Novel Chemotypes for Inhibition of Bacterial and Mammalian Carbohydrate-Binding Proteins

Inaugural-Dissertation

to obtain the academic degree

Doctor rerum naturalium (*Dr. rer. nat.*)

submitted to the Department of Biology, Chemistry, Pharmacy of

Freie Universität Berlin

by

Elena Shanina

from Saratov, Russian Federation (USSR)

Berlin, July 2021

This dissertation was contrived between February 2018 and July 2021 under the supervision of Professor Dr. Christoph Rademacher at the Department of Biomolecular Systems headed by Professor Dr. Peter H. Seeberger at The Max Planck Institute of Colloids and Interfaces in Potsdam-Golm, Germany.

1st Reviewer: Professor Dr. Christoph Rademacher
2nd Reviewer: Professor Dr. Hartmut Oschkinat

Date of defense: 07.02.2022

Acknowledgments

Finally, I would like to thank my family. Your encouragement and support has helped me to follow my curiosity, ambition and dreams.

Contents

Publications

This thesis is divided into four chapters, two of which have been published as separate papers. Other two will be published and communicated.

1. Protein-observed ^{19}F NMR of LecA from *Pseudomonas aeruginosa*.

<u>**Shanina, E.**</u>; Siebs, E.; Zhang, H.; Varón Silva, D.; Joachim, I.; Titz, A.; Rademacher, C.

Glycobiology **2021**.

https://doi.org/10.1093/glycob/cwaa057

ES wrote the manuscript and established PrOF NMR for 5FW LecA.

2. Automated glycan assembly of ^{19}F labeled glycan probes enables high-throughput NMR studies of protein-glycan interactions.

Fittolani, G.*; <u>**Shanina, E.**</u>*; Guberman, M.; Seeberger, P.H.; Rademacher, C.; Delbianco, M.

Angew Chem Int Ed Engl **2021**.

https://doi.org/10.1002/anie.202102690

ES performed NMR experiments and contributed to writing the manuscript by describing the NMR studies.

* These authors contributed equally to the work and share the first authorship.

3. Identification of druggable allosteric pockets in β-propeller lectins.

<u>**Shanina, E.**</u>; Kuhaudomlarp, S.; Lal, K.; Seeberger, P.H.; Imberty, A.; Rademacher, C.

Submitted **2021**

ES and CR designed the studies; ES wrote the manuscript and performed NMR experiments.

4. Metal binding pharmacophores as inhibitors of carbohydrate-protein interactions.

<u>**Shanina, E.**</u>; Kuhaudomlarp, S.; Siebs, E.; da Silva Figueiredo Celestino Gomes, P.; Fuchsberger, F.F.; Denis, M.; Clausen, M.H.; Seeberger, P.H.; Rognan, D.; Titz, A.; Imberty, A.; Rademacher, C.

In preparation **2021**

ES and CR designed the studies; ES wrote the manuscript and performed NMR experiments.

List of Symbols and Abbreviations

2FF	2-Deoxy-2-fluoro-L-fucose
5FI	5-Fluoroindole
5FW	5-Fluorotryptophan
AGA	Automatic glycan assembly
APC	Antigen-presenting cell
CF	Cystic fibrosis
CLR	C-type lectin receptors
CMP-Neu5Ac	Cytidine monophosphate donor
CPMG	Carr-Purcell-Meiboom-Gill pulse sequence
CRD	Carbohydrate-binding domain
CSA	Chemical shift anisotropy
CSP	Chemical shift perturbation
D−Gal	D−Galactose
Da or kDa	Dalton or kilodalton
DC-SIGN	Dendritic cell specific ICAM-3 grabbing non-integrin
DD	Dipole-dipole coupling
DNA	Deoxyribonucleic acid
ECD	Extracellular domain
EDTA	Ethylenediaminetetraaceticacid
ESKAPE	Names of six highly virulent and antibiotic resistant bacterial pathogens including: *Enterococcus faecium*, *Staphylococcus aureus*, *Klebsiella pneumoniae*, *Acinetobacter baumannii*, *Pseudomonas aeruginosa*, and *Enterobacter* species
F-glycan	Fluorinated glycan
F-Le^x or ^{y}	Fluorinated Lewis antigen x or y
FAXS	Fluorine chemical shift anisotropy and exchange for screening
FBDD	Fragment-based drug design
FDA	U.S. Food and Drug Administration
FID	Free induction decay
FP assay	Competitive fluorescence polarization assay
gp120	Glycoprotein 120
HA	Heavy atoms
HEPES	*N*-2-hydroxyethylpiperazine-*N*-2-ethane sulfonic acid
HSQC	Heteronuclear single quantum coherence spectroscopy
HTS	High-throughput screening

IC_{50}	Half maximal inhibitory concentration
iDC	Immature dendritic cell
INEPT	Insensitive nuclei enhanced by polarization transfer
IPTG	Isopropyl-β-D-thiogalactoside
ITC	Isothermal titration calorimetry
K_d	Dissociation constant
LE	Ligand efficiency
MBP	Metal binding pharmacophore
MeGal	Methyl-α-D-galactose
MeFuc	Methyl-α-L-fucopyranoside
mM	Millimolar
µM	Micromolar
MS	Mass spectrometry
MW	Molecular weight
Neu5Ac	*N*-Acetyl-neuraminic acid
nM	Nanomolar
NMR	Nuclear magnetic resonance spectroscopy
PAINS	Pan-assay interference compound
Pmα23ST	α(2,3)-sialyltransferase
Pmα26ST	α(2,6)-sialyltransferase
pNPGal	p-Nitrophenyl β-D-galactoside
PrOF	Protein-observed ^{19}F NMR
PRR	Pattern recognition receptor
SAHA	Suberanilohydroxamic acid
SAR	Structure-activity relationship
Siglecs	Sialic acid-binding immunoglobulin-like lectins
SNFG	Symbol Nomenclature For Glycans
SPR	Surface plasmon resonance spectroscopy
STD	Saturation transfer difference
TLRs	Toll-like receptors
Tris	Tris(hydroxymethyl)aminomethane
TROSY	Transverse relaxation optimized spectroscopy
WT	Wild type

List of Figures

List of Tables

List of Schemes

Summary

The carbohydrate-binding proteins (lectins) emerged as viable targets to combat viral as well as bacterial pathogens. Therefore, drugs targeting lectins are desired; however their identification and development is challenging and is currently primarily focused on carbohydrate−based inhibitors. Therefore, new strategies and sensitive methods are required. Fragment-based drug design (FBDD) has proven to be a promising strategy for approaching difficult targets such as lectins. To address the current limitations in design of drug-like inhibitors for lectins, non- and metal-dependent bacterial or mammalian lectins are used. First, bacterial lectins from the opportunistic human pathogens *Pseudomonas aeruginosa* (LecA (PA-IL) and LecB (PA-IIL)) and *Burkholderia ambifaria* (BambL) were employed as models to establish ligand- (F-glycan) and protein-observed ^{19}F NMR (PrOF) methods for drug discovery. To demonstrate the utility of these methods for fragment-based drug discovery (FBDD), a druggable pocket in BambL was uncovered as a potential target site for allosteric inhibitors. Finally, these methods were employed as well as other biophysical (X-ray, SPR), computational and biochemical techniques to discover a novel class of drug-like molecules for targeting the carbohydrate-binding site of metal-dependent bacterial and mammalian lectins. Together, the ^{19}F NMR-based methods and discovery of metal-binding pharmacophores (MBPs) as novel chemotypes will support the development of small molecule inhibitors for metal-dependent lectins and bacterial lectins as new therapeutic approaches against antibiotic-resistant pathogens.

Zusammenfassung

Die kohlenhydratbindenden Proteine (Lektine) haben sich als Zielproteine zur Bekämpfung von viralen und bakteriellen Pathogenen herausgestellt. Daher sind die Medikamente, die auf die Lektine abzielen, erwünscht. Ihre Identifizierung und Entwicklung ist jedoch eine Herausforderung und ist basiert derzeit hauptsächlich auf kohlenhydrathaltige Inhibitoren. Daher werden neue Strategien und sensitive Methoden benötigt. Die fragmentbasierte Wirkstoffentwicklung (FBDD) hat sich als eine vielversprechende Strategie für die schwierige Zielproteine wie Lektine erwiesen. Zunächst wurden bakterielle Lektine aus den opportunistischen Humanpathogenen *Pseudomonas aeruginosa* (LecA (PA-IL) und LecB (PA-IIL)) und *Burkholderia ambifaria* (BambL) als Modellproteine untersucht, um Liganden- (F-Glykan) und proteinbasierte ^{19}F NMR (PrOF) Methoden für die Entwicklung von wirkstoffähnlichen Liganden zu etablieren. Die Nützlichkeit dieser Methoden für die fragmentbasierte Wirkstoffsuche (FBDD) wurde demonstriert, indem eine Wirkstofftasche in BambL als potentielle Zielstelle für allosterische Inhibitoren aufgedeckt wurde. Schließlich wurden diese und andere biophysikalische Methoden (Röntgenkristallographie und SPR) eingesetzt, um eine neue Klasse von wirkstoffähnlichen Molekülen für die Ca^{2+}-bindende bakteriellen und Säugetier-Lektinen zu entdecken. Im Zusammenfassung, die ^{19}F NMR-basierten Methoden wurden zur Entdeckung von metallbindenden Pharmakophoren (MBPs) als neuartige Klasse der wirkstoffähnlichen Liganden für metallbindende Lektine etabliert, die neue therapeutische Ansätze gegen antibiotikaresistente Krankheitserreger unterstützt werden.

1. Introduction

1.1. Fragment−based drug design

Discovering new drug candidates is challenging. To obtain hit molecules against a particular protein target, large libraries of chemical compounds are commonly screened to identify starting points for further drug development. In the past 20 years, fragment-based drug design (FBDD) has emerged as an alternative to conventional drug discovery methods such as high-throughput screening (HTS) and structure-guided drug discovery, which were fruitful for established classes of targets.[1-2] However, FBDD is frequently applied to find lead molecules for proteins lacking deep or well-defined druggable pockets, so-called "undruggable" targets.[2-3] For example, Vemurafenib was the first inhibitor that originated from an FBDD screen, which was approved by U.S. Food and Drug Administration (FDA) in 2011.[4] Many other examples for the successful use of fragment-based campaigns to discover clinical candidates followed.[5-7] Finally, fragment-screening approaches are relatively inexpensive, which facilitates their extensive use in industry as well as academia.[8]

1.1.1. Theory of fragment−based drug design

Fragment screening has several advantages over conventional HTS. FBDD libraries are composed of small molecules with a molecular weight (hereafter, MW) less than 300 Da (**Figure 1.1-1a**), whereas HTS is focused on complex drug-like compounds with a higher MW above 500 Da. In this respect, the advantage of the fragments is that small molecules cover the chemical space of the protein targets more adequately.[9] In particular, HTS hits frequently have parts of a molecule not contributing to binding, whereas fragments are more likely to optimally bind to protein pockets resulting in more 'high quality' interactions (**Figure 1.1-1b**). Moreover, the structure-activity relationship (SAR) studies of the fragment hits move forward faster, because analogues can be synthesized more easily or are commercially available due to their lower complexity.[10] Finally, screening of fragment libraries could be used to reflect the protein druggability.[11-12]

The aim of protein druggability assessment is to evaluate if it is possible to develop a drug-like molecule to modify the activity of a target protein.[3] The drug-like molecules follow the 'rule of five', which evaluates the druglikeness of a chemical compound that would likely make it orally available: MW ≤ 500 Da, H-bond donors ≤ 5 and acceptors ≤ 10, the lipophilicity (cLogP) ≤ 5, and the total number of nitrogen and oxygen atoms ≤ 10.[13] Consequently, all compounds of a fragment library used to be examined for

these drug-like properties. However, the analysis of hit compounds from early fragment screens in the mid-2000s led to new principles for the design of a fragment library, namely to the 'rule of three'.[14-17] The 'rule of three' proposes that molecules are considered fragments if they have a MW ≤ 300 Da, ≤ 3 hydrogen bond donors and acceptors, a cLogP < 3 and ≤ 3 rotatable bonds. Even though there are ambiguities in how donors and acceptors are defined, total MW and number of the atoms except for hydrogen called heavy atoms (HA) are probably the most important factors in selection of fragments for FBDD libraries.[18] Together, these rules allow assessing which fragments to obtain for FBDD campaigns.

The main challenge of the FBDD approach is that fragments have a poor affinity for the target due to the limited number of the interactions they can engage in. However, they may bind tightly enough relative to their size and number of HAs.[19] Therefore, the concept of ligand efficiency (LE) has been introduced in addition to the binding affinity.[20] LE quantifies the activity of a fragment relative to its HAs. This is usually derived from the experimental K_d or IC_{50} values. Moreover, K_d can be used to calculate the free energy of fragment binding ($\Delta G_{binding}$), which is divided by the number of HAs in the fragment. This is the most common metric used in FBDD for the evaluation of the quality of the fragment-protein interactions as opposed to the absolute binding efficiency. As a rule of thumb, hits with a high LE of ~0.3 kcal mol^{-1} HA^{-1} serve as good starting points for drug development.[21] Hereby, any gains in activity can be normalized to the increase in molecular weight. If a LE of 0.3 kcal mol^{-1} HA^{-1} is kept during the fragment evolution from its low molecular weight until its drug-like size of around 400 Da, it will result in K_d in the ten-nanomolar range. Thus, hits from fragment screening likely have improved LE values over HTS hits, because they form 'higher-quality' intermolecular interactions with the target proteins at an early stage of the drug development.[22]

Altogether, identifying a suitable starting point is generally a bottleneck for drug discovery, which is even more difficult for challenging target proteins. However, fragments continue to prove itself as promising molecules in the drug development,[5] where most HTS-derived drugs are difficult to optimize.[23-24] Consequently, FBDD campaigns are routinely used for challenging proteins in addition to HTS in academia and industry.[25]

1.1.2. Metal−binding pharmacophores in fragment−based drug design

During a fragment screen, small molecules frequently cause false-positive effects. Such compounds are known in the literature as 'bad actors' or pan-assay interference compounds (PAINS).[8,26-27] PAINS are small, aggregation-prone and insoluble

molecules, reactive covalent modifiers, redox active species and metal chelators that can bind proteins nonspecifically. Metal chelators are classified as PAINS based on empirical experience. For instance, metal chelators are commonly contaminated with metal ions, which may introduce a transition or heavy metal ions into a biological assay and thus, should be used carefully at the high concentrations.[28-29] Cumulatively, these factors led to the exclusion of metal chelators from FBDD libraries as their presence can partially or fully compromise the screen.[27,30]

Given the presence of various proteins in the genome requiring metal ions for their activity and stability, it is rather incorrect to exclude all metal coordinating fragments as PAINS.[31] For instance, Fesik et al. reported the use of hydroxamates for metalloenzyme inhibitor development for the first time in 1997.[32] Here, the authors targeted stromelysin-1 (MMP-3), which is a metalloenzyme with a Zn^{2+} ion in the active site required for catalytic activity. Interestingly, the only hit, acetohydroxamic acid, was found to inhibit MMP-3 by coordinating Zn^{2+} with the dissociation constant (K_d) of 17 mM. Future structure-activity relationship (SAR) studies improved the scaffold by introducing a biphenyl group (K_d=20 µM), which resulted in the first MMP-3 inhibitor with an IC_{50} value of 15 nM, interacting with Zn^{2+} in the active and neighboring sites.[33] Later, Cohen et al. proposed a concept of FBDD libraries using metal-binding pharmacophores (MBPs, **Figure 1.1-1c**),[34] where fragments were effectively used in the target-directed FBDD campaigns for metalloenzymes.[35-36] As result, several drugs have been marketed for metalloproteins.[37-40] However, the concept of a targeted FBDD library is controversial, as FBDD libraries are designed to be universal to cover a larger, unbiased chemical space towards the targets.[8] Nevertheless, a few reported studies using target-directed FBDD libraries were more successful than using general libraries.[41-43]

Taken together, MBPs are not well suitable fragments for being incorporated into the general libraries. However, it has proved its utility in target-oriented FBDD campaigns as it has been shown for metalloenzymes. Interestingly, despite the presence of various proteins in the genome requiring metal ions for their activity and stability, MBPs have never been reported for other clinically relevant protein targets, such as lectins.

1.1.3. Methods for examining fragment−protein interactions

The detection of weak fragment-protein interactions requires highly sensitive methods. Frequently, various biochemical fluorescence-based assays are widely used in high-throughput screenings due to their high sensitivity, which is challenging for low-affinity binders such as fragments.[44] In contrast, biophysical methods as ligand- and protein-

observed nuclear magnetic resonance spectroscopy (NMR), X-ray crystallography[45], surface plasmon resonance[46] and mass spectrometry (MS)[47] are regularly used to detect weak interactions and thus, to screen fragment libraries.[48] In particular, NMR is the most suitable fragment screening technique, because it detects fragment binding in the high µM−low mM range, which are frequently the only hits found for challenging targets.[49-50] Moreover, FBDD fragment screenings by NMR do not require prior knowledge of protein structure, function or natural binding partners, which is frequently needed in enzymatic assays. Notably, NMR is frequently used to quantify binding affinities for the identified hits in order to establish an SAR study. Finally, as it directly observes binding events, NMR does not usually lead to false-positive hits due to the light scattering artifacts, which can interfere with other fluorescence-based screening techniques.[51]
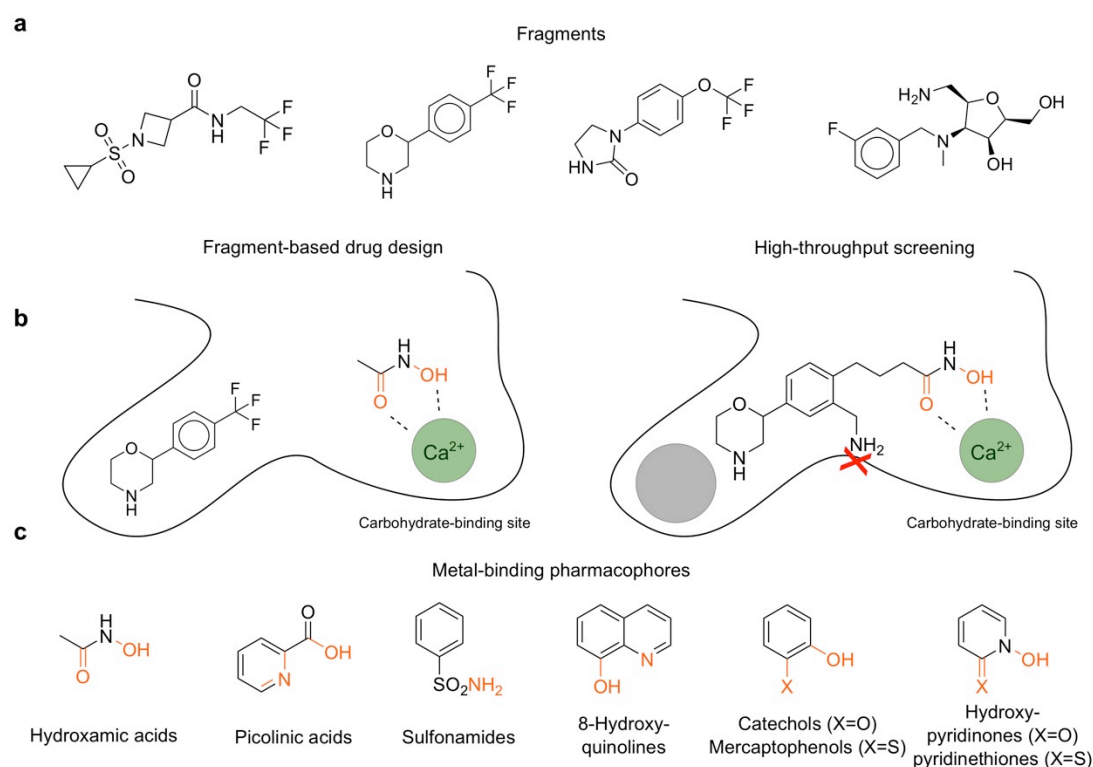
Figure 1.1-1 Fragment-based drug design.

a Example structures of drug-like fragments that are used in fragment-based drug design (FBDD). **b** Comparison of FBDD *vs* high-throughput screening (HTS) for discovery of drug-like molecules targeting challenging proteins such as carbohydrate-binding proteins. Hydroxamic acid coordinates a metal ion, whereas 2-(4-trifluoromethyl)phenylmorpholine binds the pocket near the carbohydrate-binding site. Merging both fragments could result in a potent inhibitor (*left*). In contrast, HTS hit gains binding through suboptimal interactions. Therefore, the core scaffold of the HTS hit requires a lot of modifications to remove nonessential functional groups (*right*). **c** Shown are the examples of metal-binding pharmacophores.

To characterize fragment-protein interactions by NMR, changes in the spectra of the fragments or the protein are monitored and are defined as ligand- or protein-observed NMR, respectively.[10] Ligand-observed NMR methods such as ^{1}H saturation transfer difference (STD)[52] and ^{19}F NMR[53] are well-suited for screening of large fragment libraries, and consume low amounts of protein and ligand. Complimentary, protein-observed NMR methods provide structural information on the binding site of fragment hits similar to X-ray crystallography. The details of both ligand- and protein-observed NMR methods are discussed in the **subchapters 1.1.4.** and **1.1.5**.

1.1.4. Application of ^{19}F NMR in fragment−based drug design

Incorporating fluorine atoms into a small molecule or the protein target provides a lot of benefits for studying protein-protein or ligand-protein interactions in NMR. The isotope ^{19}F has nearly the same sensitivity as [1]H (~83%). It has a spin 1/2 nucleus, is stable, has a natural abundance of 100% and is nearly absent in biological systems, thereby delivering background-free NMR spectra.[54] Unlike ^{1}H NMR, ^{19}F NMR has a wide chemical shift range and is sensitive to changes in local chemical environment. Altogether, these factors render ^{19}F NMR spectra fast and easy for acquisition and data analysis. Therefore, ^{19}F NMR is frequently applied in medicinal chemistry.

The first report on ^{19}F NMR from 1967 studied a small fluorinated molecule interacting with chymotrypsin.[55] In 2020, nearly 25% of FDA-approved drugs contain fluorine.[56] In medicinal chemistry, the ^{19}F atom is frequently present in many drug-like molecules for several reasons. First, ^{19}F can participate in hydrophobic and dipolar interactions with carbonyl groups of proteins and improve metabolic features of drugs by tuning acidity and lipophilicity for a better membrane permeability.[57-58] Furthermore, the chemical shift range and background-free ^{19}F NMR spectra enable multiplexed high-throughput screening of fluorinated fragment libraries with up to 50 compounds in a single measurement, which is difficult to achieve using ^{1}H NMR based methods due to the spectral complexity (**Figure 1.1-2a**). If a fragment binds to the target, the chemical environment of the fragment changes upon the interaction resulting in a chemical shift perturbation (CSP) or a line broadening of the ^{19}F signal. In particular, the transverse relaxation time (T_2) of the ^{19}F spin is dominated by chemical shift anisotropy (CSA), which is sensitive to the rotational correlation time of the ^{19}F spin. Therefore, ^{19}F NMR provides sharp signals for small molecules such as fragments, whereas the ^{19}F signal changes for fragments in a protein-bound state. When the ligands have a high affinity, the fluorine resonance is in a slow exchange on the NMR timescale. Therefore, two discrete resonances can be resolved in a simple and fast ^{19}F NMR spectrum of the ligand free *vs* protein-bound states. In this case, ^{19}F ligand-observed NMR can be

applied to derive the affinity (K_d) of the protein-ligand interaction. However, fragments are weak ligands as only a low amount of a fragment can be protein-bound in ^{19}F NMR screening. Therefore, fragments are frequently found in a fast exchange on NMR time scale demonstrating a CSP, where a single ^{19}F resonance represents the average of ligand-free and protein-bound populations. In this case, changes in CSPs of ^{19}F resonance can be used to derive the K_d value. However, CSPs are observed rather rarely with fragment and thus, more sensitive methods are needed.[59]

Most ligand-observed NMR techniques in medicinal chemistry focus on the relaxation rate effects. In contrast to fragments, proteins have longer rotational correlation times resulting in shorter transverse relaxation times (T_2). Consequently, a ^{19}F resonance of a fragment in a protein-bound state will show a stronger line broadening compared to the protein-free state. This effect is utilized in the spin-echo Carr-Purcell-Meiboom-Gill (CPMG) pulse sequence, which is a valuable method for detection of the weak fragment-protein interactions.[60-62] Briefly, a 90° pulse is applied resulting in a change of the equilibrium magnetization from z- down to y-axis. The system returns to the equilibrium, where two relaxation rates are distinguished: spin-lattice (longitudinal, T_1) and spin-spin (transversal, T_2). Notably, T_2 relaxation rate is accelerated if the rotational diffusion rate decreases and the rotational correlation time of a molecule increases due to protein-ligand binding.[63] The CPMG sequence is applied to isolate T_2 from T_1 relaxation, which results in a decrease of ^{19}F signal of fragments binding to the protein. The optimal length of the spin-echo filter is around 400 ms, which enhances the effect allowing the detection of the weak fragment-protein interactions in ^{19}F CPMG NMR screening.[64]

Fluorinated ligands having a low affinity for a protein are particularly well suited as reporter molecules in ^{19}F NMR allowing the detection of fragments binding to the protein site of interest. C. Dalvit and co-workers referred to this approach as fluorine chemical shift anisotropy and exchange for screening (FAXS, **Figure 1.1-2b**).[65-66] In this assay, the ^{19}F signal of the reporter molecule undergoes a line broadening as the reporter is in the protein-bound state. Once a higher affinity ligand is present, it replaces the reporter molecule from its binding site restoring the line width of the ^{19}F signal. Furthermore, the affinity of the hit compound can be determined if binding constant of the ^{19}F reporter molecule is known. Cumulatively, the competitive ^{19}F NMR provides the information on the fragment binding site and its affinity if the ^{19}F reporter molecule is used. Therefore, the development of new ^{19}F reporter molecules for lectins is of a great value.[67]

Similar to ligands, proteins can be ^{19}F-labeled as well. Protein-labeling for ^{19}F NMR followed in 1974 by Sykes et al. with the first report of a fluorine-labeled alkaline

phosphatase.[68] In protein-observed ^{19}F (PrOF) NMR the protein side chains are detected as broad resonances at low µM protein concentrations, where the line broadening increases in larger proteins (> 86 kDa) due to the chemical shift anisotropy (CSA) effect on fluorine nuclei.[69-71] This method has benefited from the commercial availability of many fluorinated aromatic amino acids such as 5-fluorotryptophan (5FW). Moreover, incorporation of fluorinated amino acids does not lead to major structural and functional perturbations.[70-72] For example, 5FW had only a minor impact on protein structure and dynamics in bacterial lectin from Ralstonia solanacearum
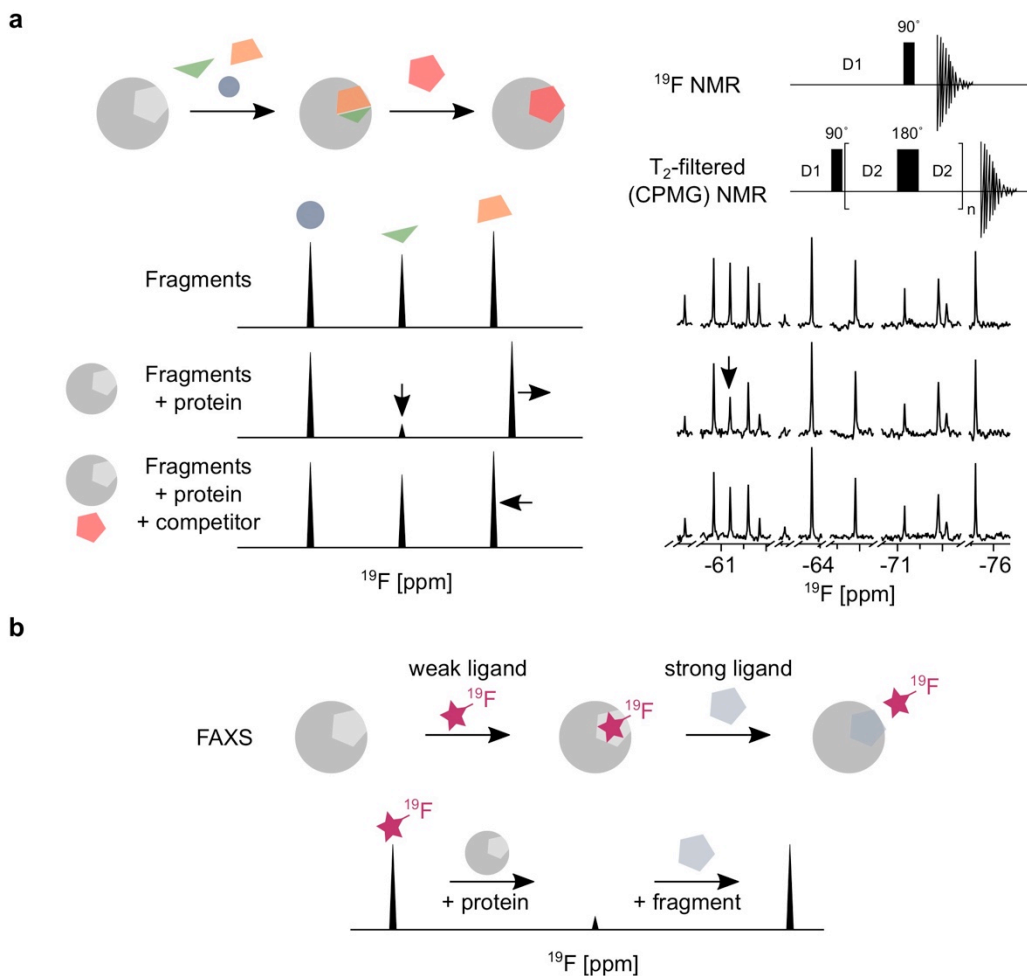
Figure 1.1-2 Ligand-observed ^{19}F NMR methods.

a Shown are the schemes for ^{19}F and CPMG NMR screenings of fluorinated fragments (*left top*) and pulse sequences (*right top*). Once a fragment is bound to a protein, a competition with a higher affinity ligand for the binding site of interest is performed to identify fragments targeting the site-of-interest (*left bottom*). Thereby, the ^{19}F signal either shows a chemical shift perturbation or a change in line width, which is competed with the stronger ligand, as shown on example of 'in-house' fluorinated fragment library and LecA using methyl-α-D-galactose (MeGal) as a competitor (*right bottom*). **b** The scheme of a FAXS experiment is shown, where a weak fluorinated ligand is a reporter in ^{19}F NMR for the protein site of interest and is competed with a stronger ligand.

lectin (RSL).[73] However, a high cost is associated with the synthesis of fluorine-labeled amino acids, whereas its precursors such as 5-fluoroindole (5FI) are cheaper, but more difficult to incorporate into proteins due to the antimicrobial properties of indole derivatives.[74-75] Despite the early reports on PrOF NMR, the protein labeling strategy has only recently been improved and applied to screen small drug-like molecules.[71,74,76] This method is not only suitable for screening, but also provides information on the fragment binding site and affinities. Compared to ^{1}H-^{15}N HSQC/TROSY NMR (see **subchapter 1.1.5**), PrOF NMR generates simpler spectra that can be quickly recorded and analyzed. Usually, hits identified in PrOF NMR screening are validated in orthogonal biophysical or biochemical methods.[71]

Taken together, both ligand- and protein-observed ^{19}F NMR are highly complimentary methods that provide structural information and binding affinities of ligands, protein-conformational effects, and affinities. Due to its high sensitivity, ^{19}F NMR is unique method in the fields of medicinal chemistry and chemical biology.

