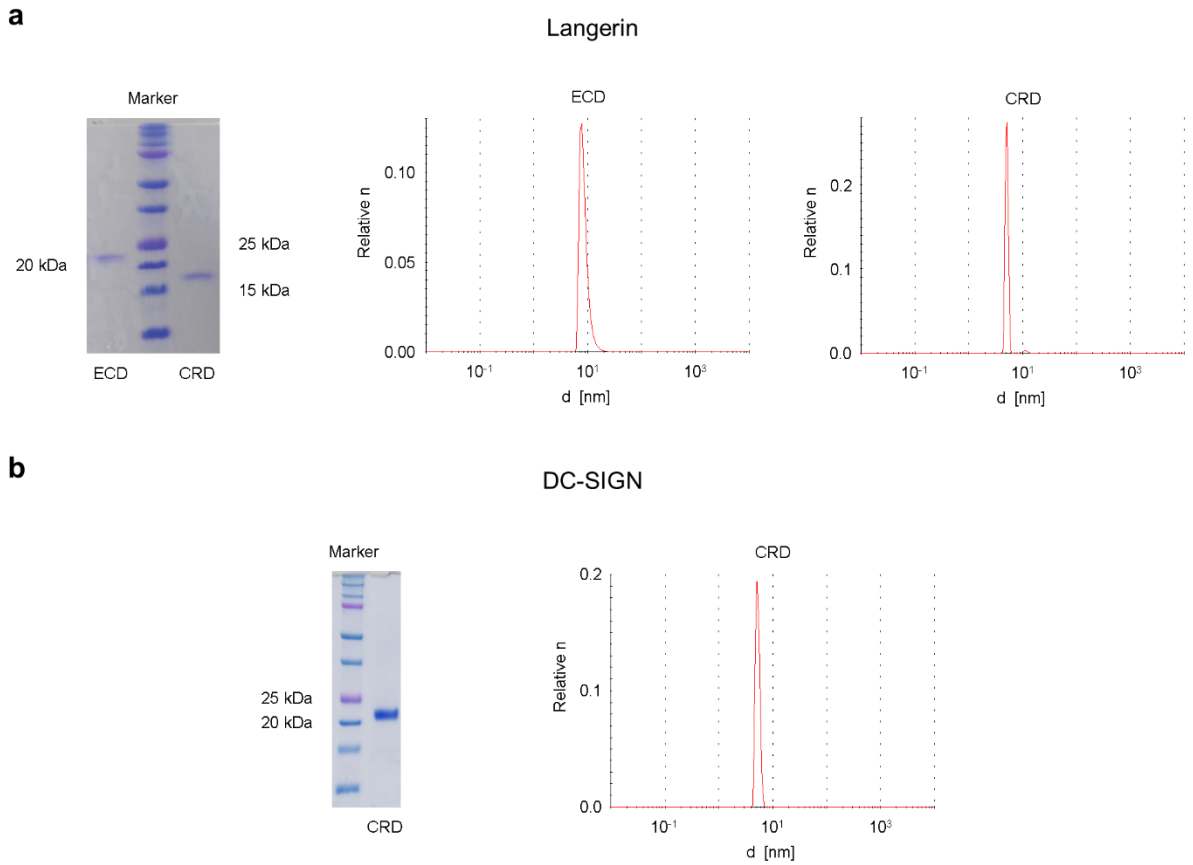


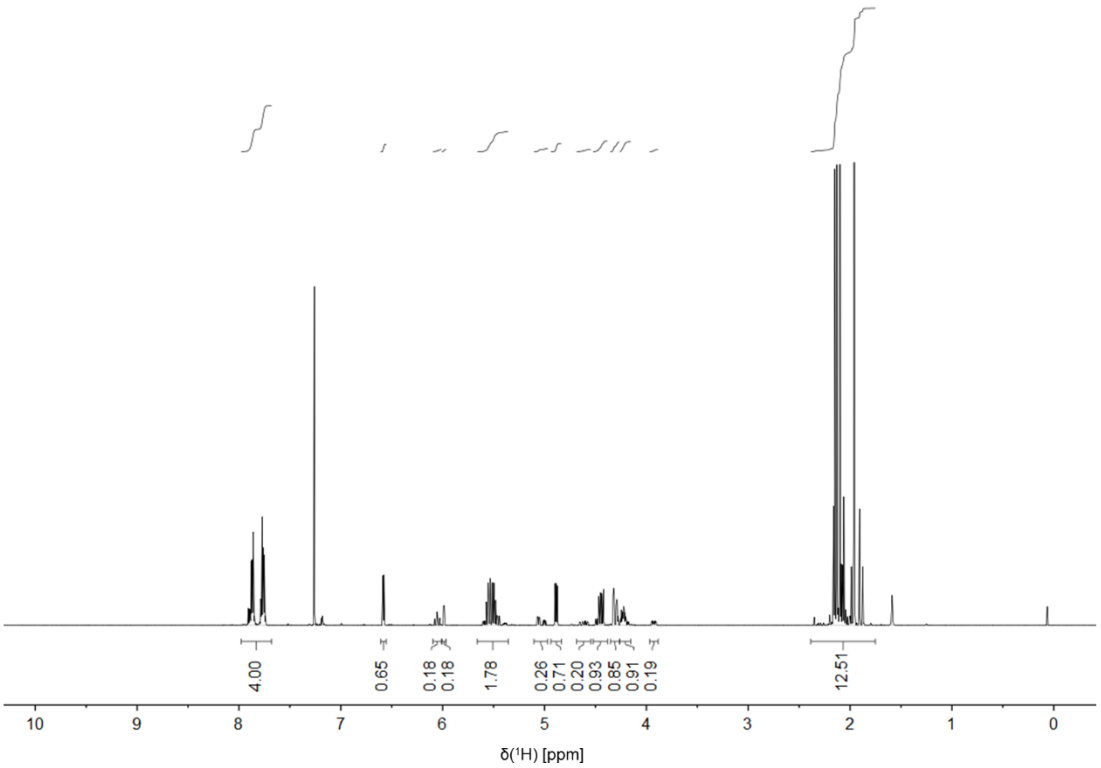
Figure A.1. Quality Controls for Receptor Samples.

a. and b. SDS-PAGE (left) and DLS (right) experiments served as quality controls for Langerin and DC-SIGN samples utilized in NMR experiments.

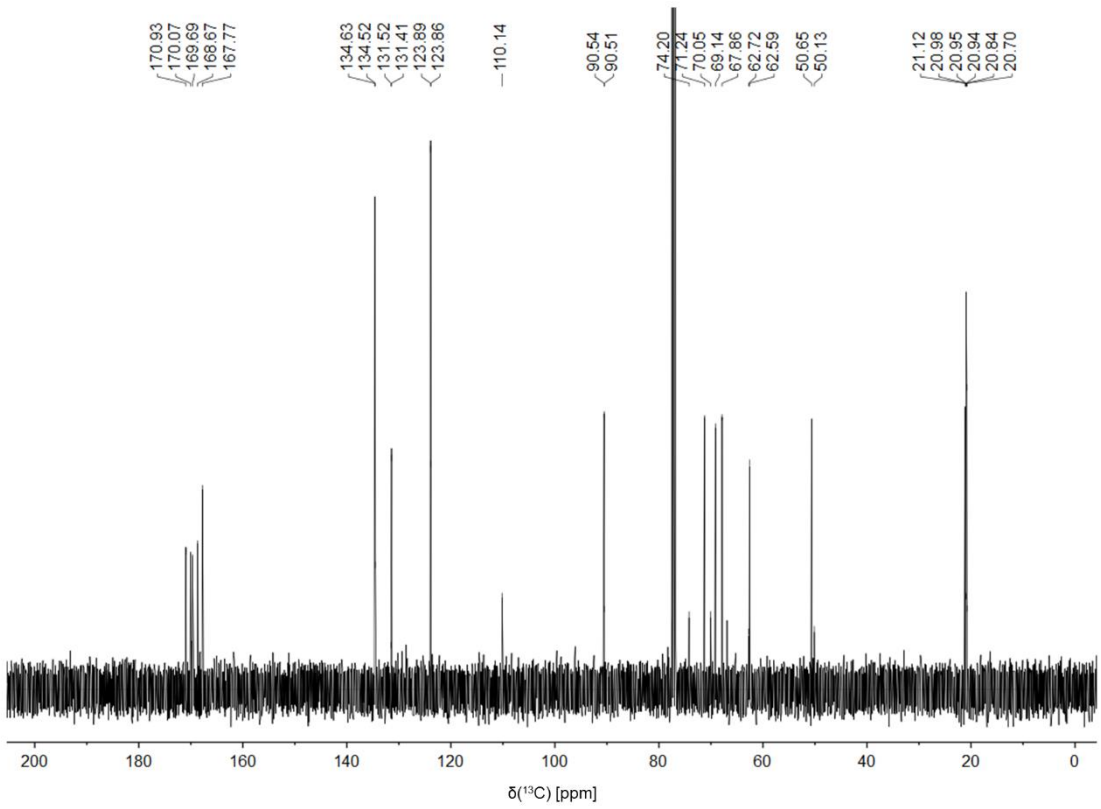
A.2. ^{1}H and ^{13}C NMR Spectra for Monosaccharide Analogs

1,3,4,6-tetra-O-acetyl-2-deoxy-2-phthalimido-D-mannopyranose (2)

^{1}H NMR (400.0 MHz, $CDCl_3$)

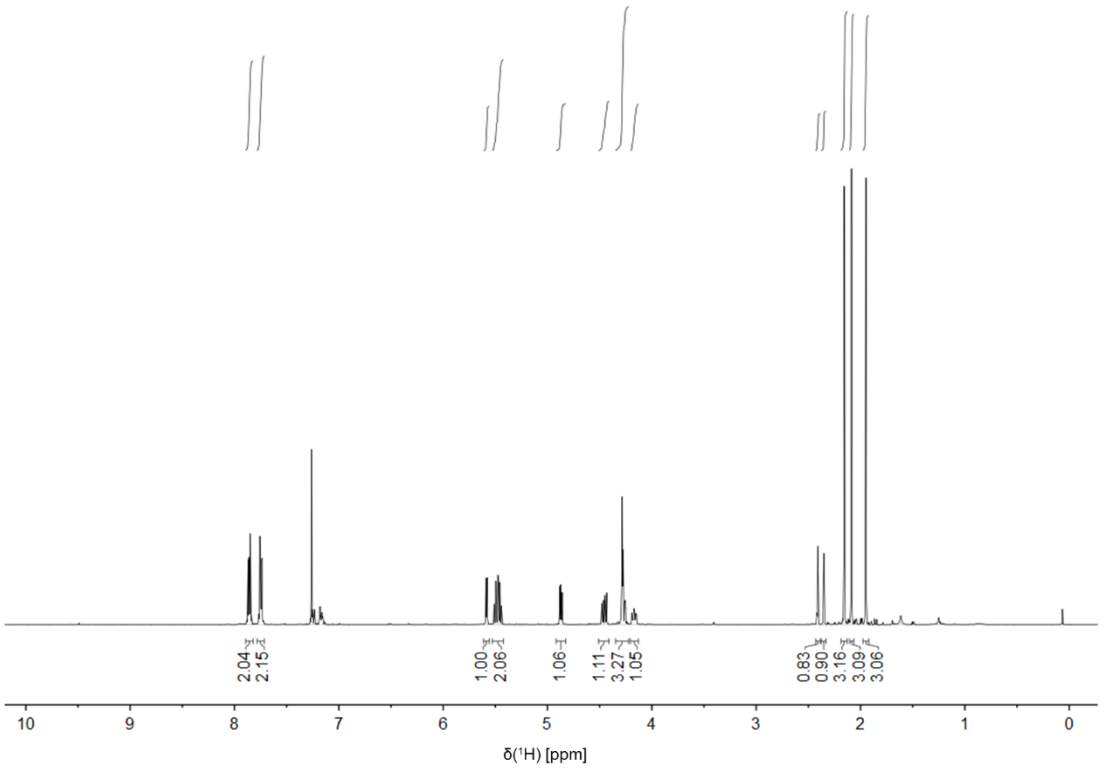


^{13}C NMR (100.6 MHz, $CDCl_3$)

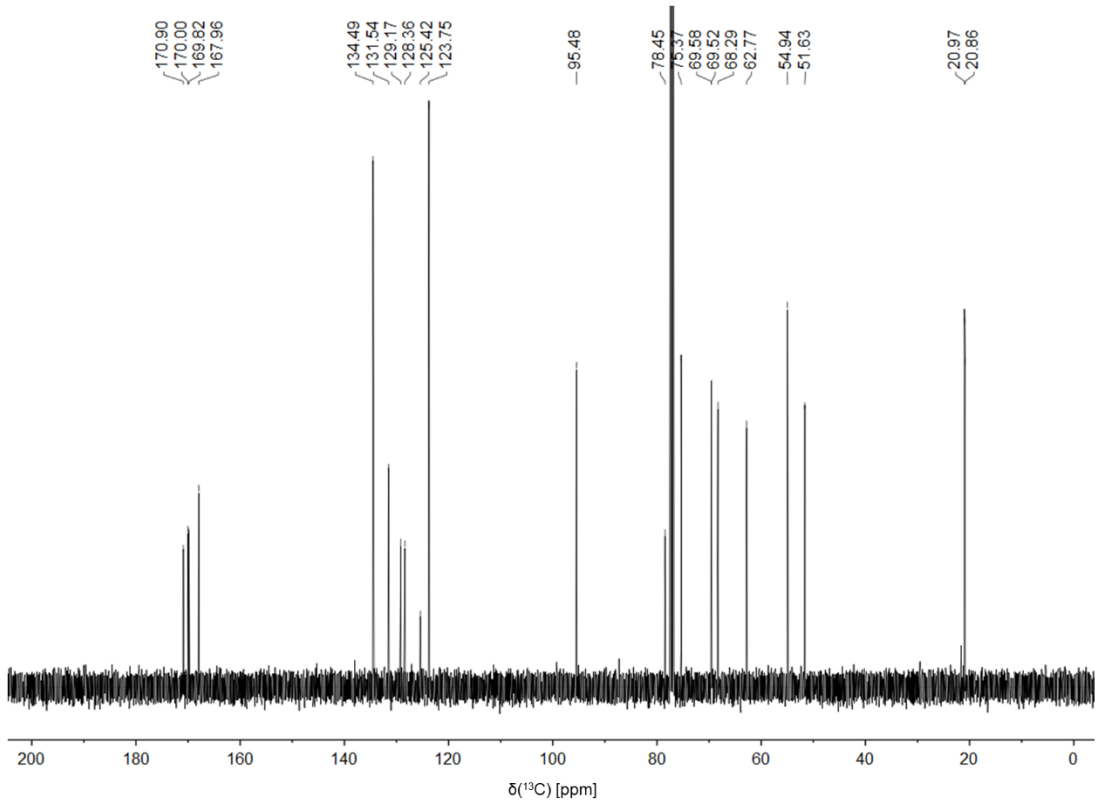


Propargyl-3,4,6-tri-O-acetyl-2-deoxy-2-phthalimido-α-D-mannopyranoside (3)

^{1}H NMR (400.0 MHz, $CDCl_3$)

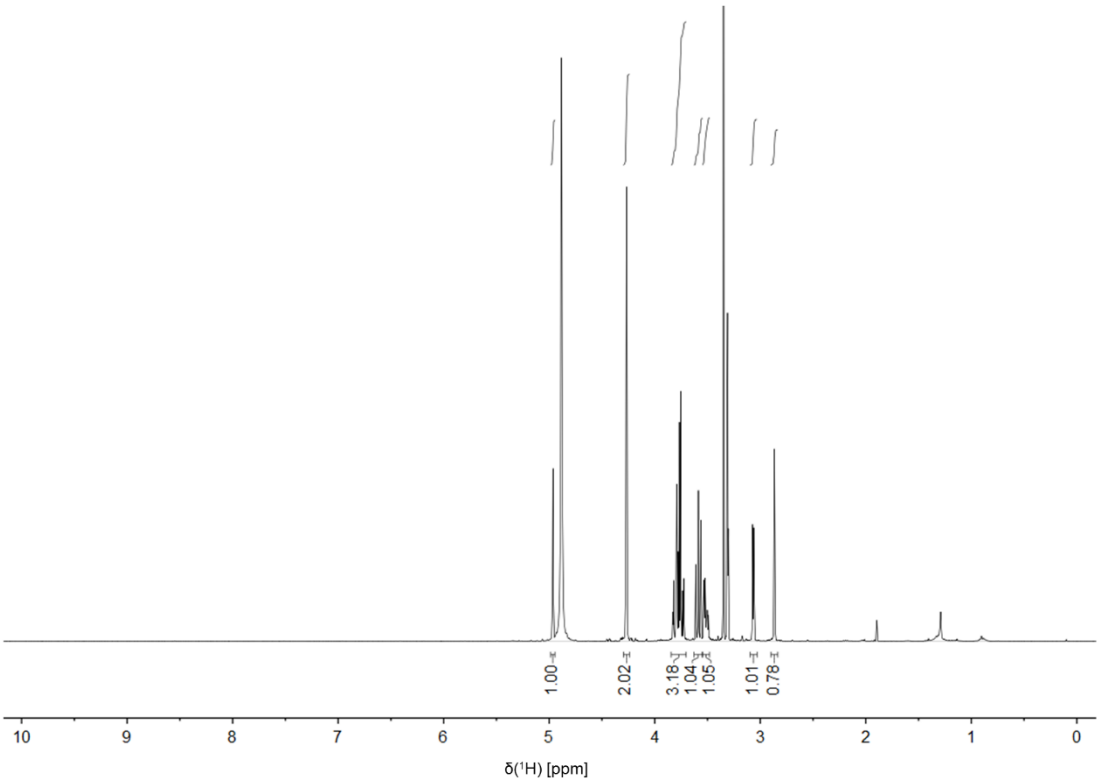


^{13}C NMR (100.6 MHz, $CDCl_3$)

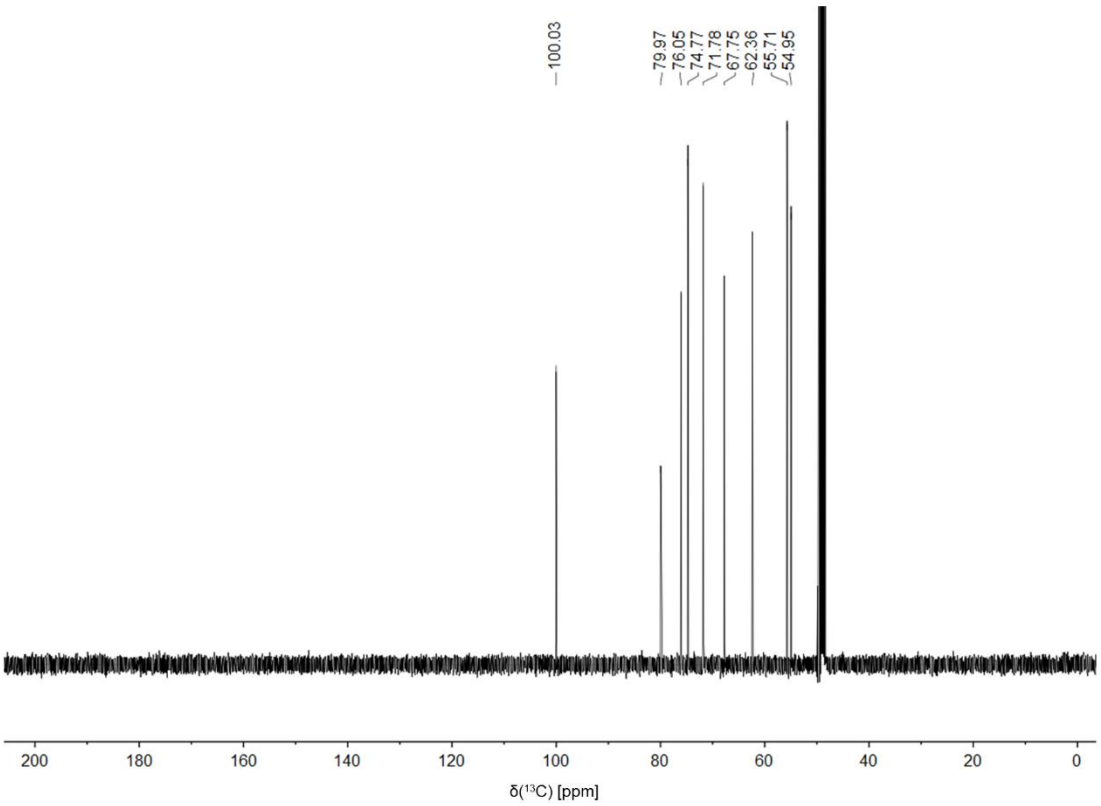


Propargyl-2-deoxy-2-amino-α-D-mannopyranoside (4)

^{1}H NMR (400.0 MHz, MeOD)

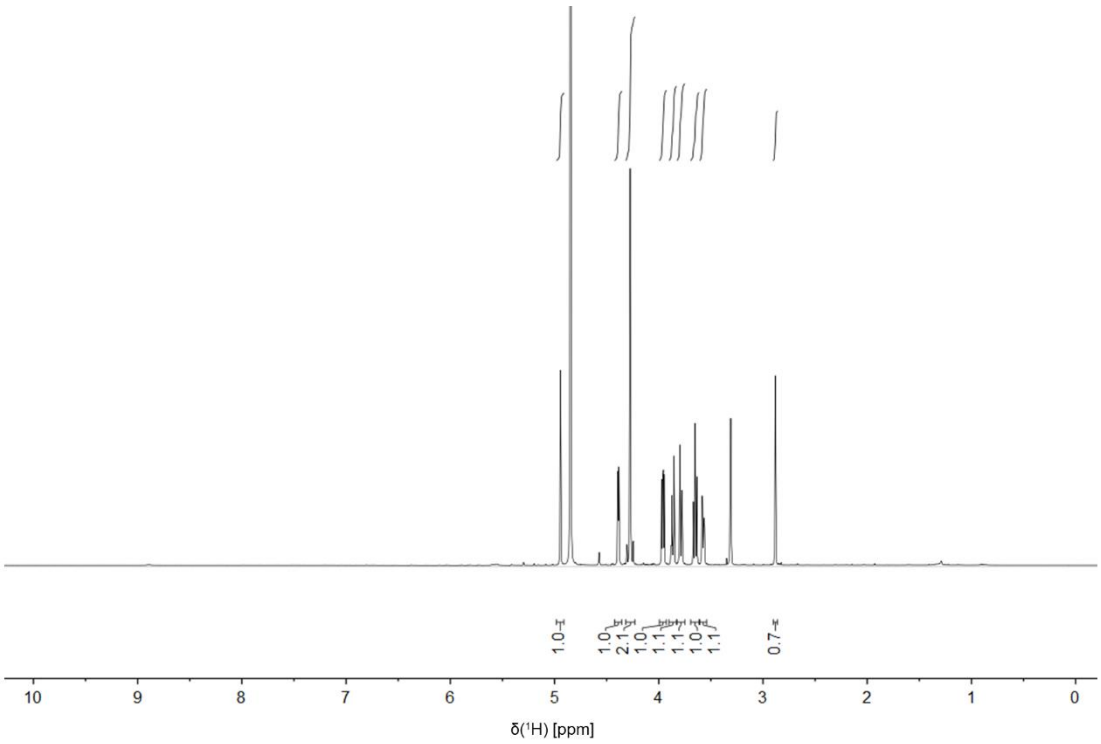


^{13}C NMR (100.6 MHz, MeOD)

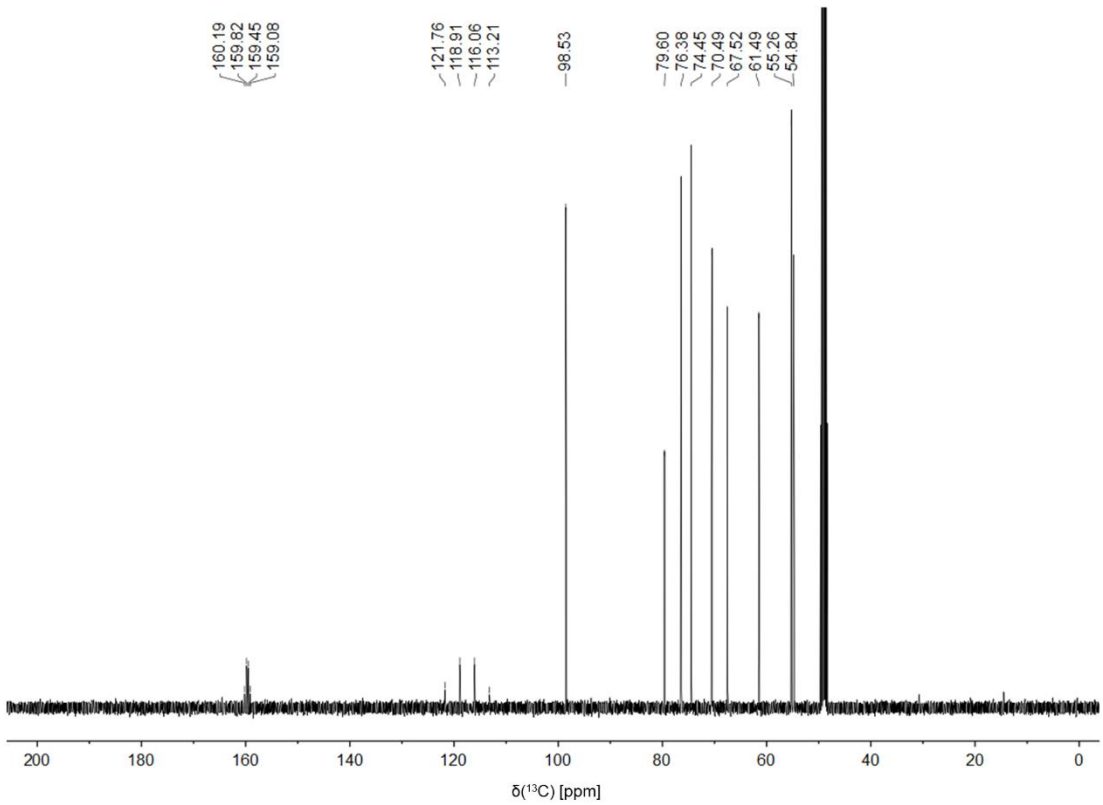


Propargyl-2-deoxy-2-2',2',2'-(trifluoro)acetamido-α-D-mannopyranoside (5.1)

^{1}H NMR (600.0 MHz, MeOD)

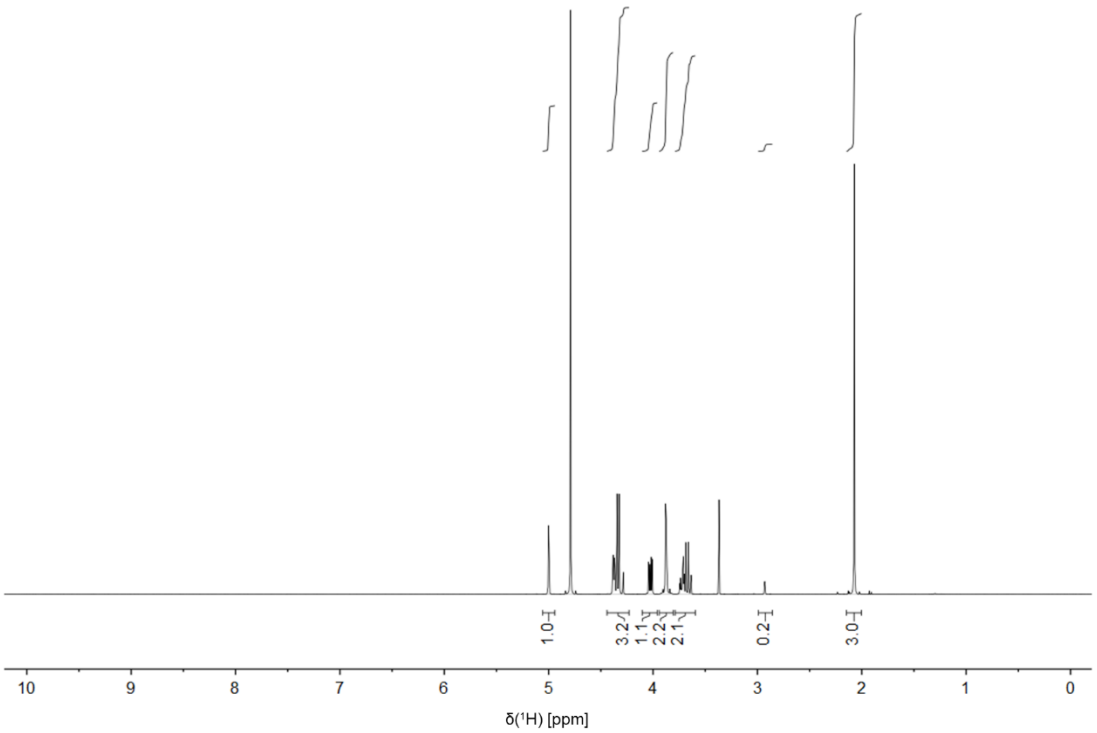


^{1}H NMR (100.6 MHz, MeOD)

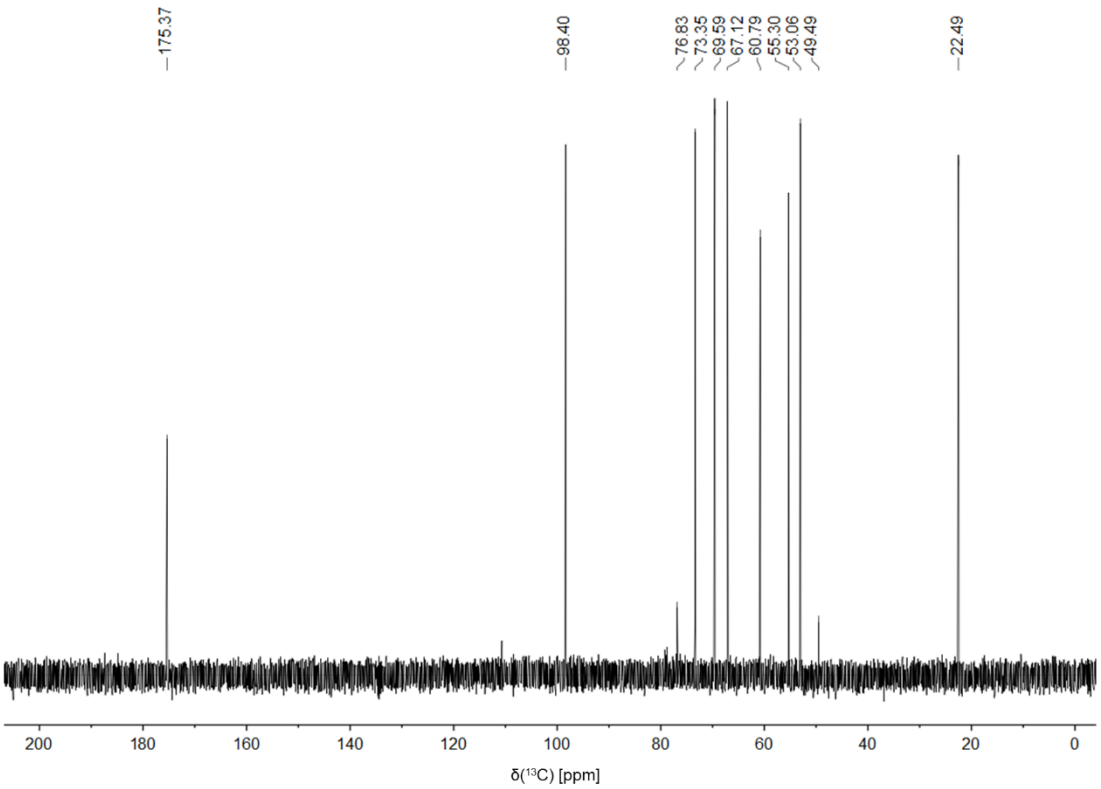


Propargyl-2-deoxy-2-acetamido-α-D-mannopyranoside (5.2)