1.1.5. ^{1}H-^{15}N HSQC/TROSY NMR in fragment−based drug design

The ^{1}H-^{15}N heteronuclear single quantum coherence (HSQC)[77] or transverse-relaxation optimized (TROSY)[78] NMR experiments are frequently applied to gain the information on the protein structure and the fragment binding sites.[10] Briefly, the ^{1}H-^{15}N HSQC NMR experiment is based on the magnetization transfer from proton (^{1}H) to nitrogen (^{15}N) using the INEPT (Insensitive Nuclei Enhanced by Polarization Transfer) method. In the conventional HSQC pulse sequence (**Figure 1.1-3a**), the ^{1}H spin is excited and the first INEPT coherence transfer forwards the magnetization to ^{15}N. Next, the magnetization progresses on ^{15}N during the evolution time t_1, and refocuses back on ^{1}H by applying the second INEPT transfer. Finally, the FID is recorded on ^{1}H while ^{15}N is decoupled during the evolution time t_2. However, the application of this pulse sequence is limited to the molecular weight of proteins as larger molecules have longer rotational correlation times. This results in shorter T_2 and thus, leads to a line broadening of peaks in HSQC NMR spectrum. This is owned the fact that the T_2 relaxation time of large proteins is dominated by two effects decreasing the signal-to-noise ratio at a high magnetic field: the dipole-dipole (DD) coupling and CSA. Consequently, TROSY pulse sequence was designed to circumvent these effects through the constructive canceling of transverse relaxation caused by CSA and DD mechanisms at the high magnetic fields starting from 700 MHz (**Figure 1.1-3a**).

The main difference between TROSY and conventional HSQC pulse sequences is that the proton is not decoupled during the ^{15}N evolution and ^{15}N is not decoupled during ^{1}H acquisition in TROSY NMR. As result of ^{1}H non-decoupling, one peak appears as

multiples with different widths and TROSY sequence selects for a component with the narrowest line width leading to an enhanced resolution and sensitivity (**Figure 1.1-3b**). Therefore, the implication of TROSY NMR revolutionized the NMR field in 1997. As the magnetic field strength of NMR spectrometers increases, larger proteins and protein complexes (>100 kDa) can be recorded in TROSY-based 2D NMR.

Finally, HSQC and TROSY NMR are suitable for the detection of fragment binding. If a fragment binds to a protein, the chemical environment of the proton and nitrogen changes promoting a change in a chemical shift or line width of the resonance in the HSQC/TROSY NMR spectrum (**Figure 1.1-3c**). Hereby, high affinity ligands promote a slow-intermediate exchange on the NMR time scale, whereas weak ligands such as fragments result in the fast exchange. Owing this, ^1H-^{15}N HSQC/TROSY NMR allows the determination of the fragment affinity and even provides the structural information for the binding site of a ligand.[79-80] However, ^{15}N-isotope labeling of proteins, requirement of high field NMR spectrometers and sample preparation make 2D NMR experiments less throughput and thus, can be applied mostly for validation of fragment binding.
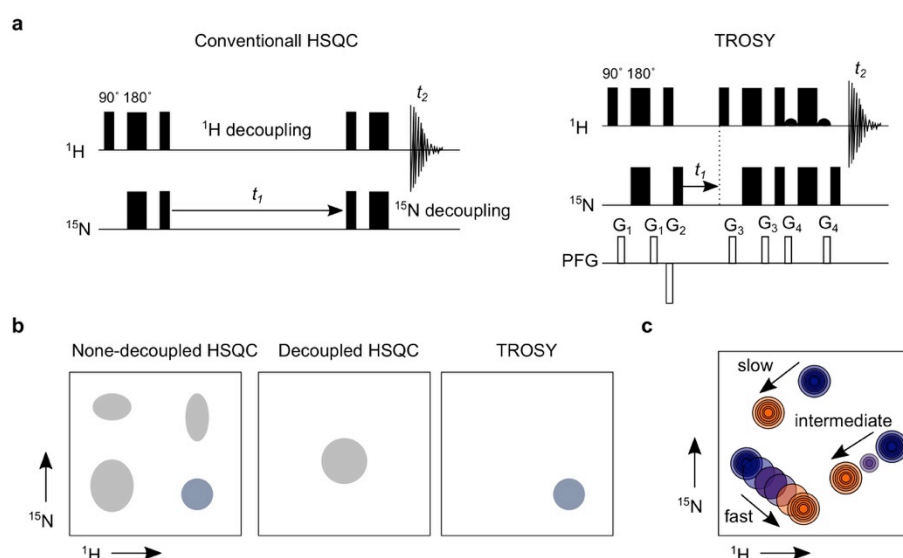
Figure 1.1-3 Protein-observed 2D NMR methods.

a Shown are schematic representations of a conventional ^1H-^{15}N HSQC and HSQC-TROSY pulse sequences (modified and adopted from [77-78]). PFG is a pulsed field gradient. **b** Shown are different relaxation rates (line widths) for each of four components as result of ^1H-^{15}N correlation in a none-decoupled HSQC, whereas a decoupled HSQC results in a single component for small molecules. At high magnetic field, the narrowest peak can be observed in TROSY NMR due to cancelling the transverse relaxation caused by CSA and DD. Therefore, TROSY NMR can be used to extend the limits of protein-observed NMR methods. **c** Shown are three binding modes of ligands in HSQC or TROSY NMR (*blue*: free state, *orange*: bound state).

1.2. Carbohydrate−binding proteins as promising drug targets

The surface of mammalian cells is decorated with a complex matrix of glycosylated proteins and lipids that are divergent from those present on pathogens. Notably, a particular glycosylation pattern can be recognized through carbohydrate-binding proteins called lectins. Innate and adaptive immune systems employ lectins in many aspects such as 'self'-'non-self' differentiation (e.g. pathogens or cancer cells) and cell-cell communication.[81-82] Interestingly, many pathogens developed strategies to exploit these mechanisms for host invasion or to modulate immune cell responses.[83] Therapeutics interfering with such interactions have a big potential as interventions against bacterial[84] and viral[85] infections, as well as cancer[86] and autoimmune diseases.[87] In an attempt to understand the nature of the pathogenesis and to develop new therapeutics, it is of a great importance to develop novel methods for the characterization of the molecular and biological mechanisms of the carbohydrate-lectin interactions.

1.2.1. Strategies for targeting lectins

Targeting lectins is a promising strategy to treat bacterial and fungal infections. However, discovering a suitable starting point for the inhibitor development has been a bottleneck for drug discovery due to the low druggability index of lectins compared to many other protein targets.[3] This is majorly due to the intrinsic properties of the carbohydrate-lectin interactions. The carbohydrate-binding sites are shallow, hydrophilic, and the specificity and the affinity are achieved through H-bonds, van-der-Waals contacts and hydrophobic binding of aromatic acid residues.[88-89] Together, it makes it difficult to approach lectins with small drug-like molecules, which lead to the vast underrepresentation of inhibitors for lectins in the drug space.[90]

1.2.2. Glycomimetics for targeting the carbohydrate−binding site

Monovalent carbohydrates often show a low affinity towards lectins. However, to reach higher avidities a multivalent display of the carbohydrates, lectins or both is utilized.[91] The vast majority of lectin inhibitors use carbohydrates as starting points for inhibitor development. For this, two main directions have been explored: 1) design of monovalent molecules to mimic the carbohydrates (glycomimetics)[92] and 2) multivalent presentation of carbohydrate epitopes (glycodendrimers,[93] glycopeptides94 and glycoclusters[95]). For instance, high affinity glycomimetic inhibitors have been reported for multiple bacterial (LecA[96-98], LecB[99] and BambL[100]) and mammalian lectins (Langerin[101], DC-SIGN[102]). Notably, this approach was applied successfully in drugs like Tamiflu® and Relenza® as influenza therapeutics. However, it requires a lot of

knowledge of the protein structure and thus, glycomimetic design is not trivial. Moreover, the carbohydrate-based glycomimetics have several disadvantages over the drug-like pharmacophores. First, a large-scale synthesis of these antagonists is still challenging. Secondly, monovalent carbohydrates usually show fast off-rates as the carbohydrate-protein interactions show low affinities. Using multivalent carbohydrate-based antagonists can improve the binding, but at the cost of lowering pharmacokinetic properties. Consequently, such large hydrophilic molecules show low tissue permeability, which impedes their oral bioavailability and thus, complicates the clinical development.[90] Taken together, the recent progress in lectin inhibitors has been focused either on the improvement of the carbohydrate-based mimetics or the discovery of drug-like pharmacophores.

1.2.3. Secondary sites for design of lectin inhibitors

As an alternative to the conventional drug design aiming to target the orthosteric site of proteins, targeting distant druggable secondary (allosteric) sites has gained a lot of attention in the past years. In 1965, Changeux introduced the concept of allostery in proteins, which states that ligands binding to the sites located away from the orthosteric site can modulate the protein activity.[103] Therefore, binding of a drug to an allosteric site can influence the protein conformation, which either enhances (positive modulation) or slows (negative modulation) the reaction at the orthosteric site.[104] Therefore, this concept might overcome the challenges of the drug selectivity and protein druggability in drug design. In particular, the allosteric modulators offer a way to overcome the limitations associated with targeting the carbohydrate-binding site of lectins. Recently, this approach has been explored for mammalian lectins.[105-106] Several druggable, allosteric pockets have been discovered for the mammalian C-type lectin receptor, DC-SIGN (CD209).[107] Moreover, an intra-domain allosteric network that modulates Ca^{2+} affinity of Langerin (CD207) has been described. This followed by the discovery of drug-like allosteric inhibitors for Langerin supporting the allosteric communication in mammalian lectins.[105,108-109] Cumulatively, these discoveries paved the way for further search of potential allosteric pockets in lectins.

1.2.4. Role of lectins in microbial pathogenicity

Lectins are found in many microorganisms such as bacteria, viruses, fungi and protozoa.[110-111] Moreover, these lectins often display a high affinity for mammalian carbohydrates compared to their mammalian counterparts, likely deriving from co-evolution.[81] Thus, pathogens take advantage of these interactions to adhere and to infect the host. Viruses use lectins for the same purpose. For instance, influenza virus

expresses hemagglutinin for binding to the sialic acid on the host cells, which results in the endosomal uptake of the virus into the cell and thus, the pH-dependent release of the viral RNA into the cytosol for viral replication. Interestingly, parasites use similar pathways for the host invasion, e.g. *Plasmodium falciparum* (*malaria*) uses EBA-175-glycan for entry into the red blood cells.[112]

Nathan Sharon and coworkers described for the first time in the 1970's how bacteria use lectins for cell adhesion and recognition to initiate bacterial infections.[113] Similar to viruses and parasites, bacterial lectins are involved in cell adhesion followed by invasion of the host. Moreover, bacterial lectins contribute to bacterial pathogenicity through biofilm formation. For instance, LecA (PA-IL) and LecB (PA-IIL) from Gram-negative *Pseudomonas aeruginosa* (*P. aeruginosa*) and the β-propeller lectin BambL from *Burkholderia ambifaria* are well-known examples of lectins from the pathogenic bacteria.[114-115] These bacteria can cause chronic infections and exhibit multidrug antibiotic resistance worldwide. They frequently affect immunocompromised patients, as well as those suffering from cystic fibrosis (CF) and can cause life-threatening pneumonia, respiratory failure and bacteremia.[116]

Together, microbial infections are increasingly difficult to treat as antibiotic resistance is rising worldwide. Therefore, identifying novel targets for design of anti-microbials or anti-adhesives became a promising approach for their treatment.[117-118]

1.2.5. LecA and LecB from *Pseudomonas aeruginosa*

The Gram-negative bacterium *P. aeruginosa* belongs to the group of ESKAPE pathogens, causes chronic infections and establishes a protective antibiotic−resistant biofilm environment in the lungs of immunocompromised patients.[119] This pathogen uses two carbohydrat-binding proteins LecA (PA-IL, **Figure 1.2-4a**) and LecB (PA-IIL, **Figure 1.2-4b**), which play important roles in establishing antibiotic-resistant biofilm and chronic infections.[114]

The crystal structures of LecA and LecB revealed two homotetramers with one and two calcium (II) ions (Ca^{2+}) in the carbohydrate recognition domain (CRD), respectively.[120-121] The Ca^{2+} ions are essential for the protein stability as well as the carbohydrate binding. For this, one Ca^{2+} ion in LecA coordinates 3-OH and 4-OH of D-galactose, whereas two Ca^{2+} ions in LecB allow it to establish interactions with L-fucose through 2-OH, 3-OH and 4-OH. Binding studies revealed that LecA has a low micromolar affinity for D-galactose (K_d=88 µM[122]), whereas LecB has exceptionally high affinity for L-fucose (K_d=2.9 µM[123]) and a lower affinity for D-mannose (methyl-α-D-mannoside, K_d=71 µM[124]).[125-127] Given these preferences, LecB was found to interact with the fucosylated antigens (ABH, Lewis, P and I systems) on human tissues and in

human milk with the highest affinity for Lewis[a] epitope (K_d=212 nM).[123,128] Moreover, high fucosylation of epithelial glycoproteins is found in airways of CF patients, which favors LecB-mediated attachment of *P. aeruginosa*.[129-130]

Both lectins are located in the cytoplasm of planktonic cells, but can also be found extracellularly in the biofilm.[131] The *P. aeruginosa* biofilms are composed of a matrix of extracellular polymers such as exopolysaccharides, extracellular DNA and various proteins such as LecA and LecB.[98,115,132] The proof of this concept for a critical role of lectins in biofilm formation was demonstrated in 2005 by showing that a LecB-deficient *P. aeruginosa* strain lost its ability to grow a biofilm.[133] Following a similar observation has been made with galactosides that have high affinities for LecA, such as IPTG (isopropyl-β-D-thiogalactoside) and p-nitrophenyl β-D-galactoside (pNPGal), which resulted in dispersion of mature biofilms.[114] Consequently, LecA and LecB have a high therapeutic interest and thus, a lot of attempts have been undertaken to design inhibitors for these lectins.

Several low molecular weight inhibitors have been reported for LecA, which are based on the β-linked galactosides with aromatic aglycon, e.g. pNPGal or GalAG2 with K_d=14.1 μM and K_d=0.1 μM, respectively (**Figure 1.2-4c**).[96-97,122,134] This remarkable improvement in binding affinity was achieved by addition of a CH-π interaction between the aromatic ring of the binding ligand and H50 in the protein. Interestingly, low affinity for monovalent inhibitors is also circumvented by multivalent display of D-galactose residues that can crosslink the neighboring carbohydrate-binding sites with K_d values in the nM range, such as GalAxG3 (K_d=2.5 nM).[92,135] Due to a better LE value, pNPGal (LE=0.31 kcal mol^{-1} HA^{-1}) has an advantage over GalAG2 (LE=0.03 kcal mol^{-1} HA^{-1}) and thus, only pNPGal serves as a good starting point for the inhibitor development. Finally, its poor pharmacokinetic properties are a large drawback for the clinically approval in future underscoring the importance of drug-like inhibitors.

Compared to LecA, design of monovalent inhibitors for LecB is more difficult due to its high affinity for fucosylated ligands. Therefore, more improvement has been achieved only with di- and oligovalent or dendrimer-based inhibitors.[136] As such, a tetravalent glycopeptide (*IC$_{50}$*=140 nM, **Figure 1.2-4c**) proved to be 10-fold more potent per fucose residue compared to its monovalent glycopeptide epitope resulting in inhibition of biofilm formation and its disruption at 50 μM.[99] Similar as for multivalent inhibitors of LecA, LecB inhibitors did not provide favorable pharmacokinetic properties for future medical applications.
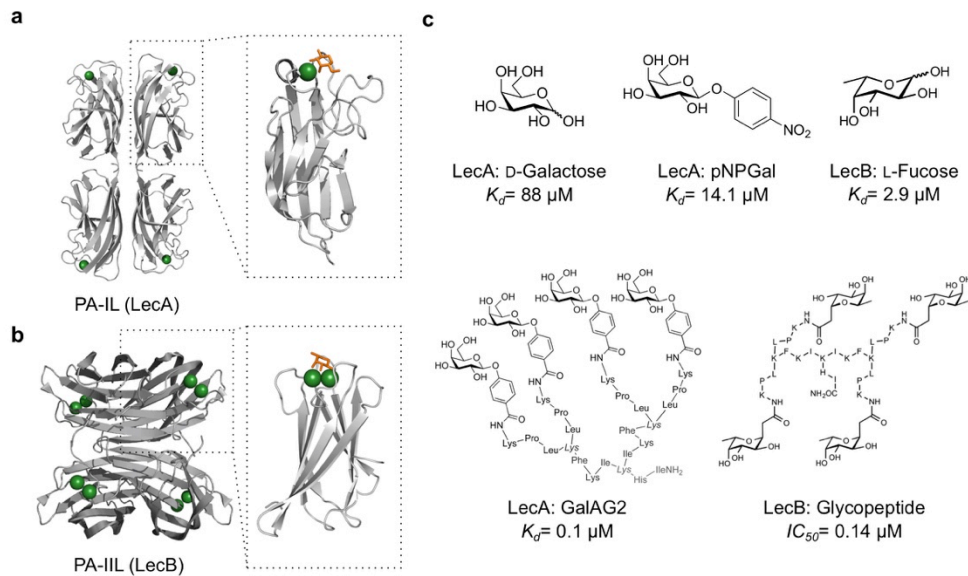
Figure 1.2-4 Bacterial lectins LecA and LecB from *P. Aeruginosa*.

a Homotetramer of LecA (*left panel*, PDB: 1OKO, *ligand not shown*) and its monomer bound to D-galactose (right panel, orange) with two Ca^{2+} ions (*green spheres*) in the carbohydrate-binding site. **b** Homotetramer of LecB (PDB: 1OXC, *ligand not shown*). The monomer contains two Ca^{2+} ions (*green spheres*) in the carbohydrate-binding site for binding to L-fucose (*orange*). **c** Shown are the structures and binding affinities (K_d) of natural ligands of LecA (D-galactose) and LecB (L-fucose). pNPGal and GalAG2 are mono- and multivalent inhibitors of LecA, whereas the glycopeptide is a multivalent inhibitor of LecB. The K_d values and the structures were taken from [123,136].

1.2.6. BambL from *Burkholderia ambifaria*

This Gram-negative bacterium pathogen *B. ambifaria* belongs to a group of closely related bacterial strains, the *Burkholderia cepacia complex*. These bacterial strains can cause chronic infections in CF patients as well as sporadic outbreaks, but its epidemiology remains largely elusive.[137-138] Interestingly, bacteria are suspected to use fucose-binding lectins for the adhesion to the lung tissue of CF patients.[139-140] *B. ambifaria* expresses a β-propeller lectin called BambL (**Figure 1.2-5a**), which has a strong affinity for α-L-fucosylated monosaccharides (methyl-α-L-fucopyranoside (MeFuc), K_d=1 µM) and complex carbohydrates (e.g. H type 2 tetrasaccharide, K_d=7.5 µM, **Figure 1.2-5b**). The crystal structure of BambL showed that the 6-bladed β-propeller consists of two similar domains and trimerizes forming a donut shape with six fucose-binding sites on one face of the protein.[141-142] Interestingly, several studies point to an underestimated role of BambL in affecting host cellular processes, which go beyond an adhesion to the human lung epithelium.[138,143]

Compounds able to inhibit the BambL-carbohydrate interaction may have therapeutic potential as antiadhesives. Similar to LecA, the design of inhibitors has been focused on using carbohydrates as a starting point.[141] Indeed, this approach has yielded potent

BambL inhibitors with 4 to 6 fucose or monovalent aryl-α-O-fucoside analogues that improved the selectivity and the affinity towards BambL with K_d ranging between 10 to 80 nM (**Figure 1.2-5b**).[144-147] However, the molecular size of such complex carbohydrate-based inhibitors remained being a limitation for BambL inhibitors as well. Consequently, discovering small, orally bioavailable drug-like molecules targeting BambL is still an unexplored research area.
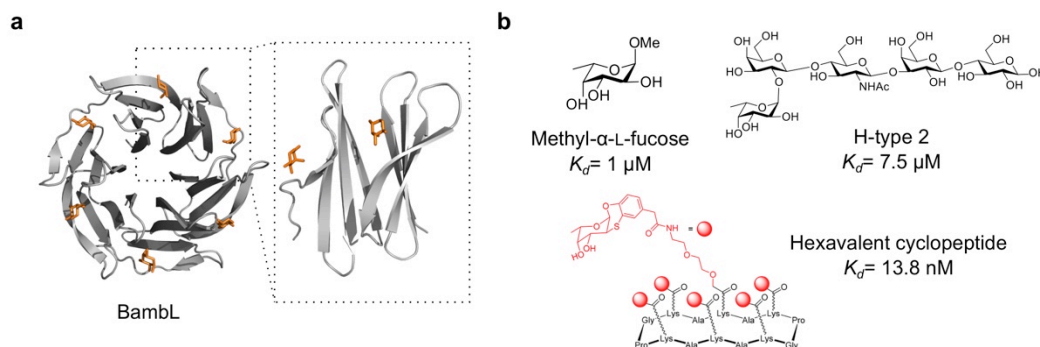
Figure 1.2-5 Bacterial lectin BambL from *B. ambifaria*.

a Homohexamer structure of BambL (PDB: 3ZZV) with six carbohydrate-binding sites for binding to α-L-fucopyranosides (*orange*), whereas one carbohydrate-binding site is within the monomer and the second one is between two monomers. **b** Structures and binding affinities of mono- and multivalent ligands of BambL. The structure of the hexavalent cyclopeptide was adopted from [145] and modified.

1.2.7. Mammalian C-type lectins

Our innate immune system has evolved to identify pathogens using pattern recognition receptors (PRRs) such as toll-like receptors (TLRs), sialic acid-binding immunoglobulin-like lectins (Siglecs) and C-type lectins (CLRs). CLRs are members of the largest mammalian lectin family and many are expressed on the surface of antigen-presenting cells (APCs), where they help APCs to recognize and endocytose the pathogens for processing and later presentation to the immune system. CLRs recognize the carbohydrates on the pathogens in a Ca^{2+}-dependent manner.[148] Interestingly, some pathogens learned to escape the endocytic pathway and thus, exploit CLRs to invade the host. Therefore, targeting CLR-carbohydrate interactions is a promising approach for antiviral therapy[149] or vaccine delivery.[150]

Langerin (CD207) and the dendritic cell specific ICAM-3 grabbing non-integrin (DC-SIGN, CD209) are well-known examples of CLRs, which are expressed on Langerhans cells and immature dendritic cells (iDCs)/macrophages, respectively.[151-153] Originally, DC-SIGN was identified as a receptor mediating T cell clustering in DC-dependent manner through binding to ICAM-3 that initiates contact with APCs.[154-155] Later reports demonstrated the role of Langerin and DC-SIGN in the recognition and

uptake of bacteria (*Mycobacterium tuberculosis*[156]) and various viruses (e.g. HIV type-1[157], Ebola[158], SARS-CoV[159] and SARS-CoV-2[160-161]), which may use CLRs for the immune escape.

Reports of the crystal structures of Langerin and DC-SIGN in combination with binding studies suggest that both CLRs recognize high-mannose oligosaccharides on the pathogens (**Figures 1.2-6a-c**).[162-163] In particular, CLRs consist of four structural subunits: 1) an intracellular signaling domain, 2) a transmembrane domain, 3) a neck domain and 4) the carbohydrate-recognition domain (CRD). Central to its function, the carbohydrate-CLR interactions require a Ca^{2+}-cofactor, e.g. Langerin shows a high affinity for Ca^{2+} (K_d=130 µM[109]). However, the CRD mediates the carbohydrate specificity by binding to the 3- and 4-OH groups of D-mannose with a low affinity (DC-SIGN: K_d=3.5 mM, Langerin K_d=6.1 mM,[164] **Figure 1.2-6d**).[162] However, the multivalent presentation of D-mannose in high-mannose structures improves the avidity of DC-SIGN to the carbohydrates, e.g. HIV gp120 (K_d=1 nM).[165] Additionally, DC-SIGN binds to the fucosylated oligosaccharides as parts of the Lewis-type epitopes, e.g, Lewis^x (Le^x).[166] Later reports implied that DC-SIGN could recognize pathogens by interaction with the carbohydrate antigen Le^x, which are found on the human pathogens, such as *Helicobacter pylori*, *Leishmania mexicana*, and *Schistosoma*.[167-168] Notably, several carbohydrate-based inhibitors (glycomimetics) against Langerin[101] and DC-SIGN[169-170] are available (**Figure 1.2-6d**), but only a few drug-like inhibitors have been discovered (**Figure 1.2-6e**).[171-173] Finally, drug-like inhibitors targeting the carbohydrate-binding site by coordinating the Ca^{2+} ion are highly desired, but have not been reported. Cumulatively, DC-SIGN and Langerin are uptake receptors for viral and bacterial pathogens. Consequently, both CLRs are promising targets for the development of 'entry' inhibitors against viral and bacterial pathogens.
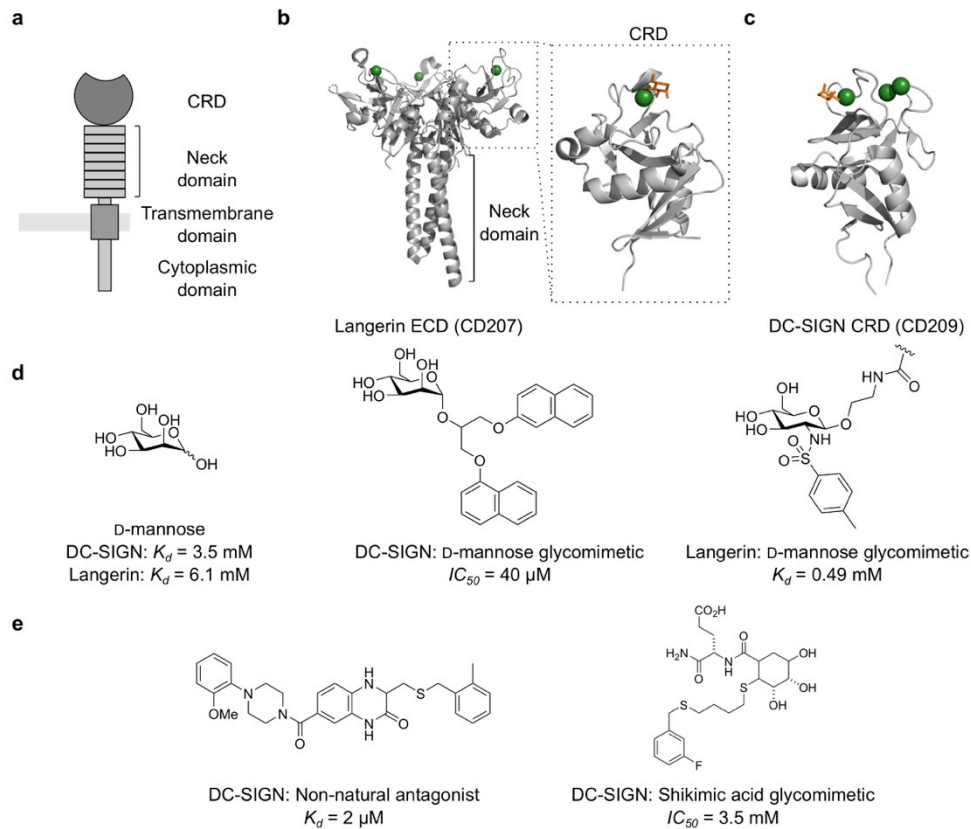
Figure 1.2-6 C-type lectin receptors Langerin and DC-SIGN.

a Schematic representation of domains in CLRs. **b** Shown is a homotrimer of Langerin (PDB: 3KQG) with one Ca^{2+} ion (*green sphere*) and D-mannose (*orange*) in the carbohydrate-binding domain (CRD, PDB: 3P5D). **c** Carbohydrate-binding domain of DC-SIGN with three Ca^{2+} ions (PDB: 1SL4). **d** Shown are the structures of D-mannose and carbohydrate-based monovalent glycomimetics of DC-SIGN and Langerin. **e** Shown are the structures a non-natural antagonist and a noncarbohydrate-based glycomimetic of DC-SIGN.

3. Aim of the thesis

Drug-like inhibitors for lectins are desired in academia as tools to learn more about the carbohydrate-protein interactions, as well as in healthcare industry for antibacterial therapy and vaccines. However, its design has been challenging and is currently limited to the carbohydrate-based inhibitors. Employing FBDD for targeting orthosteric and allosteric pockets has proven promising for approaching the difficult targets such as lectins. However, the shallow binding sites and weak affinities of fragment-lectin interactions complicate the inhibitor design. Cumulatively, discovery of new chemotypes and establishing sensitive methods are important for development of drug-like lectin inhibitors.

In this work, the aim was to establish broadly applicable protein- and ligand-observed ^{19}F NMR methods, which can improve our understanding of the carbohydrate-lectin interactions and accelerate the discovery of new drug-like molecules for lectins. First, protein-observed ^{19}F (PrOF) NMR was established and investigated for discovery of the drug-like molecules targeting the carbohydrate-binding site of LecA from *P. aeruginosa* (**Chapter 1**). Next, the goal was to expand the set of ligand-observed ^{19}F NMR methods for lectins. For this, complex fluorinated carbohydrates (hereafter, F-glycans) were shown as highly valuable reporter molecules, which became more affordable owing to advances in automated glycan-assembly (AGA). In this study, five lectins were employed to demonstrate that the carbohydrate-lectin interactions could be screened in a high-throughput manner, used to derive binding affinities and to characterize enzymatic reactions by ^{19}F NMR (**Chapter 2**). Finally, both methods as well as other computational, biophysical (X-ray, SPR) and biochemical techniques were employed to discover drug-like molecules for targeting the carbohydrate-binding and allosteric sites of lectins. In particular, a bacterial β-propeller lectin BambL from the pathogen *Burkholderia ambifaria* was investigated in fragment screening for the druggable secondary sites and the allostery, which will support the future work aiming to design drug-like allosteric inhibitors for bacterial and fungal lectins (**Chapter 3**). Moreover, the discovery of metal binding pharmacophores (MBPs) interacting with the carbohydrate-binding site provides a good starting point for the development of the drug-like inhibitors for mammalian and bacterial metal (Ca^{2+})-dependent lectins (**Chapter 4**).

4. Results

4.1. Protein-observed ^19F NMR of LecA from *P. aeruginosa*

Elena Shanina, Eike Siebs, Hengxi Zhang, Daniel Varón Silva, Ines Joachim, Alexander Titz, Christoph Rademacher

OXFORD

Structural Biology

Protein-observed 19F NMR of LecA from *Pseudomonas aeruginosa*

Elena Shanina [2,3], **Eike Siebs** [4,5,6], **Hengxi Zhang** [2,3],
Daniel Varón Silva [2,3], **Ines Joachim** [4,5,6], **Alexander Titz** [4,5,6], **and**
Christoph Rademacher [1,2,3]

[2]Max Planck Institute of Colloids and Interfaces, Department of Biomolecular Systems, Am Mühlenberg, 14424 Potsdam, Germany, [3]Free University of Berlin, Department of Biochemistry and Chemistry, 14195 Berlin, Germany, [4]Chemical Biology of Carbohydrates, Helmholtz Institute for Pharmaceutical Research Saarland, Helmholtz Centre for Infection Research, 66123 Saarbrücken, Germany, [5]Saarland University, Department of Pharmacy, 66123 Saarbrücken, Germany, and [6]German Center for Infection Research, Hannover-Braunschweig, Germany

[1]To whom correspondence should be addressed: Tel: +49 (0) 331-567-9358; Fax +49 (0) 331-567-9302;
e-mail: Christoph.Rademacher@mpikg.mpg.de

Abstract

The carbohydrate-binding protein LecA (PA-IL) from *Pseudomonas aeruginosa* plays an important role in the formation of biofilms in chronic infections. Development of inhibitors to disrupt LecA-mediated biofilms is desired but it is limited to carbohydrate-based ligands. Moreover, discovery of drug-like ligands for LecA is challenging because of its weak affinities. Therefore, we established a protein-observed 19F (PrOF) nuclear magnetic resonance (NMR) to probe ligand binding to LecA. LecA was labeled with 5-fluoroindole to incorporate 5-fluorotryptophanes and the resonances were assigned by site-directed mutagenesis. This incorporation did not disrupt LecA preference for natural ligands, Ca^{2+} and D-galactose (D-Gal). Following NMR perturbation of W42, which is located in the carbohydrate-binding region of LecA, allowed to monitor binding of low-affinity ligands such as *N*-acetyl D-galactosamine (D-GalNAc, $K_d = 780 \pm 97$ μM). Moreover, PrOF NMR titration with glycomimetic of LecA *p*-nitrophenyl β-D-galactoside (pNPGal, $K_d = 54 \pm 6$ μM) demonstrated a 6-fold improved binding of D-Gal proving this approach to be valuable for ligand design in future drug discovery campaigns that aim to generate inhibitors of LecA.

Key words: drug discovery, LecA, lectin, NMR

Introduction

Many opportunistic pathogens, such as the Gram-negative bacterium *Pseudomonas aeruginosa*, use glycan-binding proteins (lectins) to infect the host and to establish a protective antibiotic-resistant biofilm environment in the lungs of immunocompromised patients (Singh et al. 2000). LecA plays a key role in this process and has become a promising target to prevent biofilm formation and consequently disease progression (Diggle et al. 2006; Wagner et al. 2016).

LecA forms a protein homotetramer (Fig. 1A) and requires a calcium (II) ion (Ca^{2+}) to coordinate binding to its natural monosaccharide ligand, D-galactose (D-Gal) (Cioci et al. 2003). Notably, each monomer has four tryptophan residues (Fig. 1B), where W42 and W33 reside in close proximity to the D-Gal-binding region and near the hinge connecting the two domains, respectively (Fig. 1C). In the physiological context of a biofilm, the low micromolar affinity of LecA for D-Gal ($K_d = 88$ μM, (Kadam et al. 2011)) is compensated with high avidity of ligands that can crosslink the neighboring carbohydrate-binding sites, such as GalAxG3 ($K_d = 2.5$ nM, (Bergmann et al. 2016; Cecioni et al. 2015)); however, drug-like molecules do not benefit from such

159
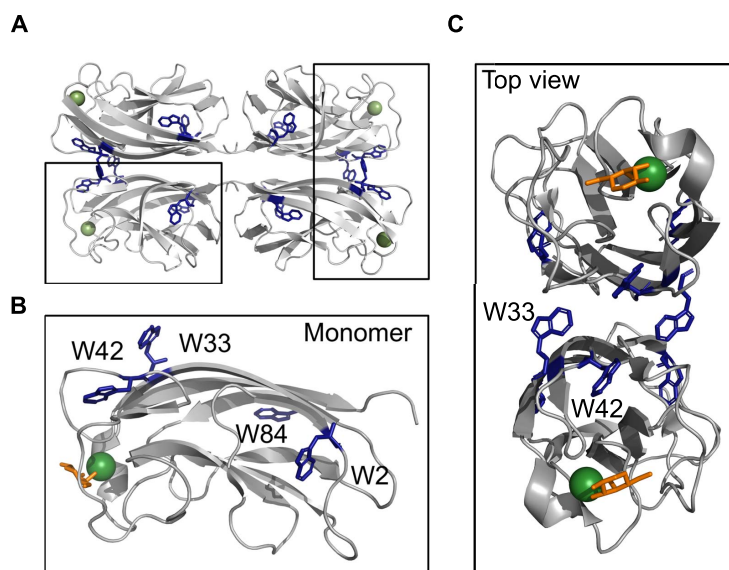
Fig. 1. Structure of LecA. Cartoon representation of tetramer LecA (PDB: 4CP9). Shown is an expansion of a LecA monomer with D-Gal (shown as sticks), Ca^{2+} ion (shown as sphere) and positions of four tryptophanes (W2, W33, W42, W84). W2 and W84 are in the protein core. Top view shows W42 and W33 being located near the carbohydrate-binding site of LecA.

multivalency. Successful approaches to find low-molecular-weight inhibitors have capitalized on the β-linked galactosides with aromatic aglycon (Garber et al. 1992), such as p-nitrophenyl β-D-galactoside (pNPGal) or GalAG0 with $K_d = 14.1$ μM and $K_d = 4.2$ μM, respectively (Kadam et al. 2011; Kadam et al. 2013; Rodrigue et al. 2013). This improvement in binding affinity is due to the additional CH–π interaction between the aromatic ring of the binding ligand and H50 at the binding pocket. Cumulatively, development of inhibitors for LecA consists of β-linked carbohydrate-based ligands (glycomimetics, (Wagner et al. 2017)) and, therefore, having drug-like ligands for LecA is crucial in future research, but discovery of such weak binders is challenging. Therefore, new methods to detect binding of weak ligands to LecA are required.

Biophysical methods are suitable to identify low-molecular-weight and low-affinity ligands (Renaud et al. 2016). 19F nuclear magnetic resonance (NMR) has proven to be valuable in the study of protein–ligand interactions for several reasons. The isotope 19F 1) has the spin 1/2 nucleus and a natural abundance of 100%, 2) is very stable and 3) nearly absent in biological systems delivering a background-free NMR spectrum (Luck and Falke 1991). In the case of 19F-labeled proteins used in protein-observed 19F (PrOF) NMR, the size of protein is not a limitation and the protein side chains are detected as broad resonances at low (25 μM) to mid-micromolar (200 μM) protein concentrations (Kitevski-LeBlanc and Prosser 2012; Liu et al. 2012). This method has benefited from the commercial availability of many fluorinated aromatic amino acids, such as 5-fluorotryptophan (5FW), 3-fluorotyrosine and 4-fluorophenylalanine. Unfortunately, these fluorine-labeled amino acids are expensive. In contrast, fluorine-labeled precursors of amino

acids, such as 5-fluoroindole (5FI), can be employed to incorporate fluorine-labeled amino acids in proteins, resulting in reduced costs (Gee et al. 2016). Moreover, incorporation of fluorinated amino acids does not lead to major structural and functional perturbations (Arntson and Pomerantz 2016; Kitevski-LeBlanc and Prosser 2012; Sharaf and Gronenborn 2015). In this context, 5FW has been shown to have only a minor impact on protein structure and dynamics in bacterial lectin from *Ralstonia solanacearum* lectin (RSL) (Tobola et al. 2018).

Here, we explored PrOF NMR using LecA labeled with 5FW (5FW LecA) to detect binding of ligands with moderate as well as low affinities. To assign 5FW resonances, we produced its wild-type (WT) and four tryptophan-to-phenylalanine mutants (W2F, W33F, W42F and W84F). In the binding studies, we determined the dissociation constants of 5FW LecA with its natural ligands Ca^{2+}, D-Gal and D-GalNAc. We compared the affinity data of LecA and 5FW LecA with other orthogonal biophysical methods, such as isothermal titration calorimetry (ITC) or competitive binding by fluorescence-polarization (FP) detection. Finally, we verified the suitability of 5FW LecA PrOF NMR for a ligand design using glycomimetics pNPGal and phenyl-β-D-galactopyranoside (Ph-β-D-Gal, (Imberty et al. 2004)).

Results and discussion

Protein expression and characterization

For the stable incorporation of 5FW in LecA we followed the workflow shown in Fig. 2A. *Escherichia coli* BL21 (DE3) cells were grown in presence of 5FI and the protein was characterized for
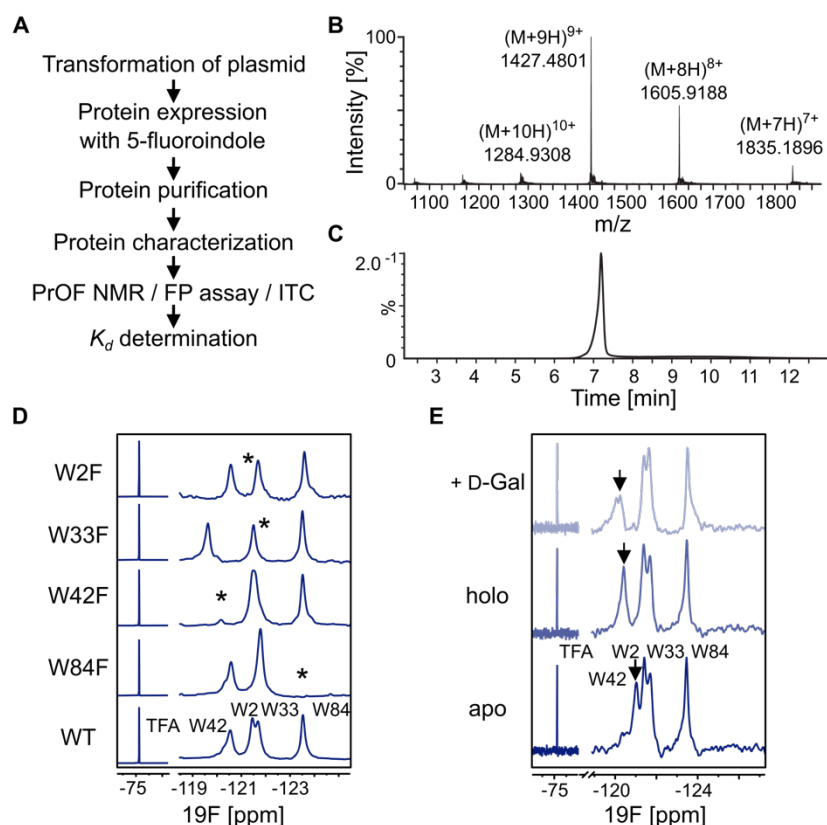
Fig. 2. PrOF NMR of 5FW LecA. **(A)** General workflow for PrOF NMR with 5FW LecA. **(B)** Chromatogram of the LC–ESI–MS analysis of 5FW LecA. **(C)** ESI-MS$^+$ spectrum of the main peak at 7.3 min $[M + H]^+$Ca = 12826.23 Da $[M + H]^+$found = 12831.34 Da corresponds to 5FW LecA. **(D)** PrOF NMR assignment of 5FW LecA WT and the mutants W84F, W42F, W33F and W2F. The tryptophanes being mutated are indicated with asterisk. All spectra were normalized and referenced to TFA. **(E)** PrOF NMR of 5FW LecA WT in Ca^{2+}-free (apo, *bottom*) and -bound (holo, *central*) forms. The W42 resonance (*black arrow*) shifted in presence of Ca^{2+} and 0.5 mM D-Gal binding verifying that protein is active.

fluorine incorporation *via* mass spectrometry (Fig. 2B and C). In the mass spectrum 5FW LecA had a dominant mass of 12831.34 Da corresponding to full incorporation of four tryptophan residues being replaced with 5FW. Protein yields as high as 45–50 mg L^{-1} using non-auxotrophic *E. coli* BL21 (DE3) cells were achieved. This compares very well to protein expression yields under non-labeling conditions (30–35 mg L^{-1}).

Establishing 5FW LecA in PrOF NMR
To optimize PrOF NMR for 5FW LecA, we varied temperature (Supplementary Fig. S1), protein concentration (Supplementary Fig. S2) and buffer (Supplementary Fig. S3). Fluorine resonances of 5FW LecA resulted in a well-resolved PrOF NMR spectrum using 100–200 μM 5FW LecA at 310 K on a 700 MHz NMR machine equipped with a cryogenic probe (Supplementary Table SI).

Next, we assigned the fluorine resonances in PrOF NMR to four tryptophanes of LecA using site-directed mutagenesis (Fig. 2D). The disappearance of fluorine resonances in LecA mutants being replaced with phenylalanine gave indication for the assignment. As result, two peaks at −120.47 and −123.24 p.p.m. could be reliably identified as W42 and W84, respectively. Two resonances at −121.43 and −121.72 p.p.m. overlapped slightly indicating that both reside in a similar chemical environment in LecA, but could be assigned as W2 and W33, respectively.

Incorporation of 5FW into LecA does not abrogate protein activity
Because in some cases the affinity of proteins to the carbohydrates might be changed due to 5FW incorporation, as it has been shown for a stronger affinity of 5FW to β-D-Gal (Hudson et al. 2015), we

compared with the K_d obtained from PrOF NMR confirming that D-GalNAc is much weaker ligand compared with Ca^{2+} or D-Gal. Moreover, our affinity data in the FP assay for ligands, in particular D-Gal, were in a close range 1230 ± 200 μM and 1991 μM for both unlabeled LecA and 5FW LecA, respectively (Supplementary Table SIV). Cumulatively, this result suggests that the affinities for D-GalNAc derived from the FP assay for LecA and 5FW LecA diverged from PrOF NMR because of higher sensitivity of 19F NMR to spot weak binders and thus, thereby shows the advantages of PrOF NMR in discovery of weak interactions.

5FW LecA PrOF NMR is sensitive to probe glycomimetics

PrOF NMR with 5FW can be useful for discovery and design of ligands for LecA. For this, we performed PrOF NMR titrations of two glycomimetics: phenyl-Ph-β-D-Gal (Supplementary Fig S6) and pNPGal (Fig. 4) to 5FW LecA resulting in K_d of 166 ± 42 μM and 54 ± 6 μM, respectively. Moreover, p-nitrophenyl group improved binding affinity of D-Gal 6-fold, which is in agreement with previous reports (Rodrigue et al. 2013). This shows that 5FW in LecA can serve as sensitive probes to follow the affinity gain to design glycomimetics using structure-activity relationship approach (Divakaran et al. 2019).

Conclusions

We have shown that 5-fluoroindole can be used as a precursor of 5FW to label LecA for PrOF NMR studies. In our binding studies with Ca^{2+}, D-Gal and D-GalNAc, PrOF NMR has proven to detect and determine the affinity of moderate as well as weak ligands. In contrast to ligand-observed NMR techniques (e.g. STD NMR; Mayer and Meyer 2001) providing information on the epitope of ligand binding, PrOF NMR provides information on the ligand-binding site in the protein.

Further studies using ITC and FP assays have demonstrated that 5FW LecA preserved its activity and the ligand preference similarly to LecA. Notably, PrOF NMR has proven more sensitive for identification of weak ligands like D-GalNAc due to chance to observe the formation of a protein–ligand complex in NMR at earlier time point compared with the FP assay. Accordingly, these results represent the first studies demonstrating the potential of 5FW LecA PrOF NMR to assess binding of weak ligands. As tryptophan is by far the most frequently found amino acid in carbohydrate binding sites of various lectins (Taroni et al. 2000), this method could prove to be a valuable tool to assess binding of fragment- and drug-like molecules targeting the carbohydrate binding site of various lectins. Together, this approach will support the future drug-discovery campaigns that aim to develop drug-like inhibitors for lectins such as LecA.

Materials and methods

Fluorinated protein expression and purification

Recombinant 5FW LecA (WT and mutants) was expressed and purified as follows: *E. coli* BL21 (DE3) cells were transformed with pET25pa1l plasmid and grown in LB medium (100 μg mL^{-1} ampicillin) at 37°C with agitation (120 rpm) until OD$_{600}$ reached 0.6. 1 L of culture was harvested by centrifugation at $2500 \times g$, 10 min and resuspended in modified minimal M9 medium (Supplementary Table SIII). It was shaken at 37°C for 60 min as

a recovery time for bacteria followed by addition of 250 μL of 5-fluoroindole (Santa Cruz, USA; 240 mg/mL in dimethyl sulfoxide [DMSO]). Protein production was induced with 250 μM IPTG at 30°C and harvested in 4 h. Cell pellets were resuspended in buffer A (20 mM Tris-HCl pH 7.4, 137 mM NaCl, 2.6 mM KCl, 25 mM CaCl$_2$) supplemented with 1 mM PMSF and DNaseI (Applichem, Darmstadt, Germany). The cells were lysed by cell disruption (Branson Digital Sonifier) at 50% power 10 s on and 40 s off pulses following removal of cell debris by centrifugation ($10,000 \times g$, 30 min, 4°C). The supernatant was loaded onto a 2 mL Pierce™ D-Gal agarose column (Thermo Fisher Scientific) that was equilibrated with 3-fold column volume of buffer A. Bound LecA was eluted with buffer B (20 mM Tris-HCl pH 7.4, 137 mM NaCl, 2.6 mM KCl, 25 mM CaCl$_2$, 100 mM D-Gal). Protein was dialyzed in MilliQ water and TBS buffer (20 mM Tris-HCl pH 7.8, 100 mM NaCl) or MES buffer (25 mM MES pH 6, 40 mM NaCl) three times for 4 h and once overnight at 4°C, respectively. The protein solution was flash frozen and stored at -80°C.

Protein-observed fluorine (PrOF) NMR of 5FW LecA

All experiments were conducted on Bruker Ascend™700 (AvanceIII HD) spectrometer equipped with a 5 mm TCI700 CryoProbe™ in 3 mm tubes (Norell S-3-800-7) with following parameters: time domain of 1972, relaxation delay 1 s, acquisition time of 0.15 s, spectral width of 10 p.p.m. and 1024 scans resulting in measurement time of 20 minutes.

For optimization, PrOF NMR with 50, 100 or 200 μM holo 5FW LecA was recorded in TBS pH 7.8 or MES pH 6 with 10% D$_2$O, 2 mM CaCl$_2$ and 100 μM TFA at 285, 298 or 310 K. We considered only changes in CSP of peaks being 2-fold greater than standard deviation of fluorine resonance upon addition of 10 mM CaCl$_2$ or 1 mM D-Gal. All data analysis, plotting and curve fitting were performed with MestReNova 11.0.0 (Mestrelab Research SL, Santiago de Compostela, Spain). All spectra were referenced and normalized to trifluoroacetic acid (TFA) as internal reference at -75.6 p.p.m. after applying the Exponential function (30 Hz) and baseline correction.

The PrOF NMR titrations of D-Gal, D-GalNAc, Ph–β-D-Gal and pNPGal were performed with 100 μM 5FW LecA in TBS pH 7.8 at 310 K. For Ca^{2+} titration, 5FW LecA was dialyzed against Chelex®-100 in MES pH 6 buffer at 4°C overnight.

The decreasing intensity of the unbound W42 in 5FW LecA was followed to determine K_d values of ligands. Here, we used these values to normalize the changes in W42 peak intensities ($I_{\text{normalized}}$) following the equation (1) resulting in values plotted on Y-axis.

$$I_{\text{normalized}} = \frac{I_0 - I_{\text{measured}}}{I_0}, \qquad (1)$$

where I_0 was unbound W42 in the reference spectrum of protein only, I_{measured} was unbound W42 peak of protein with a ligand. The K_d values were calculated according to the one-site-binding model in GraphPad Prism 8 (GraphPad Software Inc., San Diego, CA) from three independent titrations.

Isothermal titration calorimetry

ITC was performed on a Microcal ITC200 (General Electric) at 25°C. Calcium ions were removed by extensive dialysis against 1 mM EDTA pH 7.4 ($\times$4) followed by 150 mM NaCl ($\times$4) and distilled water. The protein solution was lyophilized and the solid protein stored at

A

pNPGal

B

W42

TFA

100 µM
5FW LecA

pNPGal [µM]

-76   -120   -125

^{19}F [ppm]

C

K_d = 54 ± 6 µM
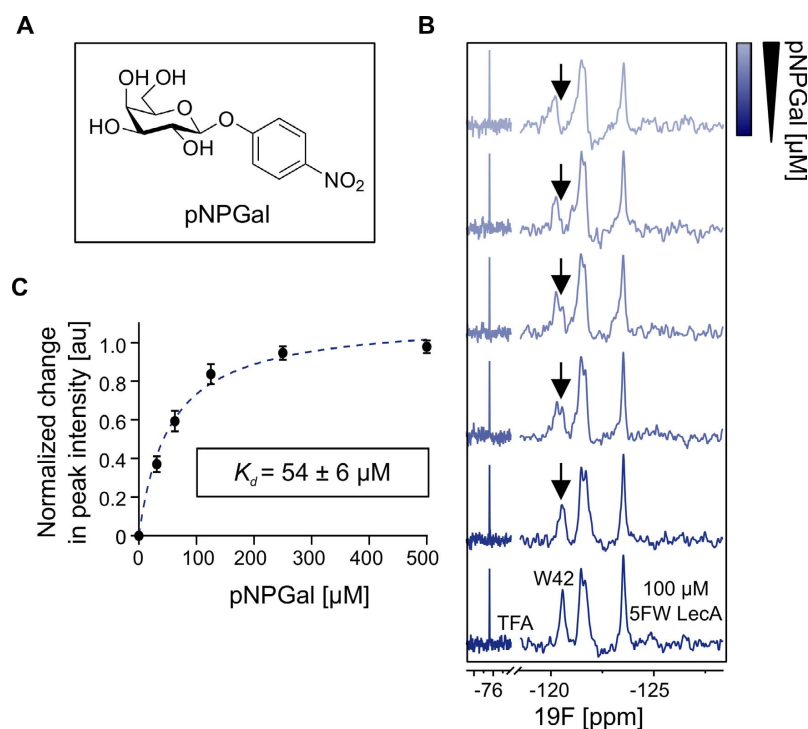
pNPGal [µM]

Fig. 4. PrOF NMR titration of a carbohydrate-based glycomimetic pNPGal to holo 5FW LecA. **(A)** The structure of pNPGal. **(B)** The PrOF NMR spectra of holo 5FW LecA (*bottom*) and titration of pNPGal (*upper*). The peak intensity of W42 resonance (*arrow*) decreased upon pNPGal addition. The change in signal intensity of free W42 peak can be followed to determine K_d. **(C)** Binding isotherm for pNPGal generated by plotting the normalized change in peak intensity of 5FW free W42 resonance as a function of ligand concentration. Data of three independent titrations were fitted to one-site-binding model to obtain K_d of 54 ± 6 µM.

−20°C. A solution of calcium chloride in TBS (20 mM Tris, 137 mM NaCl, 2.6 mM KCl at pH 7.4) was titrated into a calcium-free LecA solution in the same buffer. The data were analyzed according to the one-site-binding model using Microcal Origin software. Four independent titrations were performed using $CaCl_2$ and LecA between 2 and 3 mM and 170 and 200 µM, respectively.

Competitive binding fluorescence polarization assay

The competitive binding assay was performed as reported previously (Joachim et al. 2016). In total, 10 µL of LecA (40 µM) and *meta*-linked fluorescein-conjugate of phenyl-galactopyranoside (20 nM) in TBS pH 7.4 supplemented with 1 mM $CaCl_2$ (TBS/Ca^{2+}) and 10 µL of compound-dilution series (8 mM to 62 µM, 8% DMSO) in TBS/Ca^{2+} buffer were mixed in a 384-well plate (Greiner Bio-One, Germany) in three technical replicates. The sealed plate was centrifuged at 300 g for 1 min and incubated at room temperature with shaking. Fluorescence intensity was measured on a PheraStar FS microplate reader (BMG Labtech GmbH, Germany; ex. 485, em. 535 nm) after 1 and 16 h. Polarization was calculated and the data

were analyzed according to the four-parameter variable slope model (MARS Data Analysis Software, BMG Labtech GmbH, Germany), the top and bottom plateaus were determined from the control Me-α-D-Gal and data were reanalyzed with these values fixed.

Supplementary data

Supplementary data are available at *Glycobiology* online.

French National Research Agency [ANR-17-CE11–0048] project 'Glycomime'.

Conflict of interest statement

None declared.

References

Arntson KE, Pomerantz WC. 2016. Protein-observed fluorine NMR: A bioorthogonal approach for small molecule discovery. *J Med Chem*. 59(11):5158–5171.

Bergmann M, Michaud G, Visini R, Jin X, Gillon E, Stocker A, Imberty A, Darbre T, Reymond JL. 2016. Multivalency effects on Pseudomonas aeruginosa biofilm inhibition and dispersal by glycopeptide dendrimers targeting lectin LecA. *Org Biomol Chem*. 14(1):138–148.

Cecioni S, Imberty A, Vidal S. 2015. Glycomimetics versus multivalent glycoconjugates for the design of high affinity lectin ligands. *Chem Rev*. 115(1):525–561.

Chemani C, Imberty A, de Bentzmann S, Pierre M, Wimmerova M, Guery BP, Faure K. 2009. Role of LecA and LecB lectins in Pseudomonas aeruginosa-induced lung injury and effect of carbohydrate ligands. *Infect Immun*. 77(5):2065–2075.

Cioci G, Mitchell EP, Gautier C, Wimmerova M, Sudakevitz D, Perez S, Gilboa-Garber N, Imberty A. 2003. Structural basis of calcium and galactose recognition by the lectin PA-IL of Pseudomonas aeruginosa. *FEBS Lett*. 555(2):297–301.

Diggle SP, Stacey RE, Dodd C, Camara M, Williams P, Winzer K. 2006. The galactophilic lectin, LecA, contributes to biofilm development in Pseudomonas aeruginosa. *Environ Microbiol*. 8(6):1095–1104.

Divakaran A, Kirberger SE, Pomerantz WCK. 2019. SAR by (protein-observed) 19F NMR. *Acc Chem Res*. 52(12):3407–3418.

Garber N, Guempel U, Belz A, Gilboa-Garber N, Doyle RJ. 1992. On the specificity of the D-galactose-binding lectin (PA-I) of Pseudomonas aeruginosa and its strong binding to hydrophobic derivatives of D-galactose and thiogalactose. *Biochim Biophys Acta* 1116(3):331–333.

Gee CT, Arntson KE, Urick AK, Mishra NK, Hawk LM, Wisniewski AJ, Pomerantz WC. 2016. Protein-observed (19)F-NMR for fragment screening, affinity quantification and druggability assessment. *Nat Protoc*. 11(8):1414–1427.

Hudson KL, Bartlett GJ, Diehl RC, Agirre J, Gallagher T, Kiessling LL, Woolfson DN. 2015. Carbohydrate-aromatic interactions in proteins. *J Am Chem Soc*. 137(48):15152–15160.

Imberty A, Wimmerova M, Mitchell EP, Gilboa-Garber N. 2004. Structures of the lectins from Pseudomonas aeruginosa: Insight into the molecular basis for host glycan recognition. *Microbes Infect*. 6(2):221–228.

Joachim I, Rikker S, Hauck D, Ponader D, Boden S, Sommer R, Hartmann L, Titz A. 2016. Development and optimization of a competitive binding assay for the galactophilic low affinity lectin LecA from Pseudomonas aeruginosa. *Org Biomol Chem*. 14(33):7933–7948.

Kadam RU, Bergmann M, Hurley M, Garg D, Cacciarini M, Swiderska MA, Nativi C, Sattler M, Smyth AR, Williams P, et al. 2011. A glycopeptide dendrimer inhibitor of the galactose-specific lectin LecA and of Pseudomonas aeruginosa biofilms. *Angew Chem Int Ed Engl*. 50(45):10631–10635.

Kadam RU, Garg D, Schwartz J, Visini R, Sattler M, Stocker A, Darbre T, Reymond JL. 2013. CH-pi "T-shape" interaction with histidine explains binding of aromatic galactosides to Pseudomonas aeruginosa lectin LecA. *ACS Chem Biol*. 8(9):1925–1930.

Kitevski-LeBlanc JL, Prosser RS. 2012. Current applications of 19F NMR to studies of protein structure and dynamics. *Prog Nucl Magn Reson Spectrosc*. 62:1–33.

Liu JJ, Horst R, Katritch V, Stevens RC, Wuthrich K. 2012. Biased signaling pathways in beta2-adrenergic receptor characterized by 19F-NMR. *Science*. 335(6072):1106–1110.

Luck LA, Falke JJ. 1991. 19F NMR studies of the D-galactose chemosensory receptor. 2. Ca(II) binding yields a local structural change. *Biochemistry*. 30(17):4257–4261.

Mayer M, Meyer B. 2001. 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*. 123(25):6108–6117.

Renaud JP, Chung CW, Danielson UH, Egner U, Hennig M, Hubbard RE, Nar H. 2016. Biophysics in drug discovery: Impact, challenges and opportunities. *Nat Rev Drug Discov*. 15(10):679–698.

Rodrigue J, Ganne G, Blanchard B, Saucier C, Giguere D, Shiao TC, Varrot A, Imberty A, Roy R. 2013. Aromatic thioglycoside inhibitors against the virulence factor LecA from Pseudomonas aeruginosa. *Org Biomol Chem*. 11(40):6906–6918.

Sharaf NG, Gronenborn AM. 2015. (19)F-modified proteins and (19)F-containing ligands as tools in solution NMR studies of protein interactions. *Methods Enzymol*. 565:67–95.