^{1}H NMR (400.0 MHz, D_2O)

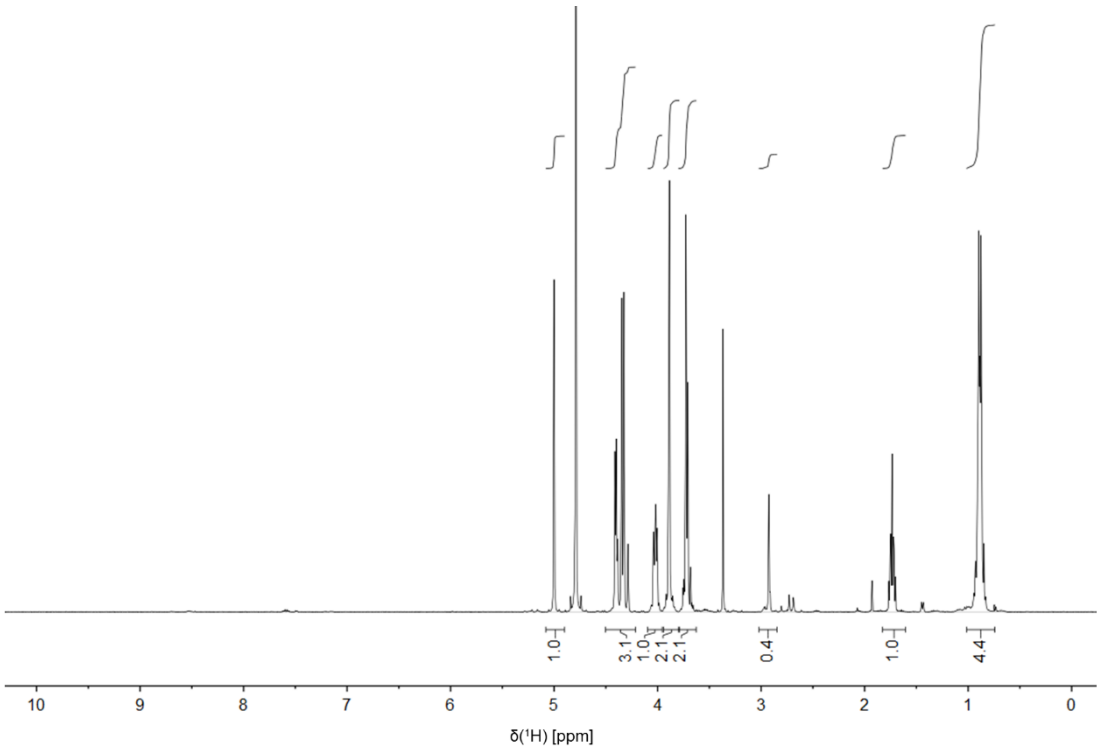


^{13}C NMR (100.6 MHz, D_2O)

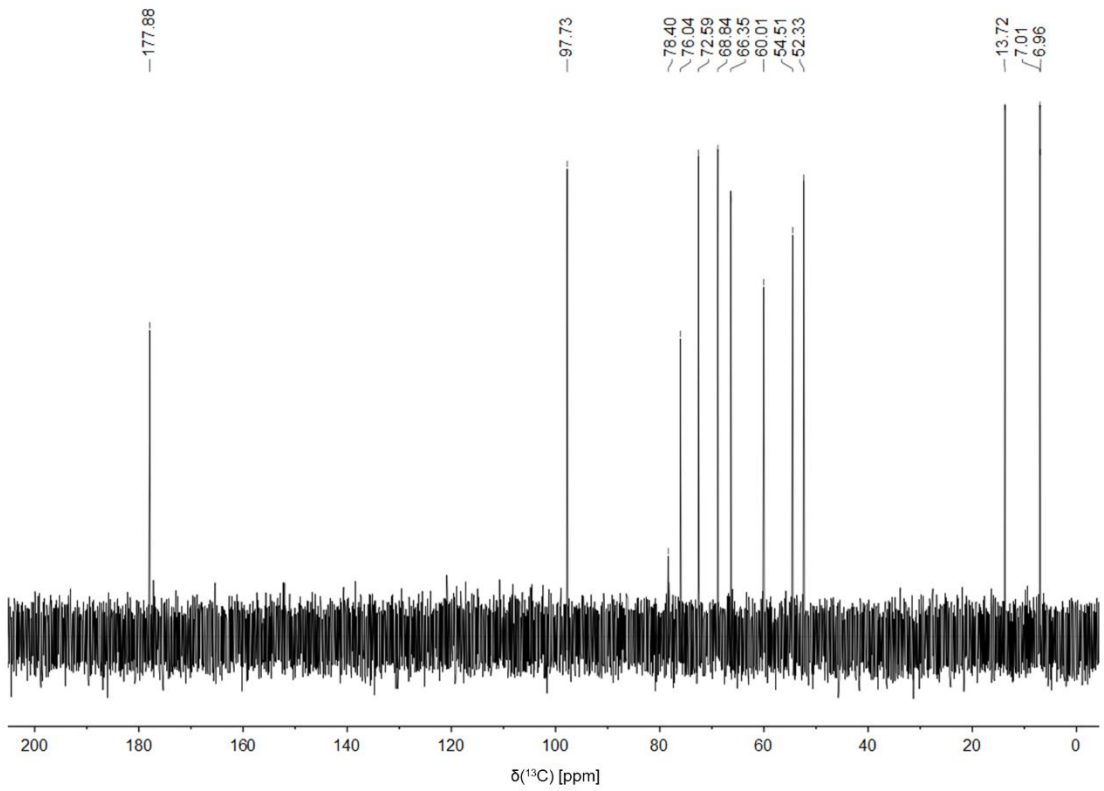


Propargyl-2-deoxy-2-cyclopropanecarboxmido-α-D-mannopyranoside (5.3)

^{1}H NMR (400.0 MHz, D_2O)

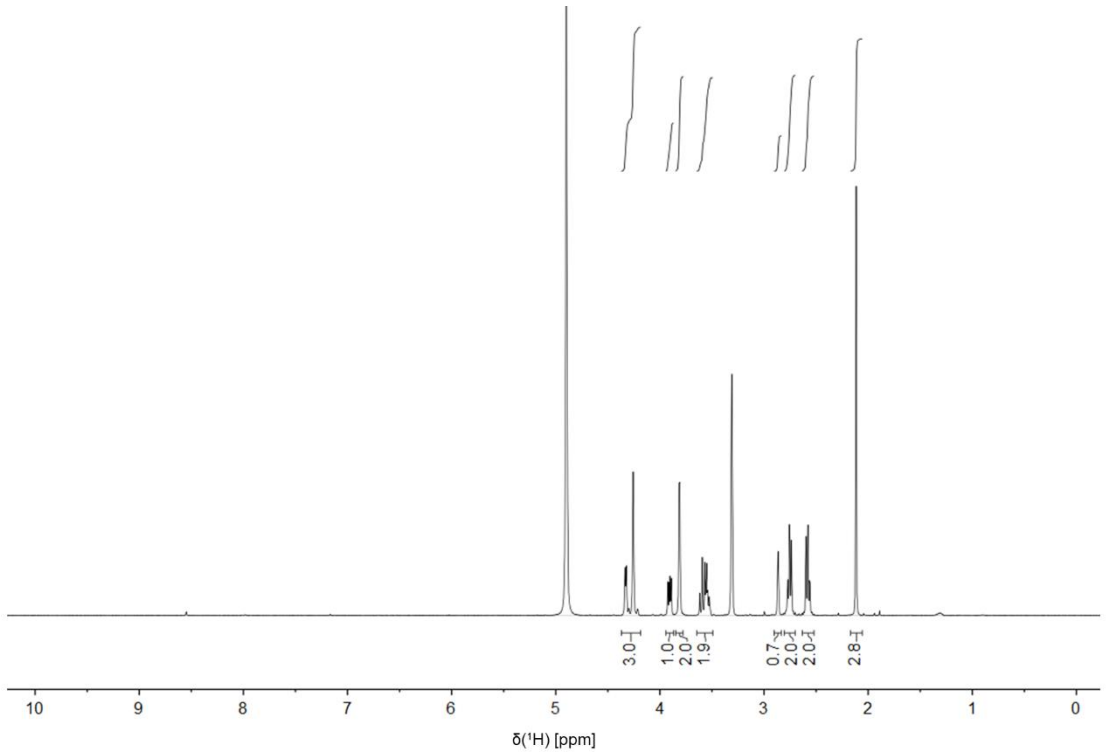


^{13}C NMR (100.6 MHz, D_2O)

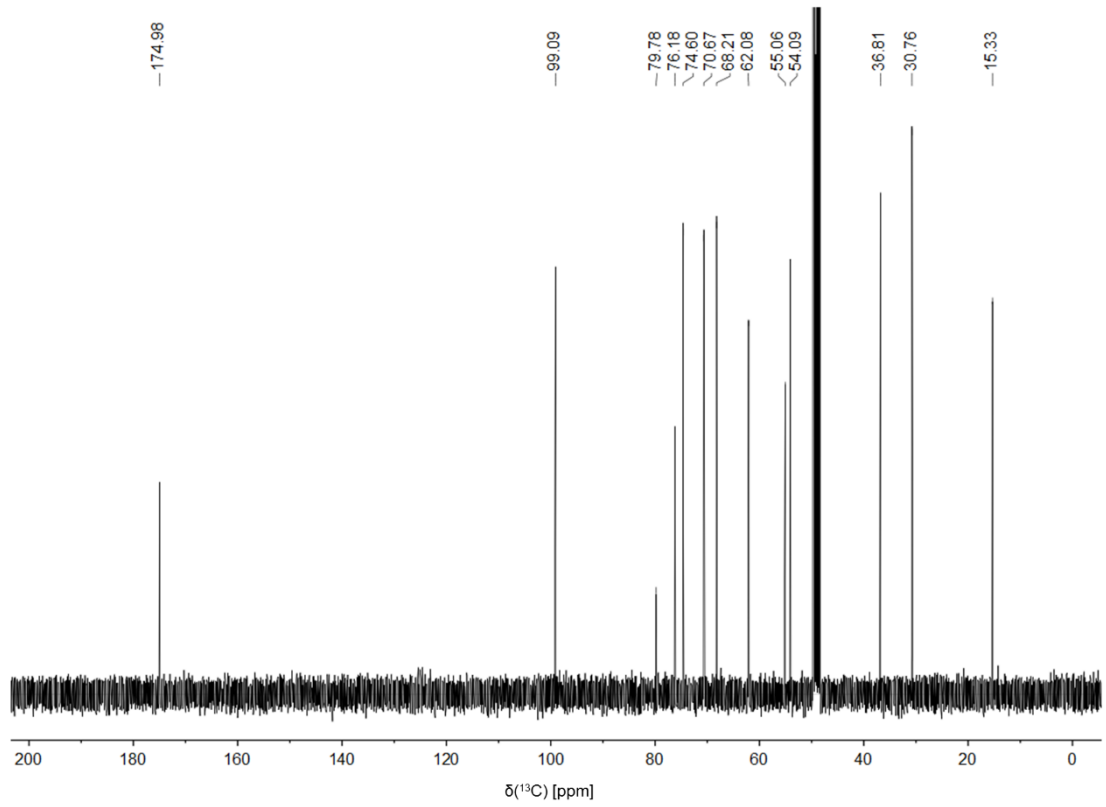


Propargyl-2-deoxy-2-3'-(methylthio)propanamido-α-D-mannopyranoside (5.4)

^{1}H NMR (400.0 MHz, MeOD)

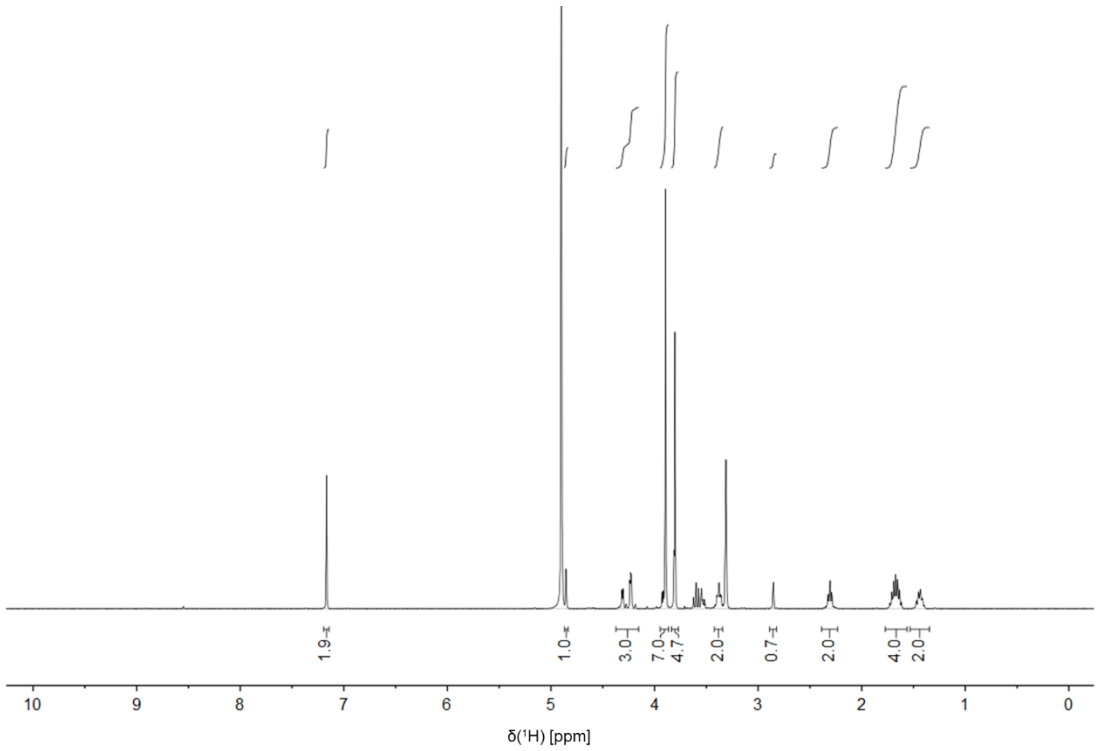


^{13}C NMR (100.6 MHz, MeOD)

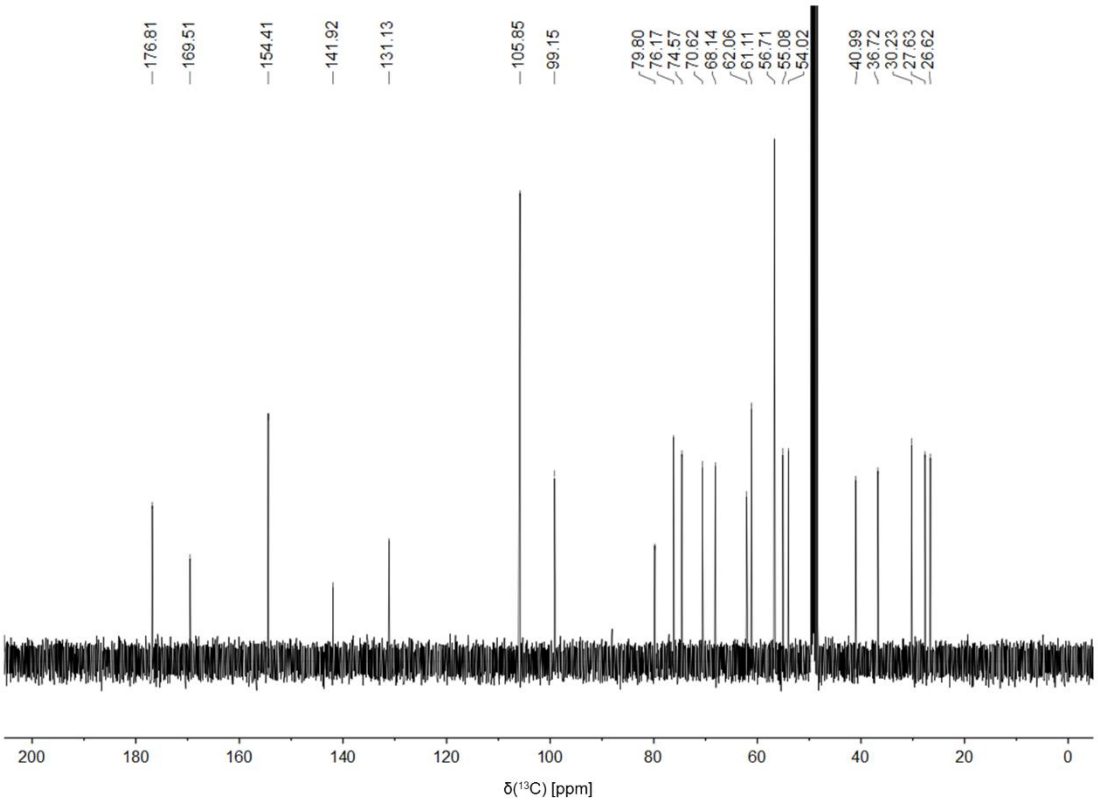


Propargyl-2-deoxy-2-6’-(3’’,4’’,5’’-(trimethoxy)benzamido)hexanamido-α-D-mannopyranoside (5.5)

^{1}H NMR (400.0 MHz, MeOD)

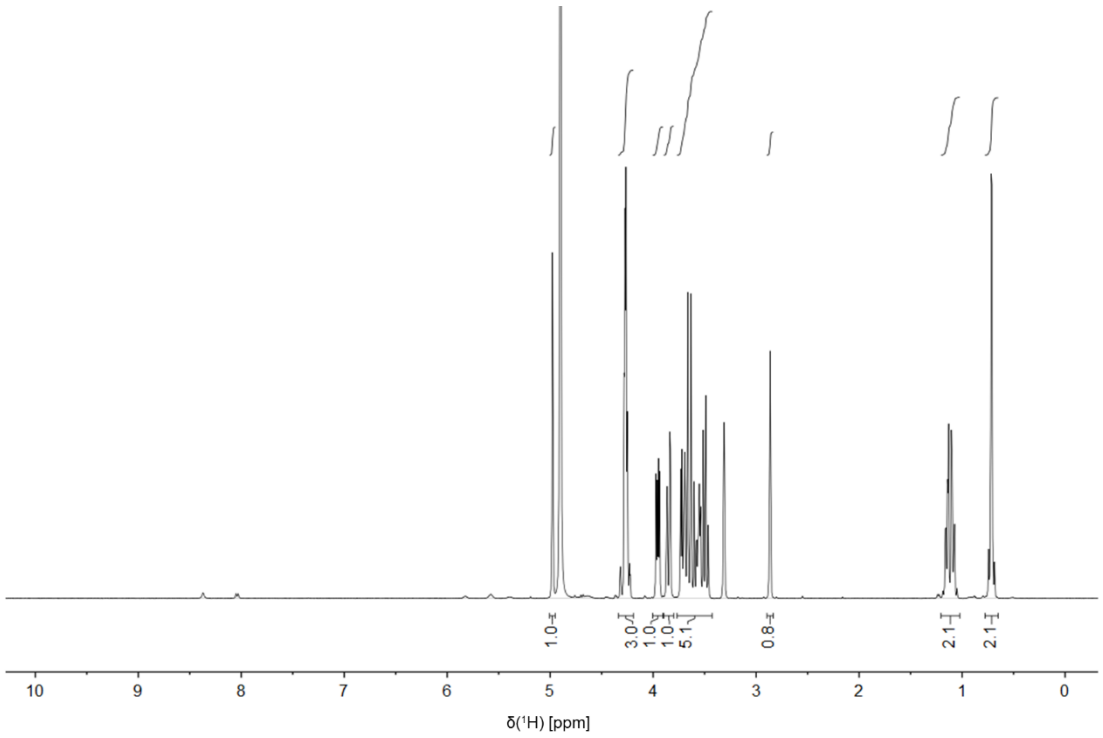


^{13}C NMR (100.6 MHz, MeOD)

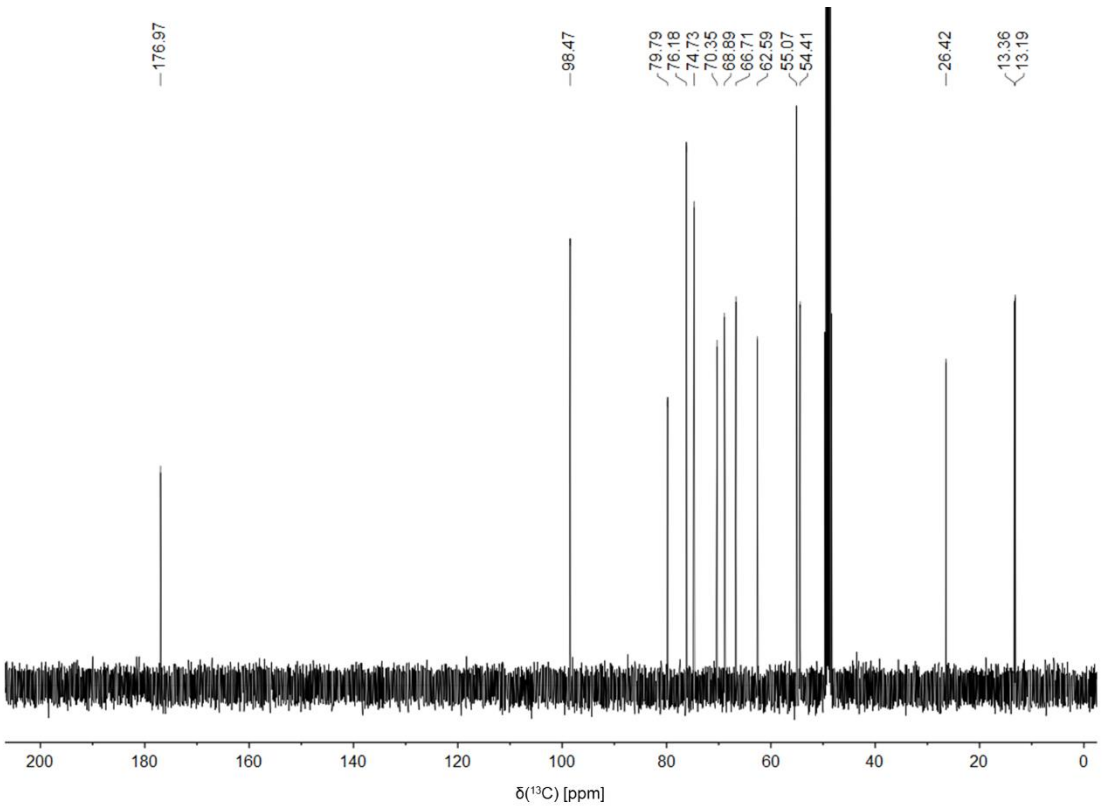


Propargyl-2-deoxy-2-1'-(hydroxymethyl)cyclopropanecarboxamide-α-D-mannopyranoside (5.6)

^{1}H NMR (400.0 MHz, MeOD)

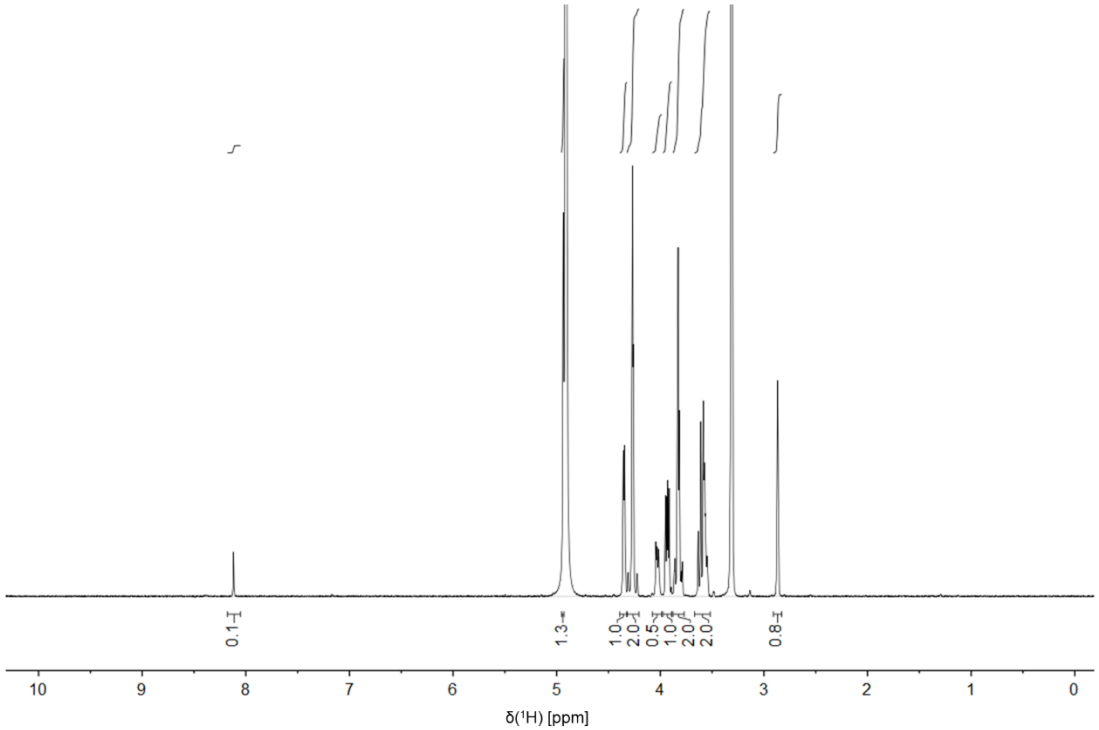


^{13}C NMR (100.6 MHz, MeOD)

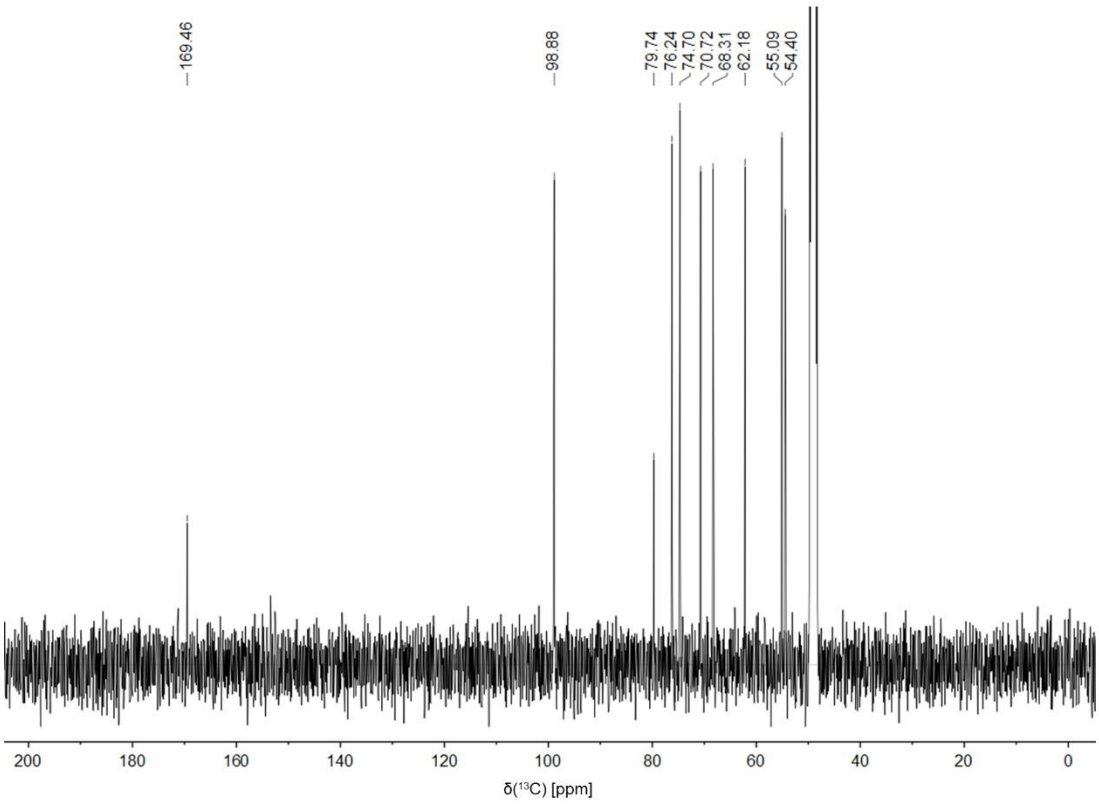


Propargyl-2-deoxy-2-2'-(*1H*-tetrazol-5'-yl)acetamido-α-D-mannopyranoside (5.7)

^{1}H NMR (400.0 MHz, MeOD)

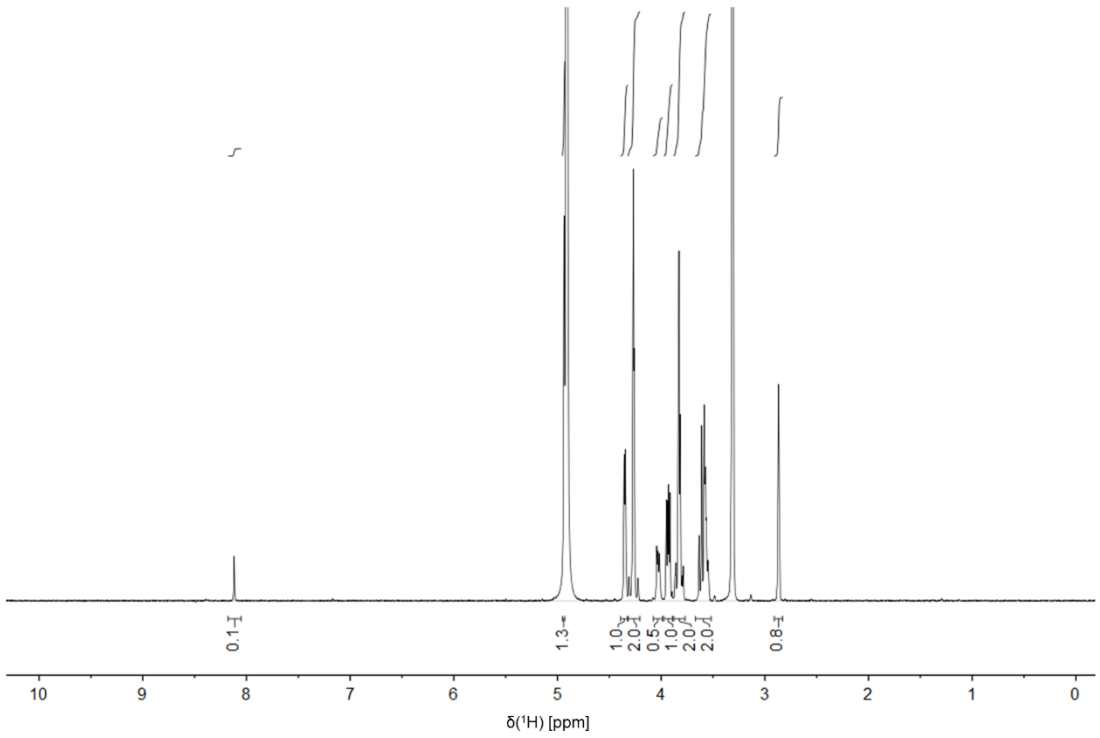


^{13}C NMR (100.6 MHz, MeOD)