Singh PK, Schaefer AL, Parsek MR, Moninger TO, Welsh MJ, Greenberg EP. 2000. Quorum-sensing signals indicate that cystic fibrosis lungs are infected with bacterial biofilms. *Nature*. 407(6805):762–764.

Taroni C, Jones S, Thornton JM. 2000. Analysis and prediction of carbohydrate binding sites. *Protein Eng*. 13(2):89–98.

Tobola F, Lelimousin M, Varrot A, Gillon E, Darnhofer B, Blixt O, Birner-Gruenberger R, Imberty A, Wiltschi B. 2018. Effect of noncanonical amino acids on protein-carbohydrate interactions: Structure, dynamics, and carbohydrate affinity of a Lectin engineered with fluorinated tryptophan analogs. *ACS Chem Biol*. 13(8):2211–2219.

Wagner S, Hauck D, Hoffmann M, Sommer R, Joachim I, Müller R, Imberty A, Varrot A, Titz A. 2017. Covalent Lectin inhibition and application in bacterial biofilm imaging. *Angew Chem Int Ed Engl*. 56(52):16559–16564.

Wagner S, Sommer R, Hinsberger S, Lu C, Hartmann RW, Empting M, Titz A. 2016. Novel strategies for the treatment of Pseudomonas aeruginosa infections. *J Med Chem*. 59(13):5929–5969.

4.2. Automated glycan assembly of ^{19}F labeled glycan probes enables high-throughput NMR studies of protein-glycan interactions

Giulio Fittolani*, **Elena Shanina***, Mónica Guberman, Peter H. Seeberger, Christoph Rademacher, Martina Delbianco

- *Shared first authorship*

Protein NMR Spectroscopy | **Hot Paper**

Automated Glycan Assembly of 19F-labeled Glycan Probes Enables High-Throughput NMR Studies of Protein–Glycan Interactions

Giulio Fittolani+, Elena Shanina+, Mónica Guberman, Peter H. Seeberger, Christoph Rademacher,* and Martina Delbianco*

Abstract: *Protein–glycan interactions mediate important biological processes, including pathogen host invasion and cellular communication. Herein, we showcase an expedite approach that integrates automated glycan assembly (AGA) of 19F-labeled probes and high-throughput NMR methods, enabling the study of protein–glycan interactions. Synthetic Lewis type 2 antigens were screened against seven glycan binding proteins (GBPs), including DC-SIGN and BambL, respectively involved in HIV-1 and lung infections in immunocompromised patients, confirming the preference for fucosylated glycans (Lex, H type 2, Ley). Previously unknown glycan–lectin weak interactions were detected, and thermodynamic data were obtained. Enzymatic reactions were monitored in real-time, delivering kinetic parameters. These results demonstrate the utility of AGA combined with 19F NMR for the discovery and characterization of glycan–protein interactions, opening up new perspectives for 19F-labeled complex glycans.*

Introduction

Glycans are a highly diverse class of biomolecules involved in several processes such as cellular communication and recognition and play important structural and modulatory roles.[1] Pathogens invade the host by mimicking or exploiting host glycans present on endothelial cells. This process is often mediated by lectins, a class of glycan-binding proteins (GBPs) expressed by both pathogens and hosts. Typically, mammalian glycans have low affinity for mammalian receptors, while showing higher affinity for bacterial lectins.[2] Profiling glycan–lectin interactions is a crucial step towards the understanding of the biological functions of glycans. Still, the extreme complexity and diversity of glycans pose a severe bottleneck to the characterization of these generally weak and promiscuous interactions.

Synthetic glycans are valuable probes to dissect glycan–protein interactions. However, lengthy synthetic protocols hampered their systematic and widespread use in glycobiology. Automated glycan assembly (AGA) enables fast access to complex and well-defined glycans.[3,4] With AGA, glycans are typically assembled in an overnight run, permitting the production of broad collections of glycans for systematic screenings.[5]

An additional challenge to the study of glycan–protein interactions is the need for highly sensitive methods able to detect the often inherently low affinities. Several analytical techniques have been developed to quantitatively describe these interactions at the molecular level and in a high-throughput manner.[6–8] Most of these strategies rely on immobilized glycans (e.g. microarray technology)[6–10] or require large amounts of samples and analysis time (ITC,[11] SPR,[12] or X-ray crystallography[13]). In contrast, NMR allows for the detection of protein–glycan interactions in solution in a fast and reliable manner, providing information on the binding mode in a homogeneous assay format in absence of immobilization protocols.[14,15]

NMR active labels are commonly introduced to simplify NMR analysis.[15,17] Among all, the 19F nucleus stands out due to its unique properties such as: i) high sensitivity to local chemical environment, ii) short acquisition times, iii) simple spectra, iv) broad chemical shift range, and v) absence in

[*] G. Fittolani,[+] E. Shanina,[+] Dr. M. Guberman, Prof. P. H. Seeberger, Prof. C. Rademacher, Dr. M. Delbianco
Department of Biomolecular Systems
Max Planck Institute of Colloids and Interfaces
Am Mühlenberg 1, 14476 Potsdam (Germany)
E-mail: christoph.rademacher@univie.ac.at
martina.delbianco@mpikg.mpg.de

G. Fittolani,[+] E. Shanina,[+] Prof. P. H. Seeberger, Prof. C. Rademacher
Department of Chemistry and Biochemistry
Freie Universität Berlin
Arnimallee 22, 14195 Berlin (Germany)

Prof. C. Rademacher
Current address: Department of Pharmaceutical Chemistry
University of Vienna
Althanstrasse 14, 1080 Vienna (Austria),
and
Current address: Department of Microbiology

Immunobiology and Genetics, Max F. Perutz Labs
Campus Vienna Biocenter 5, 1030 Vienna (Austria)

Dr. M. Guberman
Current address: Medicinal Chemistry
Leibniz-Forschungsinstitut für Molekulare Pharmakologie
Robert-Rössle Strasse 10, 13125 Berlin (Germany)

[+] These authors contributed equally to this work.

Supporting information and the ORCID identification number(s) for the author(s) of this article can be found under: https://doi.org/10.1002/anie.202102690.

biological systems (no background signal).[18,19] Even though ^{19}F NMR has enabled the description of peptide (mis)folding, real-time in vivo events,[18–22] protein–ligand interactions, and high-throughput ligand screening,[23,24] the use of fluorinated glycans to investigate protein binding[25] and enzymatic reactions[26–28] is just at the beginning. The labor-intensive multistep synthesis of ^{19}F-labeled glycans represents the main bottleneck and has limited these studies to small collections of short and relatively simple glycans.[14,29–32] Still, ^{19}F-labeled glycans have the potential to dissect protein–glycan interactions.[33,34]

Herein, we present a high-throughput NMR-based approach for the screening and characterization of protein–glycan interactions using ^{19}F-labeled glycans. AGA enabled quick access to a collection of ^{19}F-labeled Lewis type 2 complex glycans. Lewis type 2 antigens are involved in several physiological and pathological processes, including cancer, where they act as cell adhesion or recognition mediators.[35,36] Subtle differences in the fucosylation pattern strongly impact their interaction with proteins and ultimately can lead to host immune system elusion.[37–40] The ^{19}F-labeled glycan probes (hereafter F-glycans) were screened against mammalian and bacterial lectins as well as enzymes. Among mammalian lectins, we selected Langerin[41] and the dendritic cell specific ICAM-3 grabbing non-integrin (DC-SIGN)[42] both of which are known to bind high-mannose N-glycans. DC-SIGN also selectively recognizes specific fucosylated glycans,[43] playing a crucial role in the biology of viral pathogens (e.g. HIV). In addition, we screened soluble lectins produced by some opportunistic pathogens responsible for lung infections, such as *Pseudomonas* (LecA and LecB)[44] and *Burkholderia*

(BambL)[45] species. Finally, we selected two different sialyl-transferases and screened their interactions with Lewis antigens, given the importance and widespread occurrence of terminal sialylation in Lewis antigens.[46,47] The labeled glycan probes in combination with ^{19}F NMR proved to be valuable for detecting binding events in real-time, identifying new weak protein–glycan interactions, and determining affinities (K_d) as well as kinetics of enzymatic reactions.

Results and Discussion

Automated Synthesis of F-Glycans

Recently, an elegant procedure to access a collection of Lewis type 2 antigens by AGA was reported.[48] We envisioned a similar approach to produce a set of ^{19}F-labeled analogs to screen protein binding in a simple ^{19}F NMR assay. Since the position of the ^{19}F reporter is thought to be crucial to obtain valuable information, F-glycans (**F-Lac**, **F-nLac₄**, **F-Leˣ**, **F-H type 2**, and **F-Leʸ**) were designed with the ^{19}F reporter in the lactose inner core subunit (Figure 1 A). This position is distal from the binding site (i.e. non reducing end) to minimize the effect of the fluorine atom during the binding event.[49,50] We hypothesize that labeling of the inner core glucose unit should maintain sensitivity to the binding event due to overall changes in the correlation time of the glycan in the bound state, reporting changes in the ^{19}F NMR signal.[20]

[19]F-labeled analogs of Lewis type 2 antigens were assembled on a solid support (functionalized Merrifield resin, **L1**)
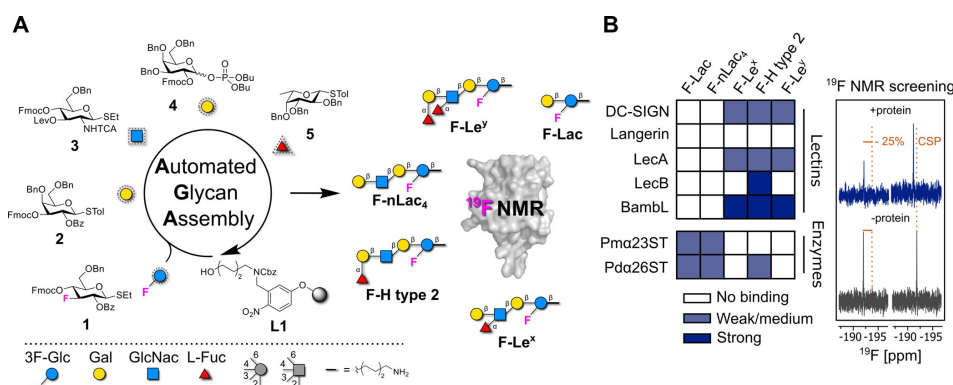
Figure 1. Integrated approach for the preparation of ^{19}F-labeled Lewis type 2 glycans by AGA and screening against lectins and enzymes. A) BBs **1**–**5**, including BB **1** bearing the ^{19}F reporter, were employed for the AGA of a collection ^{19}F-labeled Lewis type 2 antigen analogs represented following the Symbol Nomenclature For Glycans (SNFG).[16] B) The F-glycans were screened against proteins, including mammalian and bacterial lectins, as well as enzymes. The enzymes were screened in the absence of donor (i.e. CMP-Neu5Ac) to probe binding to the substrate. The binding strength was defined depending on the changes observed in the NMR after addition of the protein (right panel). Strong binding (blue) is defined as a decrease in peak intensity higher than −25% or a chemical shift perturbation (CSP) bigger than 0.01 ppm in the ^{19}F NMR. Weak/medium binding (light blue) is defined as a decrease in peak intensity higher than −25% in the CPMG-filtered ^{19}F NMR. No binding (white) is defined as a decrease in peak intensity lower than −25% in CPMG-filtered ^{19}F NMR.

using building blocks (BBs) **1–5** (Figure 1A). The BBs are equipped with a thioether or a dibutylphosphate reactive leaving group. Orthogonal cleavage of the 9-fluorenylmethoxycarbonyl (Fmoc) and levulinoyl (Lev) temporary protecting groups permits regioselective chain elongation. Benzyl (Bn), benzoyl (Bz), and *N*-trichloroacetyl (TCA) groups protect the remaining functionalities. β-Stereoselectivity during glycosylation with BBs **1–4** is ensured by anchimeric assistance of the protecting groups at C-2, while α-stereoselectivity with BB **5** was verified in previous studies.[48] BB **1** is labeled with the ^{19}F reporter at the C-3 position.[51] Each oligosaccharide was assembled in an overnight run following previously reported conditions for unlabeled analogs (see SI).[48] Post-AGA manipulations included solid-phase methanolysis,[51] photocleavage[52] from the solid support, and hydrogenolysis (see SI). A single final purification step afforded the target F-glycans in overall yields of 5% to 16% over 7 to 15 steps.

^{19}F NMR Screening of F-Glycan Library

A ^{19}F and CPMG NMR screening was performed to probe the interactions of five F-glycans (**F-Lac**, **F-n-Lac₄**, **F-Leˣ**, **F-H type 2**, and **F-Leʸ**) with mammalian (Langerin, DC-SIGN) and bacterial (LecA, LecB, BambL) lectins and enzymes (α(2,3)-sialyltransferase from *Pasteurella multocida* (Pmα23ST)[53] and α(2,6)-sialyltransferase from *Photobacterium damsela* (Pdα26ST)[54]) (Figure 1B). Upon protein binding, the molecular tumbling rate of the glycan is drastically affected resulting in a decrease of the ^{19}F signal intensity.[20] Monitoring ^{19}F chemical shift perturbation (CSP) or change in peak intensity upon addition of protein allowed us to qualitatively evaluate the strength of the interaction. A decrease in peak intensity or a CSP in ^{19}F NMR indicates strong binding. Application of a CPMG-based spin echo filter allows us to detect weak binders. As a result, bacterial (LecA, LecB, and BambL) and mammalian (DC-SIGN ECD) lectins preferred fucosylated glycans (Figures S2A, S2B, S2C, and S2E). No binding to F-glycans was observed in presence of Langerin ECD (Figure S2D), in agreement with previous reports.[55] In contrast, the enzymes showed much weaker interactions and a slight preference for shorter non-branched glycans (Figure S3).

Reporter Position on F-Glycans Does Not Affect Binding to Mammalian and Bacterial Lectins

DC-SIGN recognizes cellular ligands and pathogens that express Lewis antigens. In particular, Leˣ and Leʸ present on *Schistosoma mansoni*[56] and *Helicobacter pylori*[43] or endothelial cells,[57] respectively, are known binding partners for DC-SIGN.[58] The strong preference of DC-SIGN for fucosylated ligands has also been elucidated with the crystal structure of the carbohydrate-binding site of DC-SIGN bound to Leˣ.[59] The qualitative CPMG NMR screening of mammalian lectins confirmed the interaction of DC-SIGN with fucosylated glycans **F-Leˣ**, **F-H type 2**, and **F-Leʸ** (Fig-
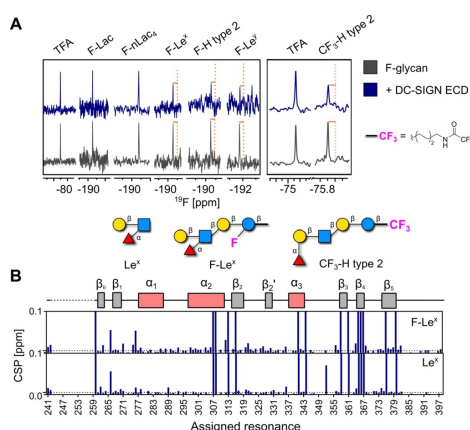
Figure 2. Mammalian lectin (DC-SIGN) binding to F-glycans and study on the reporter position. A) CPMG NMR screening of F-glycans alone (gray) and in presence of DC-SIGN ECD (blue). DC-SIGN ECD binds to **F-Leˣ**, **F-H type 2**, and **F-Leʸ** as shown by a decrease in peak intensity in presence of protein (orange lines, left panel). CPMG NMR spectra of **CF₃-H type 2** alone (gray) and in presence of DC-SIGN ECD (blue; right panel). B) Cartoon of assigned domains of DC-SIGN CRD (unassigned resonances in dashed line) and CSP plot of assigned resonances in presence of **F-Leˣ** and Leˣ showing that **F-Leˣ**-perturbed resonances similarly to unlabeled **Leˣ**.

ure 2A), as indicated by changes in the NMR peak intensity of the reporter molecule. This effect is maximized with a protein-to-ligand ratio of 2:1 (Figure S4A).

First, we explored the role of the ^{19}F reporter in F-glycan binding to DC-SIGN. We performed protein-observed ^{15}N HSQC NMR and recorded an HSQC NMR spectrum of DC-SIGN CRD in the presence of **F-Leˣ** and Leˣ. Both ligands promoted similar changes in the backbone of DC-SIGN CRD (Figure 2B and S4B). Next, we investigated the effect of the reporter's position on the ability to reveal binding events. We conjugated a CF₃ moiety to the remote end of the aminopentyl linker on **H type 2** (**CF₃-H type 2**), far from the carbohydrate-binding site, and tested the new ligand in ^{19}F and CPMG NMR. Remarkably, its binding was observed with both mammalian (DC-SIGN, Figure 2A) and bacterial lectins (BambL, Figure S5). These results indicate that the positioning of the ^{19}F reporter on the Glc unit does not affect the binding of F-glycans with proteins. Furthermore, the ^{19}F reporter can be remote to the glycan binding site to avoid any interference with the binding event, while preserving excellent sensitivity. However, the functionalization of the amino linker with a CF₃ moiety prevents any further conjugation of the glycan (e.g. to protein, surface, liposome).

We further investigated the interactions of DC-SIGN CRD with **F-Leʸ** and **F-H type 2** in ^{15}N HSQC NMR (Figure 3A and S6A). Even though Leʸ is known for its interaction with DC-SIGN, structural data are lacking.[57]
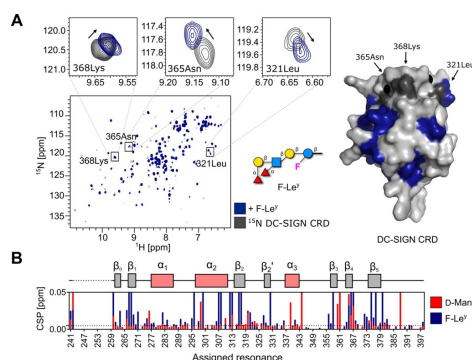
Figure 3. Mammalian lectin (DC-SIGN) binding to **F-Le^y**.
A) HSQC NMR (left) shows the interaction of **F-Le^y** with ^15N-labeled DC-SIGN CRD and the perturbed residues were mapped on a structure of DC-SIGN CRD (blue). Surface diagram of the crystal structure of DC-SIGN CRD (PDB: 1sl4; right). **F-Le^y** targets the carbohydrate-binding site of DC-SIGN CRD based on changes in resonances (e.g. 321Leu, 365Asn and 368Lys, gray). B) Cartoon of assigned domains of DC-SIGN CRD (unassigned resonances in dashed line) and CSP plot showing that **F-Le^y**-perturbed resonances similarly to D-mannose (red, positive control). The magnitude of **F-Le^y**-promoted CSPs is higher compared to D-mannose. CSPs exceeding the threshold (dashed line at 0.005 ppm) and intensities decreasing by more than 50% were used for mapping the binding site of **F-Le^y** on a structure of DC-SIGN CRD.

Both ligands promoted CSPs of the residues located in the carbohydrate-binding and remote sites of DC-SIGN CRD. Binding to **F-Le^y** promoted larger changes in DC-SIGN CRD than **F-H type 2** or the monosaccharide positive control D-mannose (Figure 3B and S6B). This result proved that the avidity effect plays a crucial role in the interactions between DC-SIGN and Lewis type 2 antigens, as similarly noted for high-mannose structures.[60] The CSPs observed in remote parts of the protein suggest allosteric binding, a known mechanism for C-type lectins such as DC-SIGN.[61–63] Cumulatively, we believe these probes are valuable tools for the description of the interaction mechanisms between DC-SIGN and fucosylated blood antigens.

Binding Affinity of F-Glycans to Bacterial Lectins

Bacterial lectins show a remarkably high affinity for fucosylated blood group antigens.[35,64] The interaction of BambL from *Burkholderia ambifaria* with **H type 2** has been thoroughly investigated and two binding sites were identified in a crystal structure of the complex (Figure 4A).[35] We set out to verify this interaction for F-glycans in ^19F and protein-observed NMR.

First, we performed ^19F NMR screening and titration experiments with fucosylated F-glycans. ^19F NMR experiments allowed us to confirm the interaction and obtain

affinity constants for **F-H type 2** ($K_d = 9 \pm 2$ μM, Figure 4B and 3C) and **F-Le^y** ($K_d = 14 \pm 2$ μM, Figure S7A). Given that BambL has two binding sites available for glycan binding, we applied one- and two-binding site models to derive the affinities for both sites. Both models resulted in matching K_d values, in agreement with values reported by ITC.[35] Even though we did not observe a difference in the affinities between the two sites in ^19F NMR, we showed that ^19F NMR can be applied reliably to derive affinities while considerably reducing the amount of ligand needed for ITC.

We verified the interaction of **F-H type 2** (Figure 4D) and **F-Le^y** (Figure S7B) with ^15N-labeled BambL in protein-observed ^15N TROSY NMR. Changes in protein backbone similar to the one obtained with α-Me-L-fucose indicate that the α-L-fucose branch was mainly responsible for the binding (Figures S7C and 4D). To derive affinities, we titrated both ligands and followed the changes in peak intensities and CSPs for the peaks in slow (**F-H type 2**: $K_d = 12 \pm 8$ μM, Figure 4F and **F-Le^y**: $K_d = 17 \pm 3$ μM, Figure S7D), and fast (**F-H type 2**: $K_d = 94 \pm 33$ μM, Figure 4G and **F-Le^y**: $K_d = 245 \pm 29$ μM, Figure S7E) exchange regimes, respectively. However, protein-observed NMR is not well suitable for the determination of K_d for ligands with high affinities and thus, it hampered the accurate derivation of the K_d.[65] This underscores the advantage of the ^19F NMR ligand-observed approach.

In addition to the known strong interactions of LecB and BambL with fucosylated glycans,[66] CPMG NMR screening revealed weak interactions between LecA and fucosylated F-glycans. To confirm this observation, we performed ^19F R_2-filtered, protein-observed ^19F (PrOF) and ^15N TROSY NMR experiments. **F-H type 2** showed a faster relaxation in presence of protein, indicating a weak interaction with LecA (Figure S8B). Protein-observed NMR experiments with 5-fluorotryptophan (5FW, Figure S8C) and ^15N-labeled LecA (Figure S8D and S8E) confirmed that this interaction takes place in the canonical carbohydrate-binding site of LecA, as indicated by perturbation of W42 and CSPs promoted in a similar manner to D-galactose, respectively. To the best of our knowledge this is the first report of such weak binding detected using a biophysical method.[67,68] These results demonstrate that F-glycans serve as probes for the affinity determination and discovery of new interactions using low amounts of protein and ligand.

Enzyme Binding and Real-Time Kinetics with F-Glycans

The ^19F NMR assay allowed us to monitor the binding of F-glycans (**F-Lac** and **F-nLac_4**) to enzymes. Two sialyltransferases (Pmα23ST[53] and Pdα26ST[54]) were screened in the absence of donor (i.e. CMP-Neu5Ac) and revealed weak binding to the glycan substrate (Figure 1B and S3). This is particularly relevant because binding sites of transferases usually have a very low affinity for the acceptors, making these interactions difficult to detect. Shorter non-branched glycans (**F-Lac** and **F-nLac_4**) showed stronger binding than longer branched ones. **F-Le^x** did not show any binding with Pmα23ST or Pdα26ST, matching its known poor reactivity as acceptor (Figure S3).[69] In contrast, Pdα26ST showed weak
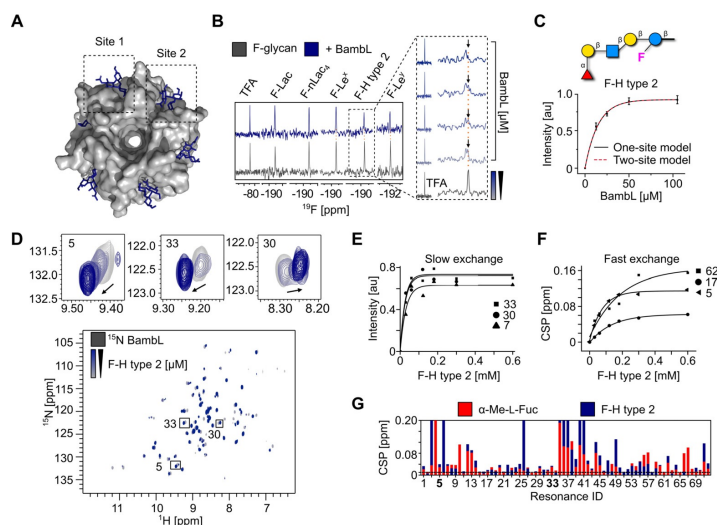
Figure 4. Bacterial lectin (BambL) binding to F-glycans. A) Surface diagram of the crystal structure of BambL in complex with **H-F type 2** (PDB: 3zzv). Sites 1 and 2 correspond to the carbohydrate-binding sites within a monomer and between two monomers, respectively. B) ^{19}F NMR screening of F-glycans alone (gray) and in presence of BambL (blue). BambL binds **F-Le^x**, **F-Le^y**, and **F-H type 2** strongly as shown by CSP in presence of protein (orange line). The ^{19}F NMR titration spectra shows **F-H type 2** undergoing slow exchange on the chemical shift timescale upon increase of BambL concentration. C) The K_d of **F-H type 2** was calculated from the changes in peak intensity and fitted to one- and two-site models resulting in a K_d of 9 ± 2 μm.
D) TROSY NMR verified **F-H type 2** binding to ^{15}N-labeled BambL. Given that BambL has two binding sites, peaks showing a slow (30, 7, and 33), intermediate and fast exchange (5, 17, and 62) on the chemical shift timescale have been observed upon titration of **F-H type 2**. One-site model for slow (E) and fast exchange (F) peaks was applied to derive the K_d values of 12 ± 8 μm and 94 ± 33 μm, respectively. G) CSP plot showing the resonances perturbed in presence of α-Me-ʟ-fucose and **F-H type 2**.

binding to **F-H type 2**, in agreement with previously reported enzymatic activity (Figure S3).[54] This simple assay could be envisioned as screening platform to identify acceptor substrates for known enzymes and for the discovery of new glycosyltransferases.[70,71]

The high sensitivity of the ^{19}F reporter to subtle modifications in its chemical environment offers a valuable tool for real-time monitoring of enzymatic reactions. The possibility to place the ^{19}F reporter on a carbohydrate unit in proximity to the functionalization site is crucial for detecting a chemical shift perturbation. We selected two enzymes (β-galactosidase[72] and Pmα23ST[53]) and we monitored their activity on a model substrate, **F-Lac**. Glycosidic bond cleavage, mediated by β-galactosidase, was followed by ^{19}F NMR. Cleavage of the terminal β-galactose induced a chemical shift perturbation and real-time ^{19}F NMR tracking allowed for derivation of the K_M of the enzymatic reaction (Figure 5A). Next, glycosidic bond formation promoted by Pmα23ST[53] was monitored in real-time. *N*-Acetyl-neuraminic acid (Neu5Ac) is transferred from an activated cytidine monophosphate donor (CMP-Neu5Ac) to the C-3 OH of the terminal galactose unit of **F-Lac** to yield **F-sLac**. The electron-withdrawing nature of

Neu5Ac induced a chemical shift perturbation of 0.2 ppm on the ^{19}F-labeled acceptor, allowing to track in real-time the enzymatic sialylation process (Figure 5B). When the ^{19}F reporter was positioned remotely to the reactive site of the acceptor (>3 sugar units away, **F-nLac$_4$**), no chemical shift perturbation was noticed, despite the success of the enzymatic transformation (Figure S10). Thus, in contrast to what is observed for protein binding, the position of the ^{19}F reporter is key for monitoring enzymatic reactions.

Conclusion

AGA enabled the fast assembly of ^{19}F-labeled Lewis type 2 antigens for the high-throughput screening of protein binding. Mammalian and bacterial lectins as well as enzymes were analyzed. ^{19}F NMR screening of F-glycans permitted a quick qualitative evaluation as well as a reliable quantification of lectin binding (K_d). The assay does not require labeled proteins or complex 2D NMR experiments. All NMR experiments can be performed in an extremely small scale (few nmol of glycan and protein per experiment). Enzymatic reactions, including sialylation, were monitored in real-time, demonstrating that ^{19}F-labeled glycans hold a great potential as molecular probes to uncover enzymatic processes and for high-throughput screening.[27] Protocols for the selective ^{19}F-labeling of monosaccharides are available;[73–75] the implementation of these novel BBs in AGA will fuel the production of new classes of glycan probes. Given the high dispersion of ^{19}F NMR signals, libraries of F-glycans with diverse chemical shifts can be designed to increase the high throughput of this approach.[76] The ability of ^{19}F glycan probes to reveal binding or enzymatic transformation in solution and in real-time could open the way to in cell NMR applications, often hampered by high background signals.[14,77,78] Overall, these probes are valuable tools for a better molecular understanding of the interactions of complex glycans with protein receptors.

GDCh | *Research Articles* | **Angewandte** *International Edition* *Chemie*
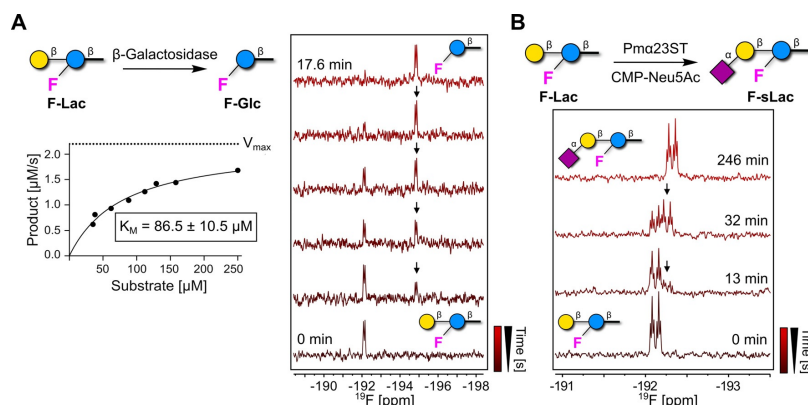
Figure 5. Real-time enzyme kinetics by ^{19}F NMR using F-glycans. A) ^{19}F NMR of **F-Lac** incubated with β-galactosidase. ^{19}F NMR real-time tracking of product formation (black arrows) upon incubation of **F-Lac** with β-galactosidase (right). Kinetic data were derived plotting the product formation rate as a function of the substrate concentration. The best fit of the experimental data provides a K_M value of 86.5 ± 10.5 μM according to the Henry-Michaelis–Menten equation (left). B) ^{19}F NMR of **F-Lac** incubated with Pmα23ST in presence of CMP-Neu5Ac. The formation of **F-sLac** (black arrows) can be followed by ^{19}F NMR in real-time. Product formation was confirmed by HPLC (Figure S9).