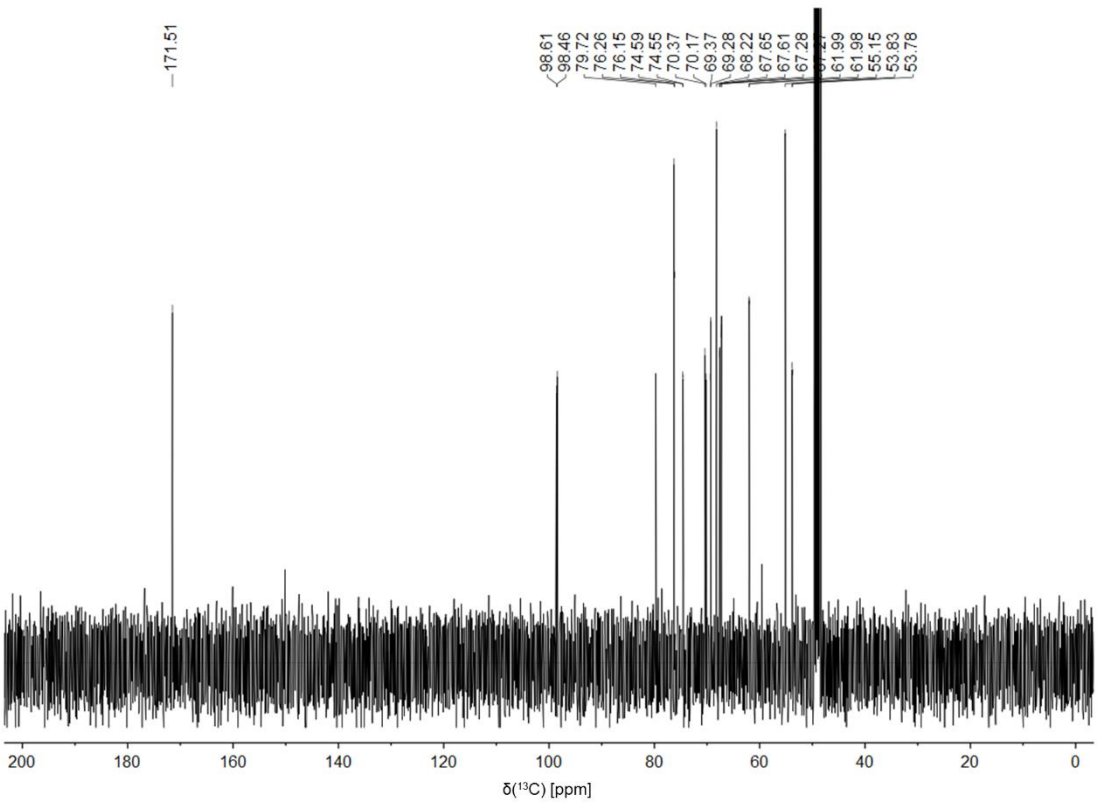


Propargyl-2-deoxy-2-1',4'-dioxane-2'-carboxamido-α-D-mannopyranoside (5.8)

^{1}H NMR (400.0 MHz, MeOD)

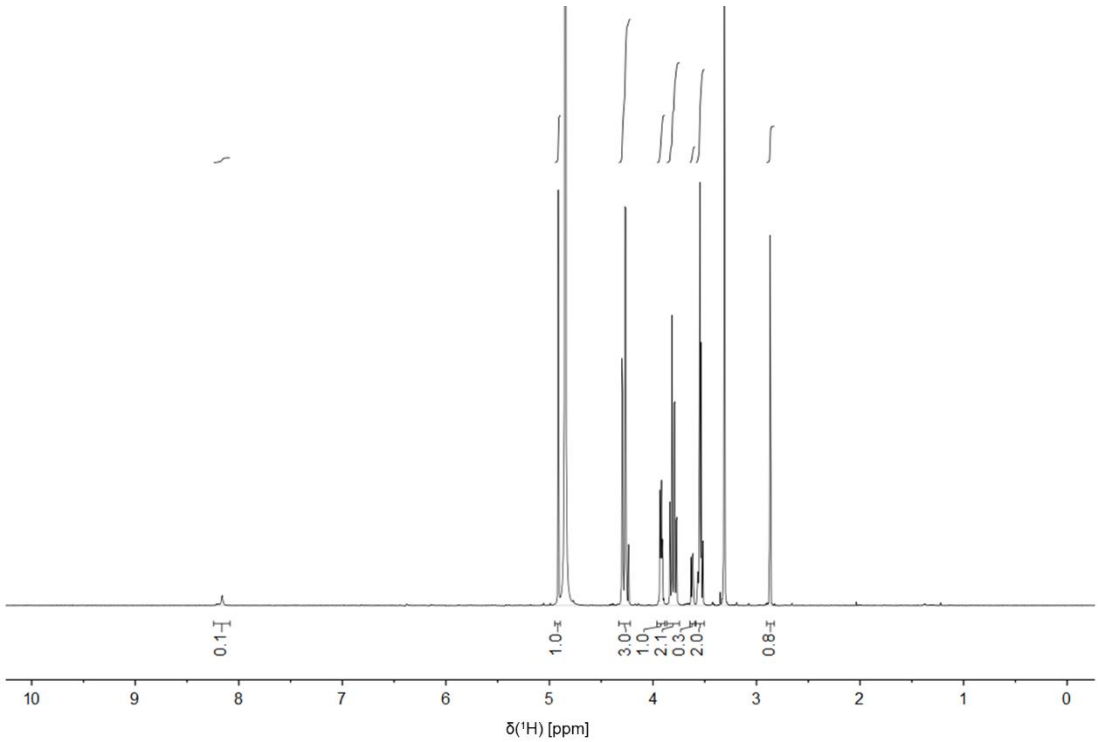


^{13}C NMR (100.6 MHz, MeOD)

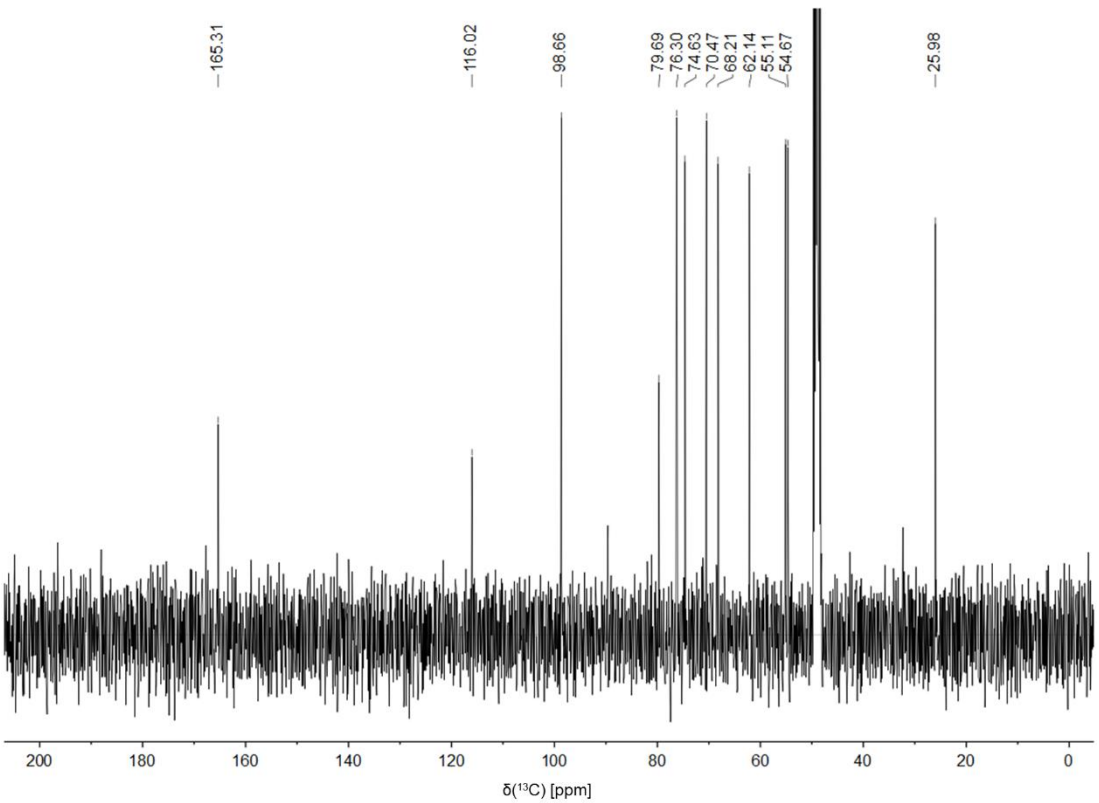


Propargyl-2-deoxy-2-2’-cyanoacetamido-α-D-mannopyranoside (5.9)

^{1}H NMR (600.0 MHz, MeOD)

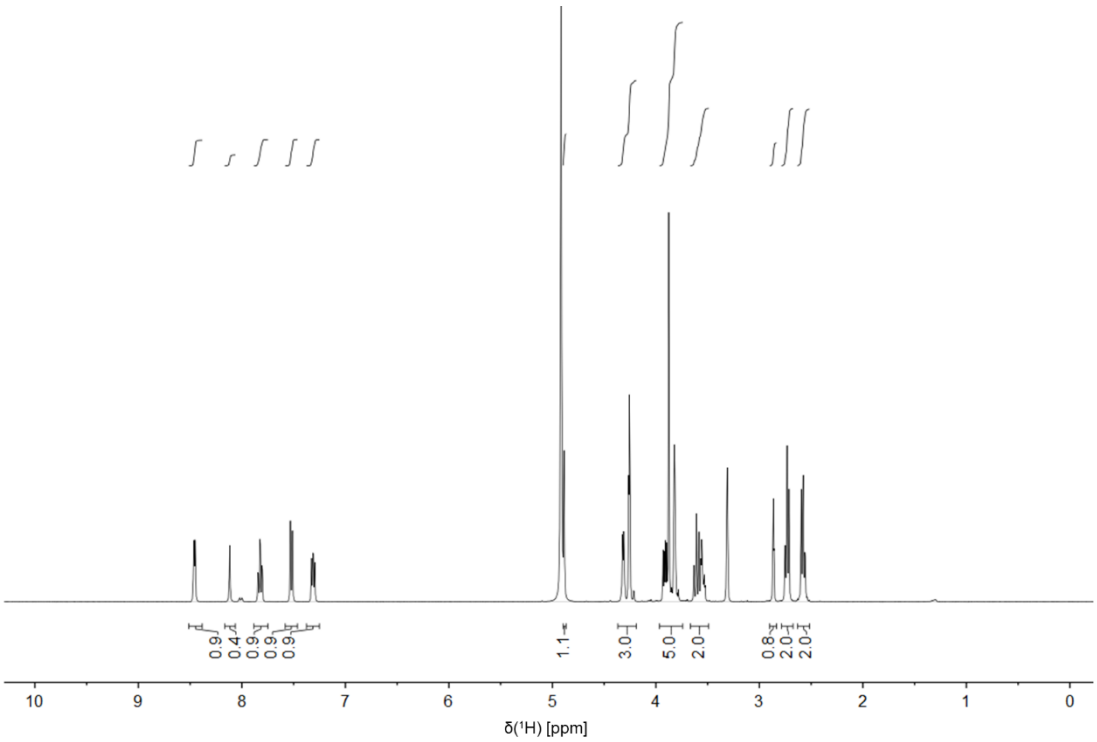


^{13}C NMR (100.6 MHz, MeOD)

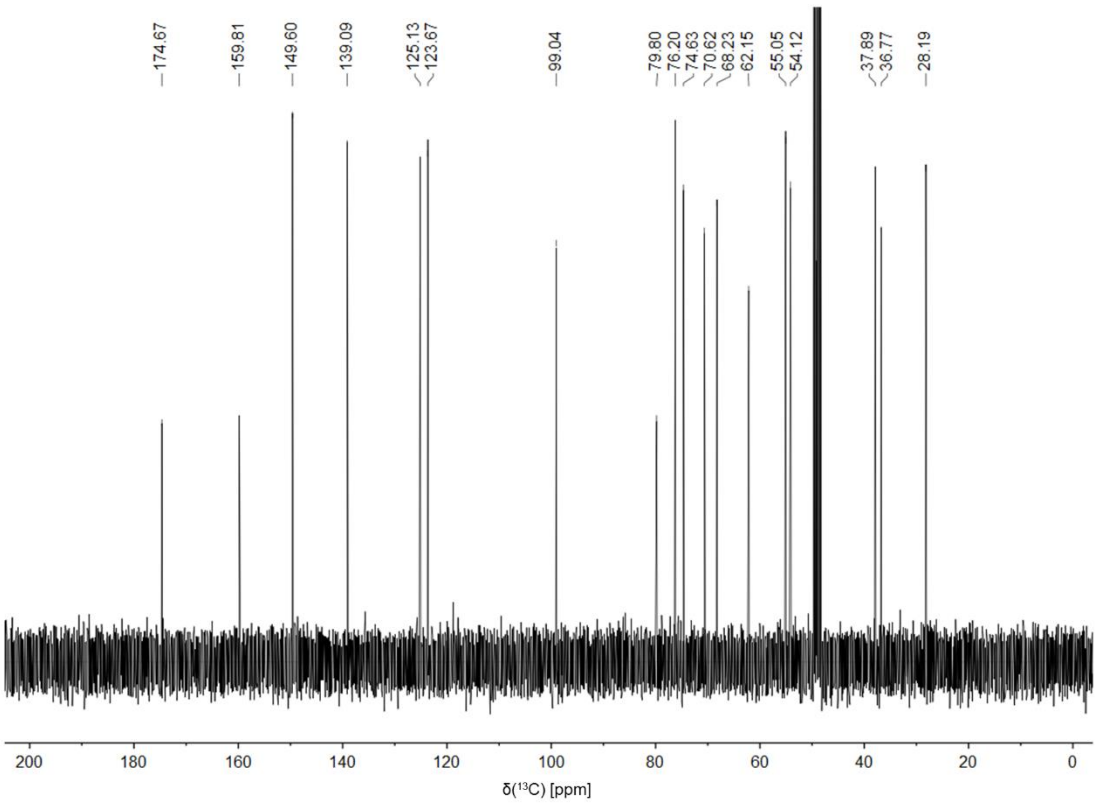


Propargyl-2-deoxy-2-3'-((pyridin-2'-ylmethyl)thio)-propanamido-α-D-mannopyranoside (5.10)

^{1}H NMR (400.0 MHz, MeOD)

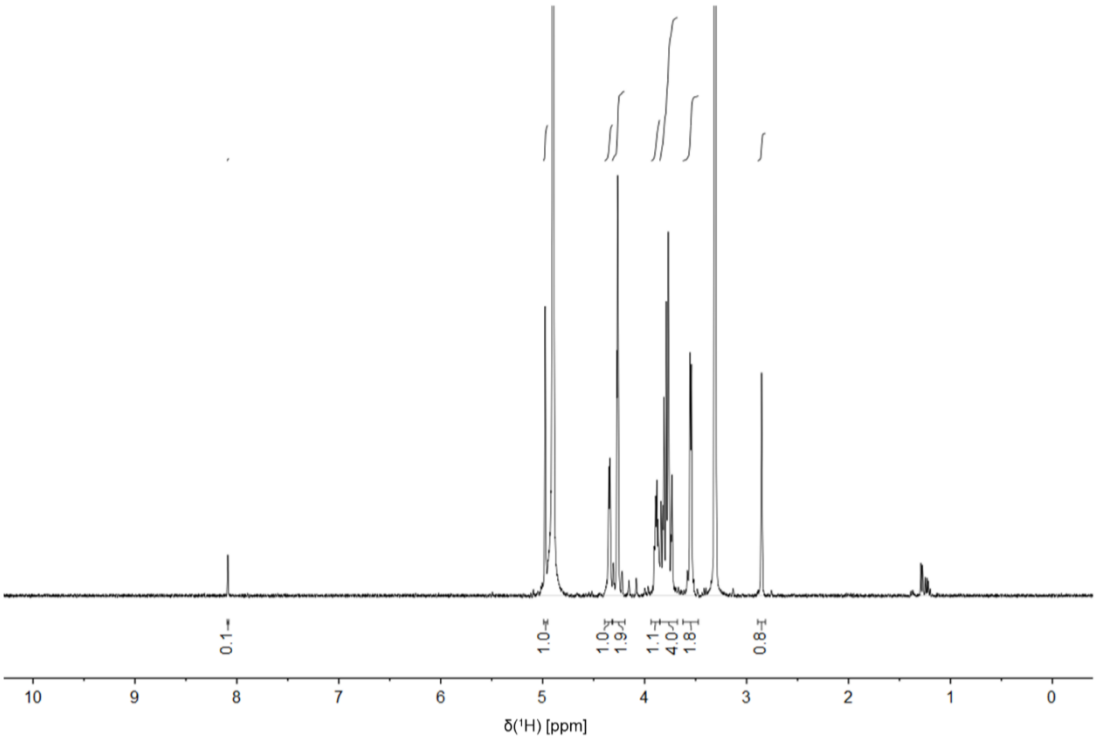


^{13}C NMR (100.6 MHz, MeOD)

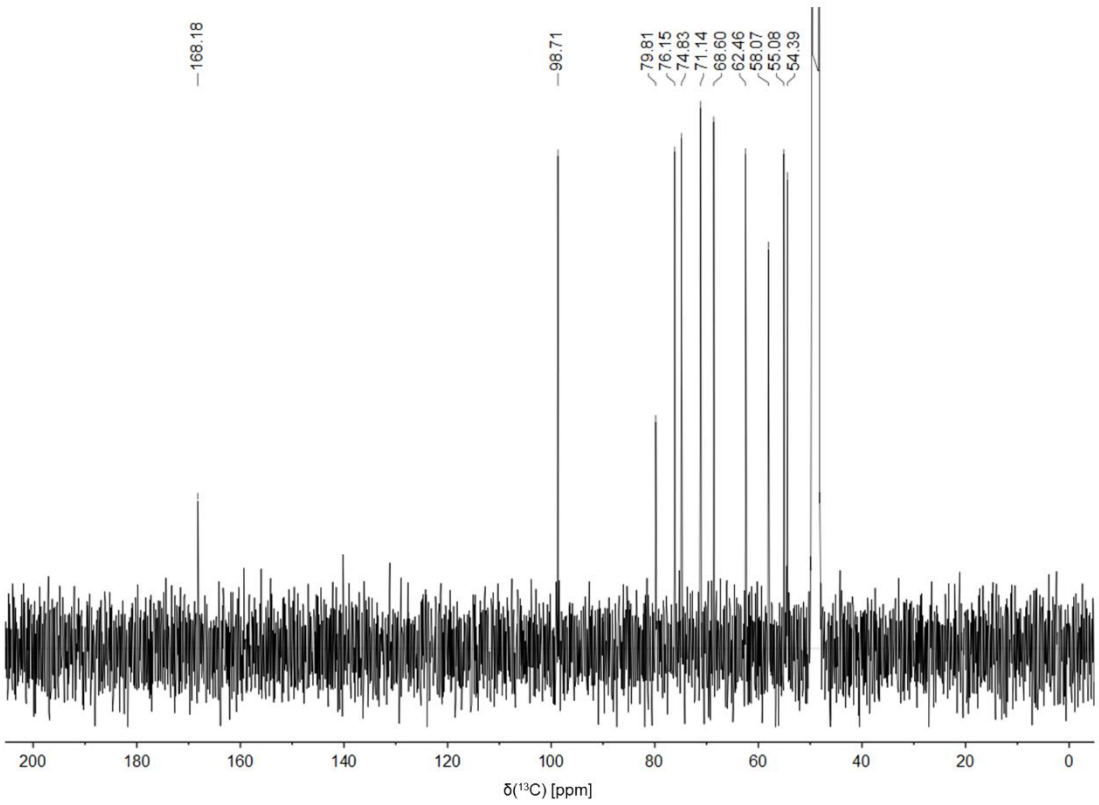


Propargyl-2-deoxy-2-2'-sulfoacetamido-α-D-mannopyranoside (5.11)

^{1}H NMR (600.0 MHz, MeOD)

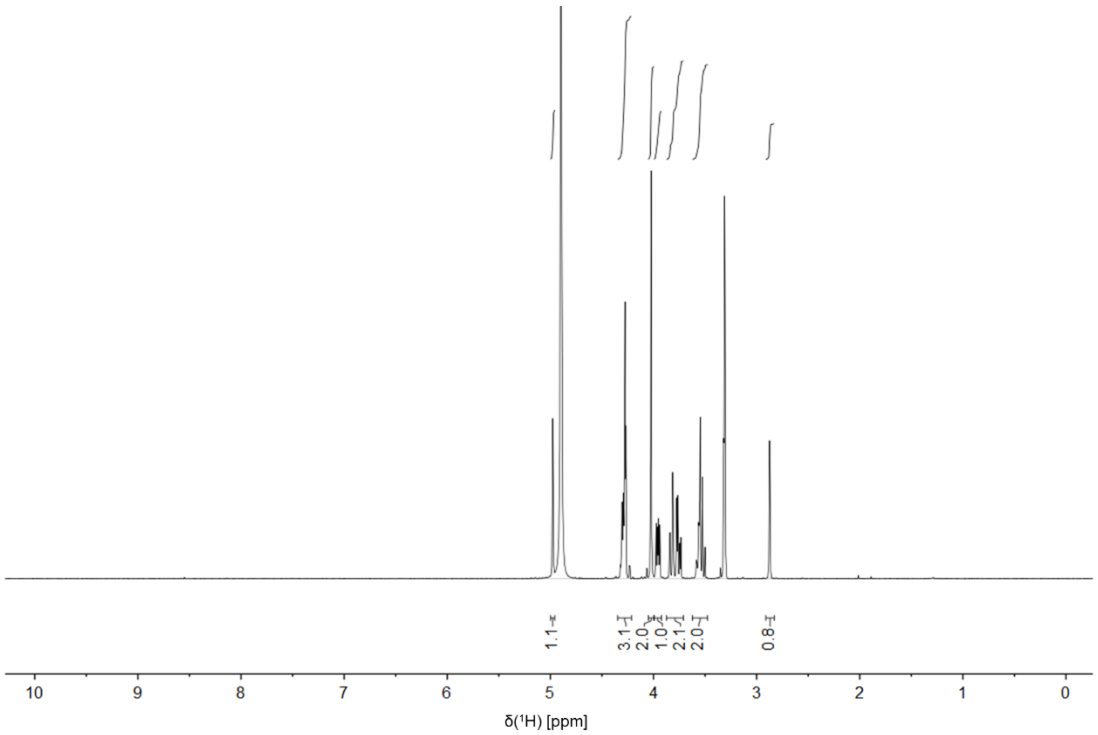


^{13}C NMR (100.6 MHz, MeOD)

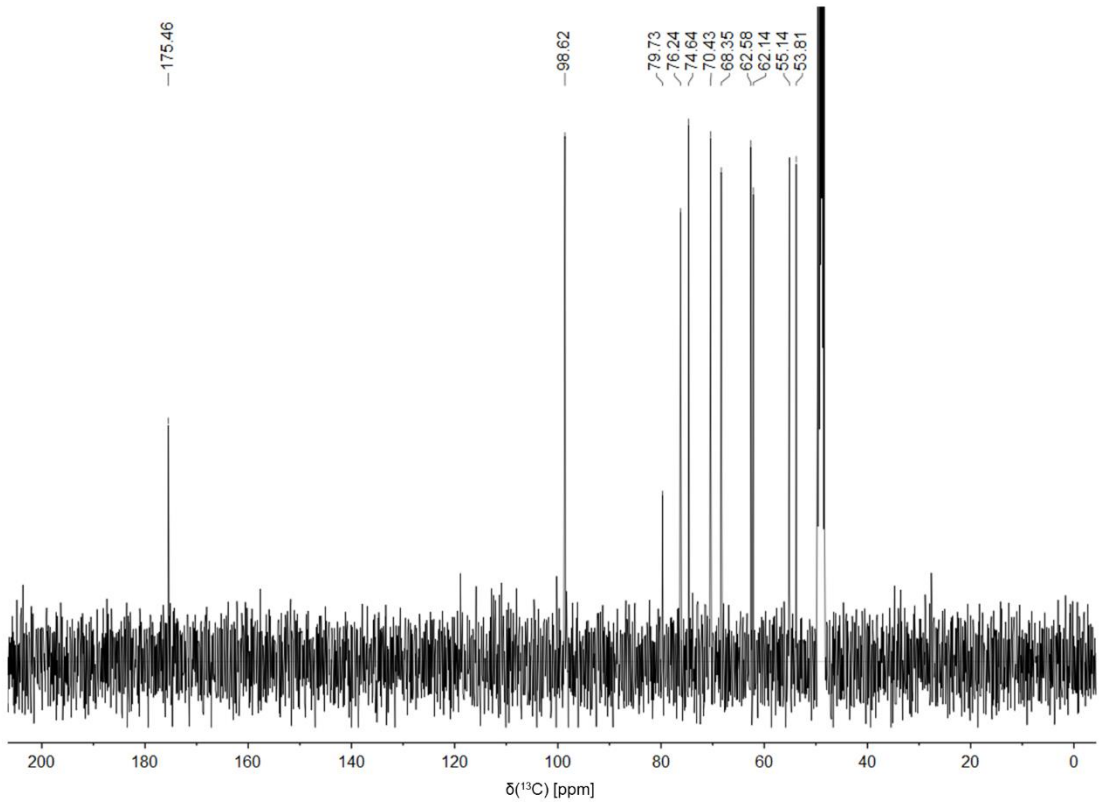


Propargyl-2-deoxy-2-2'-(thiazol-2''-yl)acetamido-α-D-mannopyranoside (5.12)

^{1}H NMR (600.0 MHz, MeOD)

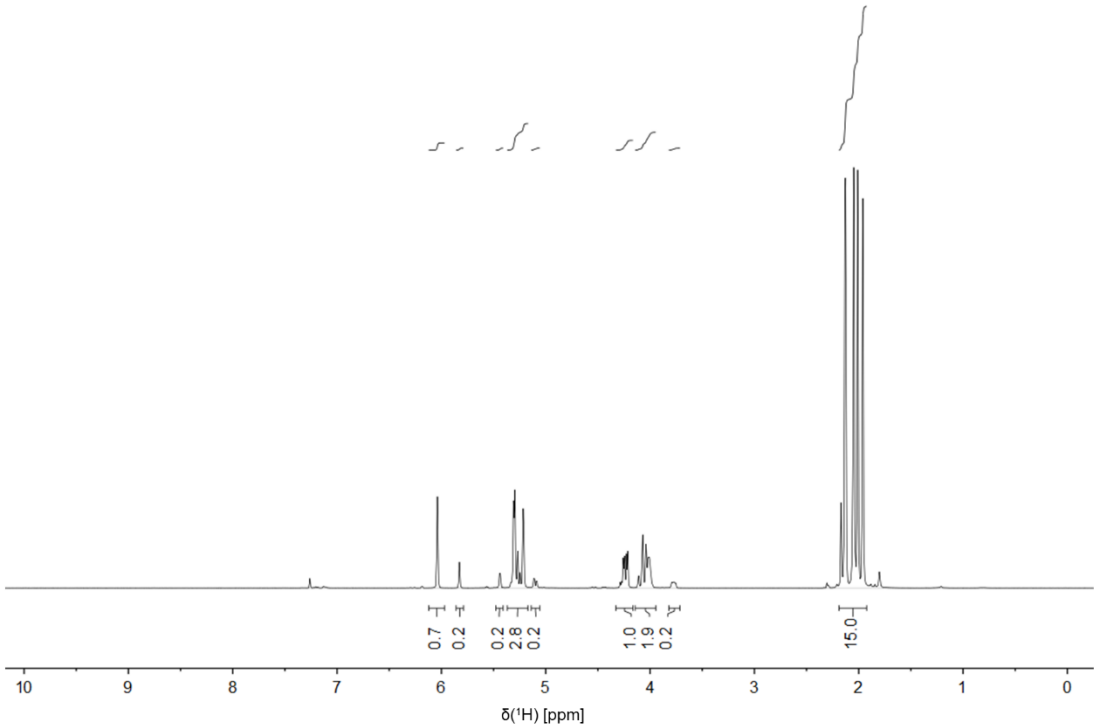


^{13}C NMR (100.6 MHz, MeOD)

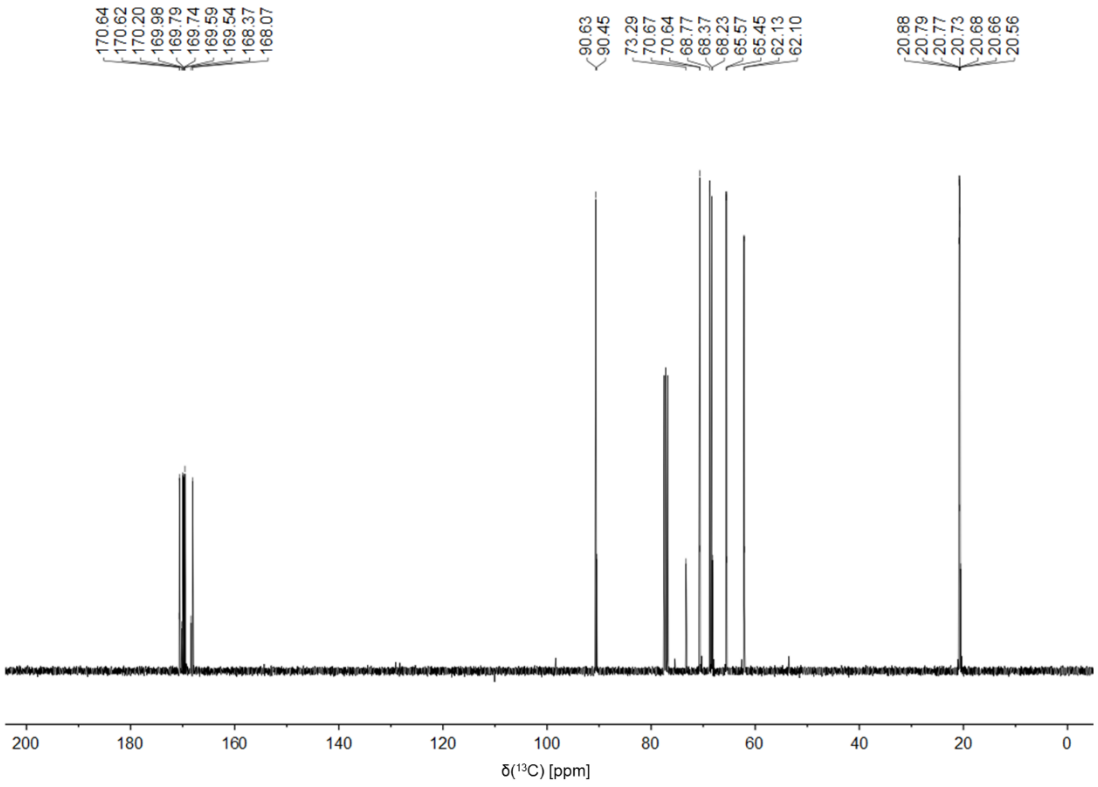


1,2,3,4,6-penta-O-acetyl-D-mannopyranose (10)