Conflict of interest

The authors declare no conflict of interest.

[1] A. Varki, R. D. Cummings, J. D. Esko, P. Stanley, G. W. Hart, M. Aebi, A. G. Darvill, T. Kinoshita, N. H. Packer, J. H. Prestegard, R. L. Schnaar, P. H. Seeberger, *Essentials of Glycobiology*, 3rd Ed., CSH Press, New York, **2017**.

[2] P. Gagneux, A. Varki, *Glycobiology* **1999**, *9*, 747–755.

[3] M. Panza, S. G. Pistorio, K. J. Stine, A. V. Demchenko, *Chem. Rev.* **2018**, *118*, 8105–8150.

[4] M. Guberman, P. H. Seeberger, *J. Am. Chem. Soc.* **2019**, *141*, 5581–5592.

[5] M. Delbianco, A. Kononov, A. Poveda, Y. Yu, T. Diercks, J. Jiménez-Barbero, P. H. Seeberger, *J. Am. Chem. Soc.* **2018**, *140*, 5421–5426.

[6] C. D. Rillahan, J. C. Paulson, *Annu. Rev. Biochem.* **2011**, *80*, 797–823.

[7] J. Stevens, O. Blixt, J. C. Paulson, I. A. Wilson, *Nat. Rev. Microbiol.* **2006**, *4*, 857–864.

[8] M. Mende, V. Bordoni, A. Tsouka, F. F. Loeffler, M. Delbianco, P. H. Seeberger, *Faraday Discuss.* **2019**, *219*, 9–32.

[9] D. J. Valles, Y. Naeem, A. Y. Rozenfeld, R. W. Aldasooky, A. M. Wong, C. Carbonell, D. R. Mootoo, A. B. Braunschweig, *Faraday Discuss.* **2019**, *219*, 77–89.

[10] C.-H. Liang, S.-K. Wang, C.-W. Lin, C.-C. Wang, C.-H. Wong, C.-Y. Wu, *Angew. Chem. Int. Ed.* **2011**, *50*, 1608–1612; *Angew. Chem.* **2011**, *123*, 1646–1650.

[11] T. K. Dam, C. F. Brewer, *Chem. Rev.* **2002**, *102*, 387–430.

[12] E. A. Smith, W. D. Thomas, L. L. Kiessling, R. M. Corn, *J. Am. Chem. Soc.* **2003**, *125*, 6140–6148.

[13] E. J. Toone, *Curr. Opin. Struct. Biol.* **1994**, *4*, 719–728.

[14] P. Valverde, J. I. Quintana, J. I. Santos, A. Ardá, J. Jiménez-Barbero, *ACS Omega* **2019**, *4*, 13618–13630.

[15] A. Gimeno, P. Valverde, A. Ardá, J. Jiménez-Barbero, *Curr. Opin. Struct. Biol.* **2020**, *62*, 22–30.

[16] A. Varki, R. D. Cummings, M. Aebi, N. H. Packer, P. H. Seeberger, J. D. Esko, P. Stanley, G. Hart, A. Darvill, T. Kinoshita, et al., *Glycobiology* **2015**, *25*, 1323–1324.

[17] M. D. Battistel, H. F. Azurmendi, B. Yu, D. I. Freedberg, *Prog. Nucl. Magn. Reson. Spectrosc.* **2014**, *79*, 48–68.

[18] S. L. Cobb, C. D. Murphy, *J. Fluorine Chem.* **2009**, *130*, 132–143.

[19] H. Chen, S. Viel, F. Ziarelli, L. Peng, *Chem. Soc. Rev.* **2013**, *42*, 7971.

[20] C. Dalvit, *Prog. Nucl. Magn. Reson. Spectrosc.* **2007**, *51*, 243–271.

[21] S. Tsukiji, M. Miyagawa, Y. Takaoka, T. Tamura, I. Hamachi, *Nat. Chem. Biol.* **2009**, *5*, 341–343.

[22] E. N. G. Marsh, Y. Suzuki, *ACS Chem. Biol.* **2014**, *9*, 1242–1250.

[23] M. Garavís, B. López-Méndez, A. Somoza, J. Oyarzabal, C. Dalvit, A. Villasante, R. Campos-Olivas, C. González, *ACS Chem. Biol.* **2014**, *9*, 1559–1566.

[24] N. S. Troelsen, E. Shanina, D. Gonzalez-Romero, D. Danková, I. S. A. Jensen, K. J. Śniady, F. Nami, H. Zhang, C. Rademacher, A. Cuenda, et al., *Angew. Chem. Int. Ed.* **2020**, *59*, 2204–2210; *Angew. Chem.* **2020**, *132*, 2224–2230.

[25] B. Linclau, A. Ardá, N.-C. Reichardt, M. Sollogoub, L. Unione, S. P. Vincent, J. Jiménez-Barbero, *Chem. Soc. Rev.* **2020**, *49*, 3863–3888.

GDCh

Research Articles

Angewandte
International Edition *Chemie*

[26] T. J. Kieser, N. Santschi, L. Nowack, A. Axer, G. Kehr, S. Albrecht, R. Gilmour, *ACS Chem. Neurosci.* **2020**, *11*, 2129–2136.

[27] T. Keenan, F. Parmeggiani, J. Malassis, C. Q. Fontenelle, J.-B. Vendeville, W. Offen, P. Both, K. Huang, A. Marchesi, A. Heyam, et al., *Cell Chem. Biol.* **2020**, *27*, 1199.

[28] C. E. Council, K. J. Kilpin, J. S. Gusthart, S. A. Allman, B. Linclau, S. S. Lee, *Org. Biomol. Chem.* **2020**, *18*, 3423–3451.

[29] A. Ardá, J. Jiménez-Barbero, *Chem. Commun.* **2018**, *54*, 4761–4769.

[30] S. A. Allman, H. H. Jensen, B. Vijayakrishnan, J. A. Garnett, E. Leon, Y. Liu, D. C. Anthony, N. R. Sibson, T. Feizi, S. Matthews, et al., *ChemBioChem* **2009**, *10*, 2522–2529.

[31] S.-J. Richards, M. I. Gibson, M. A. Fascione, B. Linclau, M. C. Galan, A. N. Baker, H. Ledru, C. S. Webster, C. E. Council, M. Walker, et al., *Chem. Sci.* **2021**, *12*, 905–910.

[32] P. Valverde, S. Delgado, J. D. Martínez, J.-B. Vendeville, J. Malassis, B. Linclau, N.-C. Reichardt, F. J. Cañada, J. Jiménez-Barbero, A. Ardá, *ACS Chem. Biol.* **2019**, *14*, 1660–1671.

[33] E. Matei, A. André, A. Glinschert, A. S. Infantino, S. Oscarson, H.-J. Gabius, A. M. Gronenborn, *Chem. Eur. J.* **2013**, *19*, 5364–5374.

[34] L. P. Calle, B. Echeverria, A. Franconetti, S. Serna, M. C. Fernández-Alonso, T. Diercks, F. J. Cañada, A. Ardá, N.-C. Reichardt, J. Jiménez-Barbero, *Chem. Eur. J.* **2015**, *21*, 11408–11416.

[35] A. Audfray, A. Varrot, A. Imberty, *C. R. Chim.* **2013**, *16*, 482–490.

[36] J. E. Heggelund, A. Varrot, A. Imberty, U. Krengel, *Curr. Opin. Struct. Biol.* **2017**, *44*, 190–200.

[37] A. Gimeno, S. Delgado, P. Valverde, S. Bertuzzi, M. A. Berbís, J. Echavarren, A. Lacetera, S. Martín-Santamaría, A. Surolia, F. J. Cañada, et al., *Angew. Chem. Int. Ed.* **2019**, *58*, 7268–7272; *Angew. Chem.* **2019**, *131*, 7346–7350.

[38] A. D. Srivastava, L. Unione, M. A. Wolfert, P. Valverde, A. Arda, J. Jiménez-Barbero, G.-J. Boons, *Chem. Eur. J.* **2020**, *26*, 15605–15612.

[39] C. Dalvit, C. Invernizzi, A. Vulpetti, *Chem. Eur. J.* **2014**, *20*, 11058–11068.

[40] A. Gimeno, N.-C. Reichardt, F. J. Cañada, L. Perkams, C. Unverzagt, J. Jiménez-Barbero, A. Ardá, *ACS Chem. Biol.* **2017**, *12*, 1104–1112.

[41] L. de Witte, A. Nabatov, M. Pion, D. Fluitsma, M. A. W. P. de Jong, T. de Gruijl, V. Piguet, Y. van Kooyk, T. B. H. Geijtenbeek, *Nat. Med.* **2007**, *13*, 367–371.

[42] J. Geurtsen, N. N. Driessen, B. J. Appelmelk, in *Microb. Glycobiol.*, Elsevier, Amsterdam, **2010**, pp. 673–695.

[43] E. van Liempt, C. M. C. Bank, P. Mehta, J. J. García-Vallejo, Z. S. Kawar, R. Geyer, R. A. Alvarez, R. D. Cummings, Y. van Kooyk, I. van Die, *FEBS Lett.* **2006**, *580*, 6123–6131.

[44] A. Imberty, M. Wimmerová, E. P. Mitchell, N. Gilboa-Garber, *Microbes Infect.* **2004**, *6*, 221–228.

[45] A. Audfray, J. Claudinon, S. Abounit, N. Ruvoën-Clouet, G. Larson, D. F. Smith, M. Wimmerová, J. Le Pendu, W. Römer, A. Varrot, et al., *J. Biol. Chem.* **2012**, *287*, 4335–4347.

[46] K. T. Nam, S. A. Shelby, P. H. Choi, A. B. Marciel, R. Chen, L. Tan, T. K. Chu, R. A. Mesch, B.-C. Lee, M. D. Connolly, et al., *Nat. Mater.* **2010**, *9*, 454–460.

[47] A. Varki, *Nature* **2007**, *446*, 1023–1029.

[48] M. Guberman, M. Bräutigam, P. H. Seeberger, *Chem. Sci.* **2019**, *10*, 5634–5640.

[49] Q. Li, W. Jiang, J. Guo, M. Jaiswal, Z. Guo, *J. Org. Chem.* **2019**, *84*, 13232–13241.

[50] J. Guo, W. Jiang, Q. Li, M. Jaiswal, Z. Guo, *Carbohydr. Res.* **2020**, *492*, 107999.

[51] Y. Yu, T. Tyrikos-Ergas, Y. Zhu, G. Fittolani, V. Bordoni, A. Singhal, R. J. Fair, A. Grafmüller, P. H. Seeberger, M. Delbian-

co, *Angew. Chem. Int. Ed.* **2019**, *58*, 13127–13132; *Angew. Chem.* **2019**, *131*, 13261–13266.

[52] M. Hurevich, J. Kandasamy, B. M. Ponnappa, M. Collot, D. Kopetzki, D. T. McQuade, P. H. Seeberger, *Org. Lett.* **2014**, *16*, 1794–1797.

[53] H. Yu, H. Chokhawala, R. Karpel, H. Yu, B. Wu, J. Zhang, Y. Zhang, Q. Jia, X. Chen, *J. Am. Chem. Soc.* **2005**, *127*, 17618–17619.

[54] Y. Kajihara, T. Yamamoto, H. Nagae, M. Nakashizuka, T. Sakakibara, I. Terada, *J. Org. Chem.* **1996**, *61*, 8632–8635.

[55] H. Feinberg, M. E. Taylor, N. Razi, R. McBride, Y. A. Knirel, S. A. Graham, K. Drickamer, W. I. Weis, *J. Mol. Biol.* **2011**, *405*, 1027–1039.

[56] I. van Die, S. van Vliet, A. Kwame Nyame, R. D. Cummings, C. M. C. Bank, B. Appelmelk, T. B. H. Geijtenbeek, Y. van Kooyk, *Glycobiology* **2003**, *13*, 471–478.

[57] J. J. García-Vallejo, E. van Liempt, P. da Costa Martins, C. Beckers, B. van het Hof, S. I. Gringhuis, J.-J. Zwaginga, W. van Dijk, T. B. H. Geijtenbeek, Y. van Kooyk, *Mol. Immunol.* **2008**, *45*, 2359–2369.

[58] B. J. Appelmelk, I. van Die, S. J. van Vliet, C. M. J. E. Vandenbroucke-Grauls, T. B. H. Geijtenbeek, Y. van Kooyk, *J. Immunol.* **2003**, *170*, 1635–1639.

[59] Y. Guo, H. Feinberg, E. Conroy, D. A. Mitchell, R. Alvarez, O. Blixt, M. E. Taylor, W. I. Weis, K. Drickamer, *Nat. Struct. Mol. Biol.* **2004**, *11*, 591–598.

[60] H. Feinberg, R. Castelli, K. Drickamer, P. H. Seeberger, W. I. Weis, *J. Biol. Chem.* **2007**, *282*, 4202–4209.

[61] J. Hanske, S. Aleksić, M. Ballaschk, M. Jurk, E. Shanina, M. Beerbaum, P. Schmieder, B. G. Keller, C. Rademacher, *J. Am. Chem. Soc.* **2016**, *138*, 12176–12186.

[62] J. Aretz, H. Baukmann, E. Shanina, J. Hanske, R. Wawrzinek, V. A. Zapol'skii, P. H. Seeberger, D. E. Kaufmann, C. Rademacher, *Angew. Chem. Int. Ed.* **2017**, *56*, 7292–7296; *Angew. Chem.* **2017**, *129*, 7398–7402.

[63] B. G. Keller, C. Rademacher, *Curr. Opin. Struct. Biol.* **2020**, *62*, 31–38.

[64] G. Beshr, R. Sommer, D. Hauck, D. C. B. Siebert, A. Hofmann, A. Imberty, A. Titz, *MedChemComm* **2016**, *7*, 519–530.

[65] M. P. Williamson, *Prog. Nucl. Magn. Reson. Spectrosc.* **2013**, *73*, 1–16.

[66] S. Perret, C. Sabin, C. Dumon, M. Pokorná, C. Gautier, O. Galanina, S. Ilia, N. Bovin, M. Nicaise, M. Desmadril, et al., *Biochem. J.* **2005**, *389*, 325–332.

[67] N. Gilboa-Garber, D. Sudakevitz, M. Sheffi, R. Sela, C. Levene, *Glycoconjugate J.* **1994**, *11*, 414–417.

[68] I. Joachim, S. Rikker, D. Hauck, D. Ponader, S. Boden, R. Sommer, L. Hartmann, A. Titz, *Org. Biomol. Chem.* **2016**, *14*, 7933–7948.

[69] G. Sugiarto, K. Lau, J. Qu, Y. Li, S. Lim, S. Mu, J. B. Ames, A. J. Fisher, X. Chen, *ACS Chem. Biol.* **2012**, *7*, 1232–1240.

[70] L. Ban, N. Pettit, L. Li, A. D. Stuparu, L. Cai, W. Chen, W. Guan, W. Han, P. G. Wang, M. Mrksich, *Nat. Chem. Biol.* **2012**, *8*, 769–773.

[71] W. Shao, R. Sharma, M. H. Clausen, H. V. Scheller, *Plant Methods* **2020**, *16*, 99.

[72] G. R. Craven, E. Steers, C. B. Anfinsen, *J. Biol. Chem.* **1965**, *240*, 2468–2477.

[73] A. Baumann, S. Marchner, M. Daum, A. Hoffmann-Röder, *Eur. J. Org. Chem.* **2018**, 3803–3815.

[74] K. Dax, M. Albert, J. Ortner, B. J. Paul, *Carbohydr. Res.* **2000**, *327*, 47–86.

[75] C. Bucher, R. Gilmour, *Angew. Chem. Int. Ed.* **2010**, *49*, 8724–8728; *Angew. Chem.* **2010**, *122*, 8906–8910.

[76] J. D. Martínez, A. I. Manzano, E. Calviño, A. De Diego, B. Rodriguez de Francisco, C. Romanò, S. Oscarson, O. Millet, H.-

J. Gabius, J. Jiménez-Barbero, et al., *J. Org. Chem.* **2020**, *85*, 16072 – 16081.

[77] Z. Serber, V. Dötsch, *Biochemistry* **2001**, *40*, 14317 – 14323.

[78] M. J. Smith, C. B. Marshall, F.-X. Theillet, A. Binolfi, P. Selenko, M. Ikura, *Curr. Opin. Struct. Biol.* **2015**, *32*, 39 – 47.

34

4.3. Identification of druggable allosteric pockets in β-propeller lectins

Elena Shanina, Sakonwan Kuhaudomlarp, Kanhaya Lal, Peter H. Seeberger, Anne Imberty and Christoph Rademacher

Status: submitted

4.3.1. Introduction

Bacterial infections, especially those involving biofilm formation, are becoming increasingly difficult to treat as antibiotic resistance is rising worldwide. Therefore, identifying new protein targets and designing anti-adhesives is a promising approach for future treatment of bacterial infections. Given that carbohydrate-binding proteins (lectins) are found in many pathogenic microorganisms and are involved in recognition of host, adhesion, and biofilm formation, targeting lectins from pathogens is an attractive alternative strategy to treat bacterial and fungal infections.[1] In contrast to lectins from plants, lectins from pathogens often display a high affinity for mammalian carbohydrates, likely deriving from co-evolution.[2] Thus, bacteria take advantage of these interactions to adhere and infect the host. A well-known example is the β-propeller lectin BambL from the Gram-negative bacterium *Burkholderia ambifaria*.[3] This opportunistic pathogen belongs to a group of closely related bacterial strains, the *Burkholderia cepacia complex*, causing chronic infections and exhibiting multidrug antibiotic resistance. *B. ambifaria* affects immunocompromised patients as well as those suffering from cystic fibrosis (CF) and can cause pneumonia, respiratory failure and bacteremia.[4] Moreover, *B. ambifaria* can cause sporadic outbreaks, but its epidemiology remains elusive.[5] Several studies point to an underestimated role of BambL in affecting host cellular processes, which go beyond an adhesion to the human lung epithelium.[6] Therefore, blocking BambL-carbohydrate interactions is a potential avenue to treat chronic infections, but strategies for design of inhibitors are required.

The crystal structure of BambL revealed that the protein consists of two similar domains and trimerizes to form a 6-bladed β-propeller with 6 fucose-binding sites.[3] Bacterial and fungal β-propeller is an efficient carbohydrate-binding fold, presenting all binding sites on one face of the donut shape.[7] In recent years, several inhibitors for BambL have been reported. Given the strong affinity of BambL to α-L-fucosylated monosaccharides (methyl α-L-fucopyranoside (MeFuc), dissociation constant K_d=1 µM) and complex carbohydrates (H type 2 tetrasaccharide, K_d=7.5 µM), the design of inhibitors has been focused using carbohydrates as a starting point.[3] Indeed, this approach has yielded potent BambL monovalent aryl-α-O-fucoside inhibitors with an affinity comparable to MeFuc.[8] Moreover, multivalent compounds with 4 to 6 fucose or aryl-α-O-fucosyl analogues improved selectivity and affinity towards BambL with K_d ranging between 10 to 80 nM.[8a, 9] However, the main limitation of such complex carbohydrate-based inhibitors is their molecular size. This limits their oral bioavailability and thus, complicates the future clinical approach. Consequently, discovering small, orally bioavailable drug-like molecules targeting bacterial lectins is

highly desired, but it has not been reported. Lectins have been associated with a low druggability due to their hydrophilic and solvent-exposed carbohydrate-binding sites.[10] To overcome these limitations, we have previously explored the concept of allosteric modulators for mammalian lectins.[11,12] Allosteric modulators do not bind to the orthosteric (carbohydrate)-binding site, but target an alternative (allosteric) pocket that affects the orthosteric site and *vice versa*. Several druggable, allosteric pockets have been discovered for the mammalian lectins such as DC-SIGN (CD209).[13] Notably, an intra-domain allosteric network that modulates Ca^{2+} affinity of Langerin (CD207) has been described. This has been followed by the discovery of drug-like allosteric inhibitors for Langerin supporting the allosteric communication in mammalian lectins.[11,14] Altogether, these discoveries paved the way for further search of potential allosteric pockets in lectins.

Motivated by these previous reports, we assessed the druggability of a β-propeller bacterial lectin BambL using 350 fluorinated (^{19}F) fragments. Competitive ^{19}F and T_2-filtered (CPMG) NMR allowed us to distinguish drug-like fragments binding to lectins in the carbohydrate-binding region or the secondary sites. To narrow the number of hits, compounds were counter-screened by surface plasmon resonance spectroscopy (SPR) and protein-observed ^1H-^{15}N HSQC/TROSY NMR (hereafter, TROSY NMR). The affinity potential modulatory properties of the most promising hits were derived in three orthogonal NMR experiments (TROSY NMR, PrOF and ^{19}F R_2-filtered NMR). Computational analysis was applied to predict potential druggable binding sites in BambL and validated experimentally by site-directed mutagenesis and NMR. Finally, we assessed the presence of secondary druggable sites in other β-propeller bacterial and fungal lectins.

4.3.2. Results and Discussion

Fragment screening reveals druggability of a bacterial β-propeller lectin BambL
Ligand-observed ^{19}F NMR is a highly sensitive method to screen for weak fragment-protein interaction. This is owned by the changing molecular tumbling rate of a ^{19}F fragment in solution upon its binding to the protein.[15] Therefore, ^{19}F NMR screening of fragment mixtures is frequently used in drug discovery to estimate the druggability of protein targets. Previously, we successfully applied our diversity-oriented fragment library and ^{19}F NMR to discover drug-like molecules for mammalian lectins.[10a, 11, 13, 16] Encouraged by this discovery, we applied this approach to assess the druggability of β-propeller lectins. For this, BambL was screened against 350 fluorinated fragments with MW<300 Da to assess its druggability. Herein, we carefully monitored

perturbations of the chemical shifts or changes in peak intensities of the ^{19}F resonance in presence of BambL using ^{19}F and T_2-filtered (CPMG) NMR, respectively. As a consequence of the high sensitivity, we observed an unusually high hit rate of 48% for BambL, as shown for fragment **24** as an example (**Figure 4.3-1b**). Compared to mammalian C-type lectins, the same library has given rise to 10-15% in previous screenings.[10a] To further narrow down the number of potential hits, we subsequently defined the fragments targeting the orthosteric site by competition with MeFuc. Surprisingly, only 2 fragments (<1%) were competitive with MeFuc in ^{19}F and ^{19}F CPMG NMR, indicating the presence of secondary druggable sites in BambL.[17]

Identification of druggable binding sites in BambL

Given the large number of ^{19}F NMR hits, we followed up 111 hits with the strongest effects in ^{19}F and ^{19}F CPMG NMR experiments (**Figure 4.3-S1a**). Thereby, 13 compounds were removed due to their poor solubility resulting in 98 hits subjected to the orthogonal screening using SPR and protein-observed TROSY NMR. Briefly, we confirmed binding of 78 out of 91 compounds in SPR (**Figures 4.3-1c−1d** and **4.3-S1b**). Therefore, a 'golden standard' for hit validation TROSY NMR was applied to narrow number of hits.[18] For this, we chose 39 compounds with the strongest effects in ^{19}F NMR and SPR resulting, subsequently, in 10 hits being positive in ^{19}F NMR, SPR and TROSY NMR (**Figure 4.3-S1c**).

To rank further the 10 hits, we performed titration experiments in TROSY NMR. Compounds **10**, **12** and **24** promoted the strongest dose-dependent chemical shift perturbations in [15]N BambL (**Figures 4.3-S2a−S2c**). Therefore, we derived their affinities (K_d) and ligand efficiency (LE)[19] allowing us to estimate whether these compounds serve as good starting points for lead development. As result, fragment **24** showed a two-fold stronger affinity (K_d=0.4±0.2 mM) than **10** and **12** (both K_d=0.8±0.1 mM, **Figure 4.3-S2d**), and a better LE value of 0.29 kcal mol^{-1} HA^{-1} due to its smaller molecular weight (**Figure 4.3-1e**). Given a better binding affinity of **24**, the interaction between BambL and **24** was verified in an orthogonal ligand-observed ^{19}F R_2-filtered NMR assay (**Figures 4.3-1f−1g** and **4.3-S3**), which revealed that the affinity of **24** was in the similar range to that obtained by protein-observed TROSY NMR (K_d=0.3±0.1 mM, LE=0.3 kcal mol^{-1} HA^{-1}, **Figure 4.3-1h**).

To reveal the potential binding sites of fragment **24**, we applied computational pocket prediction algorithm using SiteMap[20] on the crystal structures of BambL in complex with α-L-fucose (PDB ID: 3ZW0) and H-type 2 tetrasaccharide (PDB ID: 3ZZV).[3] SiteMap identified one secondary site per monomer (**Figures 4.3-2a** and **4.3-S4a**) that could potentially host the drug-like molecules. On the whole lectin, three such

symmetry-related predicted sites are located at the interface between the monomers, close to the C-terminus, and form narrow channels involving residues T18, N20, K23, T25, G67, T69, G86 and L87 (**Figure 4.3**-**S5b**). The binding sites are surrounded by hydrophilic residues, which make them suitable to accommodate ligands with polar groups. K23 (in three sites) and L87 (in one site) illustrate differences in side chain orientation, which slightly changes the shape and the size of the predicted sites. Nonetheless, these sites were top ranked by SiteMap for their propensities to bind drug-like molecules. Although the three sites present slight differences in the crystal structure 3ZW0 selected for docking, this is only due to differences in side chain orientation in the crystal structure and they are identical in solution. We selected only one of them for the rest of the docking study. Docking of compounds **24**, **10** and **12** could be performed successfully with the program Glide (version 7.8) and resulted binding in the predicted site (**Figures 4.3**-**2b, 4.3**-**S5** and **4.3**-**S6**). The results from the docking show that the residues T18, K23, T25, G67, Y84 and L87 play key role in ligand binding. Compound **24** binds with almost identical pose and indicates only a minor difference in orientation of morpholine ring in multiple binding poses. Likewise, compounds **12** and **10** were also accommodated in the site showing H-bond interactions with the identified key residues.

To support this prediction experimentally, we quantified the chemical shift perturbations in TROSY NMR spectra of ^{15}N BambL in presence of fragments **10**, **12** and **24**. Despite the lack of protein backbone assignment, we observed that **10**, **12** and **24** perturbed the same resonances in ^{15}N BambL, suggesting that fragments target the same binding site in BambL, which supports our computational and docking data (**Figures 4.3-2c−2e**).

Next, we confirmed that **24** targets a secondary pocket distinct from the fucose-binding site. For this, we employed a competitive ligand-observed ^{19}F CPMG NMR, using 2-deoxy-2-fluoro-L-fucose (2FF), which was reported to bind to BambL with K_d of 18 µM.[21] Here, we used **24** as a reporter molecule to test if it could compete with 2FF. Indeed, 2FF bound to BambL in a dose-dependent manner (**Figure 4.3-2f**) allowing us to derive the IC_{50} value of 0.18±0.02 mM for 2FF (**Figure 4.3-2g**). However, 2FF did not fully compete **24** for its binding site confirming that fragment bound BambL in the secondary pocket.

To analyze the impact of **24** on the carbohydrate-binding region of BambL, we employed protein-observed ^{19}F (PrOF) NMR spectroscopy. Previously, this method proved to be valuable for identification of small molecules targeting the carbohydrate-binding site of a bacterial lectin LecA.[22] Given that BambL monomer contains six tryptophan residues, we sought to apply PrOF NMR to verify the impact of **24** binding
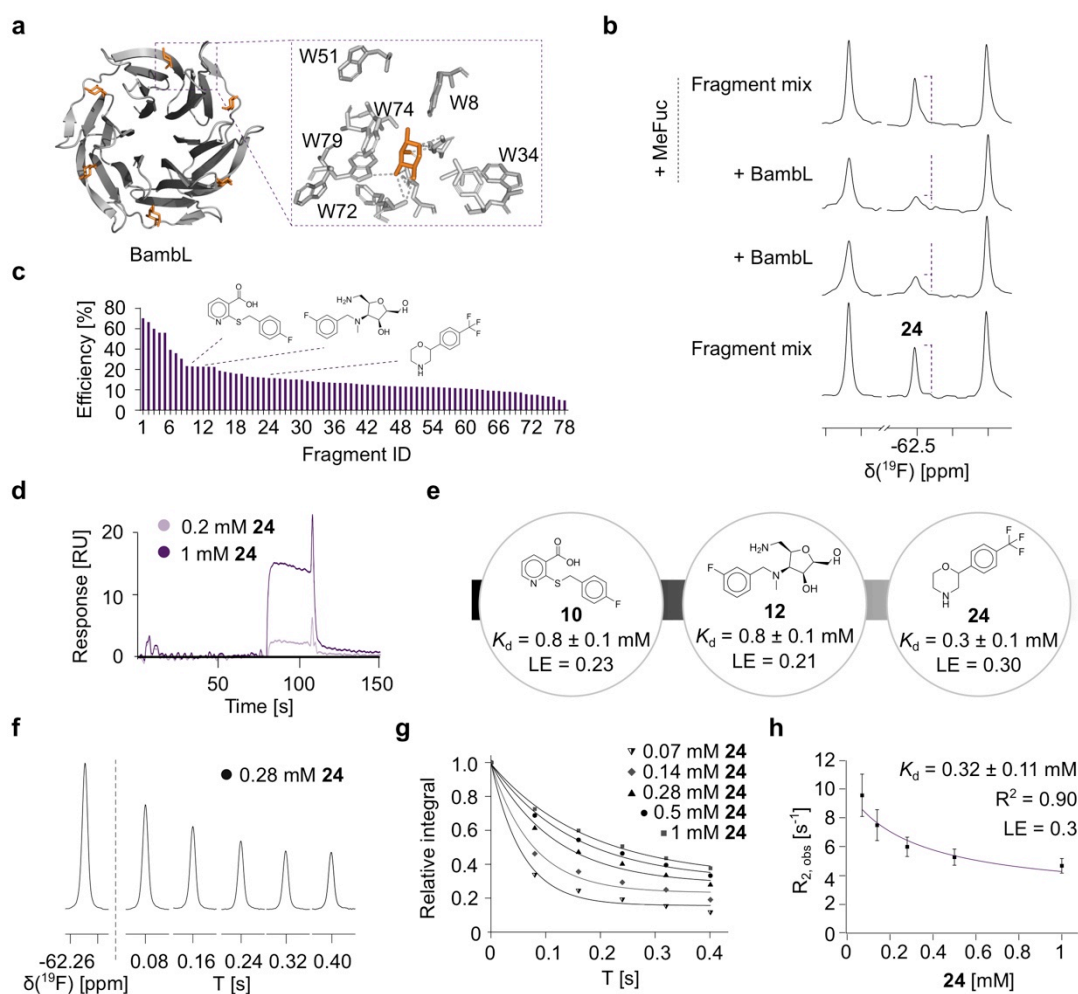
Figure 4.3-1 Druggability assessment of a bacterial lectin BambL.