^{1}H NMR (400.0 MHz, $CDCl_3$)

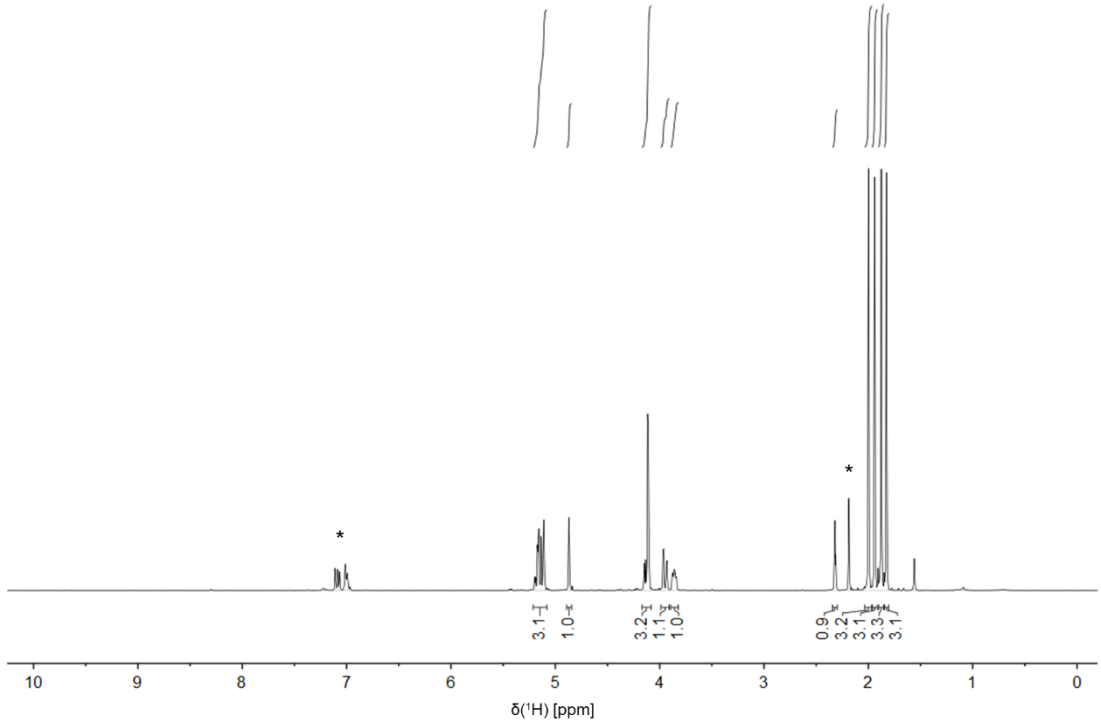


^{13}C NMR (100.6 MHz, $CDCl_3$)

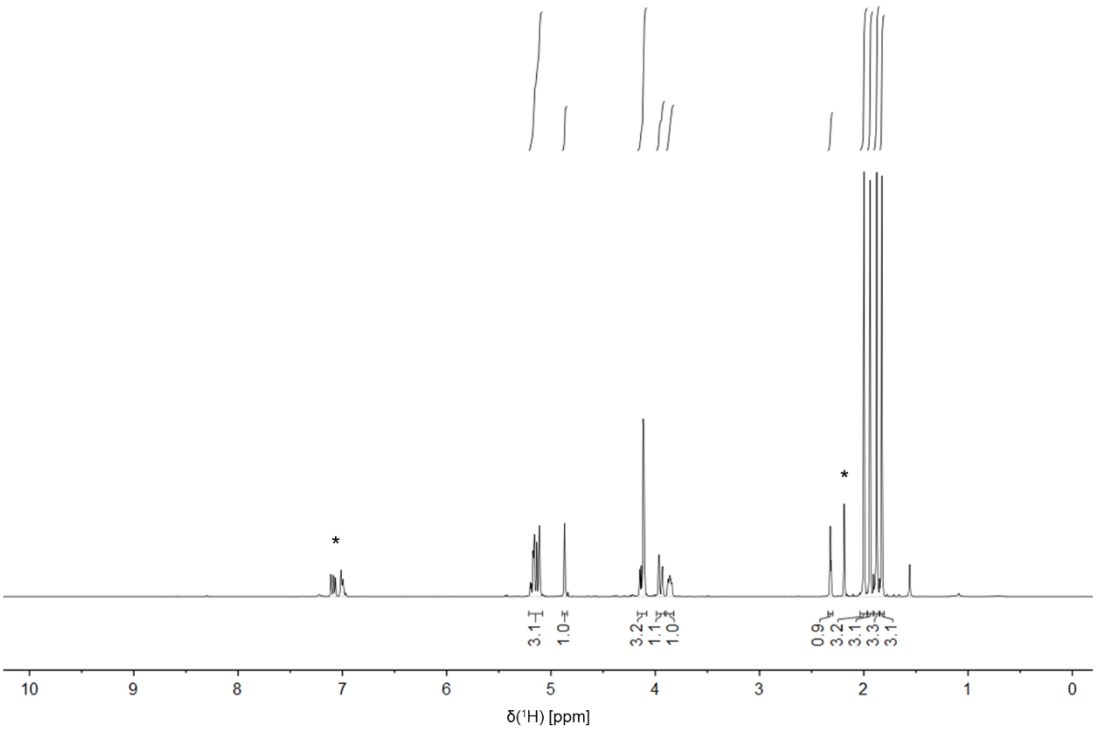


Propargyl-2,3,4,6-tetra-O-acetyl-α-D-mannopyranoside (11)

^{1}H NMR (400.0 MHz, $CDCl_3$)



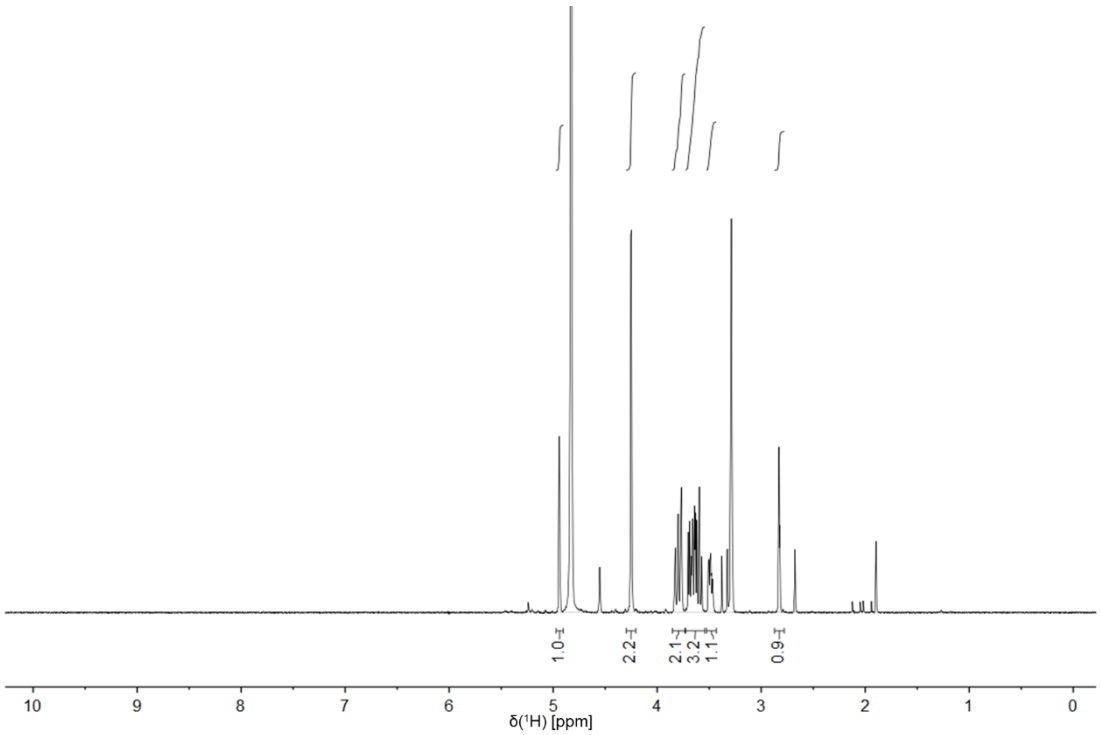
^{13}C NMR (100.6 MHz, $CDCl_3$)



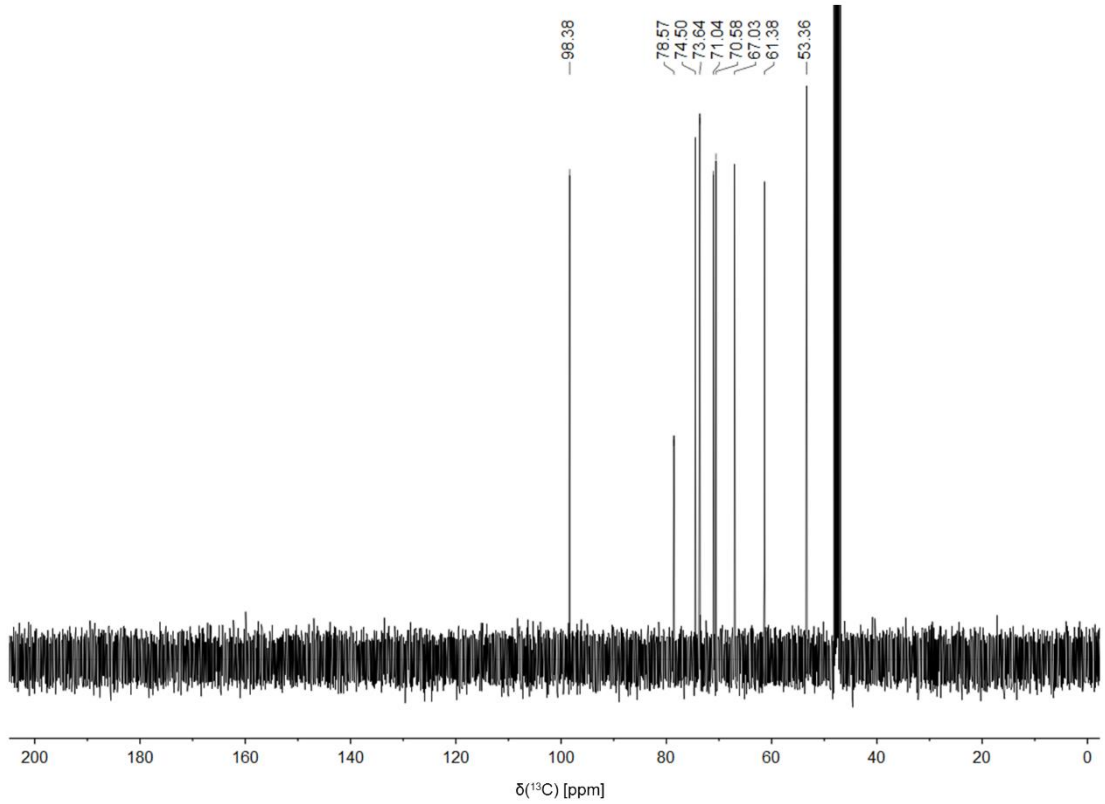
*Sample contained residual toluene. The yield was determined after removal of the solvent *in vacuo*.

Propargyl-α-D-mannopyranoside (12.1)

^{1}H NMR (400.0 MHz, MeOD)

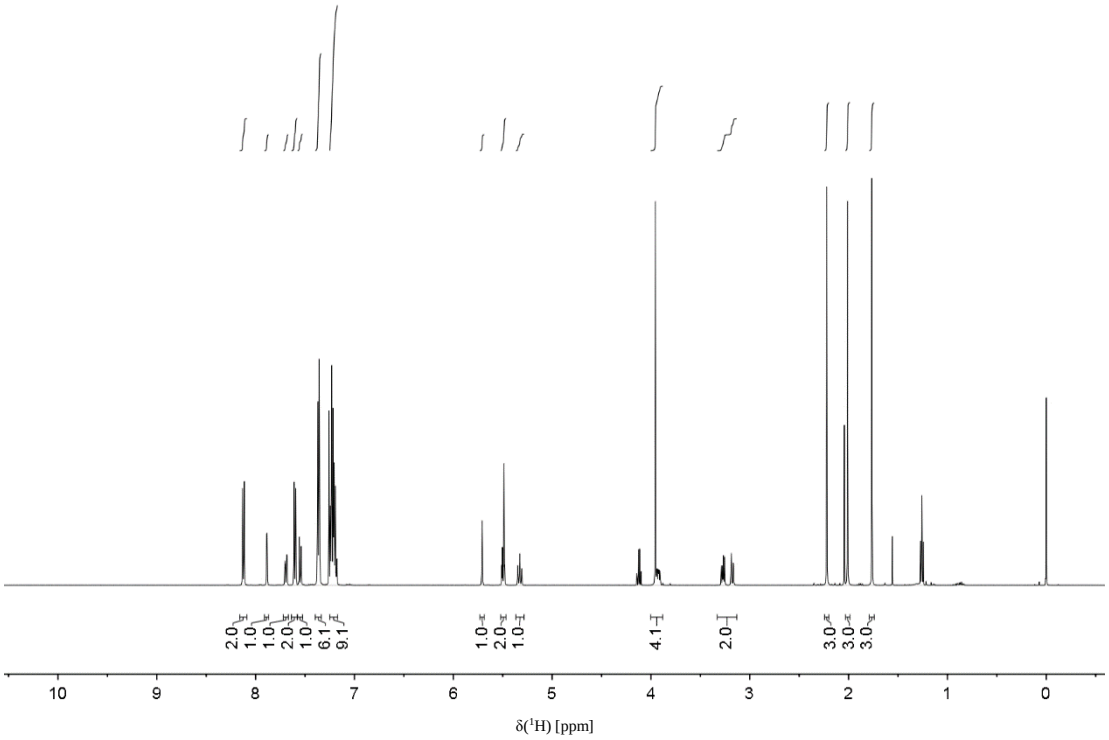


^{13}C NMR (100.6 MHz, MeOD)

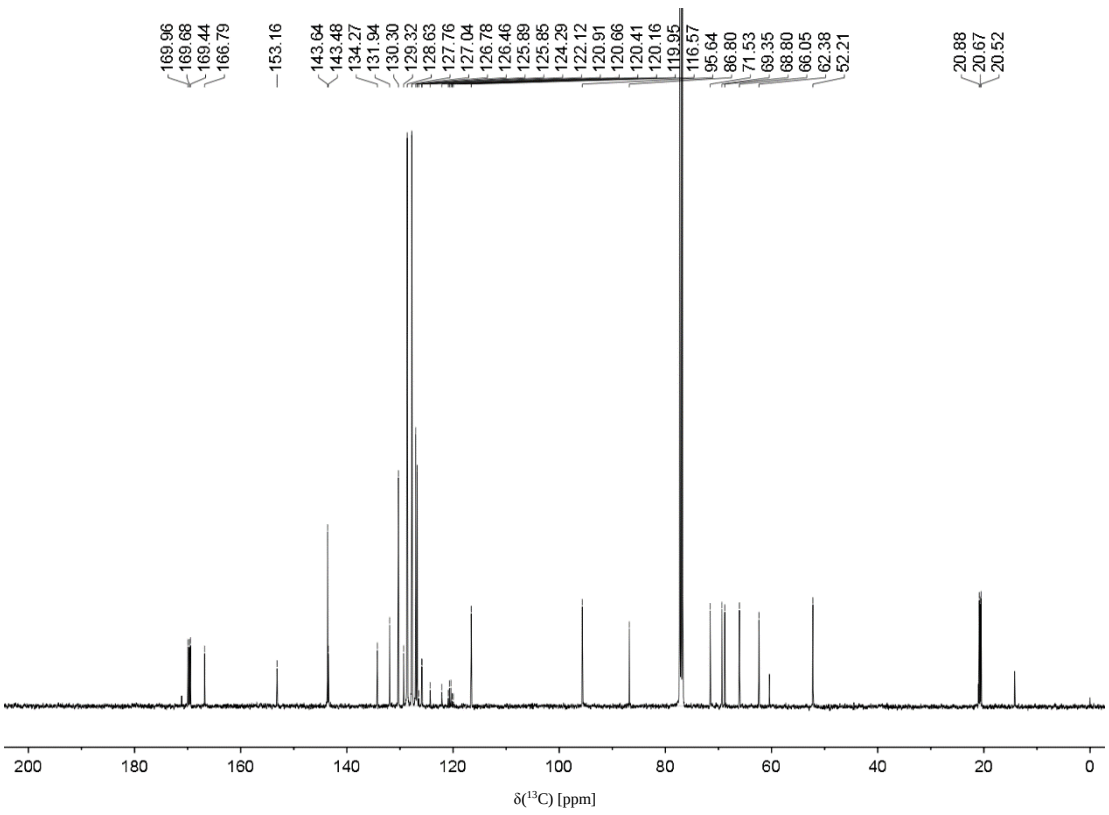


Methyl-4'-(2'',3'',4''-tetra-O-acetyl-6''-O-trityl-α-D-mannopyranosyloxy)-3'-trifluoromethylbiphenyl-4-carboxylate (20)

^{1}H NMR (500.0 MHz, $CDCl_3$)

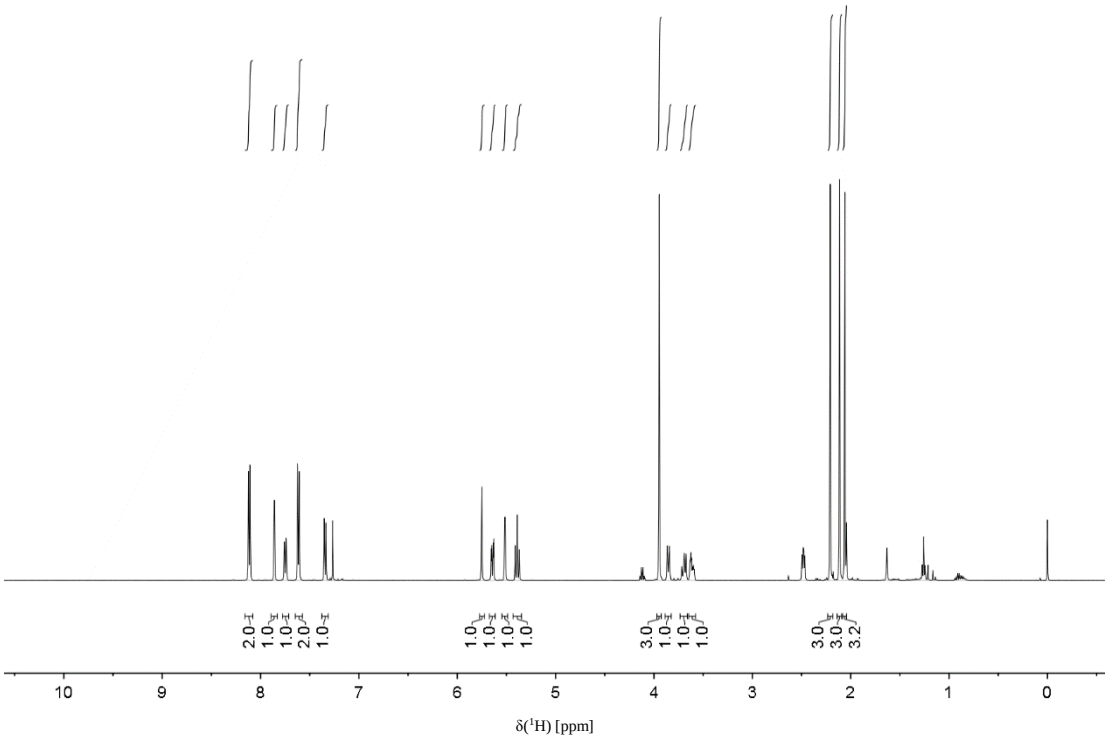


^{13}C NMR (125.8 MHz, $CDCl_3$)

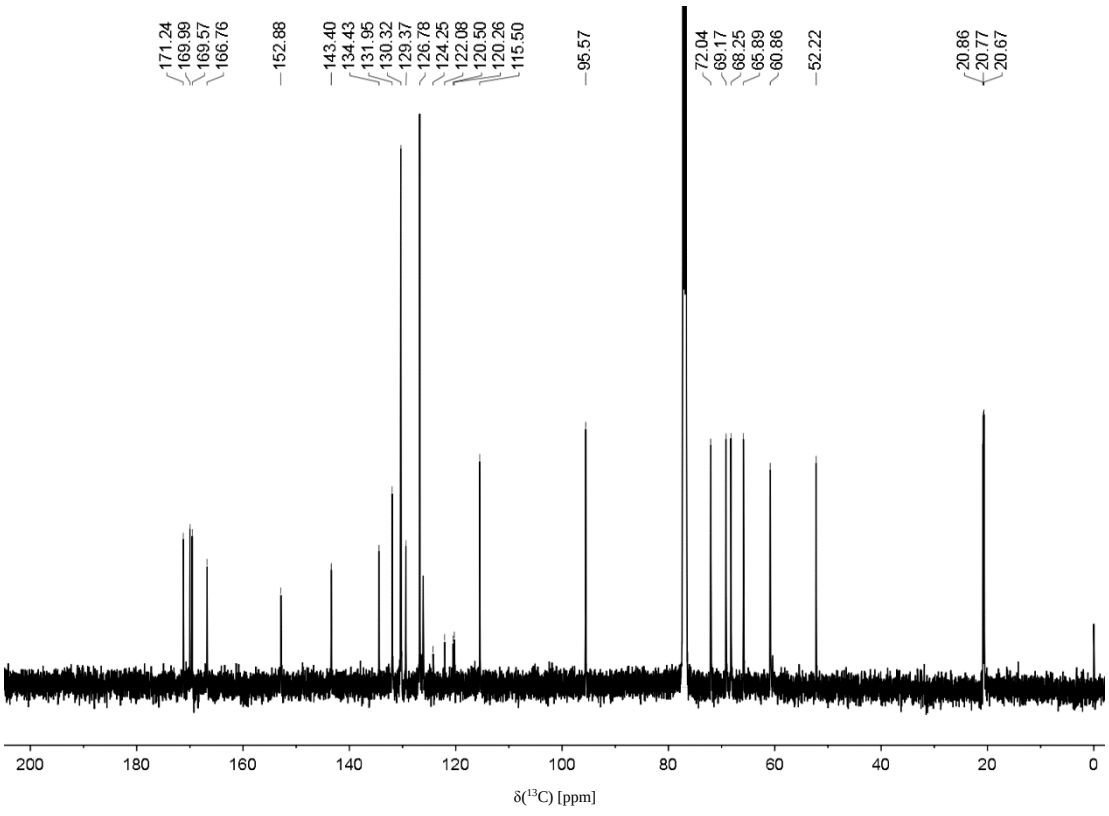


Methyl-4’-(2’’,3’’,4’’-tetra-O-acetyl-α-D-mannopyranosyloxy)-3’-trifluoromethylbiphenyl-4-carboxylate (21)

^{1}H NMR (500.0 MHz, $CDCl_3$)

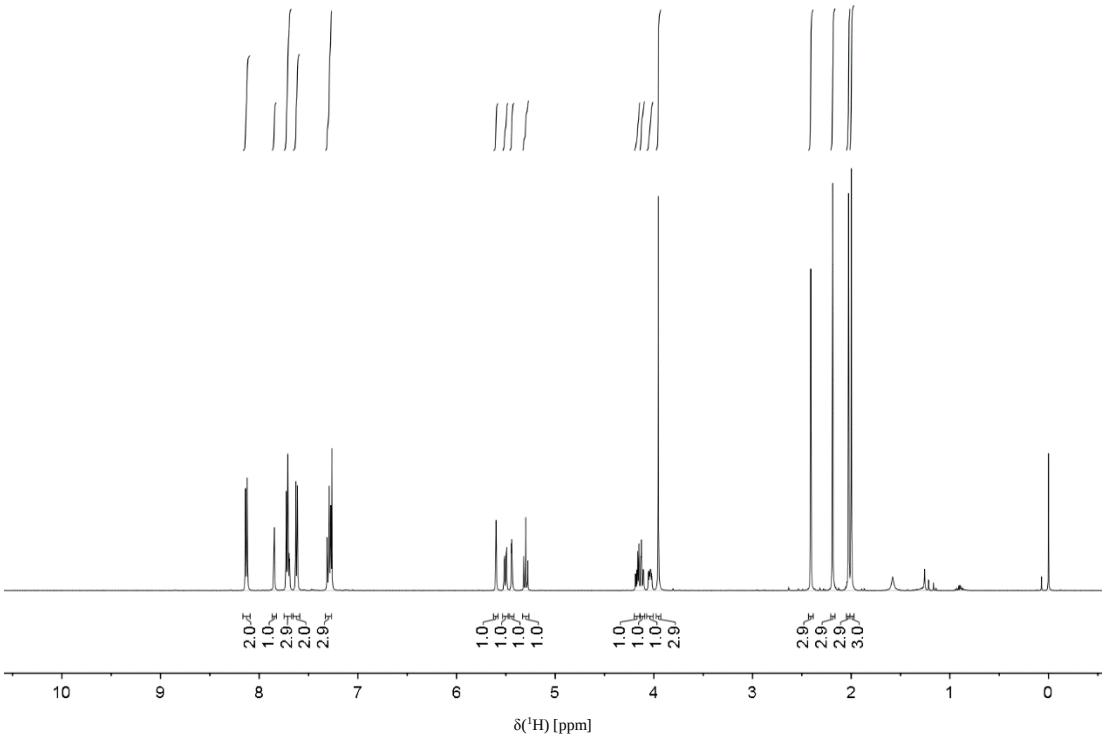


^{13}C NMR (125.8 MHz, $CDCl_3$)

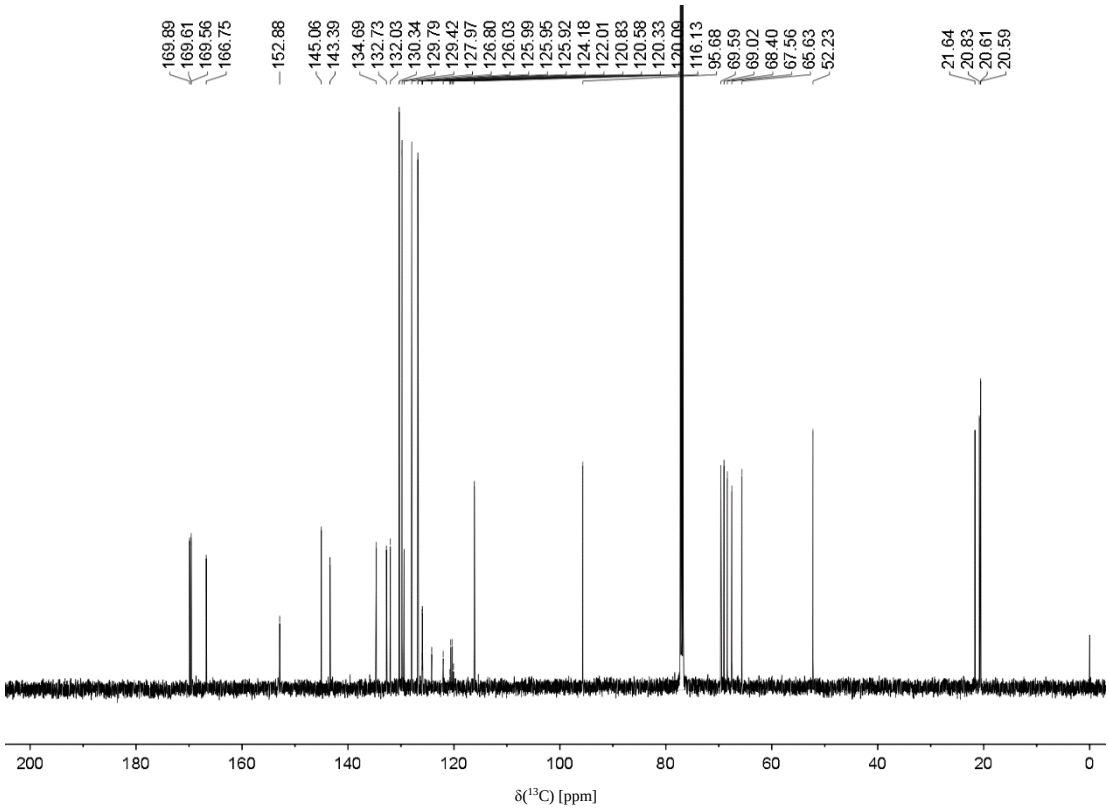


Methyl-4'-(2'',3'',4''-tetra-O-acetyl-6''-O-tosyl-α-D-mannopyranosyloxy)-3'-trifluoromethylbiphenyl-4-carboxylate (22)

^{1}H NMR (500.0 MHz, $CDCl_3$)

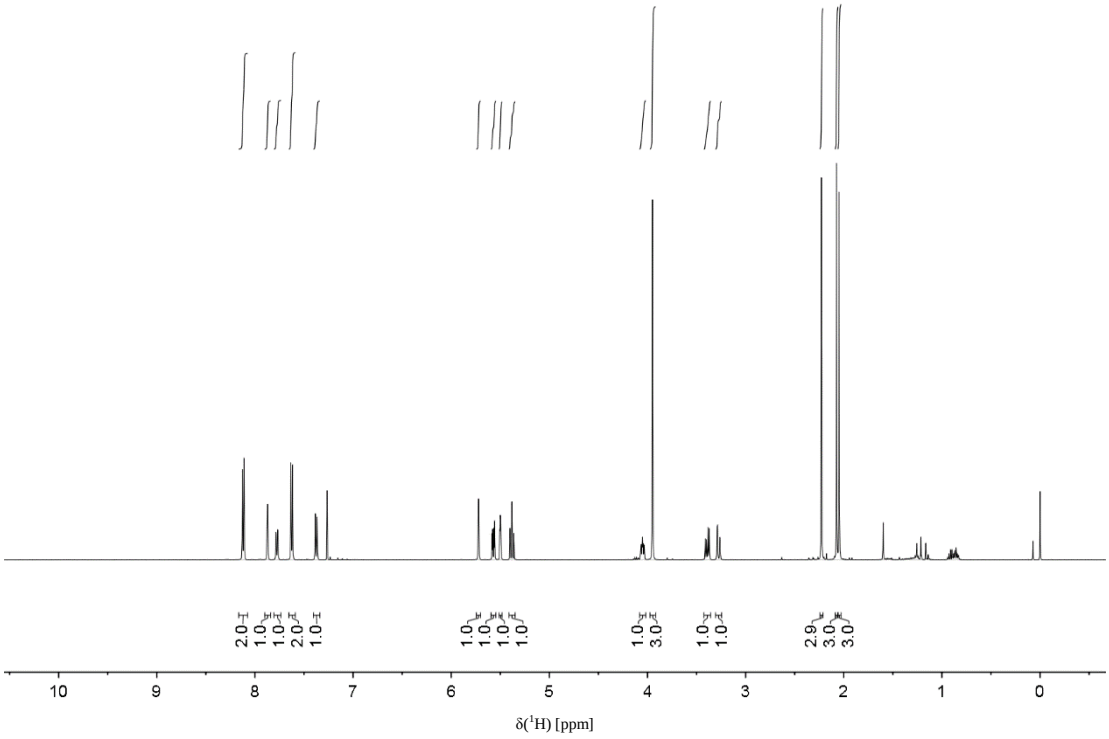


^{13}C NMR (125.8 MHz, $CDCl_3$)

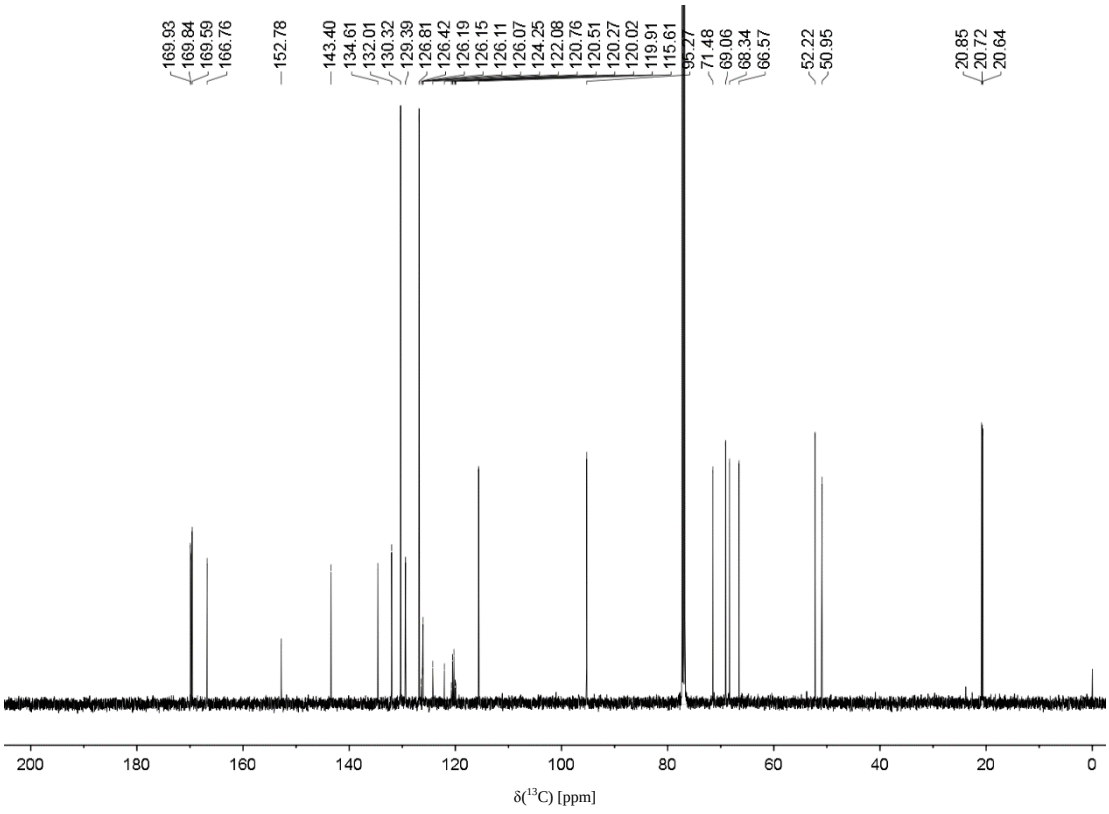


Methyl-4’-(2’’,3’’,4’’-tetra-O-acetyl-6’’-deoxy-6’’-azido-α-D-mannopyranosyloxy)-3’-trifluoromethylbiphenyl-4-carboxylate (23)

^{1}H NMR (500.0 MHz, $CDCl_3$)

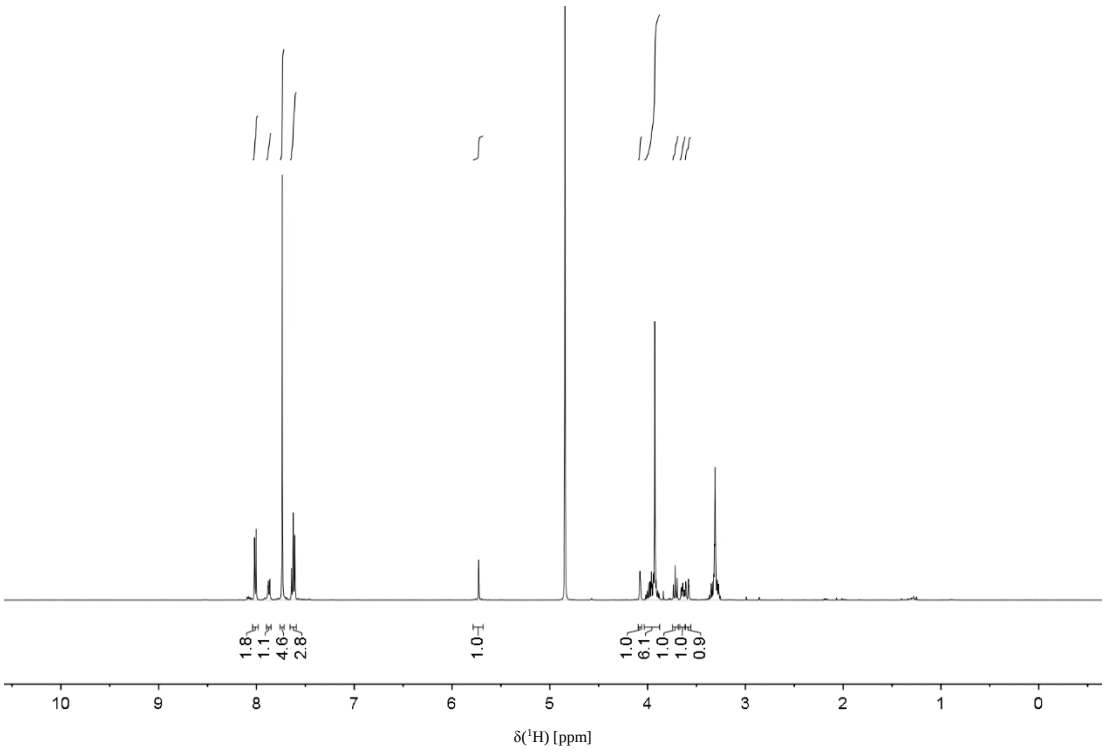


^{13}C NMR (125.8 MHz, $CDCl_3$)

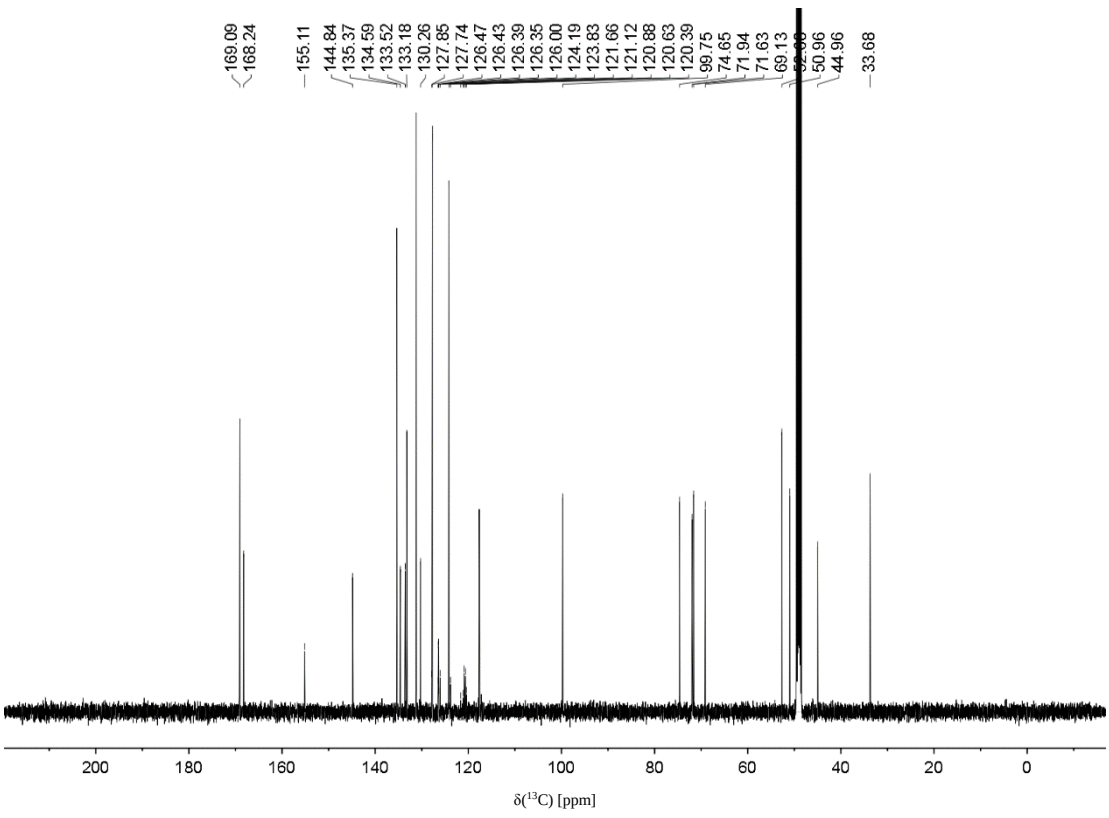


Methyl-4’-(2’’,3’’,4’’-tetra-O-acetyl-6’’-deoxy-6’’-2’’’-(phthalimido)ethylsulfonamido-α-D-mannopyranosyloxy)-3’-trifluoromethylbiphenyl-4-carboxylate (26)

^{1}H NMR (500.0 MHz, MeOD)

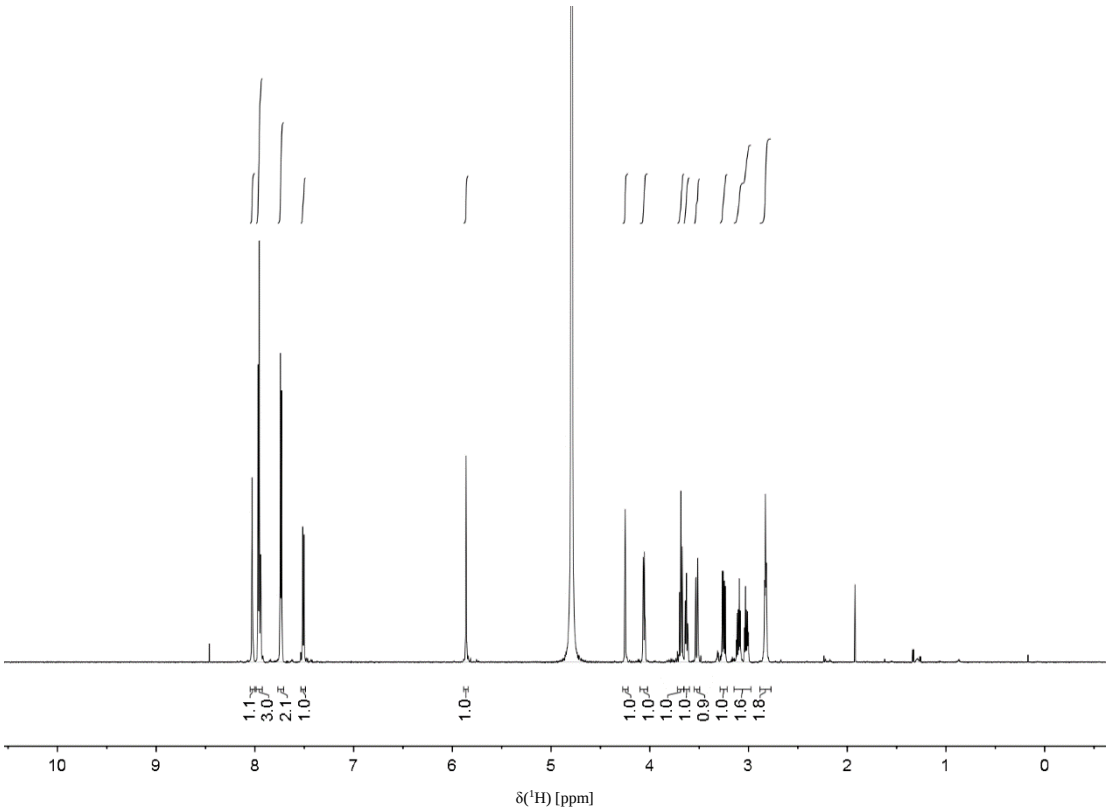


^{13}C NMR (125.8 MHz, MeOD)

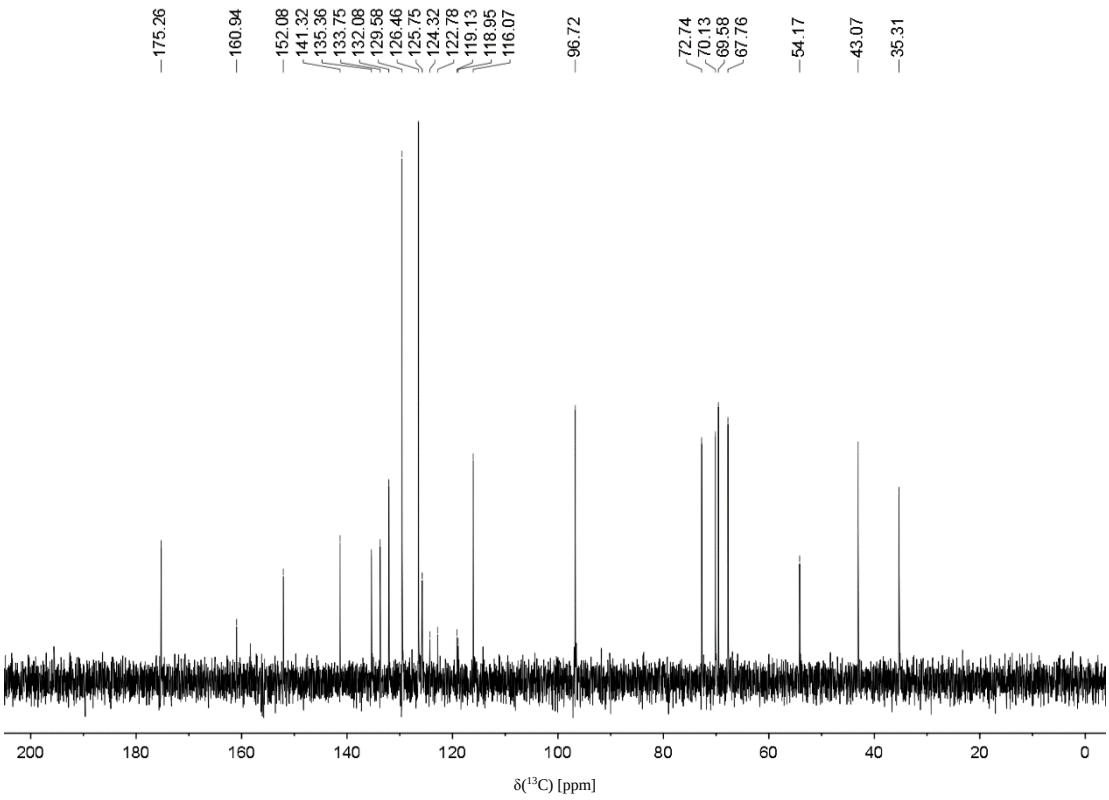


4’-(2’’,3’’,4’’-tetra-O-acetyl-6’’-deoxy-6’’-2’’’-aminoethylsulfonamido-α-D-mannopyranosyloxy)-3’-trifluoromethylbiphenyl-4-carboxylate (14.4)

^{1}H NMR (700.0 MHz, D_2O)

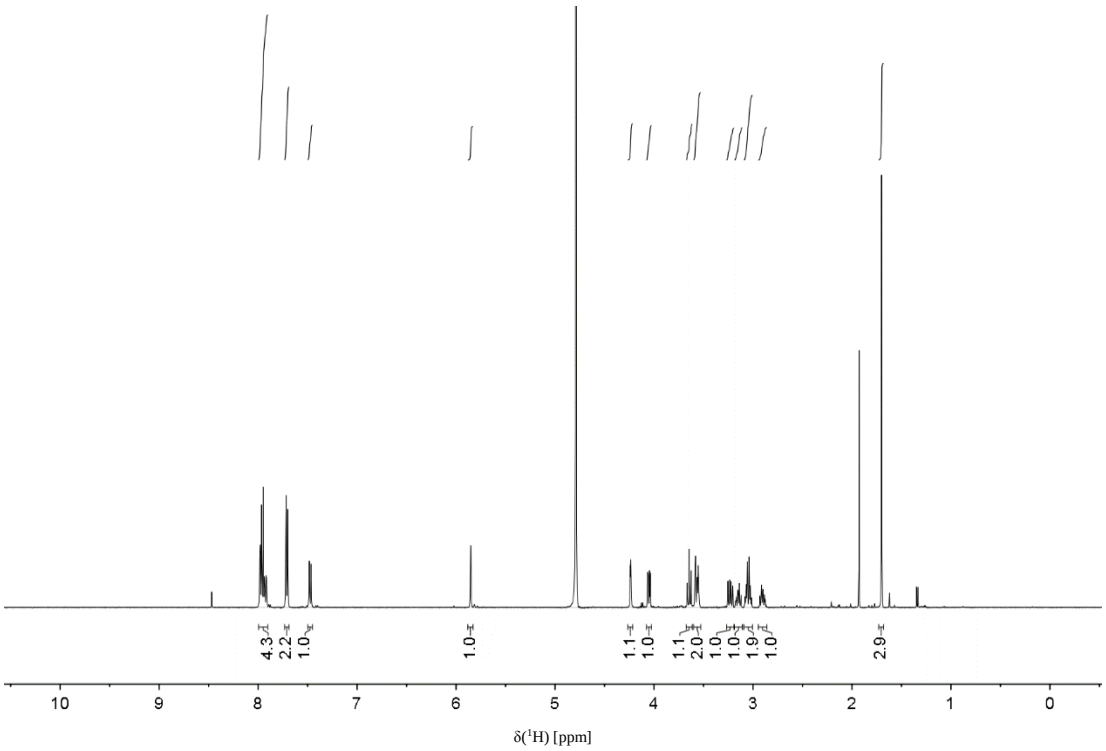


^{13}C NMR (176.0 MHz, D_2O)

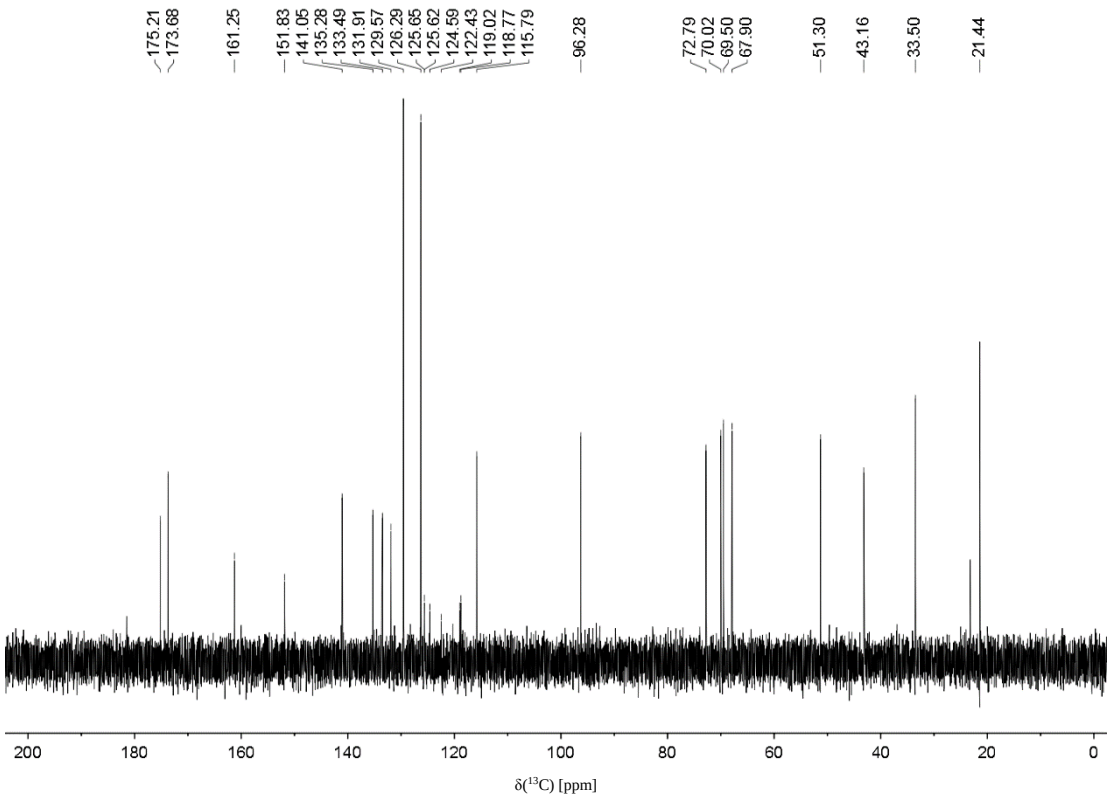


4’-(2’’,3’’,4’’-tetra-O-acetyl-6’’-deoxy-6’’-2’’’-(acetamido)ethylsulfonamido-α-D-mannopyranosyloxy)-3’-trifluoromethylbiphenyl-4-carboxylate (14.5)

^{1}H NMR (500.0 MHz, D_2O)

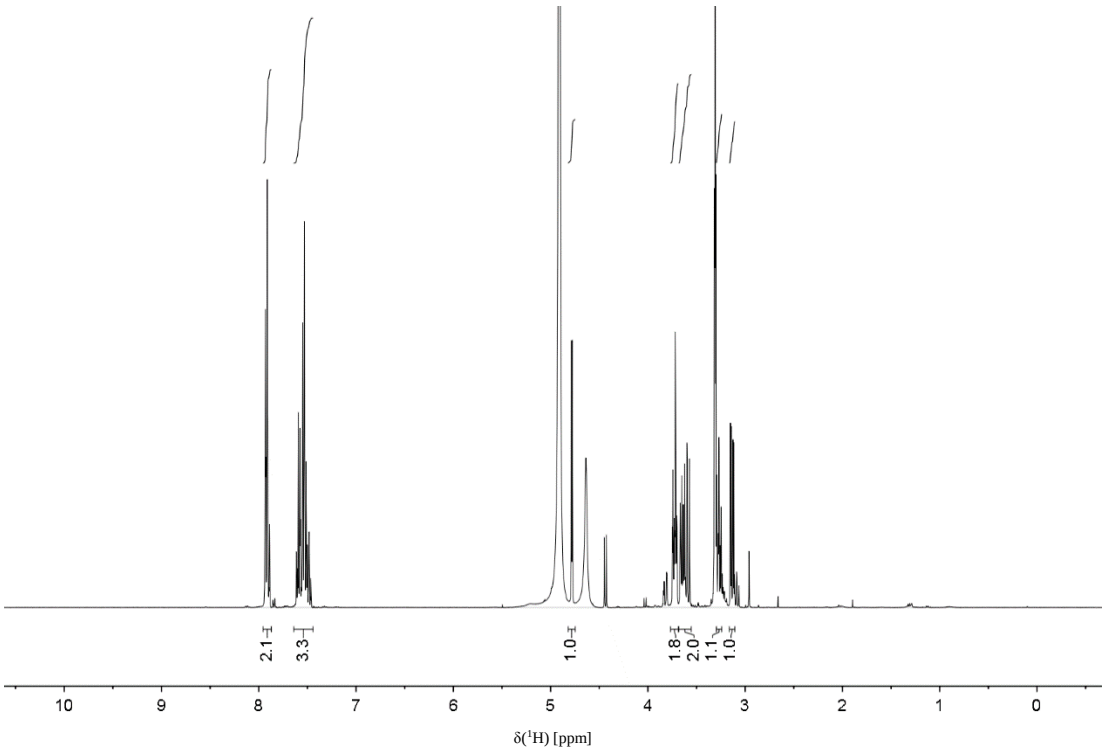


^{13}C NMR (125.8 MHz, D_3O)

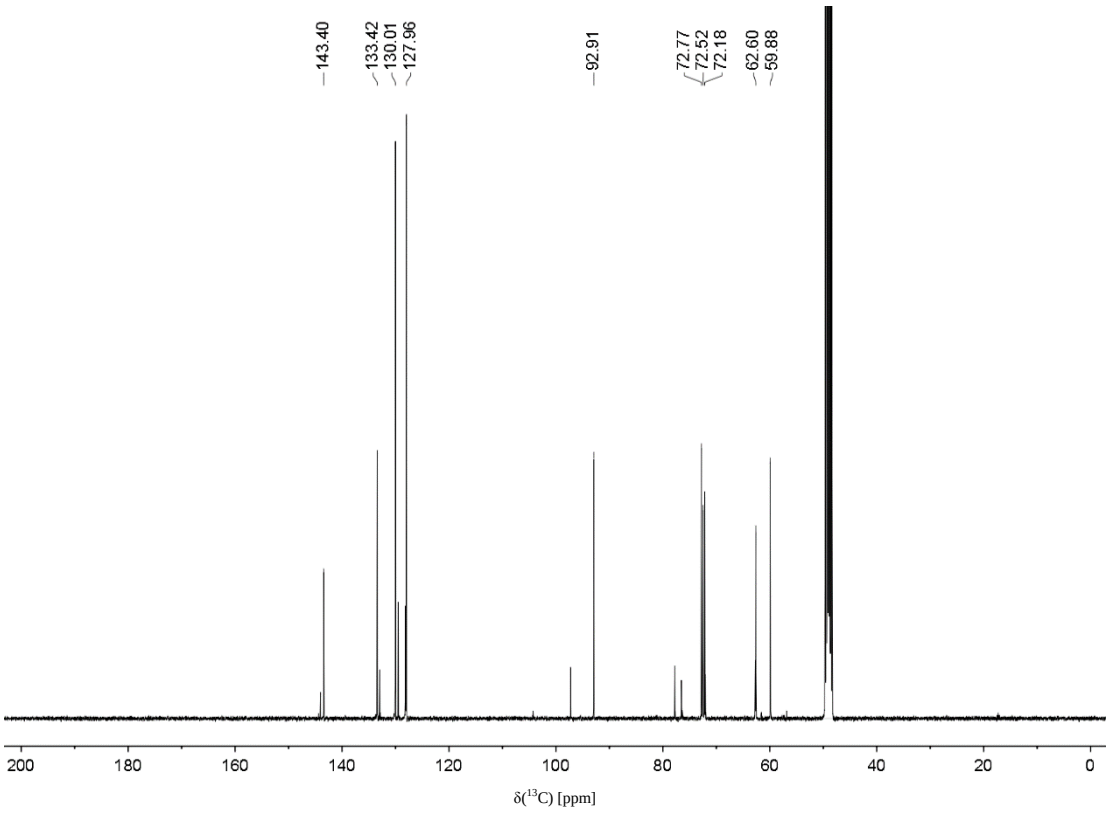


2-deoxy-2-phenylsulfonylamido-D-glucopyranose (31.1)

^{1}H NMR (400.0 MHz, MeOD)

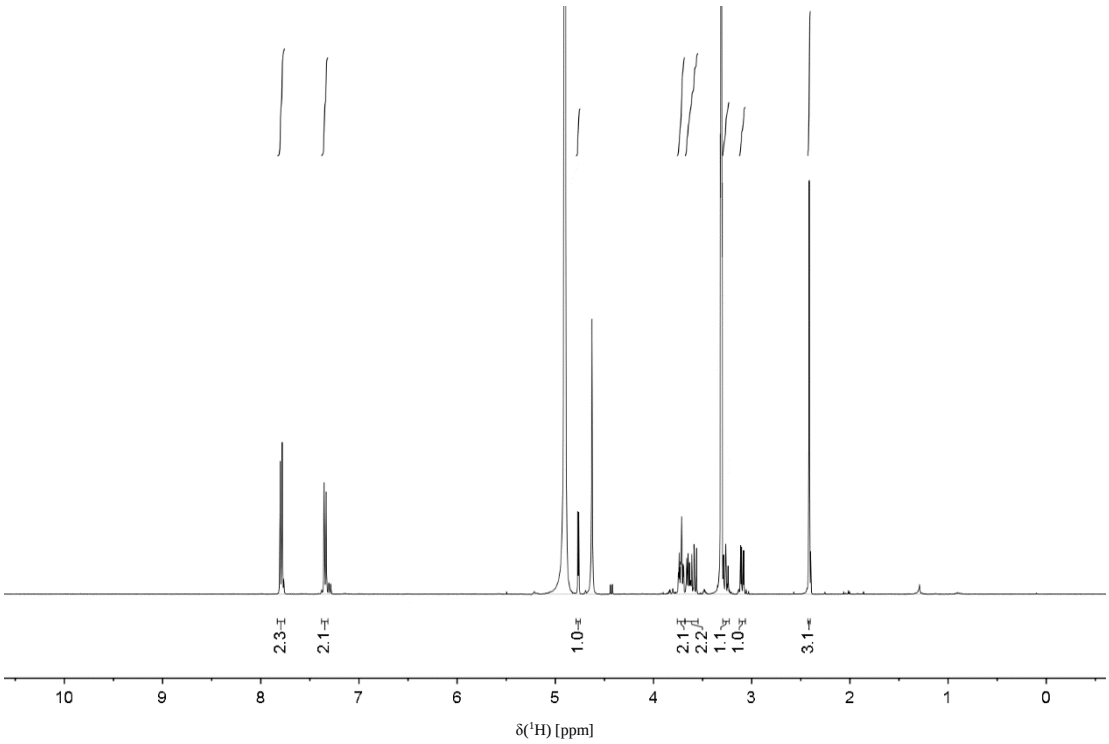


^{13}C NMR (100.6 MHz, MeOD)