Cartoon representations of a bacterial lectin crystal structure BambL in complex with L-fucose *(orange,* PDB ID: 3ZZV). A zoom-in shows six tryptophan residues in the carbohydrate-binding site. **b** Shown are ^{19}F CPMG NMR spectra of a mixture from diversity-oriented general fragment library. Screening was performed to estimate the druggability of BambL resulting in hit rates of 48% as shown on example of 0.05 mM **24**, which showed a strong line broadening effect in presence of 20 µM BambL. Competitive ^{19}F NMR using 10 mM MeFuc identified compounds targeting the orthogonal carbohydrate-binding site in BambL. **c** Hit validation by SPR verified binding of 78 fragments as shown on example of **24** SPR sensorgram binding to BambL at two doses (**d**). **e** Structures of ^{19}F NMR screening hits for BambL confirmed in SPR and TROSY NMR. **f** The interaction between **24** and 0.1 mM BambL was quantified by the relaxation rate $R_{2,obs}$ using the CPMG pulse sequence in ^{19}F NMR. **g** Representative decay curves for BambL are shown. **h** One-site fitting model revealed K_d value of 0.3±0.1 mM for BambL. $R_{2,obs}$ values were determined in duplicates.

on the carbohydrate-binding site. For this, we engineered BambL by substituting Trp residues by 5-fluorotryptophanes (5FW) and assigned the resonances by site-directed mutagenesis (**Figure 4.3**-**S7a**). Next, we confirmed protein activity with MeFuc and 2FF, where all six 5FW resonances showed a slow exchange on the NMR time scale

(**Figure 4.3**-**S7b**). This demonstrated that both MeFuc and 2FF ligands are strong binders. However, we noticed that 5FW BambL PrOF NMR was not well suitable to determine K_d values for high-affinity ligands such as 2FF and thus, it hampered the accurate derivation of the K_d (K_d=46±10.7 µM compared to reported K_d=18.8±2.3 µM,[21] **Figures 4.3**-**S7c**−**S7d**). This is not surprising, given that similar limitations were reported for protein-observed ^{1}H-^{15}N HSQC NMR.[23]

Following this, we employed 5FW BambL PrOF NMR to verify the impact of fragment (**10**, **12** and **24**) binding on the carbohydrate-binding site. Interestingly, **10**, **12** and **24** perturbed W79/W34, W51 and W72 demonstrating that fragment binding influenced the carbohydrate-binding region of BambL (**Figure 4.3**-**2h**, **Table 4.3**-**S1**). Interestingly, titration of **24** to 5FW BambL did not only affect 5FW resonances in a dose-dependent manner, but also resulted in K_d of 0.31±0.07 mM in agreement with our previous results (**Figures 4.3**-**2i** and **4.3**-**S8**). To test if fragments (**10**, **12** and **24**) could inhibit 5FW BambL interaction with carbohydrates, we performed a PrOF NMR titration with 2FF in presence of the fragments. Notably, the fragments remained bound to 5FW BambL in presence of 2FF as shown for **24** (**Figure 4.3**-**S7e**). However, we did not observe inhibition of 2FF−5FW BambL interaction resulting in a similar K_d value of 52±3 µM in presence of **24** (**Figures 4.3**-**S7f**−**S7g**). This is not surprising given the low affinity of fragments.

Taken together, computational and experimental analyses confirmed the presence of druggable secondary sites in BambL. Despite the lack of inhibitory properties, binding of fragments **10**, **12** and **24** to the secondary site perturbed an effect on the carbohydrate-binding site in BambL, which strongly suggests the presence of a communication between the orthosteric and the predicted secondary site. Given this, fragment **24** was subjected to further studies.

Structure activity relationship study of 24

In our initial SAR study, we aimed to improve the activity landscape of **24** using commercially available analogues (**Table 4.3**-**S2**, **Figures 4.3**-**3a** and **4.3**-**S9a**). For this, we employed computational and experimental TROSY NMR analyses.

Briefly, the experimental TROSY NMR analysis of 16 compounds derived from **24** revealed the importance of the morpholine group in **24** given a fully and partially abrogated binding upon its replacement with piperidine (**84**), morpholine-3-one (**91**) and tetrahydro-2H-pyran-4-ol (**92**) groups, respectively (**Figure 4.3**-**S9b**). Notably, further modification on the amine group to 4-(2-aminoethyl)-morpholine (**90**) was tolerated compared to a more hydrophobic and bulky change as 5-bromopyrimidine (**89**, **Figure 4.3**-**S9c**). Moreover, we observed that replacement (**95**) or lack (**96**) of CF_3

and changing the position of the benzyl group from 1 (**96**) to 2 (**87**) did not abrogate BambL binding either (*data not shown*). Therefore, we explored the role of the benzyl group by changing it to 1,3-dichlorobenzene (**86**), 3-methylpyrazole (**88**), methyl acetate (**98**), 2-bromo- (**85**) or thiophene (**94**) without promoting a large effect on the interaction as estimated by the total number of chemical shift perturbations (**Figure 4.3**-**S9d**). However, replacing the benzyl group to N-formylpiperidine (**97**) and tetrazole (**83**) abrogated and improved BambL binding, respectively (**Figures 4.3**-**S9e** and **4.3**-**3b**). Altogether, we successfully unraveled and improved the initial scaffold **24** to a slightly more potent binder **83** as shown in a competitive ^{19}F CPMG NMR experiment with **24** (**Figure 4.3**-**3c**). Evidently, 10 µM **24** was fully competed by 10 µM **83** proving
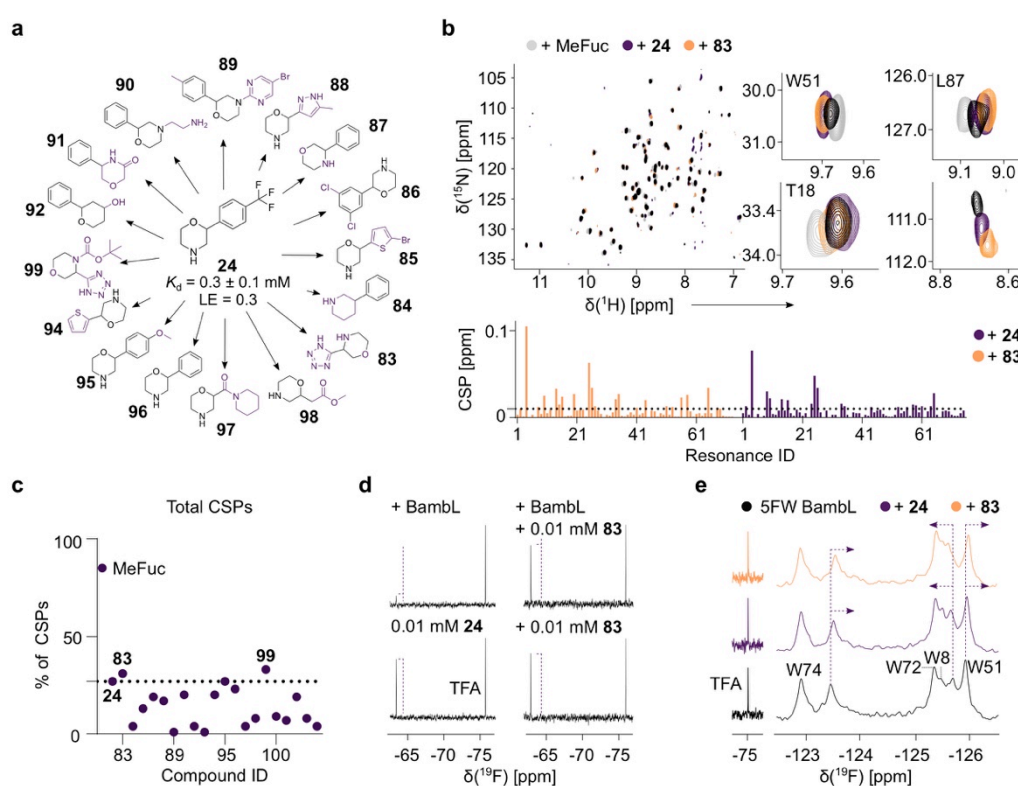
Figure 4.3-3 Structure activity relationship studies of 24.

a Shown are 16 out of 22 commercial analogues of **24**. TROSY NMR was used to rank the derivatives of **24**. **b** Example TROSY NMR (*left panel*) spectrum of ^{15}N BambL shows an improved binding of fragment **83** (*orange*) compared to **24** (*violet*). CSP plots (*right panel*) show **83** improved the magnitude of CSPs compared to **24** and partially similarly to methyl-α-L-fucose (MeFuc), which perturbed 88% of resonances. Dashed line indicates CSPs > 0.01 ppm. **c** Total % of CSPs derived in TROSY NMR shows **83** (31%) and **99** (33%) promoted more CSPs in ^{15}N BambL 31% and 33% CSPs, respectively, compared to initial hit **24** (27%). Dashed line was set at 27%. **d** ^{19}F CPMG NMR experiment with 0.1 mM **24** showed competition by 0.1 mM **83** indicating the superior binding of 83 and that both fragments bind the same pocket. **e** PrOF NMR with 0.1 mM 5FW BambL supported improved binding of 1 mM **83** compared 1 mM **24**. Moreover, **83** showed CSPs of W51 and W34/W79 demonstrating the effect of both fragments on the carbohydrate-binding site.

that **83** serves as a better affinity ligand and both fragments target the same pocket in BambL (**Figure 3d**). Next, we validated five commercially available analogues of **83** in TROSY NMR (**99**-**104**, **Table 4.3**-**S2**).

Compounds **100** and **104** proved the importance of the tetrazole and morpholine groups for binding of **83** to the secondary site, as we did not observe binding in TROSY NMR. Similar to **90**, the presence of a substituent (tert-butyl formate, **99**) on a nitrogen atom of morpholine group was tolerated well and thus, this position could serve for future fragment growing (**Figure 4.3**-**S10**).

Computational docking analysis of four structural derivatives of compounds **24** (**83**, **84**, **87**, **90** and **94**) was performed to check if the predicted pocket could accommodate these compounds (**Figure 4.3**-**S11**). Hereby, we chose these fragments to check the importance of morpholine and benzyl groups. As result, all compounds could be accommodated in the site, whereas **84** did not form electrostatic interactions. These observations were in agreement with our experimental data showing the importance of the morpholine group in **24** as shown for **87**, whereas other parts of the compound are rather interchangeable (e.g. **83** and **94**).

Finally, we used **83** in PrOF NMR to confirm our observation with the initial hit **24** promoting changes in the carbohydrate-binding site of 5FW BambL (**Figure 4.3**-**3e**). Indeed, **83** perturbed W79/W34, W51 and W72 similar to the initial hit **24** causing even larger NMR chemical shift perturbations of 5FW resonances. Together, both fragments with a similar scaffold promoted an effect on the carbohydrate-binding site in BambL.

Communication between the carbohydrate and remote sites in BambL

To further prove allosteric communication between the unknown secondary site and the carbohydrate site, we proposed that mutations in the carbohydrate and secondary sites of BambL should introduce perturbations that propagate through the network and result in similar chemical shift changes. For this, we used four (W8F, W51F, W72F and W74F) and three (T18S, T25S and L87R) mutants for the orthosteric and the predicted remote sites, respectively (**Figure 4.3**-**4a**). All mutants were folded and active as observed by TROSY NMR (**Figure 4.3**-**S12**).

Next, we quantified and compared the NMR chemical shift perturbations (CSPs) induced by mutations in their apo forms with respect to the wild type (WT). Interestingly, perturbations were not restricted to residues in the close periphery, but also affected remote residues, as shown for W51F and T18S mutants. Here, we clearly identified a chemical shift perturbation of W72 in both WT and T18S mutants moving along the same vector (**Figure 4.3**-**4b**). Quantification of NMR chemical shift perturbations of the apo WT to other apo mutant forms (W8F, W72F, W74F, L87R and

T18S) revealed the conformational changes through the same paths in ^{15}N BambL, which is typical for allosteric proteins (**Figure 4.3-4c**).[24]

To assess if mutations in the predicted pocket alter the affinity of proteins to the natural carbohydrates, we titrated 2FF and a complex carbohydrate (F-H type 2) to BambL WT and T18S (**Figures 4.3-4d** and **4.3-S13a**). Compared to BambL WT, T18S preserved its affinity for monosaccharide 2FF (**Figure 4.3-4e**, WT: K_d=7.9±0.2 µM, T18S: K_d=8.2±0.2 µM). However, BambL T18S showed nearly two−fold decrease in affinity for a complex carbohydrate, F-H type 2 (**Figure 4.3-S13b**, K_d=16.7±2.5 µM) compared to BambL WT (K_d=9±2 µM[25]). To verify the binding epitope of F-H type 2, we applied TROSY NMR using ^{15}N BambL T18S and compared it to WT (**Figure 4.3-S13c**). Overall, T18S mutation reduced the magnitude of CSPs in ^{15}N BambL suggesting a negative modulatory role of the pocket on the carbohydrate-binding site in recognizing complex carbohydrates (**Figure 4.3-S13d**). Interestingly, a discrepancy between two carbohydrate-binding sites in binding to complex carbohydrates, but not MeFuc has been reported for BambL recently.[9a] Given the lack in affinity change with 2FF, we propose that inhibition of secondary site could potentially down-regulate the affinity of BambL by tuning the orthosteric site between monomers, but not within a monomer. However, this hypothesis requires further investigations.

Finally, we investigated the impact of the pocket mutations on binding of most potent fragments **24** and **83** by ^{19}F CPMG and TROSY NMR. Notably, pocket mutations reduced **24** binding two−fold in ^{19}F CPMG (**Figure 4.3-4f**) and TROSY NMR experiments (**Figures 4.3-S14a−b**) allowing us to conclude that the mutations blocked the entrance into the predicted pocket only partially. Similarly, we observed this effect with the fragment **83** (**Figures 4.3-S14c−d**), which is in agreement with the computational docking analysis suggesting the presence of two orientations for **83** and its derivative **99** (**Figure 4.3-S15**). Interestingly, mutation in the carbohydrate-binding site (W51F) reduced **83** binding similarly to ^{15}N BambL T18S and other pocket mutants (**Figures 4.3-4g** and **4.3-S14e**), which confirms the 'end−to−end' communication of both sites.

Taken together, these data reveal the existence of a druggable allosteric site in BambL. The NMR chemical shift perturbations of backbone resonances of ^{15}N BambL pocket mutants are located in sites distal from the actual pocket and the communication extends to the carbohydrate recognition site, suggesting a propagation of conformational changes in BambL upon changes in the allosteric pocket.
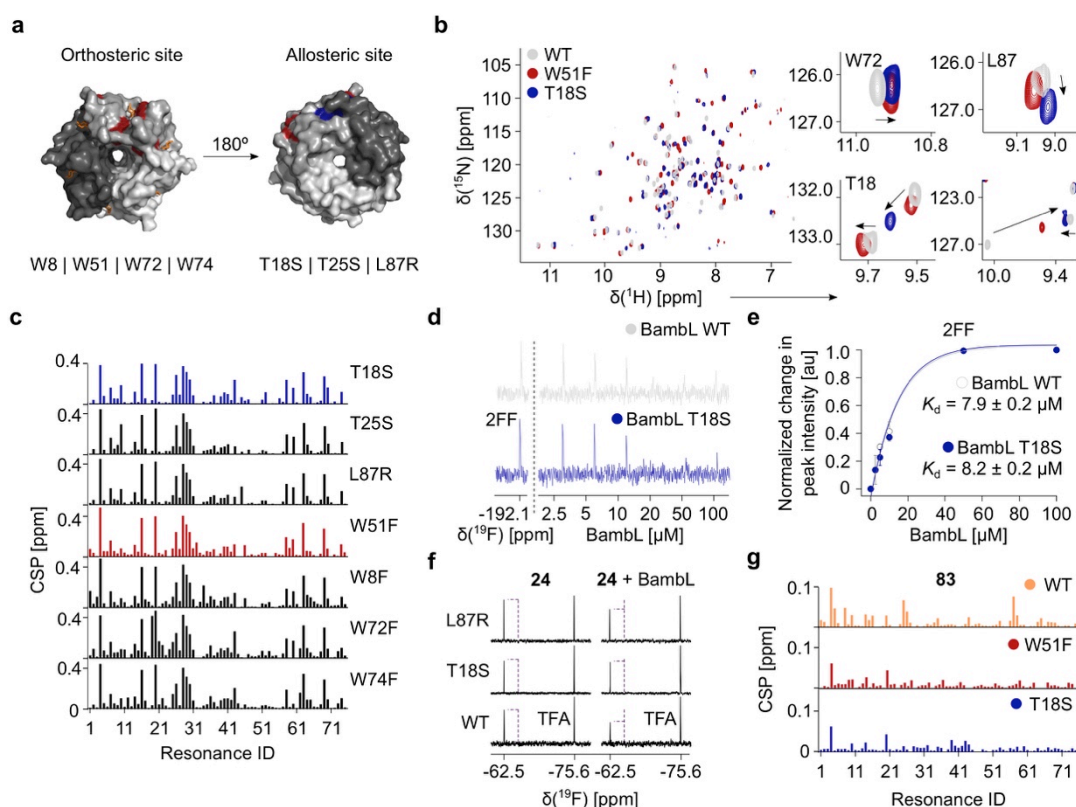
Figure 4.3-4 Characterization of secondary site in BambL.

a Top and bottom views on the surface of a crystal structure BambL in complex with L-fucose (*orange*, PDB ID: 3ZZV) show the orthosteric (*red*) and potential allosteric site (*blue*), respectively. Single-point mutations in the carbohydrate-binding and secondary site have been proposed to check the communication between the two sites. **b** Overlay of ^{15}N TROSY NMR spectra of WT, W51F and T18S as example showing conformational changes introduced by both site-directed mutations. Notably, W51F and T18S mutations promoted identical changes on other resonances in the orthosteric (W72) and secondary site (L87R) in BambL. **c** CSP studies of mutant apo forms compared to BambL WT show a preserved CSP pattern in both tryptophan and allosteric pocket mutants. **d** ^{19}F NMR spectra of 2FF in presence of BambL WT and T18S. **e** Determination of 2FF K_d values for BambL T18S revealed a preserved affinity compared to BambL WT. **f** ^{19}F CPMG NMR of **24** with BambL WT, T18S and L87R to verify its binding site. **g** Binding of **83** to W51F and T18S promoted less CSPs compared to WT supporting the existence of communication between orthosteric and remote sites.

primary cell walls of higher plants (xyloglucans) causing lethal wilt in many agricultural crops worldwide.[26a,27] Blocking these lectins without inducing antimicrobial resistance is highly desired environmentally and in health care.[1b]

Other β-propeller lectins contain secondary druggable sites

To assess if our discovery of druggable secondary sites could be applied to other β-propeller lectins, we included a bacterial (RSL) and a fungal (AFL) lectins from bacterium *R. solanacearum* and fungus *A. fumigatus*, respectively. These lectins have sequence and structure similarities with BambL (RSL: 76% sequence identity, RMSD=0.56Å, AFL: 39% sequence identity, RMSD=1.84Å), albeit with different oligomerization for AFL, and both have a low micromolar affinity for terminal α-L-fucose on animal and plant carbohydrates (AFL: K_d=76.4 µM and RSL: K_d=0.64 µM, **Figure 4.3-5a**).[26]

Similar to BambL, AFL specifically recognize α-L-fucosyl residues on lung tissue promoting lung failure in the immunocompromised and CF patients. RSL from *R. solanacearum* attaches to fucosylated carbohydrates of the primary cell walls of higher plants (xyloglucans) causing lethal wilt in many agricultural crops worldwide.[26a,27] Blocking these lectins without inducing antimicrobial resistance is highly desired environmentally and in health care.[1b]

First, we screened our diversity-oriented fragment library against both β-propeller lectins (**Figure 4.3**-**5b**). Interestingly, ^{19}F NMR screening revealed a comparable total hit rate for both RSL and AFL of 33 and 48%, respectively. Competitive ^{19}F NMR screening resulted in 17% and 5% of fully or partially competed fragments for AFL and RSL, respectively. A similar hit rate for the orthosteric site of RSL and BambL can be explained by their structural similarity, whereas the structure of AFL differs slightly from the bacterial lectins.

Similar as before, we applied computational pocket prediction algorithms using SiteMap tool on the apo and holo forms of RSL and AFL.[20] Interestingly, SiteMap calculations identified three secondary pockets in RSL trimer (apo and holo) equivalent to the newly identified pockets in BambL (**Figure 4.3**-**S5a**). However, shape and size of the predicted sites were slightly different due to differences in residues in the binding sites (**Figure 4.3**-**S16**). Finally, SiteMap identified a completely different, but also druggable region in the fungal lectin AFL, which is structurally more distant from both bacterial lectins (**Figure 4.3**-**S17**).

Taken together, a high number of hits for the secondary sites in ^{19}F NMR and our computational pocket prediction analysis strongly suggest the availability of druggable sites in other bacterial β-propeller lectins, as well as fungal lectins as shown for RSL and AFL, respectively.

4.3.3. Conclusions

We report the presence of druggable pockets in a bacterial lectin BambL, which could be used to design allosteric inhibitors. We showed binding of fragments to BambL in a ^{19}F NMR screening and validated hits using orthogonal methods: SPR and TROSY NMR. Computational pocket prediction analysis SiteMap identified three potential druggable pockets in BambL trimer. We also showed that the potential secondary binding sites could accommodate drug-like molecules (**24**, **10** and **12**). Initial SAR study of **24** (K_d=0.3±0.1 mM, LE=0.3) proved the pocket identity by confirming the predicted part of **24** scaffold responsible for its binding to the pocket in our docking study. Notably, fragment binding to the secondary site induced conformational changes in the carbohydrate-binding site of 5FW BambL in PrOF NMR. This
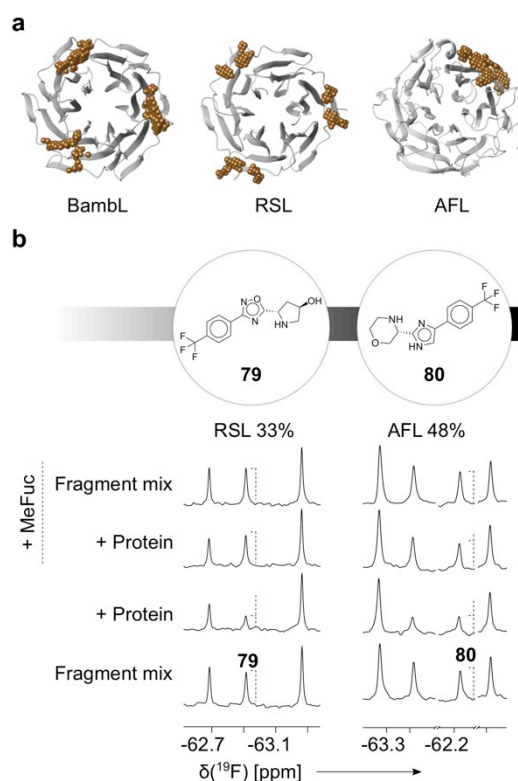
Figure 4.3-5 Other β−propeller lectins contain secondary druggable sites.

a Shown are secondary pockets (*dotted spheres*) in BambL (PDB ID: 3ZW0), RSL (PDB ID: 3Zl8 *apo*) and AFL (PDB ID: 4AGI) identified using SiteMap. **b** ^{19}F NMR screening was performed to estimate the druggability of RSL and AFL resulting in hit rates of 33% and 48% as shown on example of 0.05 mM fragments **79** and **80**, respectively. Compounds showed a strong line broadening effect in presence of 40 µM RSL and 20 µM AFL. Competitive ^{19}F NMR experiment identified hits being competed with 10 mM MeFuc.

observation allowed us to propose the presence of a communication between two spatially distant binding sites in BambL. Employing site-directed mutagenesis within the predicted site and the carbohydrate pocket, we observed conformational changes of [15]N BambL backbone resonances in TROSY NMR in distal regions from the mutation sites. Such behavior is typical for allosteric proteins. Given a fungal AFL and bacterial RSL lectins show similarities in structure and druggability hit rates to BambL, we believe the allostery could also be present in other β-propeller lectins. These observations will support future drug-discovery campaigns that aim to develop drug-like allosteric inhibitors for bacterial and fungal lectins.

4.3.4. Acknowledgements

program (MSCA 76558), ANR PIA Glyco@Alps (ANR-15-IDEX-02) and Labex Arcane/CBH-EUR-GS (ANR-17-EURE-0003).

4.3.5. References

[1] a) L. C. Breitenbach Barroso Coelho, P. Marcelino dos Santos Silva, W. Felix de Oliveira, M. C. de Moura, E. Viana Pontual, F. Soares Gomes, P. M. Guedes Paiva, T. H. Napoleão, M. T. dos Santos Correia, *Journal of Applied Microbiology* **2018**, *125*, 1238-1252; b) J. Meiers, E. Siebs, E. Zahorska, A. Titz, *Current opinion in chemical biology* **2019**, *53*, 51-67.

[2] J. R. Bishop, P. Gagneux, *Glycobiology* **2007**, *17*, 23r-34r.

[3] A. Audfray, J. Claudinon, S. Abounit, N. Ruvoën-Clouet, G. Larson, D. F. Smith, M. Wimmerová, J. Le Pendu, W. Römer, A. Varrot, A. Imberty, *The Journal of biological chemistry* **2012**, *287*, 4335-4347.

[4] a) K. A. Ramsay, C. A. Butler, S. Paynter, R. S. Ware, T. J. Kidd, C. E. Wainwright, S. C. Bell, Journal of clinical microbiology 2013, 51, 3975-3980; b) G. A. Pradenas, B. N. Ross, A. G. Torres, *Vaccines* **2016**, *4*.

[5] a) E. C. Nannini, A. Ponessa, R. Muratori, P. Marchiaro, V. Ballerini, L. Flynn, A. S. Limansky, *The Brazilian journal of infectious diseases : an official publication of the Brazilian Society of Infectious Diseases* **2015**, *19*, 543-545; b) T. Coenye, E. Mahenthiralingam, D. Henry, J. J. LiPuma, S. Laevens, M. Gillis, D. P. Speert, P. Vandamme, *International journal of systematic and evolutionary microbiology* **2001**, *51*, 1481-1490.

[6] I. Wilhelm, E. Levit-Zerdoun, J. Jakob, S. Villringer, M. Frensch, R. Übelhart, A. Landi, P. Müller, A. Imberty, R. Thuenauer, J. Claudinon, H. Jumaa, M. Reth, H. Eibel, E. Hobeika, W. Römer, *Science Signaling* **2019**, *12*, eaao7194.

[7] F. Bonnardel, A. Kumar, M. Wimmerova, M. Lahmann, S. Perez, A. Varrot, F. Lisacek, A. Imberty, *Structure* **2019**, *27*, 764-775.e763.

[8] a) D. Goyard, V. Baldoneschi, A. Varrot, M. Fiore, A. Imberty, B. Richichi, O. Renaudet, C. Nativi, Bioconjugate Chemistry 2018, 29, 83-88; b) B. Richichi, A. Imberty, E. Gillon, R. Bosco, I. Sutkeviciute, F. Fieschi, C. Nativi, *Organic & biomolecular chemistr*y **2013**, *11*, 4086-4094.

[9] a) S. Kuhaudomlarp, L. Cerofolini, S. Santarsia, E. Gillon, S. Fallarini, G. Lombardi, M. Denis, S. Giuntini, C. Valori, M. Fragai, A. Imberty, A. Dondoni, C. Nativi, *Chemical science* **2020**, *11*, 12662-12670; b) N. Galanos, Y. Chen, Z. P. Michael, E. Gillon, J.-P. Dutasta, A. Star, A. Imberty, A. Martinez, S. Vidal, *ChemistrySelect* **2016**, *1*, 5863-5868; c) C. Ligeour, A. Audfray, E. Gillon, A. Meyer, N. Galanos, S. Vidal, J.-J. Vasseur, A. Imberty, F. Morvan, *RSC Advances* **2013**, *3*, 19515-19524.

[10] a) J. Aretz, E.-C. Wamhoff, J. Hanske, D. Heymann, C. Rademacher, *Frontiers in Immunology* **2014**, *5*; b) B. Ernst, J. L. Magnani, *Nature reviews. Drug discovery* **2009**, *8*, 661-677.

[11] J. Aretz, U. R. Anumala, F. F. Fuchsberger, N. Molavi, N. Ziebart, H. Zhang, M. Nazaré, C. Rademacher, *Journal of the American Chemical Society* **2018**, *140*, 14915-14925.

[12] B. G. Keller, C. Rademacher, *Current opinion in structural biology* **2020**, *62*, 31-38.

[13] J. Aretz, H. Baukmann, E. Shanina, J. Hanske, R. Wawrzinek, V. A. Zapol'skii, P. H. Seeberger, D. E. Kaufmann, C. Rademacher, *Angewandte Chemie International Edition* **2017**, *56*, 7292-7296.

[14] a) T. T. Waldron, T. A. Springer, *Proceedings of the National Academy of Sciences of the United States of America* **2009**, *106*, 85-90; b) J. Hanske, S. Aleksić, M. Ballaschk, M. Jurk, E. Shanina, M. Beerbaum, P. Schmieder, B. G. Keller, C. Rademacher, *Journal of the American Chemical Society* **2016**, *138*, 12176-12186.

[15] C. Dalvit, A. Vulpetti, *Journal of medicinal chemistry* **2019**, *62*, 2218-2244.

[16] a) J. Aretz, Y. Kondoh, K. Honda, U. R. Anumala, M. Nazaré, N. Watanabe, H. Osada, C. Rademacher, *Chemical communications* (Cambridge, England) **2016**, *52*, 9067-9070; b) J. Schulze, H. Baukmann, R. Wawrzinek, F. F. Fuchsberger, E. Specker, J. Aretz, M. Nazaré, C. Rademacher, *ACS chemical biology* **2018**, *13*, 3229-3235.

[17] R. F. Ludlow, M. L. Verdonk, H. K. Saini, I. J. Tickle, H. Jhoti, *Proceedings of the National Academy of Sciences* **2015**, *112*, 15910-15915.

[18] A. D. Gossert, W. Jahnke, *Progress in nuclear magnetic resonance spectroscopy* **2016**, *97*, 82-125.

[19] A. L. Hopkins, C. R. Groom, A. Alex, *Drug discovery today* **2004***, 9*, 430-431.

[20] T. A. Halgren, *J Chem Inf Model* **2009**, *49,* 377-389.

[21] T. Dingjan, É. Gillon, A. Imberty, S. Pérez, A. Titz, P. A. Ramsland, E. Yuriev, *Journal of Chemical Information and Modeling* **2018**, *58,* 1976-1989.