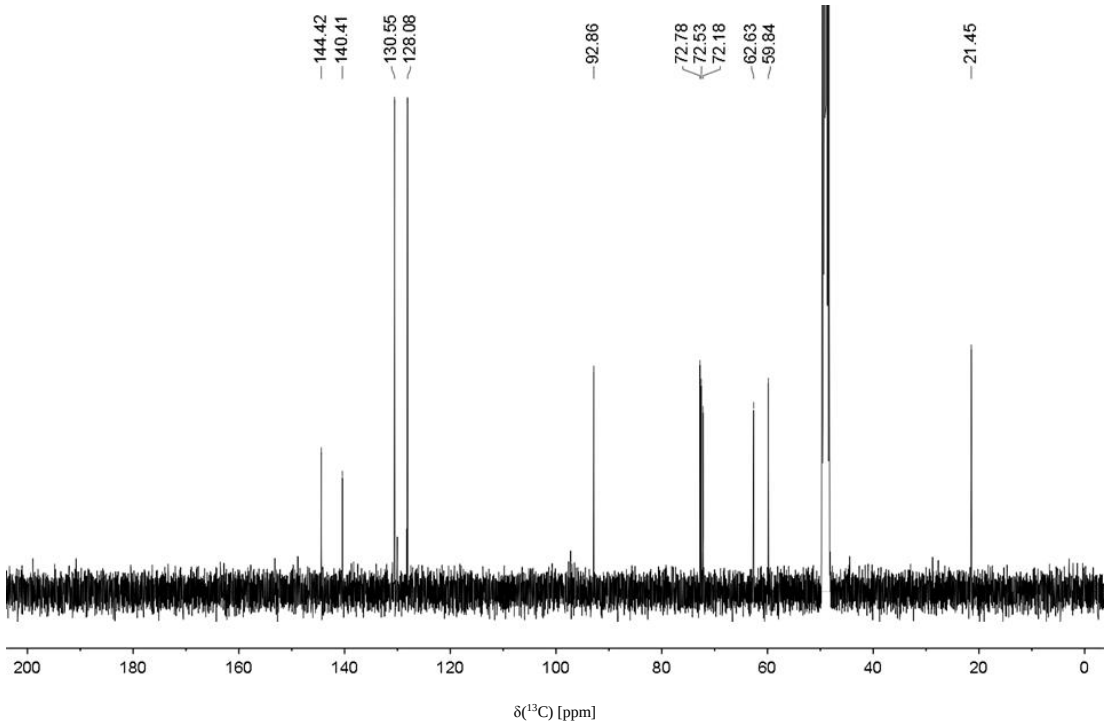


2-deoxy-2-O-tosyl-D-glucopyranose (31.2)

^{1}H NMR (400.0 MHz, MeOD)

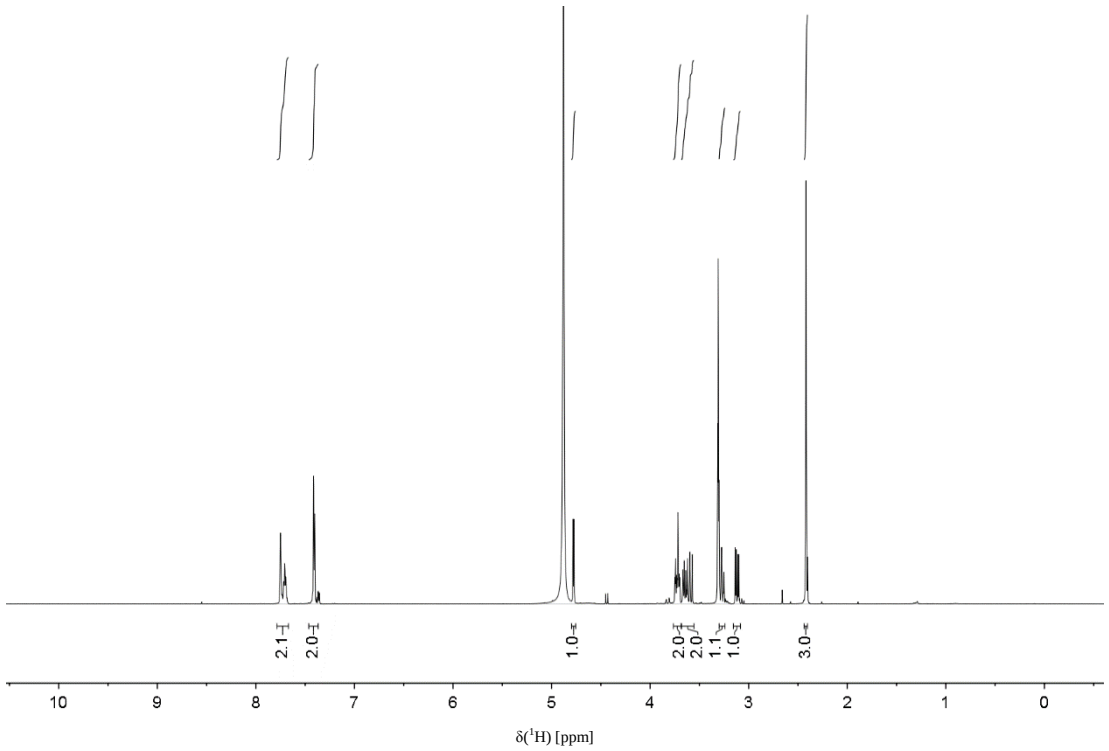


^{13}C NMR (100.6 MHz, MeOD)

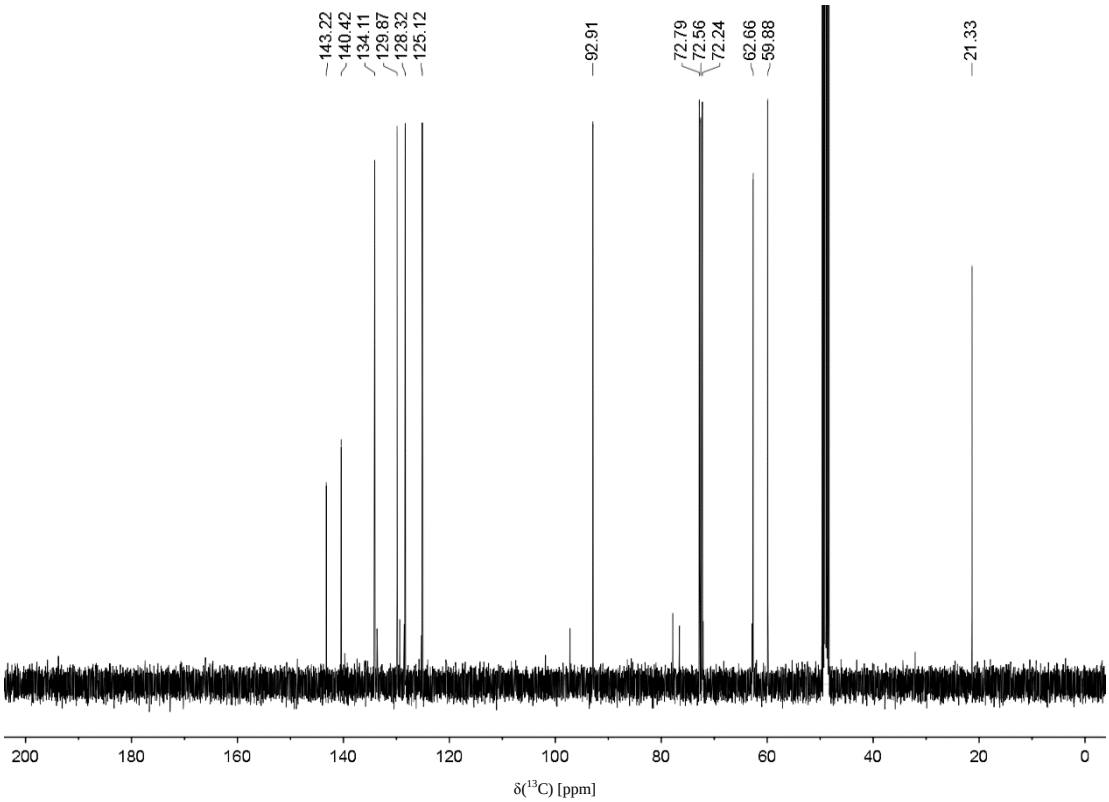


2-deoxy-2-3'-(methyl)phenylsulfonylamido-D-glucopyranose (31.3)

^{1}H NMR (400.0 MHz, MeOD)

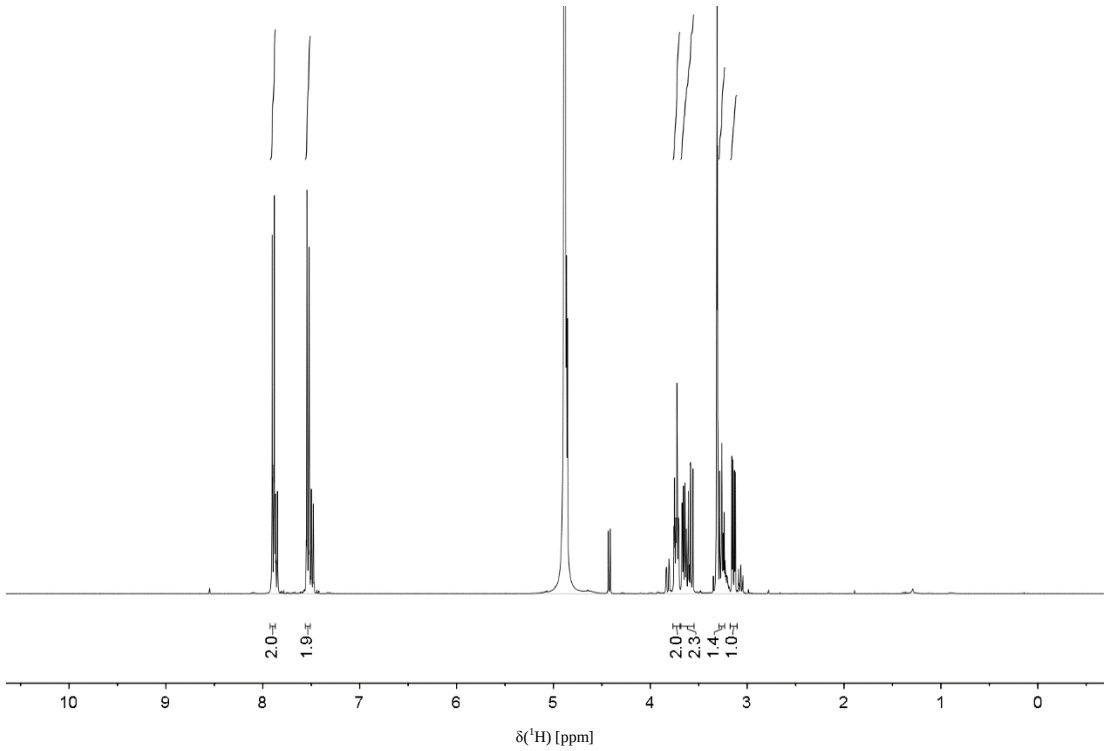


^{13}C NMR (100.6 MHz, MeOD)

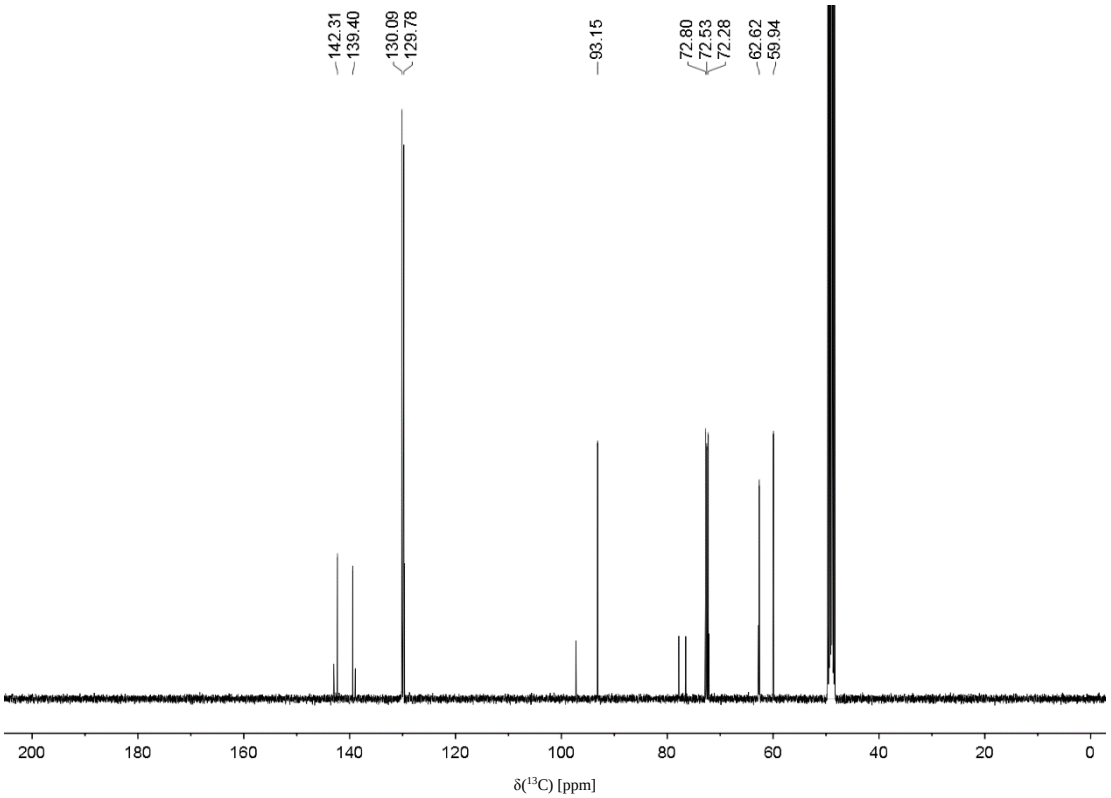


2-deoxy-2-4’-(chloro)phenylsulfonylamido-D-glucopyranose (31.4)

^{1}H NMR (400.0 MHz, MeOD)

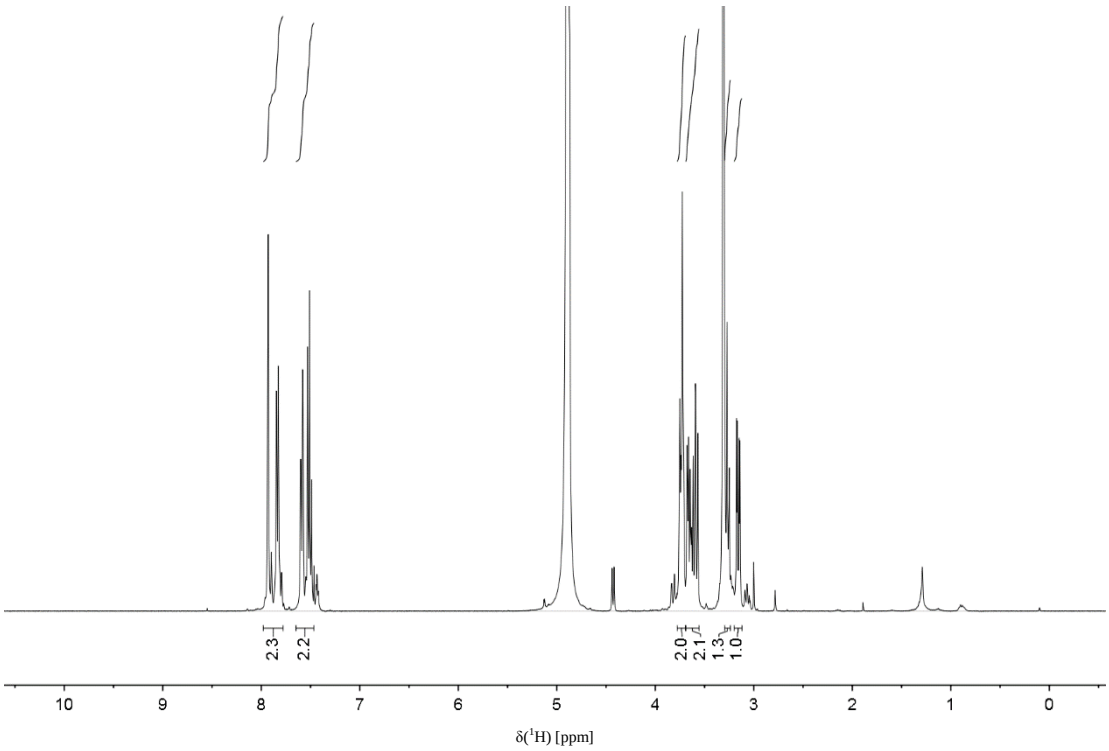


^{13}C NMR (100.6 MHz, MeOD)

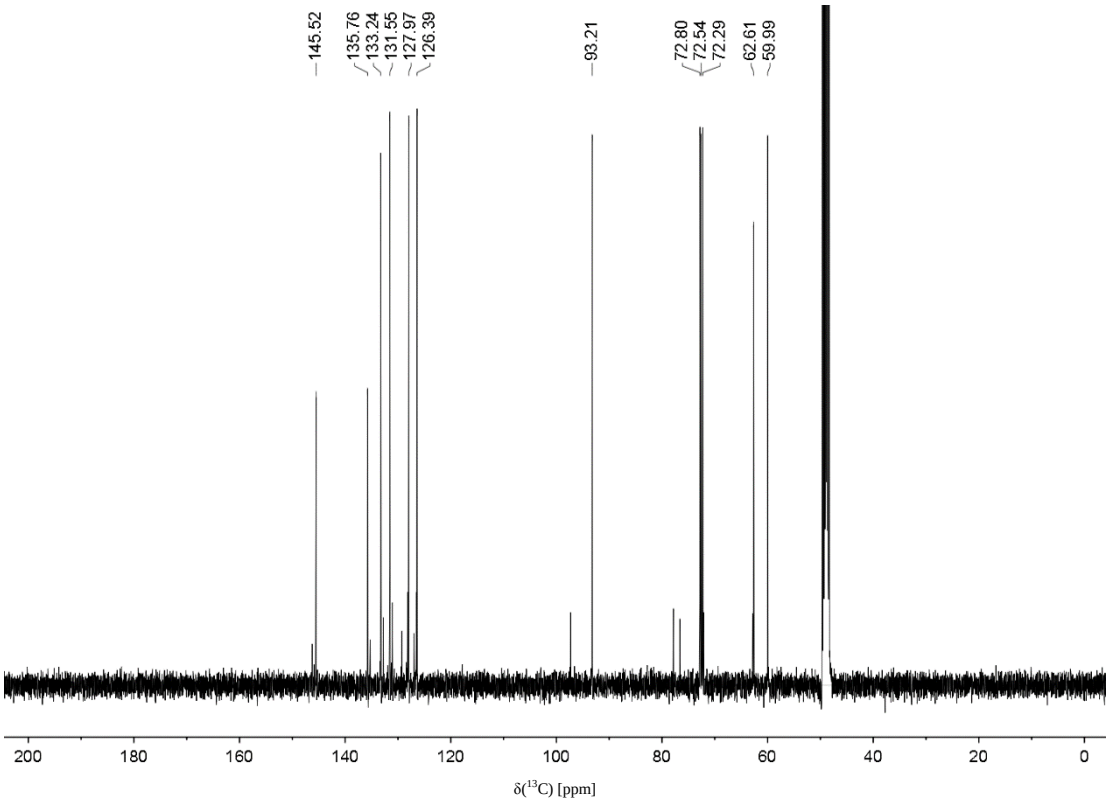


2-deoxy-2-3’-(chloro)phenylsulfonylamido-D-glucopyranose (31.5)

^{1}H NMR (400.0 MHz, MeOD)



^{13}C NMR (100.6 MHz, MeOD)



A.3. Analytical HPLC Traces for Monosaccharide Analogs

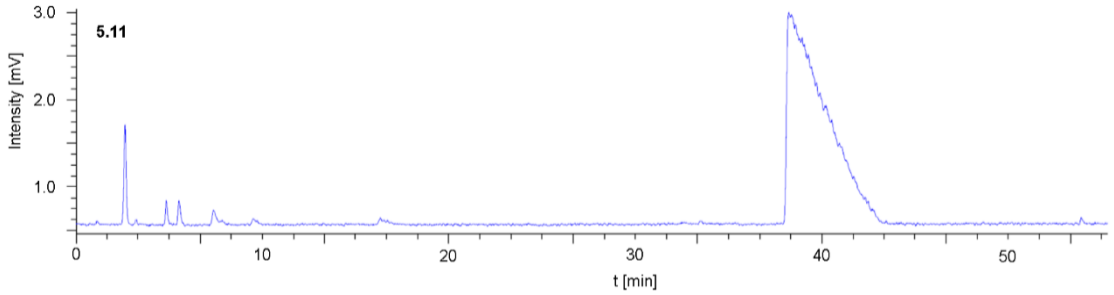


Figure A.2. Analytical HPLC trace for ManNAc analog 5.11.

The purity of ManNAc analog **5.11** was determined to be >95% *via* analytical reversed-phase HPLC utilizing a 0.1% FA in H_2O:acetonitrile gradient (0% acetonitrile for 10 min, 0% to 20% acetonitrile in 25 min, 20% to 50% acetonitrile in 7.5 min, 50% to 100% acetonitrile in 5 min and 100% acetonitrile for 5 min). The analysis was conducted on a HyperCarb column (Thermo Scientific) at a flow rate of 1.0 ml·min^{-1} and the elution of the analog was detected *via* ELSD.

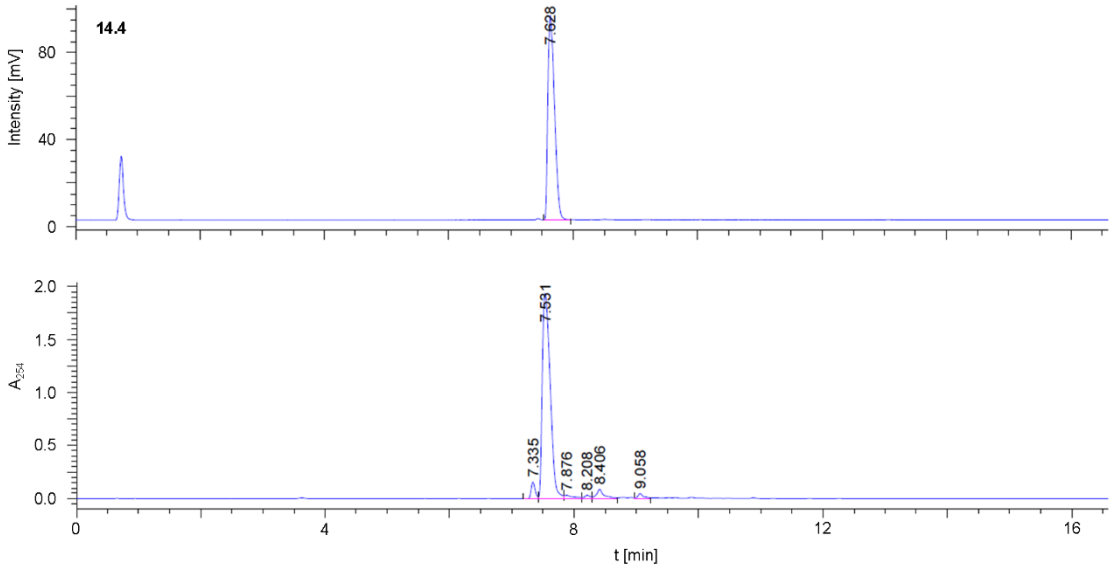


Figure A.3. Analytical HPLC trace for Man analog 14.4

The purity of ManNAc analog **14.4** was determined to be >95% *via* analytical reversed-phase HPLC utilizing an H_2O: acetonitrile gradient (5% in acetonitrile for 10 min and 5% to 95% in acetonitrile in 15 min). Both solvents contained 0.01% TFA. The analysis was conducted on an Atlantis T3 column (Waters) at a flow rate of 0.5 ml·min^{-1} and the elution of the analog was detected *via* ELSD or absorbance A_{254} measurements at 254 nm.

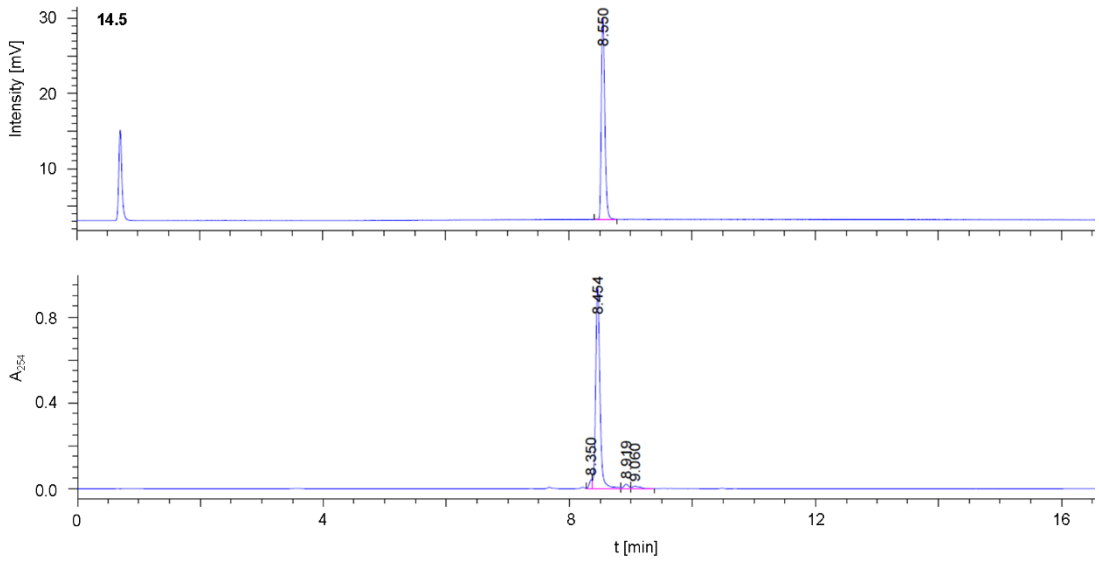
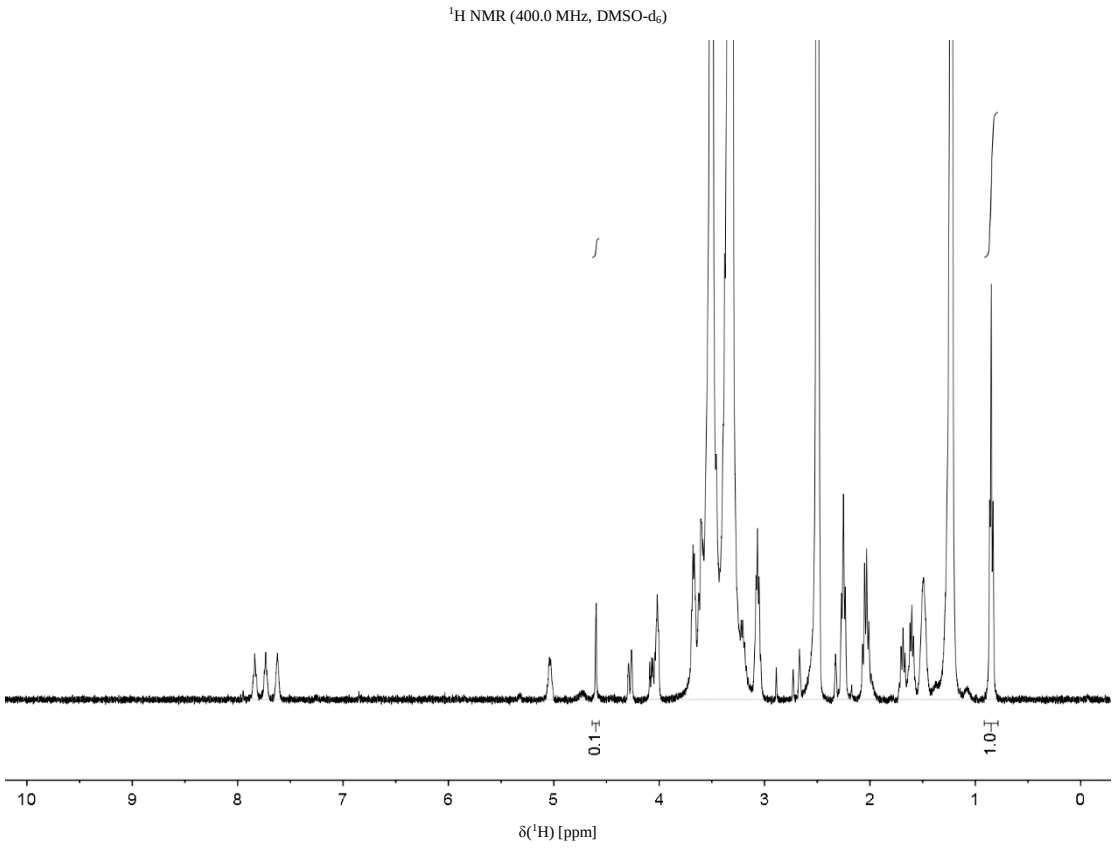


Figure A.4. Analytical HPLC trace for Man analog 14.5

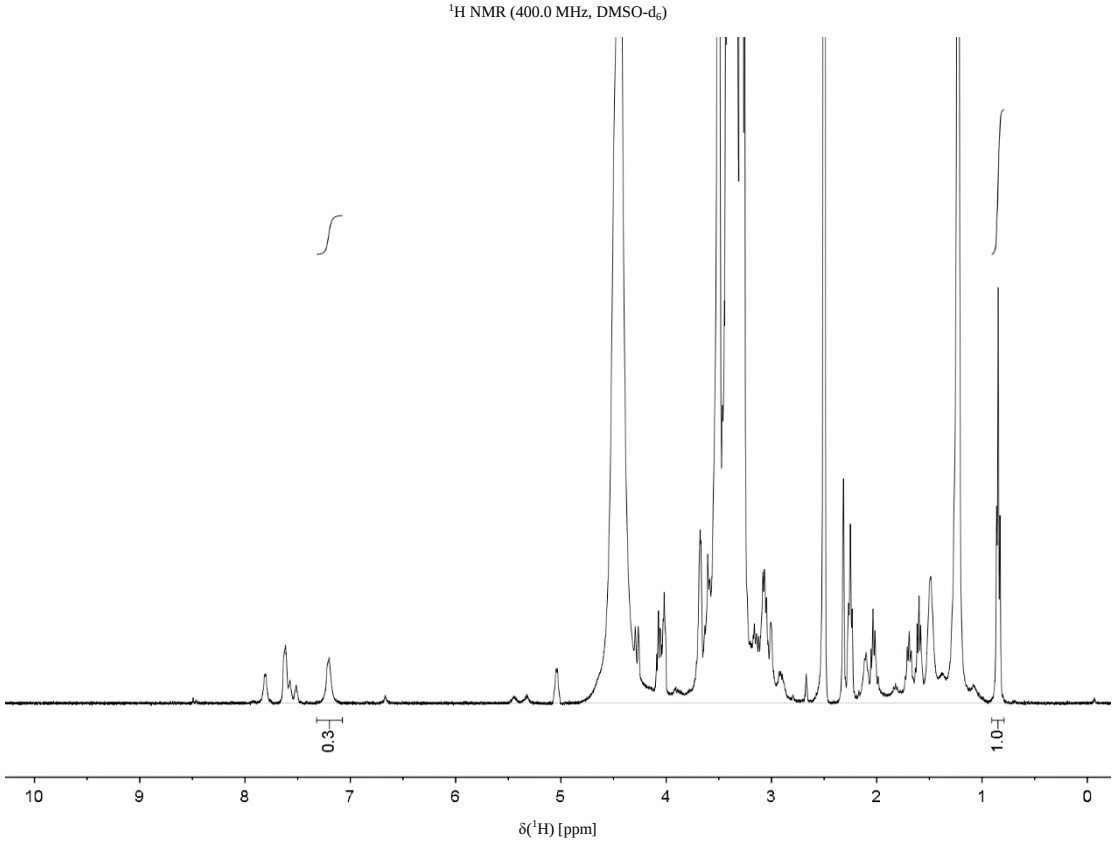
The purity of ManNAc analog **14.5** was determined to be >95% *via* analytical reversed-phase HPLC utilizing an H_2O: acetonitrile gradient (5% in acetonitrile for 10 min and 5% to 95% in acetonitrile in 15 min). Both solvents contained 0.01% TFA. The analysis was conducted on an Atlantis T3 column (Waters) at a flow rate of 0.5 ml·min^{-1} and the elution of the analog was detected *via* ELSD or absorbance A_{254} measurements at 254 nm.

A.4. ^{1}H NMR Spectra for Glycolipids

12.29-PEG-DSPE (42)



41.1-PEG-DSPE (43)



A.5. Composition of the Fragment Library

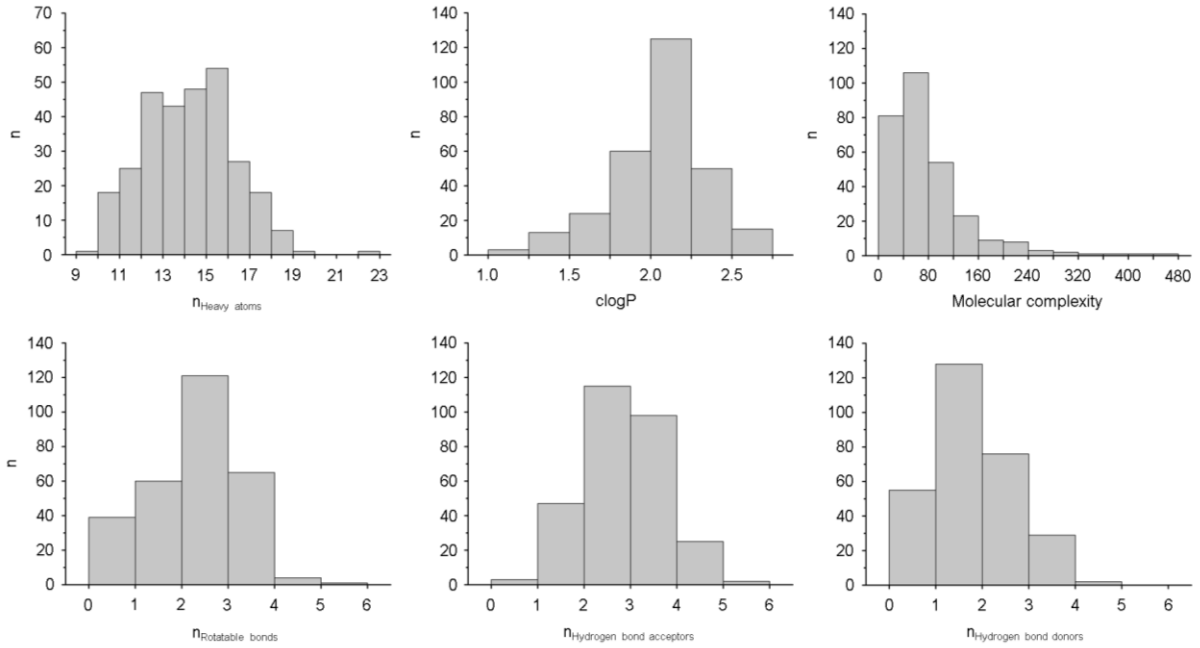


Figure A.5. Composition of the fragment library.

290 fragments were randomly selected from our in-house fragment library and screened against Langerin with the ^{19}F R_2-filtered NMR assay. Relevant descriptors such as clogP, molecular complexity as well as the number of heavy atoms, non-terminal rotatable bonds, hydrogen bond acceptors and hydrogen bond donors are compliant with published guidelines for fragment library design (Chen and Hubbard, 2009; Keseru *et al.*, 2016; Mannhold *et al.*, 2009). Adapted from (Wamhoff *et al.*, 2016).

List of Identifiers for Fragment Library

01A05, 01A08, 01A09, 01B03, 01B04, 06C08, 06C10, 06C12, 06D01, 06D02, 06D04, 06D05, 06D06, 06D07, 06D08, 06D10, 08B01, 08B02, 08B03, 08B04, 08B05, 08B07, 08B08, 08B09, 08B10, 08B11, 08C01, 08C03, 08C04, 08C05, 08C06, 08C11, 08C12, 08D01, 08D03, 08D05, 08D06, 08D08, 08D09, 08D10, 08D11, 08E03, 08E04, 08E05, 08E07, 08E08, 08E10, 08E12, 08F01, 08F02, 08F03, 08F04, 08F05, 08F06, 08G09, 08G10, 08G11, 08G12, 08H01, 08H07, 08H08, 08H09, 08H10, 08H12, 09A01, 09A02, 09A03, 09A04, 09A06, 09A07, 09A09, 09A10, 09A11, 09B01, 09B02, 09B03, 09B04, 09B05, 09B08, 09B11, 09B12, 09C01, 09C03, 09C04, 09C05, 09C06, 09C07, 09C08, 09C09, 09C10, 09C12, 09D01, 09E01, 09E03, 09E05, 09E08, 09E09, 09E11, 09E12, 09F01, 09F02, 09F03, 09F05, 09F08, 09F10, 09F11, 09G01, 09G02, 09H01, 09H04, 09H05, 09H06, 09H07, 09H08, 09H09, 09H11, 09H12, 10A01, 10A03, 10B07, 10B08, 10B09, 10B11, 10B12, 10C01, 10C02, 10C04, 10C05, 10C07, 10C08, 10C09, 10C10, 10C11, 10D01, 10D03, 10D04, 10D07, 10D09, 10D10, 10D12, 10E01, 10E02, 10E04, 10E05, 10E06, 10E07, 10E08, 10E09, 10E10, 10E11, 10F02, 10G06, 10G07, 10G10, 10G12, 10H01, 10H02, 10H04, 10H06, 10H07, 10H08, 10H09, 10H10, 10H11, 10H12, 11A02, 11A03, 11A04, 11A06, 11A07, 11A08, 11A10, 11A11, 11A12, 11B01, 11B02, 11B03, 11B04, 11B05, 11B06, 11B07, 11B09, 11B10, 11B12, 11D09, 11D10, 11D11, 11E01, 11E02, 11E03, 11E04, 11E05, 11E06, 11E07, 11E08, 11E09, 11E10, 11E11, 11E12, 11F01, 11F02, 11F03, 11F04,

11F05, 11F06, 11F07, 11F08, 11F09, 11F10, 11F11, 11F12, 11G01, 11G02, 11G03, 11G05, 11G06, 11G07, 11H08, 11H09, 11H10, 11H11, 12A01, 12A02, 12A03, 12A04, 12A05, 12A06, 12A07, 12A10, 12A11, 12A12, 12B01, 12B02, 12B03, 12B04, 12B05, 12B06, 12B07, 12B08, 12B09, 12B10, 12B12, 12C01, 12C02, 12C03, 12C04, 12C05, 12C06, 12C09, 12C10, 12E01, 12E03, 12E04, 12E07, 12E08, 12E09, 12E10, 12E11, 12E12, 12F01, 12F02, 12F03, 12F04, 12F05, 12F08, 12F09, 12G05, 12G06, 12G08, 12G09, 12G10, 12G11, 12H01, 12H02, 12H03, 12H04, 12H05, 12H06, 12H07, 12H08, 12H09, 12H10, 13D05, 13D07, 13D08, 13D11, 13E01, 13E05, 13E06, 13E07.

Fragment **6** corresponds to 08D08, fragment **7** corresponds to 08D06, fragment **8** corresponds to 08D05 and fragment **9** corresponds to 08D01.

B. Supplementary Data – Results and Discussion

B.1. Structure-Based Multiple Sequence Alignment of C-Type Lectin Receptors

The identities of the CLRs corresponding to the PDB codes are given in Table 1 (Chapter 2.1.2)

```
1DV8        -GKAWADADN YCRL-E-DAH LVVVTSWEEQ KFVQHHI--- GPV--NTWMG
1ESL        EAMTYDEASA YCQQ-R-YTH LVAIQNKEEI EYLNSIL--- SYSPSYYWIG
1G1S        -AYSWNISRK YCQN-R-YTD LVAIQNKNEI DYLNKVL--- PYYSSYYWIG
1H8U        --QTFSQAWF TCRRCY-RGN LVSIHNFNIN YRIQCSVS-- ALNQGQVWIG
1HUP        --MTFEKVKA LCVK-F-QAS VATPRNAAEN GAIQNLIK-- ----EEAFLG
1K9J        -QRNWHDSVT ACQE-V-RAQ LVVIKTAEEQ NFLQLQTS-- RSN-RFSWMG
1QDD        -RETWVDADL YCQN-MNSGN LVSVLTQAEG AFVASLIKES GTDDFNVWIG
1TN3        -TKTFHEASE DCIS-R-GGT LSTPQTGSEN DALYEYLRQS VGNEAEIWLG
1UV0        -PKSWTDADL ACQK-RPSGN LVSVLSGAEG SFVSSLVKSI GNSYSYVWIG
2C6U        -NLTWEESKQ YCTD-M-NAT LLKIDNRNIV EYIKARTH-- ----LIRWVG
2CL8        --NSWYGSKR HCSQ-L-GAH LLKIDNSKEF EFIESQTS-- SHRINAFWIG
2H2T        --KQWVHARY ACDD-M-EGQ LVSIHSPEEQ DFLTKRAS-- ---HTGSWIG
2OX8        -KEIFEDAKL FCED-K-SSH LVFINTREEQ QWIKKQM--- VGR-ESHWIG
2XR6        -QRNWHDSIT ACKE-V-GAQ LVVIKSAEEQ NFLQLQSSRS ---NRFTWMG
2YHF        -ESSWNESRD FCKG-K-GST LAIVNTPEKL KFLQDITD-- ---AEKYFIG
3CFW        -PMNWQRARR FCRDN--YTD LVAIQNKAEI EYLEKTL--- PFSRSYYWIG
3IKN        --KPFTEAQL LCTQ-A-GGQ LASPRSAAEN AALQQLVVAK ---NEAAFLS
3P5F        -PKTWYSAEQ FCVS-R-NSH LTSVTSESEQ EFLYKTAG-- ---GLIYWIG
3VPP        -WSIWHTSQE NCLK-E-GST LLQIESKEEM DFITGSLRK- IKGSYDYWVG
3WH3        -TKSWALSLK NCSAM--GAH LVVINSQEEQ EFLSYKK--- -PKMREFFIG
3WHD        -NKTWAESER NCSG-M-GAH LMTISTEAEQ NFIIQFL--- -DRRLSYFLG
1YPO        -SFNWEKSQE KCLS-L-DAK LLKINSTADL DFIQQAISY- -S-SFPFWMG

1DV8        LHDQN----- --GPWKWV-D GTDYETGFK- NWRP---EQP DDW-YGHGLG
1ESL        IRKVN----- --NVWVWVGT QKPLTEEAK- NWAP---GEP NNR-Q-----
1G1S        IRKNN----- --KTWTWVGT KKALTNEAE- NWAD---NEP NNKRN-----
1H8U        GRITGSGRC- --RRFQWVDG SRW---NFA- YWAA---HQP WSRG------
1HUP        ITDEKT--E- --GQFVDL-T GNRL--TYT- NWNE---GEP NNAGS-----
1K9J        LSDLNQ--E- --GTWQWV-D GSPLSPSFQR YWNS---GEP NNSG------
1QDD        LHDPKK--N- --RAWHWSSG SLV---SYK- SWGI---GAP SSVNP-----
1TN3        LNDMAA--E- --GTWVDMTG ARI---AYK- NWETEITAQP DGG-------
1UV0        LHDPTQGTEP NGEGWEWSSS DVM---NYF- AWERNPSTIS SP--------
2C6U        LSRQKSN--- --EVWKWE-D GSVISENMFE FLE-----DG -KG-------
2CL8        LSRN-QS--- -EGPWFWE-D GSAFF---PN SFQVRNAVPQ ESL-------
2H2T        LRNLDLK--- --GEFIWVD- GSHV--DYS- NWAPGE--PT SRS-------
2OX8        LTDSERE--- --NEWKWL-D GTSP--DYK- NWKA---GQP DNWGHG--HG
2XR6        LSDLNQE--- --GTWQWV-D GSPLLPSFKQ YWNRGE---P NNVG------
```

```
2YHF        LIYH-RE--- -EKRWRWI-N NSVFN---GN V-------TN QNQ-------
3CFW        IRKIG----- --GIWTWVGT NKSLTEEAE- NWGD---GEP NNKKN-----
3IKN        MTDSKTE--- --GKFTYP-T GESL--VYS- NWAP---GEP NDDGG-----
3P5F        LTKAGM---- -EGDWSWVDD TPFNKVQSAR FWIPGE---P NNAGN-----
3VPP        LSQDGH---- -SGRWLWQ-D GSSPSPGLL- ------PAER SQ-------S
3WH3        LSDQVV---- -EGQWQWV-D GTPLTKSLS- FWDVGEPNNI A---------
3WHD        LRDENA---- -KGQWRWVDQ TPFNPR-RV- FWHKNEP--- DNS-------
1YPO        LSRRNP---- -SYPWLWE-D GSPLMPHLFR ------VRGA VSQT-----Y

1DV8        GGEDCAHFTD -----D-GRW NDDVCQRPY- RWVCET-
1ESL        KDEDCVEIYI KREKDV-GMW NDERCSKKK- LALCY--
1G1S        -NEDCVEIYI KSPSAP-GKW NDEHCLKKK- HALCY--
1H8U        --GHCVALCT R----G-GYW RRAHCLRRL- PFICSY-
1HUP        -DEDCVLLLK -----N-GQW NDVPCSTSH- LAVCEFP
1K9J        -NEDCAEFS- -----G-SGW NDNRCDVDN- YWICKKP
1QDD        --GYCVSLT- -SSTGF-QKW KDVPCEDKF- SFVCKFK
1TN3        KTENCAVLSG ---AAN-GKW FDKRCRDQL- PYICQ--
1UV0        --GHCASLSR ---STAFLRW KDYNCNVRL- PYVCK--
2C6U        -NMNCAYFH- ------NGKM HPTFCENKH- YLMCER-
2CL8        -LHNCVWIHG -------SEV YNQICNTSS- YSICEK-
2H2T        QSEDCVMMRG -----S-GRW NDAFCDRKLG AWVCD--
2OX8        PGEDCAGLIY -----A-GQW NDFQCEDVN- NFICEKD
2XR6        -EEDCAEFSG -------NGW NDDKCNLAK- FWICKK-
2YHF        -NFNCATIGL ------TKTF DAASCDISY- RRICEK-
3CFW        -KEDCVEIYI KRNKDA-GKW NDDACHKLK- AALCY--
3IKN        -SEDCVEIFT -----N-GKW NDRACGEKR- LVVCEF-
3P5F        -NEHCGNIKA P----SLQAW NDAPCDKTF- LFICKRP
3VPP        ANQVCGYVKS -------NSL LSSNCDTWK- YFICEK-
3WH3        TLEDCATMR- -DSSNPRQNW NDVTCFLNY- FRICEM-
3WHD        QGENCVVLVY ---NQDKWAW NDVPCNFEA- SRICK--
1YPO        PSGTCAYIQR -------GAV YAENCILAA- FSICQK-
```

B.2. Structure-Based *in silico* Design of *N*-Acetylmannosamine Analog Library 5

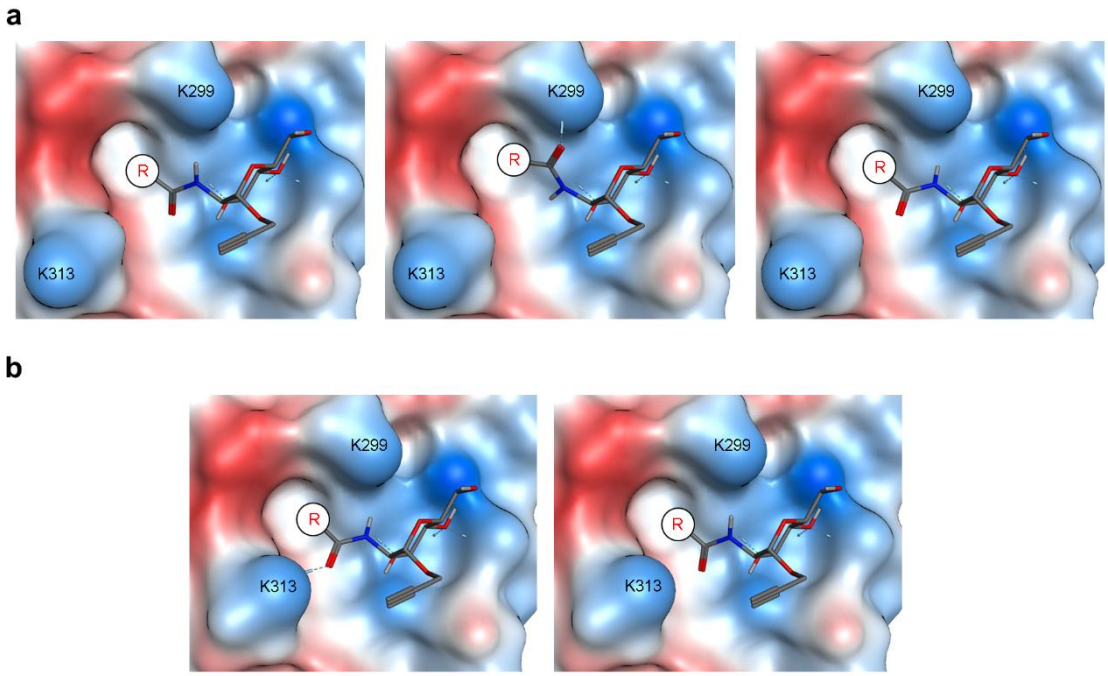


Figure B.1. Generation of carboxamide linker conformations from low mode MD simulations.

a. Different carboxamide linker conformations were selected from low mode MD simulations and utilized for the *in situ* conjugation of commercially available carboxylic acids (PDB code: 3P5F) (Feinberg *et al.*, 2011). **b.** Moreover, the alternative conformation of K313 observed for the Langerin complex with Gal-6-OS was accounted for in additional molecular docking studies (PDB code: 3P5I) (Feinberg *et al.*, 2011). The receptor surface is colored according to its lipophilicity (lipophilic: red, hydrophilic: blue). Adapted from (Wamhoff *et al.*, 2016).

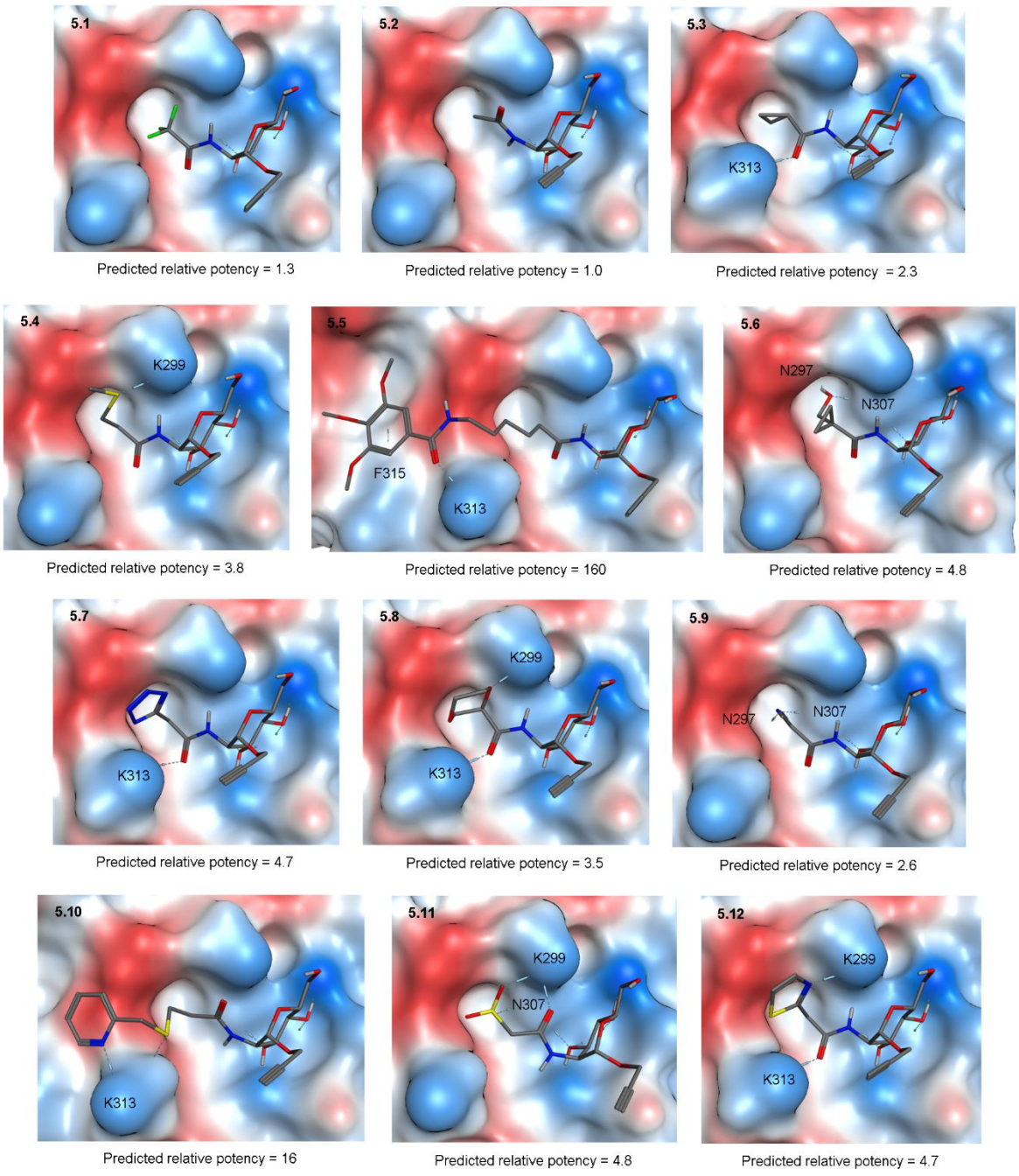


Figure B.2. Docking poses for ManNAc analog library 5.

For each ManNAc analog **5**, the docking pose displaying the highest GBVI/WSA ΔG score is depicted. The corresponding scores served to predict the affinity increase over **5.2**. Residues involved in the formation of hydrogen bonds or π-π interactions with the substituents in C2 of the Man scaffold are indicated. The receptor surface is colored according to its lipophilicity (lipophilic: red, hydrophilic: blue). Adapted from (Wamhoff *et al.*, 2016).

B.3. K_I Determination for *N*-Acetylmannosamine Analog Library 5

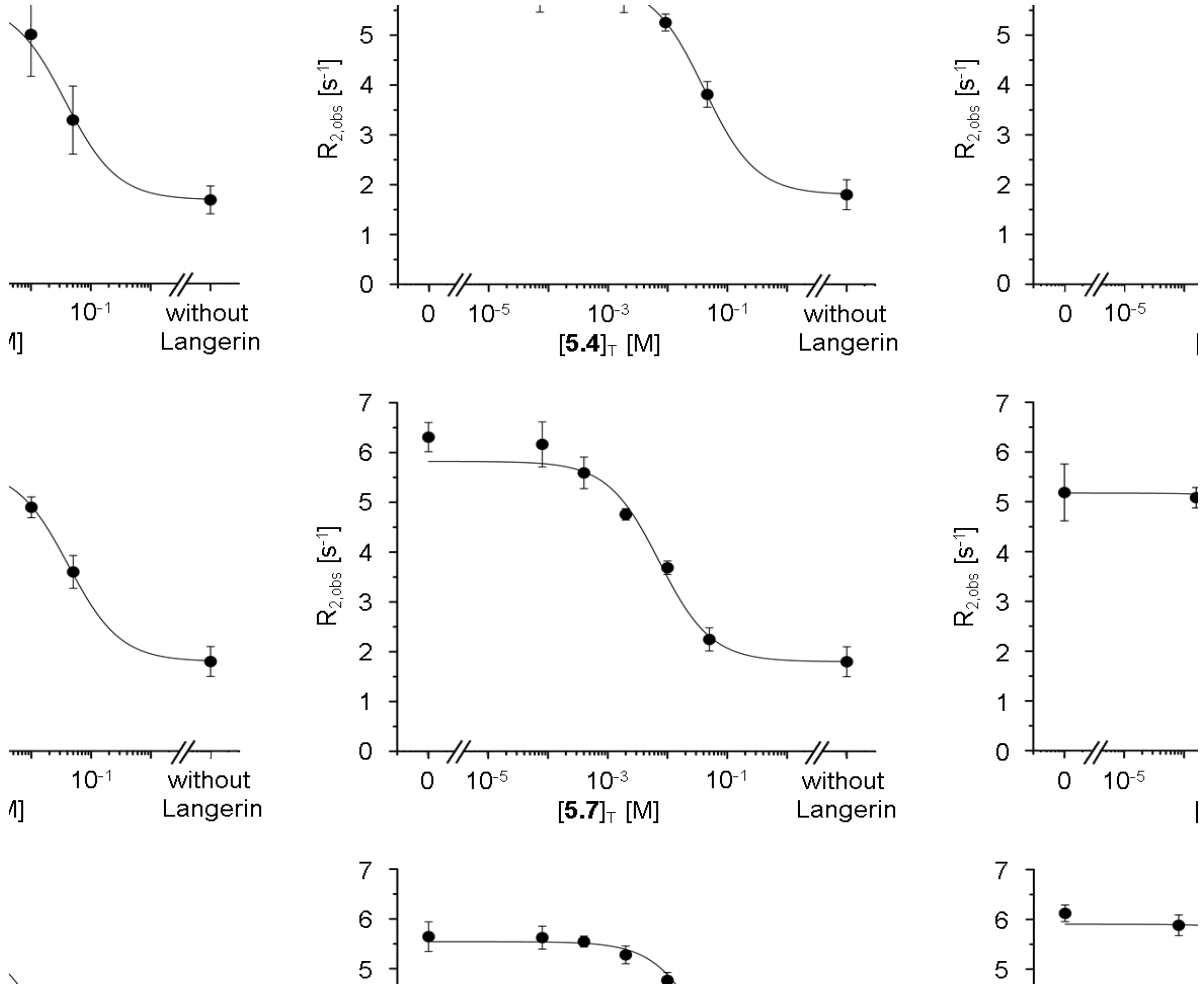
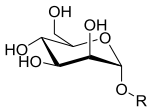
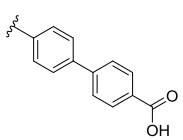


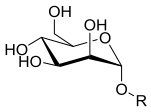
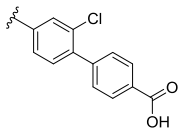
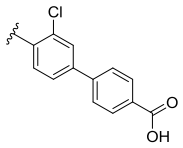
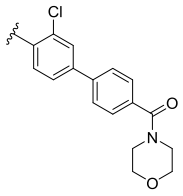
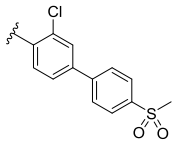
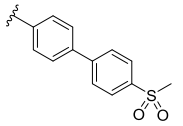
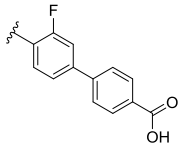
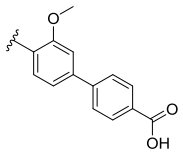
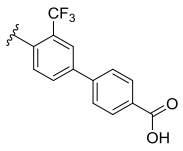
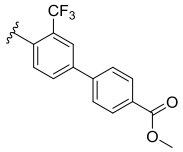
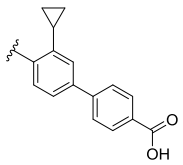
Figure B.3. K_I determination for ManNAc analog library 5.

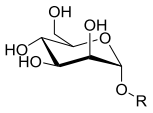
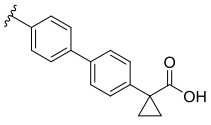
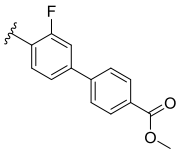
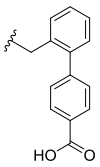
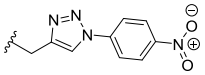
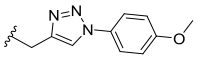
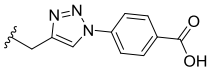
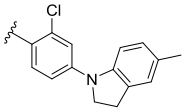
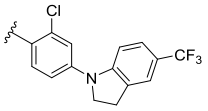
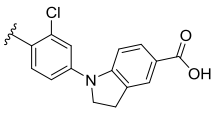
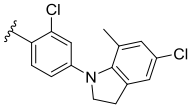
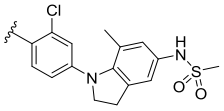
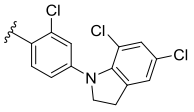
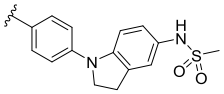
Competitive binding experiments served to determine the affinities for ManNAc analog library **5**. Obtained K_I values are given in Table 6 (Chapter 3.3.1).

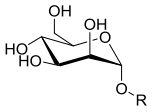
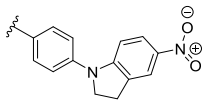
B.4. Screening of Mannose Analog Libraries 12 and 14

Table B.1. Screening of Man analog library 12.

Name	Structure	R	$[I]_T$ [mM]	$\Delta R_{2,obs}$ [Hz]	$K_{I,est}$ [mM][a]
12.2			10	1.9	12
12.3			4	3.0	9

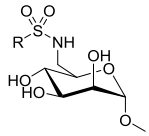
Name	Structure	R	$[I]_T$ [mM]	$\Delta R_{2,obs}$ [Hz]	$K_{I,est}$ [mM][a]
12.4			0.5	1.0	8
12.5			4	3.4	0.3
12.6			0.5	0.3	6
12.7			0.5	0.8	2
12.8			1	0.7	4
12.9			0.5	0.5	3
12.10			1	0.7	7
12.11			1	3.3	0.2
12.12			0.5	0.9	2
12.13			1	1.7	1

Name	Structure	R	$[I]_T$ [mM]	$\Delta R_{2,obs}$ [Hz]	$K_{I,est}$ [mM][a]
12.14			1	1.0	3
12.15			0.1	No competition	
12.16			1	1.9	1
12.17			1	0.7	5
12.18			1	1.2	3
12.19			0.5	1	2
12.20			1	3.2	0.4
12.21			1	2.6	5
12.22			4	3.0	0.7
12.23			1	3.5	0.3
12.24			1	2.0	1
12.25			1	2.9	0.3
12.26			1	1.8	1

Name	Structure	R	$[I]_T$ [mM]	$\Delta R_{2,obs}$ [Hz]	$K_{I,est}$ [mM][a]
12.27			0.5	0.9	1

[a]Estimated K_I values were determined *via* ^{19}F R_2-filtered NMR experiments at a single competitor concentration.