[22] E. Shanina, E. Siebs, H. Zhang, D. Varón Silva, I. Joachim, A. Titz, C. Rademacher, *Glycobiology* **2021**, *31,* 159-165.

[23] M. P. Williamson, *Progress in nuclear magnetic resonance spectroscopy* **2013***, 73,* 1-16.

[24] D. D. Boehr, J. R. Schnell, D. McElheny, S.-H. Bae, B. M. Duggan, S. J. Benkovic, H. J. Dyson, P. E. Wright, *Biochemistry* **2013**, *52,* 4605-4619.

[25] G. Fittolani, E. Shanina, M. Guberman, P. H. Seeberger, C. Rademacher, M. Delbianco, A*ngewandte Chemie International Edition* **2021**, *60,* 13302-13309.

[26] a) N. Kostlánová, E. P. Mitchell, H. Lortat-Jacob, S. Oscarson, M. Lahmann, N. Gilboa-Garber, G. Chambat, M. Wimmerová, A. Imberty, *Journal of Biological Chemistry* **2005**, *280,* 27839-27849; b) J. Houser, J. Komarek, N. Kostlanova, G. Cioci, A. Varrot, S. C. Kerr, M. Lahmann, V. Balloy, J. V. Fahy, M. Chignard, A. Imberty, M. Wimmerova, *PLOS ONE* **2013**, *8,* e83077.

[27] J. Houser, J. Komarek, G. Cioci, A. Varrot, A. Imberty, M. Wimmerova, *Acta Crystallographica Section D* **2015**, *71,* 442-453.

4.4. Metal binding pharmacophores as inhibitors of carbohydrate-protein interactions

Elena Shanina, Sakonwan Kuhaudomlarp, Eike Siebs, Priscilla da Silva Figueiredo Celestino Gomes, Peter H. Seeberger, Felix F. Fuchsberger, Maxime Denis, Mads H. Clausen, Didier Rognan, Alexander Titz, Anne Imberty and Christoph Rademacher

Status: In preparation

4.4.1. Introduction

All cells are covered in a complex matrix of carbohydrates with established roles in health and disease.[1-2] Mammalian immune system employs carbohydrate-binding proteins (GBPs) called lectins for various processes such as embryonic development, pathogen and tumor recognition.[3-4] In particular, the family of Ca^{2+}-binding C-type lectins (CLRs) is well known for its roles in immune homeostasis, cell signaling, self- or pathogen recognition and antigen presentation.[5] However, many pathogens (e.g. *Pseudomonas aeruginosa*[6], influenza[7] and SARS-CoV-2[8]) and tumors[9] exploit these mechanisms to escape or suppress immune cell responses. Consequently, lectins evolved as novel therapeutic targets for design of anti-adhesives and drug delivery systems to combat antimicrobial and autoimmune diseases.[10-11] In particular, two neuraminidase inhibitors Zanamivir (Relenza®) and Oseltamivir (Tamiflu®) are currently on the market for the treatment of influenza virus infections.[12] This example demonstrates the utility of inhibitors targeting the carbohydrate-protein interactions in antiviral therapy. Moreover, lectins can be used for the selective drug-delivery systems. This approach has been successful for Givosiran targeting the hepatocytes through the asialoglycoprotein receptor (ASGPR)[13] and the development of dendritic cell based vaccines as in the case of Dec-205 (CD205).[14] Cumulatively, such examples increased the therapeutic interest in other clinically relevant lectins.

In recent years, CLRs have emerged as therapeutic targets due to their cell-specific expression on immune cells, endocytic and immunomodulatory properties.[15] In particular, a cell-specific ICAM-3 grabbing non-integrin (DC-SIGN, CD209) and Langerin (CD207) have been well studied. DC-SIGN is responsible for the uptake and processing of various antigens in dendritic cells (DCs) or macrophages, whereas Langerin has similar functions in a subset of antigen presenting cells in skin called Langerhans cells.[16-17] In particular, pathogens such as *Mycobacterium tuberculosis*[18] or viruses (e.g. HIV type-1,[19] Ebola,[20] SARS-CoV[21] and SARS-CoV-2[22]) use their surface carbohydrates to attach to DC-SIGN and Langerin even promoting an immune escape in some cases. As for the majority of the CLRs, these cellular mechanisms require a Ca^{2+}-cofactor, e.g. Langerin shows a high affinity for Ca^{2+} (K_d=130 μM[23]). However, both CLRs show rather a low affinity for D-mannose (DC-SIGN: K_d=3.5 mM, Langerin: K_d=6.1 mM[24]) and a high avidity for high-mannose carbohydrates on pathogens, e.g. HIV gp120 (K_d=1 nM).[25-27] Reports showing the role of DC-SIGN in promoting HIV trans-infection of T cells have drawn a lot of attention to it as a potential target for the antiviral therapy.[28] Aside from the pathogen recognition, Langerin is an efficient endocytic recycling receptor with established roles in immunity and tolerance.[29] Thus, Langerin evolved as an attractive target for the design of anti-

infectives, as well as for development of new skin vaccines to modulate the immune system.[30]

Besides mammalian CLRs, bacterial lectins have also been proposed as viable targets for drug development. To establish protective antibiotic−resistant biofilms in the lungs of immunocompromised patients, opportunistic bacteria *Pseudomonas aeruginosa* and *Burkholderia ambifaria* employ the lectins PA-IL (LecA)/PA-IIL (LecB) and BambL, respectively.[31] As compared to mammalian lectins, bacterial lectins show a remarkably high affinity for multivalent carbohydrates on mammalian cells, which is frequently Ca^{2+}-dependent as well.[32] For instance, LecA and LecB bind to D-galactose (K_d=88 µM[33]) and L-fucose (K_d=0.41 µM[34]) branched carbohydrates, respectively.[34] Given the key roles LecA and LecB play in biofilm formation, both emerged as promising therapeutic targets to prevent biofilm formation and disease progression.[35]

In conclusion, therapeutics interfering with carbohydrate-lectins interactions potentially serve as interventions against bacterial[36] and viral[37] infections, as well as cancer[38] and autoimmune diseases.[39] However, inhibitors of the orthosteric site of lectins have been a bottleneck for drug discovery.[40] This is not surprising given the hydrophilic nature of carbohydrate-lectin interactions, whose specificity and affinity are achieved through H-bonds, van-der-Waals contacts and hydrophobic binding of aromatic acid residues.[41] Moreover, some lectins carry metal ions in the carbohydrate-binding site as central coordinating element. The flexibility, dynamics and electronic structure of metal-centered complexes further complicate shaping metal-fragment interactions in 'classical' drug design.[42] This overall leads to a low druggability of lectins compared to other protein targets and thus, renders this target class challenging.[40] Altogether, these difficulties have contributed to the vast underrepresentation of inhibitors for lectins in the drug space.[43]

Recent developments of lectin inhibitors focus on mimicking the carbohydrates (glycomimetics).[43] Several carbohydrate-based glycomimetics against Langerin[44] and DC-SIGN[45] are available, but only a few drug-like small molecules have been reported.[46-48] Similar to the mammalian CLRs, various carbohydrate-based glycomimetics have been designed for LecA and LecB, but drug-like inhibitors are lacking.[49] To discover new drug-like molecules for lectins, fragment-based drug discovery (FBDD) has emerged as a promising strategy for the identification of high affinity inhibitors for challenging targets.[50] At the foundation of its success lies the design of the library. Many libraries are selected based on the 'rule of three',[51] allowing further fragment evolution e.g. using ligand efficiency (LE) as a measure, which relates the affinity to the number of non-hydrogen atoms.[52] Thus, FBDD serves as good starting point for lead development. Notably, compounds likely to interfere with protein

activity assay (PAINS) are excluded from general FBDD libraries. PAINS are small aggregation-prone and insoluble molecules, reactive covalent modifiers, redox active species and metal chelators that can bind protein nonspecifically.[53] Metal chelators are classified as PAINS based on the empirical experience, e.g. these are commonly contaminated with metal ions, which may introduce transition or heavy metal ions into a biological assay and thus, should be used carefully at high concentrations.[54] Cumulatively, these factors led to the exclusion of metal-binding pharmacophores (MBPs) from FBDD libraries as its presence can partially or fully contradict the screen.[55] However, excluding all metal coordinating fragments as PAINS is inaccurate given the numerous protein targets that bind to metals. Therefore, a concept of target-directed FBDD campaigns using MBPs has been proposed previously.[56] The MBP-FBDD campaigns have been successful for identification of inhibitors for metalloenzymes leading to several drugs on the market.[57] Certainly, captoprile58 and suberanilohydroxamic acid (SAHA or Zolinza)[59] demonstrate the power of MBP-FBDD campaign in discovery drugs with a positive impact on human health. Despite the importance of metal ions in maintaining the activity and stability of a vast majority of proteins in the genome, MBPs have not been reported for other clinically relevant protein targets as metal-binding lectins.

Here, we aim to show that metal-dependent lectins are more druggable as previously anticipated. To address this, we screen four fragment libraries identifying metal-binding pharmacophores (MBPs) as novel scaffolds. To demonstrate the potential of MBPs in inhibition of carbohydrate-protein interactions, we study their effect on clinically relevant Ca^{2+}-binding mammalian lectins DC-SIGN and Langerin and three bacterial lectins, namely LecA and LecB, whereas a non-metal-dependent lectin BambL serves as a control. Employing NMR,[60] surface plasmon resonance (SPR), X-ray crystallography, biochemical competitive binding fluorescence polarization (FP[61]) and a cell-based fragment assay (CellFy[62]), we explore the structure-activity relationships of several MBPs to further improve lectin binding and selectivity.

4.4.2. Results and Discussion

Druggability Assessment of Ca^{2+}-Dependent Bacterial Lectins

To show that metal-dependent lectins are more druggable as previously anticipated, we employed fragments from virtual and three physical libraries: 3F Fsp^3-rich, a diversity-oriented general and MBP. The libraries were screened for binding to Ca^{2+}-dependent bacterial lectins (LecA and LecB), mammalian CLRs (Langerin and DC-SIGN) and compared to a non-metal-binding lectin BambL (**Figure 4.4-1a**).

In previous reports, low computational druggability scores were shown for metal-dependent mammalian CLRs identifying it challenging or even undruggable.[63] However, the druggability of bacterial lectins has not been assessed before. Aiming to identify drug-like molecules for the metal-dependent bacterial lectins, virtual screening was performed for LecA and LecB as described in **Supporting Information**. Briefly, a low hit rate (<1%) was predicted for both target proteins, which was confirmed experimentally in SPR and protein-observed ^{15}N TROSY NMR (**Figure 4.4-S1**). These results indicated that the active site of metal-dependent lectins might tolerate a limited class of chemical scaffolds.

As result of low hit rates in the virtual screening, we screened three physical libraries using NMR. The 3F library of 115 Fsp^3-rich, natural-product-like fragments with diverse shapes, was screened against the bacterial lectins using ^{19}F and T$_2$-filtered NMR.[64] A moderate hit rate (9%) of 3F library was reported for DC-SIGN previously, but the binding sites of hits were not identified. To probe whether fragments from 3F Fsp^3-rich library target the Ca^{2+}-binding sites of bacterial lectins, 3F library was screened in the presence of lectin targets (LecA, LecB or BambL) and 5 mM EDTA. Next, competition with the metal cofactor was achieved by adding 10 mM CaCl$_2$ during the next screening round. No hits from the 3F fragment library bound to bacterial lectins (**Figure 4.4-1b**).

Next, we screened a diversity-oriented general library[63] of 350 fluorinated fragments against LecA and LecB using ^{19}F and T$_2$-filtered NMR, whereas 650 non-fluorinated fragments were screened for LecA additionally in protein-observed ^{15}N TROSY NMR. Previously, we showed a moderate to high druggability of CLRs namely Langerin (15.7%) and DC-SIGN (13.5%) and non-metal-dependent β-propeller bacterial lectins such as BambL (48%),[63,65] given that hit rates from fragment screening for druggable targets are between 5–15%.[66] Interestingly, we observed similar hit rates for the Ca^{2+}-dependent bacterial lectins (LecA: 15%, LecB: 14%), where most hits for LecA and LecB were metal-coordinating fragments (**Figures 4.4-S2a** and **4.4-S3a**). Moreover, these compounds were perturbed in ^{19}F NMR in presence of 10 mM CaCl$_2$ suggesting metal chelation (**Figures 4.4-1c** and **4.4-S3b**). Moreover, these compounds were perturbed in ^{19}F NMR in presence of 10 mM CaCl$_2$ suggesting metal chelation (**Figures 1c** and **4.4-S3b**). Next, we evaluated hits for LecA and LecB in ^{15}N TROSY NMR. For LecB, out of 24 fragments **3a** was the most potent scaffold from our general library as it perturbed resonances similarly to the positive control methyl-α-L-fucose (hereafter, MeFuc; **Figures 4.4-S3c**-**d**). For LecA, we confirmed two non- (**1** and **2d**) and two fluorinated (**1e**-**1f**) hits, whereas hydroxamate **1** perturbed 40% of the same ^{15}N resonances in ^{15}N LecA as the positive control methyl-α-D-galactose (hereafter,

MeGal) suggesting a similar binding site (**Figures 4.4-S2b**-**d**). Moreover, a dose-dependent binding of **1** to LecA was confirmed in SPR as well (**Figure 4.4-S2e**). Interestingly, we observed a preference of **1** binding to LecA, but not LecB demonstrating early target selectivity of the hit (**Figure 4.4-S3d**).

Altogether, our results demonstrated the druggability of metal-dependent lectins despite the limitations of *in silico* approaches. Moreover, our work highlights the importance of chemical fragment diversity in NMR screening and indicates that metal-
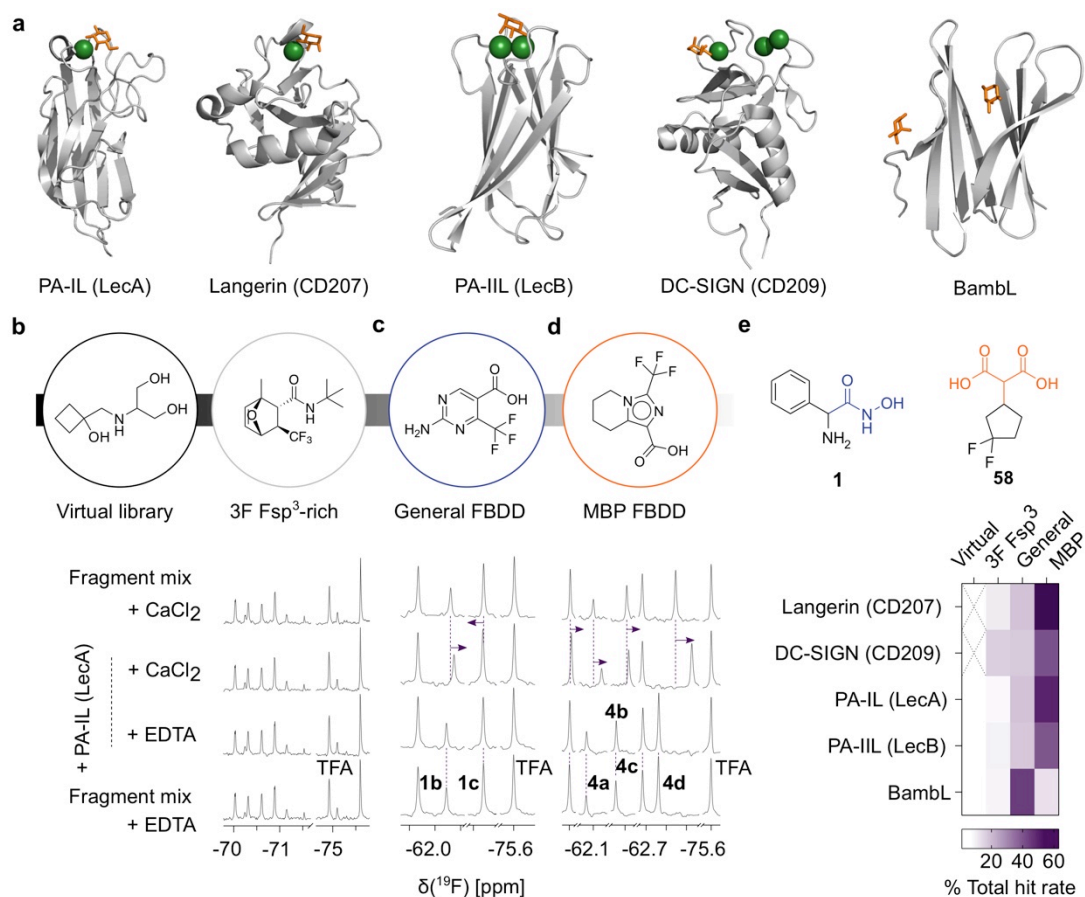
Figure 4.4-1 Fragment screening.

a Cartoon representation of monomer crystal structures showing the carbohydrate-recognition domain (CRD) bound to the monosaccharides (*orange sticks*) in a Ca^{2+}-dependent manner (*green spheres*): PA-IL (LecA, PDB: 4CP9), Langerin (PDB: 3P5D), DC-SIGN (PDB: 1SL4), LecB (PDB: 5A70). BambL does not require Ca^{2+} for carbohydrate binding (PDB: 3ZZV). All targets were screened using four libraries: **b** virtual and 3F Fsp^3, as well as **c** general and **d** MBP. Shown are examples of ^{19}F T_2-filtered NMR spectra of a fragment mixture in presence of 5 mM EDTA alone or with 20 µM LecA followed by competition with 30 mM $CaCl_2$. General and MBP-FBDD libraries revealed fragments undergoing a chemical shift perturbation (CSP) above 0.01 ppm (*violet arrow*) in presence of LecA and $CaCl_2$, indicating Ca^{2+}-dependent fragment binding. **f** Shown are the percentage values (%) of total hit rates of four libraries screened against five lectins. Compared to other libraries, MBPs improved the hit rates for Ca^{2+}-binding lectins (37-50%).importance of chemical fragment diversity in NMR screening and indicates that metal-coordinating fragments are the most prominent drug-like molecules of the Ca^{2+}-dependent lectins.

coordinating fragments are the most prominent drug-like molecules of the Ca^{2+}-dependent lectins.

Metal-Binding Pharmacophores Target Ca^{2+}-Dependent Lectins

To demonstrate the potential of metal-coordinating fragments in targeting the carbohydrate-protein interactions, we designed MBP library as described in **Supporting Information**. To this end, we subjected 142 commercial fragments for binding studies with one non- (BambL) and four metal-dependent lectins (LecA, LecB, Langerin and DC-SIGN). As expected, many MBPs showed a direct Ca^{2+} binding in ^{19}F NMR in absence of protein (**Figure 4.4-1d**, compounds **4a-d**). However, MBPs bound to metal-dependent lectins stronger in presence of protein and $CaCl_2$ as shown exemplarily for LecA (**Figure 4.4-1d**). This demonstrated MBPs target metal-dependent lectins in a Ca^{2+}-dependent manner.

Next, we derived hit rates from the fragment screening for the Ca^{2+}-dependent lectins and BambL to assess the success of our MBP-FBDD campaign (**Figure 4.4-1e**). Hit rates of 54.2% and 62.8% were observed for lectins with one Ca^{2+} ion in the carbohydrate-binding site, LecA and Langerin, respectively. Comparable hit rates were observed for lectins with two and three Ca^{2+} ions, LecB (42.8%) and DC-SIGN (42.5%). Notably, we observed a low binding of MBPs to a non-metal-dependent lectin BambL (7.7%) compared to the general library. Given that MBPs offer not only the metal complexation ability, a lower hit rate for BambL was expected. Next, we validated ^{19}F NMR hits for binding to ^{15}N-labeled LecA, LecB and DC-SIGN carbohydrate recognition domain (CRD) in ^{1}H-^{15}N HSQC/TROSY NMR experiments, identifying a malonate (**58**) as a potent scaffold for targeting LecA, LecB and DC-SIGN.

Taken together, fragment screening of four libraries demonstrated that MBP-FBDD campaign is a promising approach for targeting the carbohydrate-binding sites of Ca^{2+}-dependent lectins. As prominent hits for future development, we identified hydroxamates (**1**) and malonates (**58**) in screening of the general and MBP libraries, respectively. Further, both scaffolds were subjected to explorative structure-activity relationship (SAR) studies aiming to demonstrate the Ca^{2+}-dependency and selectivity of the scaffolds for metal-dependent lectins.

Hydroxamates as Ca^{2+}-dependent Inhibitors of LecA

Aiming to characterize hydroxamate **1**-LecA interaction and enhance its potency as well as LE, we employed five biophysical assays. A full discussion on the SAR study of **1** can be found in the **Supporting Information**. Briefly, we ranked commercial and 'in-house' synthesized derivatives of hydroxamate **1** by ^{15}N TROSY NMR. For this, we

quantified the changes in chemical shift perturbations (CSPs) of ^{15}N LecA in presence of 49 analogues using MeGal as a positive control (**Figure 4.4-S4a**). As result, hydroxamate derivatives promoted CSPs in ^{15}N LecA similar to **1** and MeGal as shown exemplarily with **35** (**Figure 4.4-2b**), suggesting that hydroxamates bound to the carbohydrate-binding site of LecA. Moreover, none of the marketed metalloproteinase inhibitors (**47**, **49**, **50**) bound to ^{15}N LecA (**Figures 4.4-S4b-c**), which indicates the presence of functional groups in marketed drugs that sterically prevent a beneficial coordination of Ca^{2+} in the carbohydrate-binding site of LecA. Taken together, we identified 18 analogues of hydroxamate **1** for targeting the carbohydrate-binding site of LecA with potentially higher affinities.

In parallel, we validated our initial ^{15}N TROSY NMR results using SPR, competitive ^{19}F NMR, fluorescence polarization (FP) and protein-observed ^{19}F (PrOF) NMR. Given a weak affinity of hydroxamates towards LecA, SPR did not allow ranking of hydroxamate derivatives (**Figure 4.4-S6**). Therefore, we designed a competitive T$_2$-filtered ^{19}F NMR using the hydroxamate derivative (**5**) as a ^{19}F reporter. To prove the utility of this assay, we evaluated the Ca^{2+}-dependency and selectivity of the hydroxamate-LecA interaction and compared it to other metal-dependent lectins, i.e. LecB, DC-SIGN and Langerin (**Figure 4.4-2c**). To test the Ca^{2+}-dependency of the **5** interactions with lectins, a T$_2$-filtered ^{19}F NMR spectrum of **5** was recorded with 2 mM EDTA, 10 mM CaCl$_2$ alone and with 10 µM lectins. Both Langerin and DC-SIGN showed a Ca^{2+}-independent binding to **5** (see **Supporting Information**), whereas only **5**-LecA interactions required Ca^{2+}, demonstrating the Ca^{2+}-dependency and early selectivity of hydroxamate-LecA interaction. Consequently, we used this assay to rank the hydroxamate derivatives as discussed in **Supporting Information**. Briefly, we observed a higher competition of **35** over the reporter **5** and other analogues in a competitive ^{19}F NMR, suggesting 35 as the most potent hydroxamate derivative in this assay (**Figure 4.4-S7**).

Next, we investigated the inhibitory properties and derived the affinities (K_d) of hydroxamates, which were active in NMR. In the FP assay the strongest inhibition compared to natural ligands of LecA, MeGal and GalNAc, was observed with linear (**35**) and cyclic (**20**) compounds (**Table S1**, **Figure 4.4-2d**).[21] To determine the binding site and to derive the affinities of hydroxamates, PrOF NMR using LecA labeled with 5-fluorotryptophanes (5FW) was employed. Since the residue W42 is located in the carbohydrate-binding site of LecA, the binding and affinities of weak ligands could be determined.[67] As result, the perturbation of W42 in presence of hydroxamates confirmed its binding to the orthosteric site of LecA (**Figure 4.4-S8a**). Moreover, changes in W42 were used to estimate the affinities (K_d) and the ligand efficiencies

(LE) of hydroxamates for LecA, as shown on example of **35** and **36** (K_d(**35**)=4.6±0.9 mM, Table S1, **Figures 2e** and **4.4-S8b-c**). Even though analogues showed a similar affinity to **35**, but many had a lower LE**.** Therefore, we selected only fragments with high affinity, LE and the % of inhibition values resulting in compounds (**35** and **36**, **Figure 4.4-2f**). Taken together, our SAR study on **1** demonstrated that hydroxamates target the carbohydrate-binding site of LecA. Furthermore, two series of scaffolds: linear (**2**, **6** and **35**) and cyclic (**20** and **21**) were identified as suitable starting points for future fragment growing strategies.

Crystal structure of LecA in complex with **35** at 1.8 Å resolution in space group $P2_12_12_1$ provides the first evidence for the interaction between a hydroxamates and a non-enzyme metal-binding protein. The electron density for **35** was detected in at the carbohydrate binding sites of LecA (**Figures 4.4-2g** and **4.4-S9**). Comparing the structures of LecA-**35** and LecA-galactose complexes illustrated galactose-mimicking properties of **35** (**Figure 4.4-2h**). In the LecA-**35** complex, two oxygen atoms in the hydroxamic acid functional group coordinate the Ca^{2+} ion via bidentate chelation, which is also typical for interactions between hydroxamate-based inhibitors and metal ions in metalloenzymes.[57] The two oxygen atoms form hydrogen bonds with N107 and D100, mimicking OH3 and OH4 of galactose. A water molecule (WAT2) was in contact with the nitrogen atom in the hydroxamic acid via a hydrogen bond. WAT2 mimics the role of galactose OH6 by forming hydrogen bonds with the side chains of H50, Q53, the main chain oxygen atom of P51 and water WAT1. Notably, the terminal benzyl ring forms CH-π interactions with P38, an interaction not observed for other LecA-targeting glycomimetics. To the best of our knowledge, this is the first report of a hydroxamate in complex with a non-metalloenzyme.

Taken together, we report the first evidence of hydroxamates targeting the carbohydrate-protein interactions. Moreover, we conducted several experiments highlighting the selectivity of hydroxamates towards LecA over mammalian Ca^{2+}-dependent lectins (Langerin/DC-SIGN). Therefore, we believe hydroxamates are promising molecules to design drug-like inhibitors against bacterial infections with a low probability of off-target effects.

Malonates Target Lectins With Multiple Ca^{2+} Ions

In the MBP-FBDD campaign, we discovered malonate **58** showing a potent binding to metal-dependent lectins namely LecA, LecB and DC-SIGN. To prove the utility of a target-oriented MBP-FBDD approach, we aimed to unravel its interactions with lectins having one, two and three Ca^{2+} ions in or close to the carbohydrate-binding site. Here, BambL served as control for a non-metal-dependent lectin expected not to interact
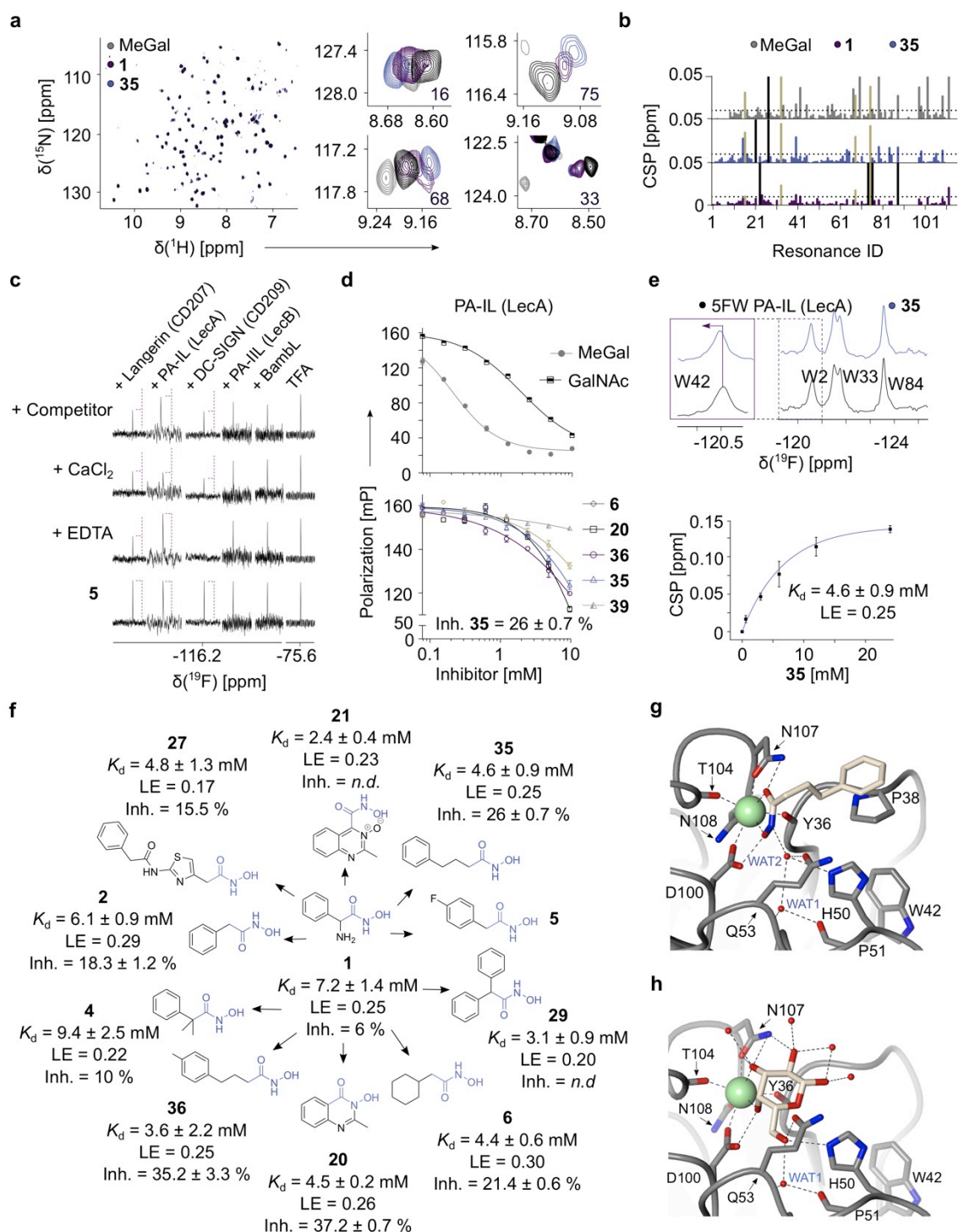
Figure 4.4-2 Hydroxamates as drug-like inhibitors of PA-IL (LecA).

a ^{15}N TROSY NMR spectra of [15]N LecA (*black*) in presence of 3 mM hydroxamates **1** (*violet*), **35** (*blue*) or 1 mM methyl-α-D-galactose (MeGal, *gray*). Hydroxamates perturbed the residues similarly to MeGal as shown for resonances 16, 33, 68 and 75. **b** Quantitative analysis of the chemical shift perturbations (CSPs) promoted by **1**, **35** and MeGal. The resonances **16**, **33**, **68** and **75** are highlighted (*green*). **c** Shown is a Ca[2+]-dependent binding of 0.01 mM LecA to **5** (*violet*) in the competitive [9]F T$_2$-filtered NMR, as evident by a decrease and recovery of the reporter **5** with (10 mM CaCl$_2$) or without (2 mM EDTA) Ca[2+], respectively. Moreover, **5**-LecA interaction showed a higher selectivity over other non- and metal-dependent lectins. **d** Competitive binding assay based on fluorescence polarization with MeGal and GalNAc (*top*) as positive controls for a strong and weak ligands, respectively. The hydroxamates **6**, **20**, **35**, **36**

and **39** (*bottom*) demonstrated LecA-dependent inhibitory properties. **e** Shown is PrOF NMR of 0.15 mM 5FW LecA with 3 mM **35** (*upper panel*) binding to the carbohydrate pocket as evident by a CSP of W42. One-site binding model of **35** (*bottom panel*) used to derive the affinity (K_d) by following CSPs of W42 (*n=3*). **f** Initial SAR study of **1** hydroxamate scaffold (*blue*). The affinities and IC_{50} values were estimated by PrOF NMR and FP assay respectively. The K_d values were used to calculate LE. **g** Crystal structure shows the interaction between LecA and **35**. Electron density for **35** is displayed at 1σ in green mesh. **h** Interaction between LecA and galactose (PDB: 1OKO).

with MBPs such as malonates. For this, the malonate **58** was subjected for docking and an initial SAR study using commercial analogues (**Figures 4.4-3a-b**, **Table 4.4-S3**). The results of both studies are discussed detailed in **Supporting Information**. Briefly, docking simulation proposed that the malonate **58** could coordinate multiple Ca^{2+} ions in DC-SIGN, LecA and LecB. Therefore, we investigated the Ca^{2+}-dependency and selectivity of malonate-lectin interactions.