Table B.2. Screening of Man analog library 14.

Name	Structure	R	$[I]_T$ [mM]	$\Delta R_{2,obs}$ [Hz]	$K_{I,est}$ [mM][a]
14.1			1	1.3	3
14.2			1	1.8	2
14.2			1	1.7	2

[a]Estimated K_I values were determined *via* ^{19}F R_2-filtered NMR experiments at a single competitor concentration.

B.5. STD NMR Build-Up Curves for Mannose Analogs 12 and 14

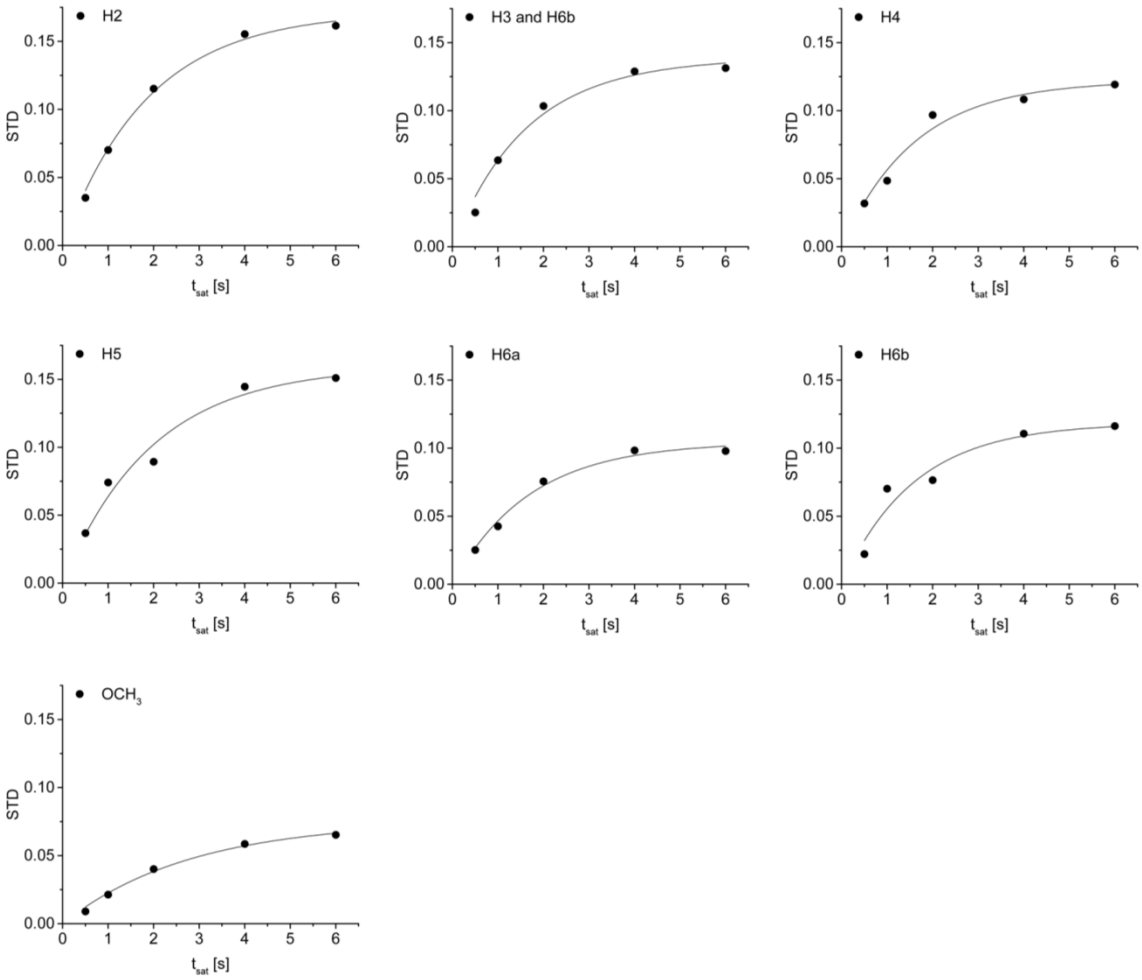


Figure B.4. STD NMR build-up curves for Man analog 12.2.

Equation 8 was fitted to STD values to calculate STD_0' values for the determination of the binding epitope of **12.2**.

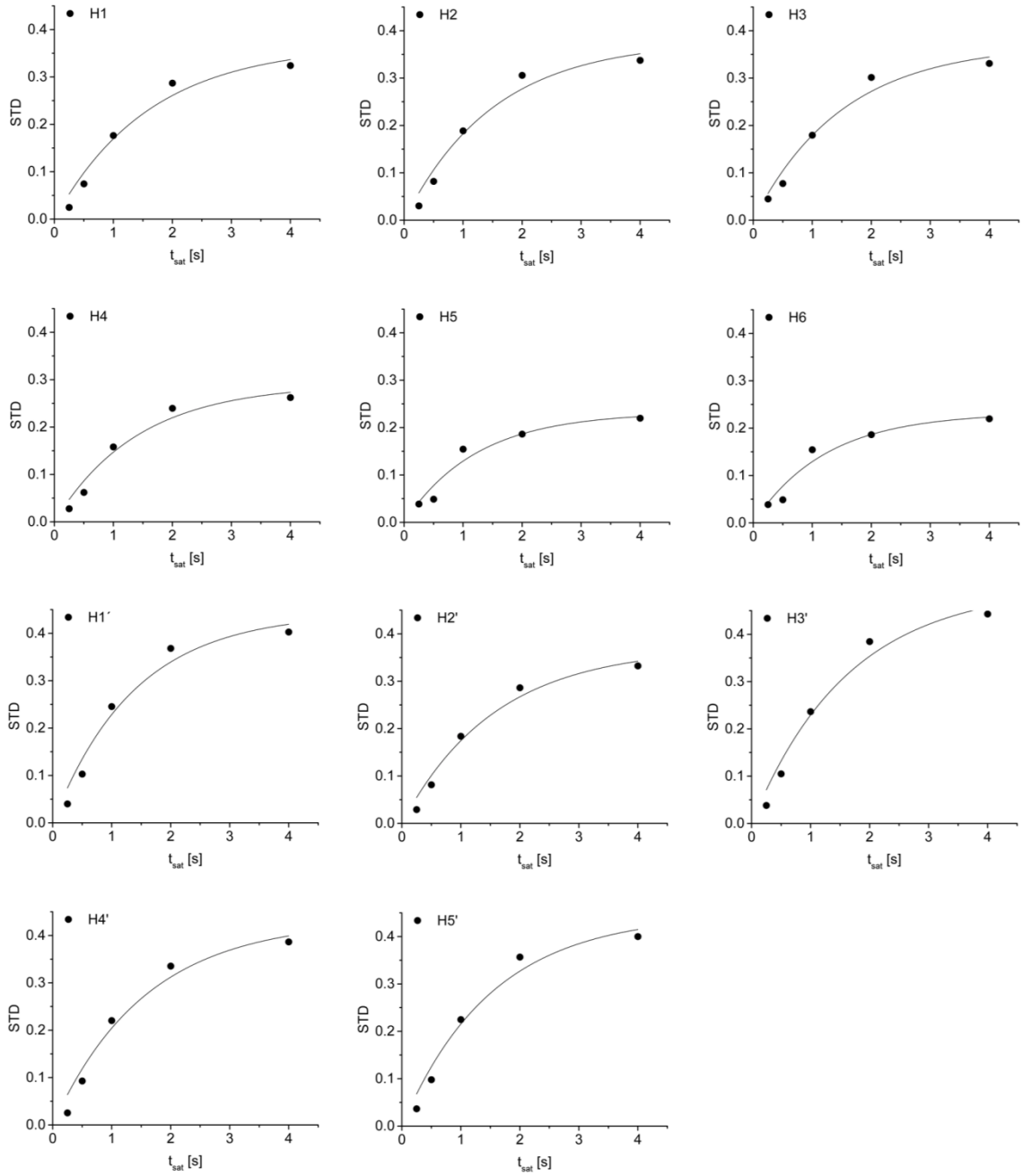


Figure B.5. STD NMR build-up curves for Man analog 12.11.

Equation 8 was fitted to STD values to calculate STD_0' values for the determination of the binding epitope of **12.2**.

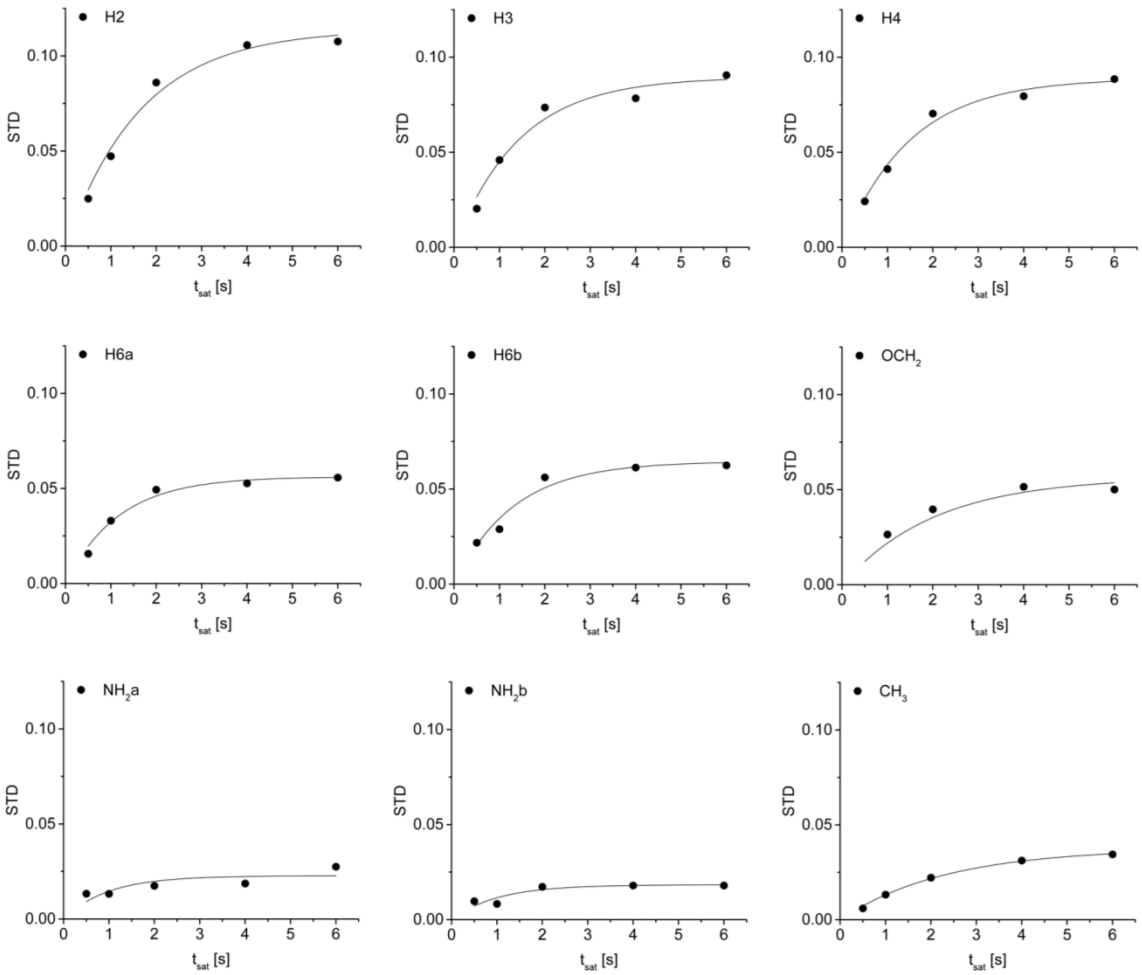


Figure B.6. STD NMR build-up curves for Man analog 12.30.

Equation 8 was fitted to STD values to calculate STD_0' values for the determination of the binding epitope of **12.30**.

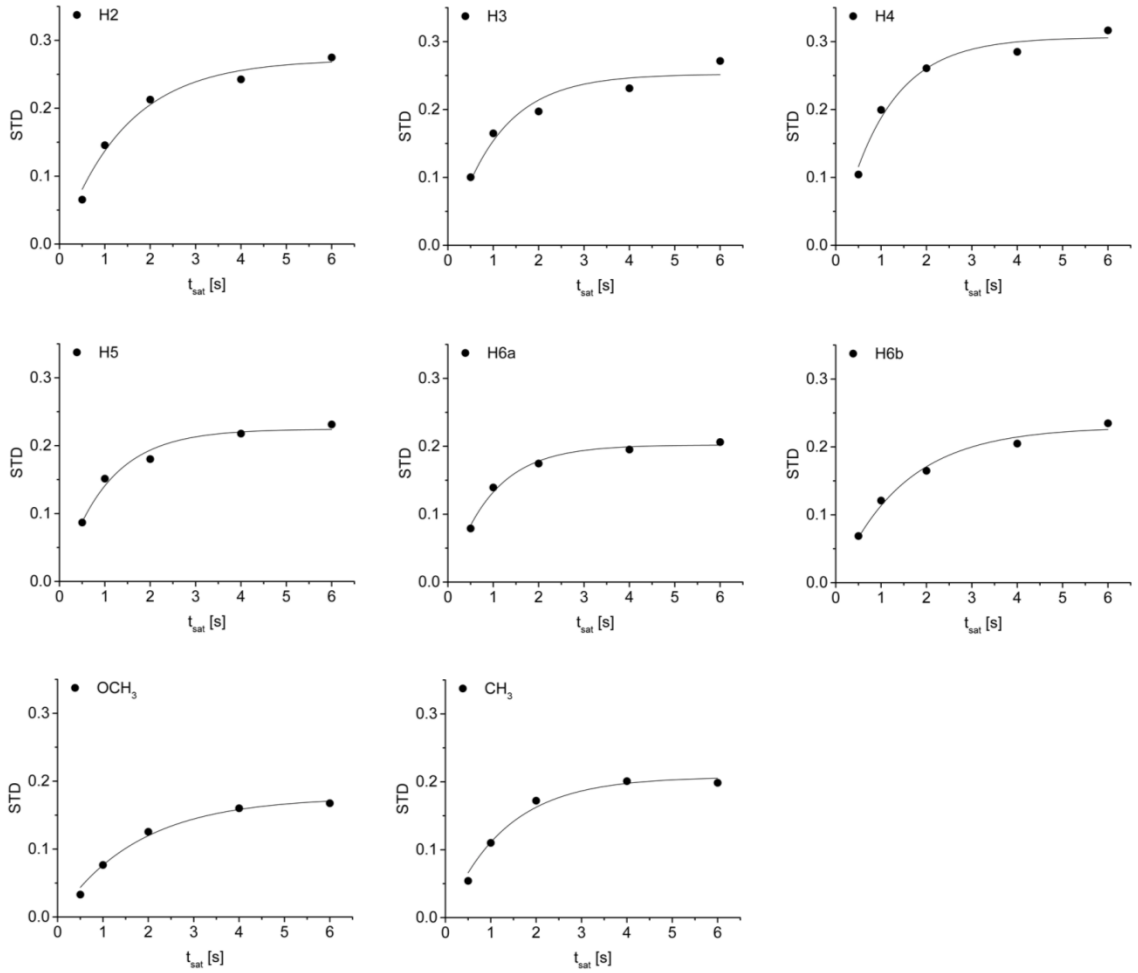


Figure B.7. STD NMR build-up curves for Man analog 14.1.

Equation 8 was fitted to STD values to calculate STD_0' values for the determination of the binding epitope of **14.1**.

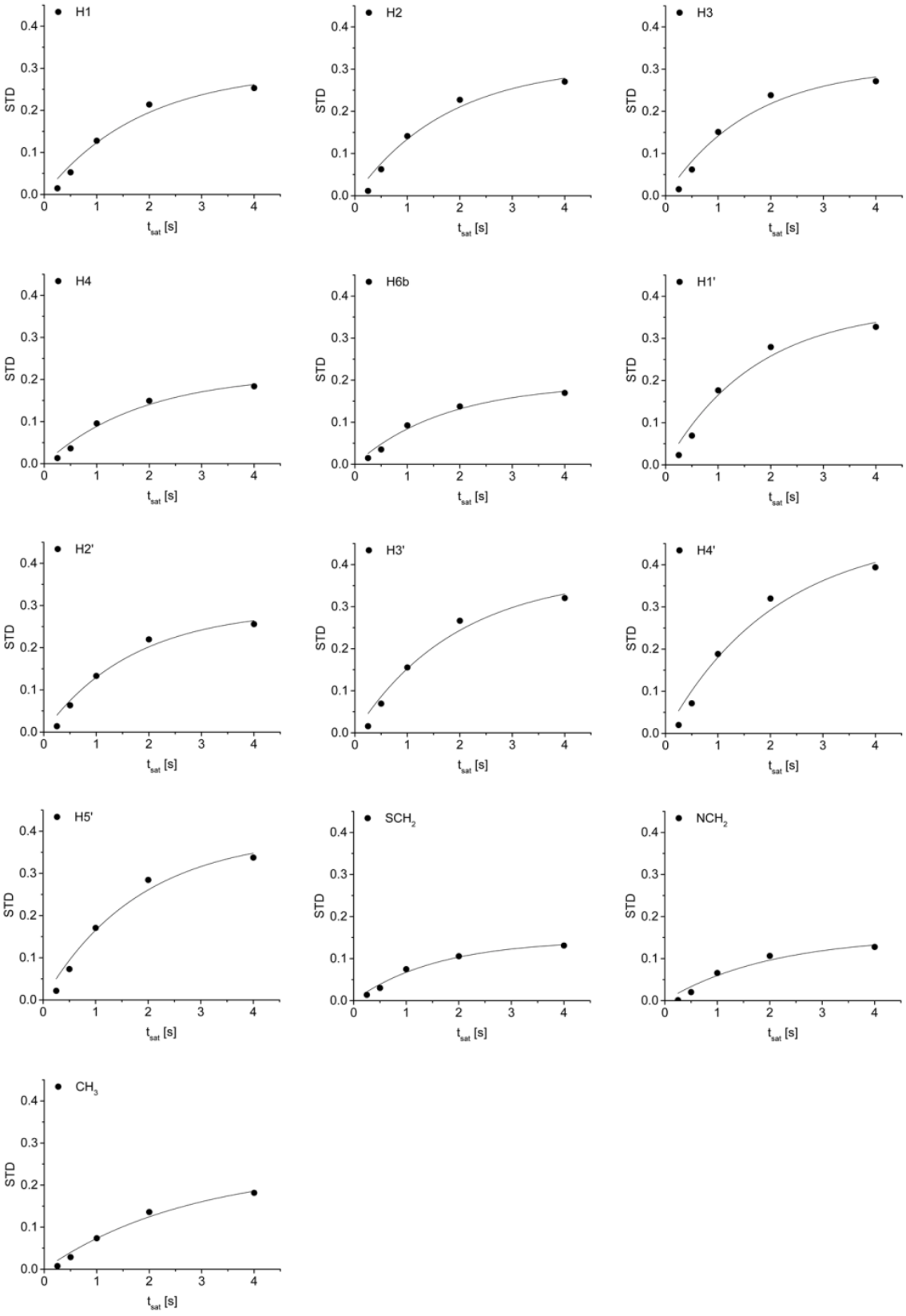


Figure B.8. STD NMR build-up curves for Man analog 14.5.

Equation 8 was fitted to STD values to calculate STD_0' values for the determination of the binding epitope of **14.1**.

B.6. K_I Determination for Glucosamine-2-Sulfate Analog Library 31

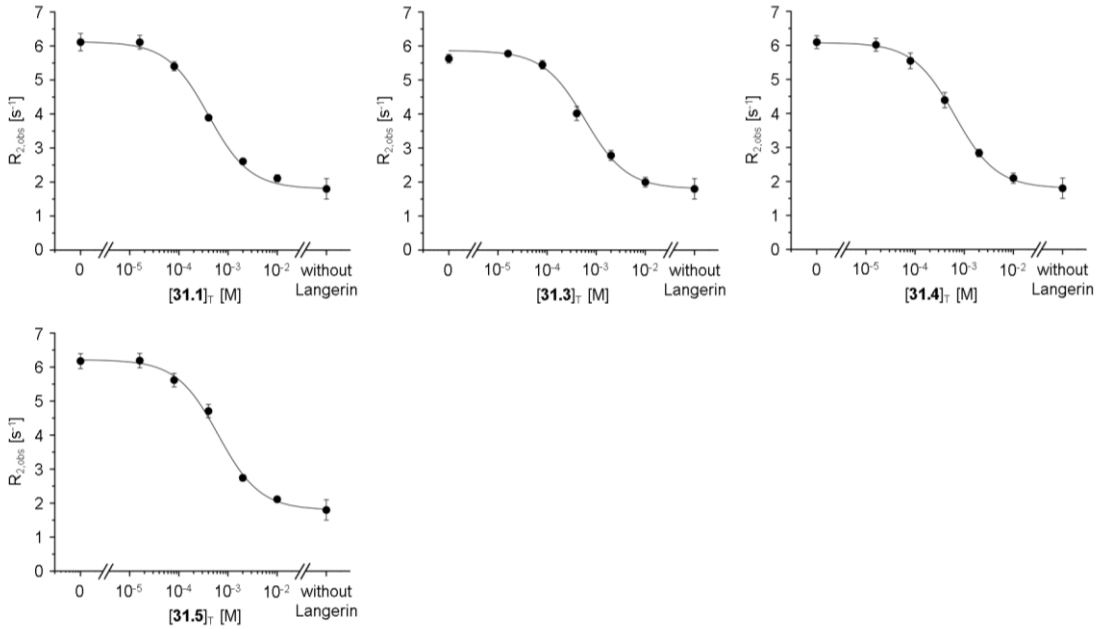


Figure B.9. K_I determination for GlcNS analog library 31.

Competitive binding experiments served to determine the affinities for GlcNS analog library **31**. Obtained K_I values are given in Table 15 (Chapter 3.4.1).

B.7. K_D Determination for Glucosamine-2-Sulfate Analog 31.2

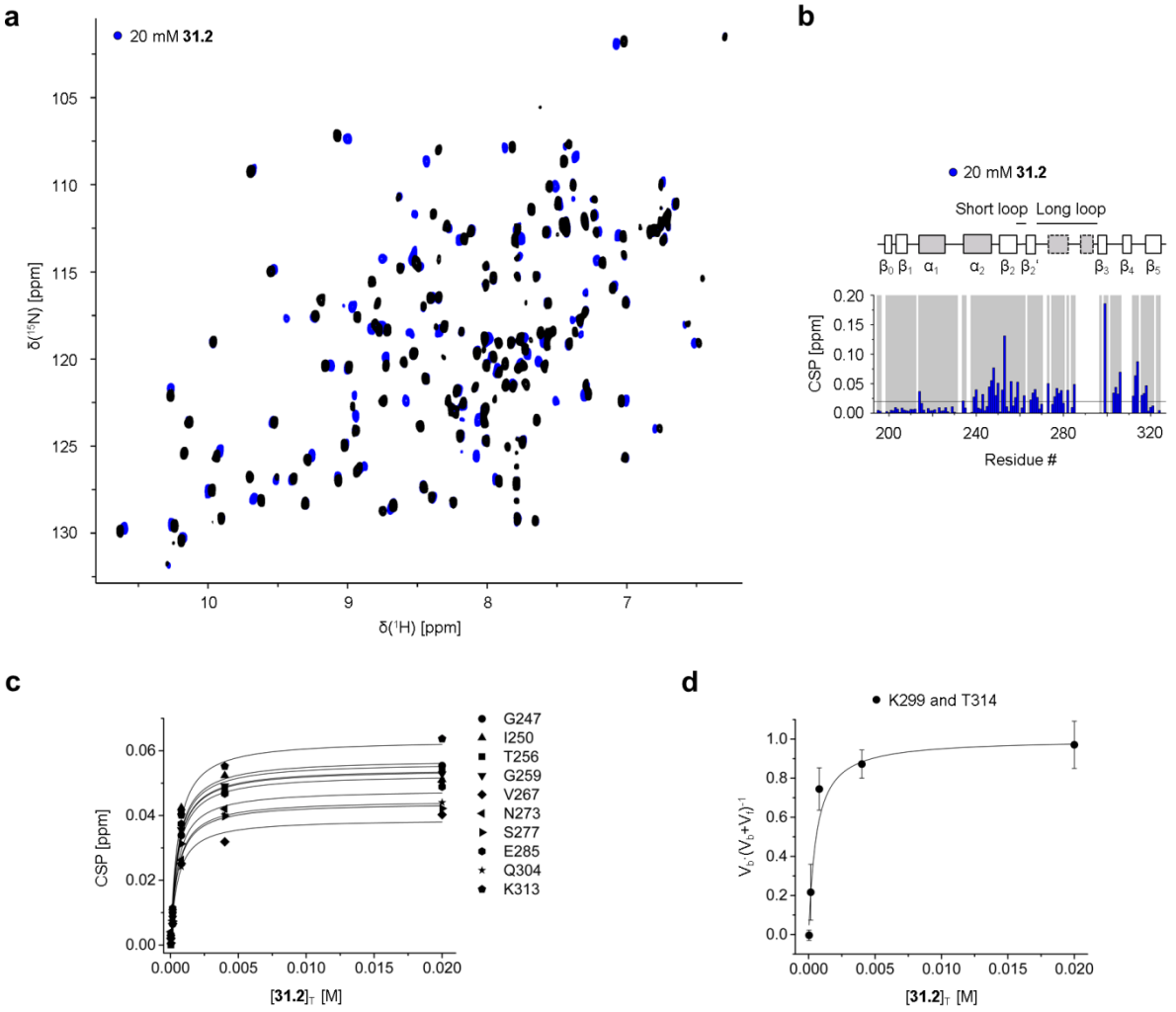


Figure B.10. K_D determination for GlcNS analog 31.2.

a and b. ^{15}N HSQC NMR experiments served to validate the obtained K_I value for **31.2**. **c.** Assigned resonances displaying fast chemical exchange and CSPs larger than 0.04 ppm were selected for the determination of K_D values. **D.** Additionally, a set of residues including K299 and T314 displayed slow exchange phenomena. For these residues, integrals V_f and V_b of resonances corresponding to the free and the bound state of the Langerin were utilized to determine K_D values. Obtained K_D values are given in Table 16 (Chapter 3.4.2).

B.8. STD NMR Build-Up Curves for Glucosamine-2-Sulfate Analog 41.2

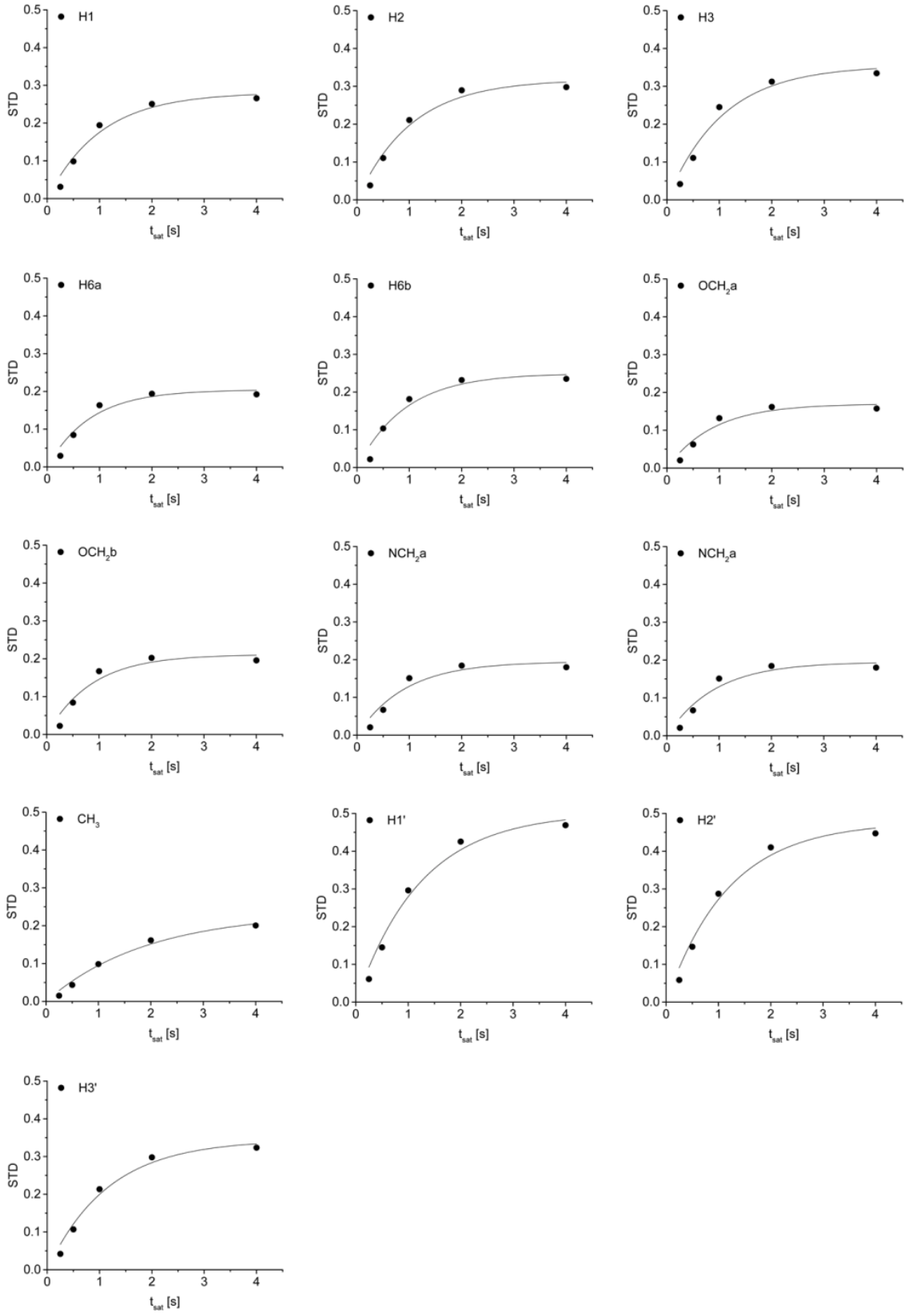


Figure B.11. STD NMR build-up curves for GlcNS analog 41.2.

Equation 8 was fitted to STD values to calculate STD_0' values for the determination of the binding epitope of **14.1**.

B.9. K_I Determination for Glucosamine-2-Sulfate Analog 31.2 with DC-SIGN

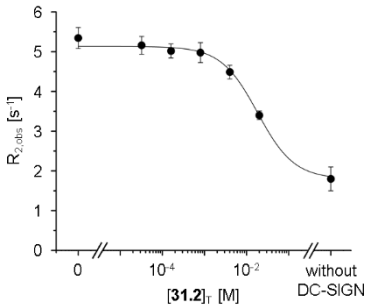


Figure B.12. K_I determination for GlcNS analog 31.2 with DC-SIGN.

Competitive binding experiments served to determine the affinities for GlcNS analog library **31**. Obtained K_I values are given in Table 17 (Chapter 3.4.3).