To rule out the off-target effect, we added a competitor (LecA: 10 mM MeGal, LecB/BambL: 10 mM MeFuc, and DC-SIGN/Langerin: 30 mM D-mannose) expecting the ^{19}F peaks to recover if malonates target the carbohydrate-binding sites of lectins. Indeed, competitors displaced **61** efficiently, showing malonates targeted the carbohydrate-binding pocket of LecA, LecB and DC-SIGN. The details of Langerin and BambL interactions with **61** are discussed in **Supporting Information**. Cumulatively, our results proposed that malonates interact with metal-dependent lectins LecA, LecB and DC-SIGN in a Ca^{2+}-dependent manner.

Next, we aimed to confirm our observations for LecA and LecB and investigated whether the number of Ca^{2+} ions in the carbohydrate-binding site plays a role for the malonate binding. For LecA, PrOF NMR with 5FW-labeled lectin was used to gain information on the binding site of **58** analogues. Similar to hydroxamates, **58** perturbed W42 in PrOF NMR demonstrating its binding to the carbohydrate-binding site of 5FW LecA (**Figure 4.4-3d**). Next, we aimed to rank the impact of 13 analogues of **58** on W42 in order to prioritize scaffolds for future fragment evolution. As result, the scaffold **61** promoted the strongest perturbation of W42 (**Figures 4.4-S10b-c**). Altogether, malonates (**58** and **61**) interact with the carbohydrate-binding site of LecA bearing one Ca^{2+} ion, similar to hydroxamates.

For LecB as a lectin with two Ca^{2+} ions, a similar tendency in malonate binding to the carbohydrate-binding site was observed in ^{15}N TROSY NMR (**Figure 4.4-3e**). Here, we observed similar changes in residues of ^{15}N LecB resonances with **58** compared to the positive control MeFuc (**Figure 4.4-3f**). Therefore, we determined the affinities (K_d) and LE of **58** analogues using ^{15}N TROSY NMR, which were in a low mM range (K_d(**58**)=1.2±0.4 mM, **Figure 4.4-3g**). Interestingly, all structural derivatives of **58** showed comparable affinities for LecB (**Figures 4.4-3b** and **4.4-S13**). Notably, only **64**
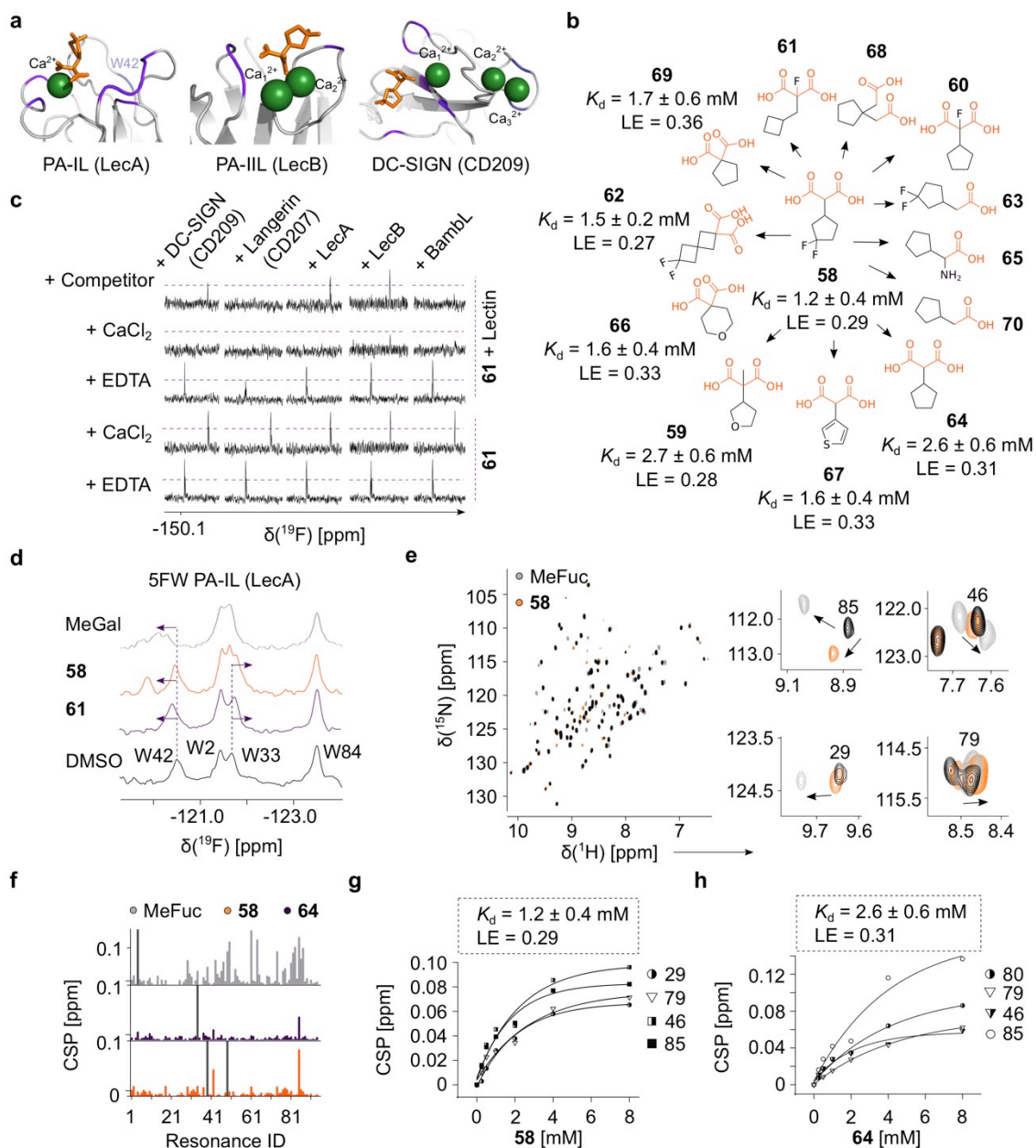
Figure 4.4-3 Malonates target lectins with one or multiple calcium ions.

a Docking poses of **58** with one and multiple Ca^{2+} in or near the carbohydrate-binding site of LecA, LecB and DC-SIGN (PDB: 4CP9, 1OXC and 2XR5). In LecA, the residues proposed to interact with lectins were highlighted (*violet*). **b** Initial SAR study of **58** malonate scaffold (*orange*), whereas its affinities for LecB were estimated in ^{15}N TROSY NMR. The K_d values were used to calculate LE. **c** To assess Ca^{2+}-dependency and selectivity of the malonate-lectin interaction, **61**-observed ^{19}F T_2-filtered NMR was used. **61** bound to 10 mM $CaCl_2$ given a perturbation of fluorine resonance. All lectins interacted with **61** in a Ca^{2+}-dependent manner. The presence of competitors: 10 mM MeGal (LecA), 10 mM MeFuc (LecB/BambL) and 30 mM D-mannose (DC-SIGN/Langerin), displaced **61** in LecA, LecB and DC-SIGN ECD verifying that malonates target the carbohydrate-binding site. **d** Shown is PrOF NMR of 0.15 mM 5FW LecA bound to 2 mM **58** or **61** as evident by CSPs of W42 and W33 (*arrow*). DMSO and MeGal served as negative and positive controls, respectively. **e** Protein fingerprint of ^{15}N LecB in ^{15}N TROSY NMR alone (*black*) and with 2 mM **58** (*orange*) or MeFuc (*gray*). **58** perturbed the resonances 85, 46, 29 and 79 similarly to MeFuc. **f** Shown are CSP plots of ^{15}N resonances in presence of 2 mM **58**, **64** and 1 mM MeFuc. **58** promoted lager CSPs in ^{15}N LecB compared to **64** supporting the docking pose 1. **g** and **h** One-site binding models of **58** and **64** derived by following CSPs of the resonances perturbed similar to MeFuc.

had a decrease in affinity (K_d(**64**)=2.6±0.6 mM) indicating the role of an electronegative group in binding to LecB (**Figure 4.4-3h**). This observation supported our docking study, where CF_2 group of **58** was predicted to interact with the protein surface through T98 (**Figures 4.4-S11a-b**). However, due to the lack of ^{15}N LecB protein assignment, co-crystallization studies are on going to support it.

Taken together, malonates interacted with lectins with one (LecA) and two (LecB) Ca^{2+} ions in the orthosteric site unlike hydroxamates. Moreover, our SAR study with LecA and LecB suggested that the selectivity of malonates is tunable as **61** had a stronger tendency to bind LecA, whereas **58** showed the preference for LecB. Finally, malonates have a higher tendency to bind to secondary sites in allosteric lectins, such as Langerin and BambL. Cumulatively, malonates offer a potential scaffold for design of orthosteric and allosteric inhibitors.

Malonates to Design Inhibitors of DC-SIGN (CD209)

Aiming to investigate the ability of malonates to bind lectins with three Ca^{2+} ions, we investigated malonate—DC-SIGN interaction using protein-observed ^{15}N HSQC NMR and a cell-based assay (cellFy).[62] Briefly, malonate **58** perturbed the resonances in the EPN motif coordinating Ca_1^{2+} and D367, whereas L321 and E324 near Ca_2^{2+} and Ca_3^{2+} showed weaker effects (**Figures 4.4-4a-b**). Quantitative analyses of the chemical shift perturbations (CSPs) caused by **58** revealed similar CSPs compared to the positive control D-mannose, suggesting that **58** mimics the carbohydrate-binding to DC-SIGN (**Figures 4.4-4c-d**). In particular, the interaction of **58** with the Ca_1^{2+}-binding site of DC-SIGN occurred through N365, N344 and F313 being in agreement with our docking results. Next, we derived the affinities (K_d) and LE values of **58** analogues for DC-SIGN CRD using ^{15}N HSQC NMR. The analogues of **58** showed a similar affinity (K_d(**58**)=1.2±0.5 mM, **Figure 4.4-4e**) and thus, three scaffold groups were defined as interchangeable (**58**, **62** and **69, Figures 4.4-4f** and **4.4-S16**). Similar to LecB, the compounds with an electronegative group on the ring (**58**, **62, 63** and **67**) were predominant and thus, in agreement with the predicted F313 interaction of DC-SIGN with CF_2 group in **58**. Compared to LecA and LecB, a methyl group in **59** was well tolerated in DC-SIGN CRD (**Figure 4.4-4g**). Therefore, this position is potentially suitable for future fragment growing to gain malonates specificity towards DC-SIGN. A full discussion on the SAR study of **58** and DC-SIGN can be found in the **Supporting Information**. Together, both computational and experimental data demonstrated that malonates could target the Ca_1^{2+} binding site of DC-SIGN similarly to D-mannose.
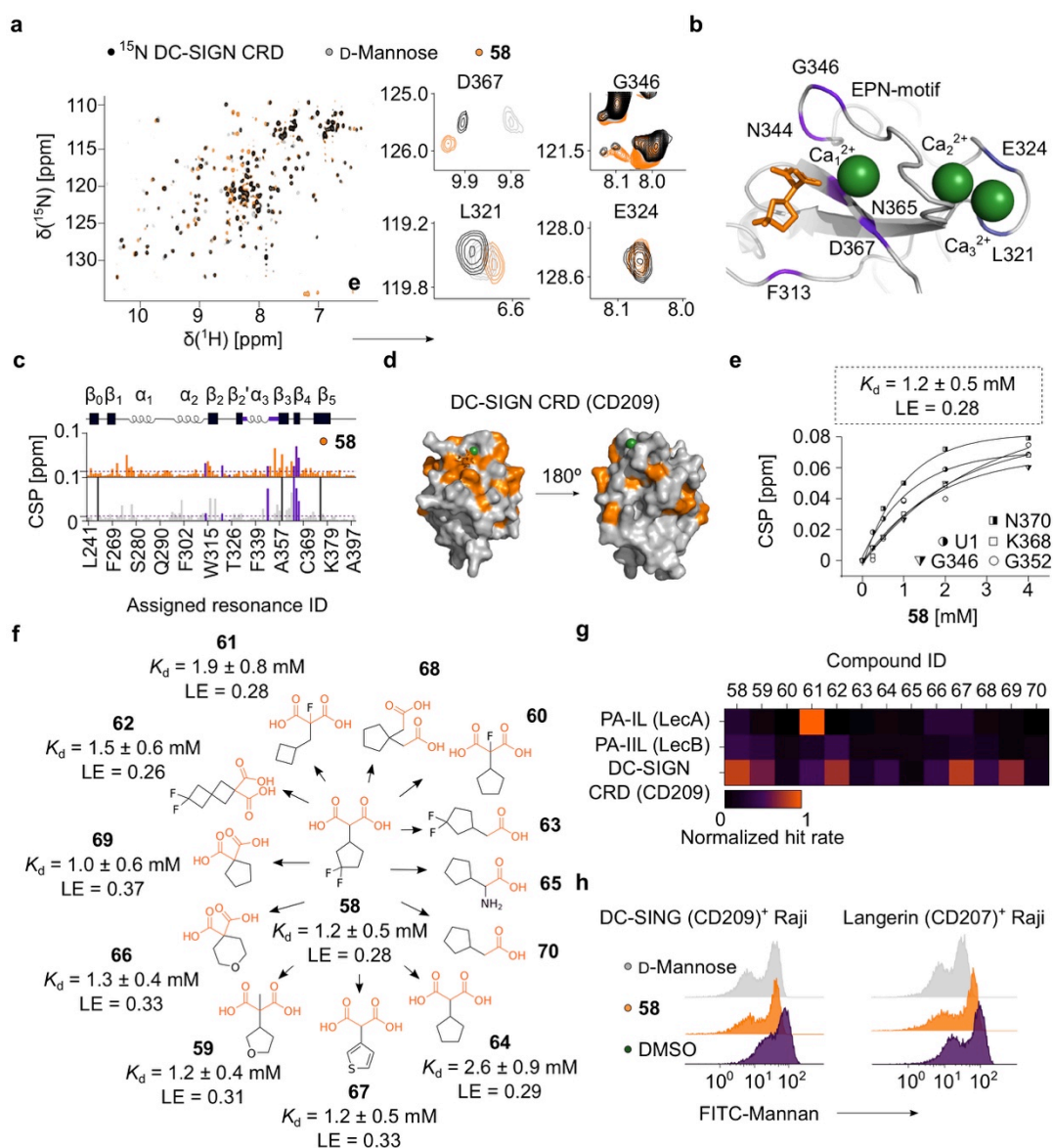
Figure 4.4-4 Malonates as inhibitors of DC-SIGN (CD209).

a Protein fingerprint of ^{15}N DC-SIGN CRD in ^{15}N HSQC NMR alone (*black*) and with 2 mM **58** (*orange*) or 30 mM D-mannose (*gray*). Similar to D-mannose, **58** perturbed the resonances D367, G346 and L321, but not E324 being in the Ca_1^{2+} and Ca_2^{2+}/Ca_3^{2+}-binding sites, respectively. These residues are highlighted in **b**, which shows a docked **58** in the Ca_1^{2+} site of DC-SIGN CRD (PDB: 2XR5). **c** Quantitative analysis of chemical shift perturbation (CSP) of ^{15}N DC-SIGN CRD in presence of **58** and D-mannose. Resonances D367, G346, L321 and E324 are highlighted (*violet*). A schematic representation of DC-SIGN CRD is shown on top. Malonates perturbed resonances similarly to D-mannose, i.e. in the long loop locating the EPN motif (*violet*). **d** The CSPs promoted in ^{15}N DC-SIGN CRD by **58** were transferred to DC-SIGN CRD (PDB: 2XR5). **e** One-site binding model of **58** to DC-SIGN CRD derived by following CSPs of the resonances perturbed in the long loop and β_4 sheet. **f** The affinities of **58** analogues for DC-SIGN CRD were estimated in ^1H-^{15}N HSQC NMR. The K_d values were used to calculate LE. **g** Heat map shows the 1:0 hit rate of **58** analogues for LecA, LecB and DC-SIGN and normalized to the effect of MeGal, MeFuc and D-mannose, respectively. **h** Shown are histograms of DC-SIGN[+] and Langerin[+] Raji cells bound to FITC-dextran in presence of **58** in a cell-based (cellFy) assay. DMSO and D-mannose used as a negative and positive controls, respectively. **58** inhibited FITC-dextran binding to DC-SIGN similarly to D-mannose, but not to Langerin.

To assess the activity of **58** in a physiologically more relevant cellular environment, we explored the ability of malonates to inhibit the carbohydrate binding in a cell-based system (cellFy).[62] For this, we assessed the binding of **58** to DC-SIGN[+] and Langerin[+] Raji cells by flow cytometry (**Figure 4.4-4h**). Indeed, we observed dose- and DC-SIGN-dependent binding of **58** to DC-SIGN[+] cells as well as negligible cytotoxicity at low concentrations. Most importantly, malonate **58** specifically bound to DC-SIGN[+] cells and not to Langerin-expressing cells, although both lectins share a common EPN motif. Together, our cellFy data demonstrated that malonates specifically inhibit DC-SIGN-carbohydrate interactions in the cellular context similar to D-mannose.

Together, our SAR study demonstrated that the selectivity of malonates is tunable. Moreover, the malonate scaffolds can serve for design of lectin inhibitors with one, two and three Ca^{2+} ions, as shown exemplarily for DC-SIGN, LecA and LecB, respectively. Thereby, **61** had a stronger tendency to interact with LecA, whereas LecB and DC-SIGN preferred **58** and **62** scaffolds, whereas the observation with **59** is valuable to gain malonates specificity towards DC-SIGN.

4.4.3. Conclusions

Identifying a suitable starting point for inhibitor development for the orthosteric site of lectins is challenging. Applying fragment screening, we approached this challenge focusing on metal-dependent lectins. As result, we identified metal-binding pharmacophores (MBPs) as novel scaffolds for targeting Ca^{2+}-dependent carbohydrate-protein interactions.

Recently, we reported the molecular basis of a catechol binding to LecA, where, two hydroxyls coordinated a Ca^{2+} ion in the carbohydrate-binding site of LecA and Langerin.[68] Here, we identified hydroxamic acids as additional potential drug-like molecules to target the Ca^{2+} ions in the carbohydrate-binding site of LecA. As hydroxamates are the most widely used inhibitors of various metalloproteins, to the best of our knowledge this is the first report demonstrating their interaction with non-metalloenzymes, namely Ca^{2+}-dependent lectins. The SAR studies with 49 analogs of **1** including marketed and commercial hydroxamates, revealed a sterically optimal presentation of a hydroxamic acid group. This knowledge will contribute to design specific hydroxamate-based inhibitors for LecA. Such inhibitors have the potential to interfere with *P. aeruginosa* biofilm integrity and increase susceptibility to antibiotic treatment of patients with cystic fibrosis.

The improved druggability rates for the orthosteric site of the Ca^{2+}-dependent lectins with catechols and hydroxamates encouraged us to initiate a target-oriented MBP-FBDD campaign. Employing target-oriented MBP-FBDD library, as we have shown a

4-fold improved hit rate for metal-dependent lectins having one (LecA and Langerin) and multiple Ca^{2+} ions (DC-SIGN and LecB). Notably, MBPs are likely to interact stronger with the proteins than with free ions in solution, as demonstrated in competitive ^{19}F NMR screening with $CaCl_2$. Among the non-fluorinated fragments, a malonate scaffold was identified as a potent drug-like ligand for Ca^{2+}-dependent lectins with one or multiple Ca^{2+}-ions. Notably, this malonate showed a low cytotoxicity (up to 10 mM) and selectivity for DC-SIGN[+], but not Langerin[+] Raji cells *in vitro* using cell-based assay (cellFy). Employing commercially available analogs of malonate **58**, we demonstrated that tuning the specificity of malonates towards lectins with one to multiple Ca^{2+} ions is possible. In particular, aside from the Ca^{2+}-dependent binding of a common malonic acid group, we discovered an additional interaction of the **58** electronegative CF_2 group with the protein surfaces of DC-SIGN (F313) and LecB by ^{15}N HSQC/TROSY NMR. This secondary interaction was not observed for LecA, which demonstrated a preference for a malonate with a different scaffold (**61**). We believe our discovery of malonate-based inhibitors will assist drug discovery campaigns aiming to target Ca^{2+}-dependent lectins.

Taken together, the potencies of both compound series to lectins are certainly lower than those observed for metalloenzymes, and still have to be improved. Nevertheless, hydroxamates and malonates already offer a promising selectivity towards LecA over other Ca^{2+}-dependent lectins. Meanwhile, selectivity of malonates is tunable towards the lectin of interest, as shown for DC-SIGN compared to LecA and LecB. Since the most potent hydroxamate (**35**) and malonate (**58**) scaffolds are still small with 179.2 and 208 Da, respectively, both serve as suitable starting points for developing the first drug-like inhibitors for the orthosteric sites of mammalian and bacterial Ca^{2+}-dependent lectins.

4.4.4. Funding

4.4.5. Acknowledgements

4.4.6. Conflict of interest statement

None declared.

4.4.7. References

1. Freeze, H. H.; Baum, L.; Varki, A., Glycans in Systemic Physiology. *In Essentials of Glycobiology*, Varki, A.; Cummings, R. D.; Esko, J. D.; Stanley, P.; Hart, G. W.; Aebi, M.; Darvill, A. G.; Kinoshita, T.; Packer, N. H.; Prestegard, J. H.; Schnaar, R. L.; Seeberger, P. H., Eds. Cold Spring Harbor Laboratory Press

Copyright 2015-2017 by The Consortium of Glycobiology Editors, La Jolla, California. All rights reserved.: Cold Spring Harbor (NY), 2015; pp 521-6.

2. Freeze, H. H.; Kinoshita, T.; Varki, A., Glycans in Acquired Human Diseases. *In Essentials of Glycobiology*, Varki, A.; Cummings, R. D.; Esko, J. D.; Stanley, P.; Hart, G. W.; Aebi, M.; Darvill, A. G.; Kinoshita, T.; Packer, N. H.; Prestegard, J. H.; Schnaar, R. L.; Seeberger, P. H., Eds. Cold Spring Harbor Laboratory Press

Copyright 2015-2017 by The Consortium of Glycobiology Editors, La Jolla, California. All rights reserved.: Cold Spring Harbor (NY), 2015; pp 583-95.

3. van Kooyk, Y.; Rabinovich, G. A., Protein-glycan interactions in the control of innate and adaptive immune responses. *Nature immunology* **2008**, *9 (6)*, 593-601.

4. Fuster, M. M.; Esko, J. D., The sweet and sour of cancer: glycans as novel therapeutic targets. Nature reviews. *Cancer* **2005**, *5 (7)*, 526-42.

5. Geijtenbeek, T. B.; Gringhuis, S. I., Signalling through C-type lectin receptors: shaping immune responses. Nature reviews. *Immunology* **2009**, *9 (7)*, 465-79.

6. Golovkine, G.; Reboud, E.; Huber, P., Pseudomonas aeruginosa Takes a Multi-Target Approach to Achieve Junction Breach. *Front Cell Infect Microbiol* **2018**, *7*, 532-532.

7. Ng, W. C.; Tate, M. D.; Brooks, A. G.; Reading, P. C., Soluble Host Defense Lectins in Innate Immunity to Influenza Virus. *Journal of Biomedicine and Biotechnology* **2012**, *2012*, 732191.

8. Gao, C.; Zeng, J.; Jia, N.; Stavenhagen, K.; Matsumoto, Y.; Zhang, H.; Li, J.; Hume, A. J.; Mühlberger, E.; van Die, I.; Kwan, J.; Tantisira, K.; Emili, A.; Cummings, R. D., SARS-CoV-2 Spike Protein Interacts with Multiple Innate Immune Receptors. *bioRxiv* **2020**, 2020.07.29.227462.

9. Munkley, J.; Elliott, D. J., Hallmarks of glycosylation in cancer. *Oncotarget* **2016**, *7 (23)*, 35478-89.

10. Valverde, P.; Ardá, A.; Reichardt, N.-C.; Jiménez-Barbero, J.; Gimeno, A., Glycans in drug discovery. *MedChemComm* **2019**, *10 (10)*, 1678-1691.

11. de Oliveira Figueiroa, E.; Albuquerque da Cunha, C. R.; Albuquerque, P. B. S.; de Paula, R. A.; Aranda-Souza, M. A.; Alves, M. S.; Zagmignan, A.; Carneiro-da-Cunha, M. G.; Nascimento da Silva, L. C.; Dos Santos Correia, M. T., Lectin-Carbohydrate

Interactions: Implications for the Development of New Anticancer Agents. *Current medicinal chemistry* **2017***, 24 (34)*, 3667-3680.

12. Rewar, S.; Mirdha, D.; Rewar, P., Treatment and Prevention of Pandemic H1N1 Influenza. *Annals of global health* **2015**, *81 (5)*, 645-53.

13. Springer, A. D.; Dowdy, S. F., GalNAc-siRNA Conjugates: Leading the Way for Delivery of RNAi Therapeutics. *Nucleic Acid Ther* **2018**, *28 (3)*, 109-118.

14. 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.; Keler, T., Induction of Antigen-Specific Immunity with a Vaccine Targeting NY-ESO-1 to the Dendritic Cell Receptor DEC-205. *Science Translational Medicine* **2014**, *6 (232)*, 232ra51-232ra51.

15. Osorio, F.; Reis e Sousa, C., Myeloid C-type Lectin Receptors in Pathogen Recognition and Host Defense. *Immunity* **2011**, *34 (5)*, 651-664.

16. Geijtenbeek, T. B.; Torensma, R.; van Vliet, S. J.; van Duijnhoven, G. C.; Adema, G. J.; van Kooyk, Y.; Figdor, C. G., Identification of DC-SIGN, a novel dendritic cell-specific ICAM-3 receptor that supports primary immune responses. *Cell* **2000**, *100 (5)*, 575-85.

17. Romani, N.; Clausen, B. E.; Stoitzner, P., Langerhans cells and more: langerin-expressing dendritic cell subsets in the skin. *Immunological reviews* **2010**, *234 (1)*, 120-41.

18. Hanske, J.; Aleksić, S.; Ballaschk, M.; Jurk, M.; Shanina, E.; Beerbaum, M.; Schmieder, P.; Keller, B. G.; Rademacher, C., Intradomain Allosteric Network Modulates Calcium Affinity of the C-Type Lectin Receptor Langerin. *Journal of the American Chemical Society* **2016**, *138 (37)*, 12176-12186.

19. Holla, A.; Skerra, A., Comparative analysis reveals selective recognition of glycans by the dendritic cell receptors DC-SIGN and Langerin. Protein engineering, design & selection : *PEDS* **2011**, *24 (9)*, 659-69.

20. Curtis, B. M.; Scharnowske, S.; Watson, A. J., Sequence and expression of a membrane-associated C-type lectin that exhibits CD4-independent binding of human immunodeficiency virus envelope glycoprotein gp120. *Proceedings of the National Academy of Sciences of the United States of America* **1992**, *89 (17)*, 8356-60.

21. Feinberg, H.; Mitchell, D. A.; Drickamer, K.; Weis, W. I., Structural Basis for Selective Recognition of Oligosaccharides by DC-SIGN and DC-SIGNR. *Science* (New York, N.Y.) **2001**, *294 (5549)*, 2163-2166.

22. Feinberg, H.; Taylor, M. E.; Razi, N.; McBride, R.; Knirel, Y. A.; Graham, S. A.; Drickamer, K.; Weis, W. I., Structural basis for langerin recognition of diverse

pathogen and mammalian glycans through a single binding site. *Journal of molecular biology* **2011**, *405 (4)*, 1027-39.

23. Geijtenbeek , T. B. H.; van Vliet , S. J.; Koppel , E. A.; Sanchez-Hernandez , M.; Vandenbroucke-Grauls , C. M. J. E.; Appelmelk , B.; van Kooyk , Y., Mycobacteria Target DC-SIGN to Suppress Dendritic Cell Function. *Journal of Experimental Medicine* **2002**, *197 (1)*, 7-17.

24. van den Berg, L. M.; Geijtenbeek, T. B., Antiviral immune responses by human langerhans cells and dendritic cells in HIV-1 infection. Advances in experimental medicine and biology 2013, 762, 45-70.

25. Alvarez, C. P.; Lasala, F.; Carrillo, J.; Muñiz, O.; Corbí, A. L.; Delgado, R., C-type lectins DC-SIGN and L-SIGN mediate cellular entry by Ebola virus in cis and in trans. *Journal of virology* **2002**, *76 (13)*, 6841-4.

26. Yang, Z.-Y.; Huang, Y.; Ganesh, L.; Leung, K.; Kong, W.-P.; Schwartz, O.; Subbarao, K.; Nabel, G. J., pH-Dependent Entry of Severe Acute Respiratory Syndrome Coronavirus Is Mediated by the Spike Glycoprotein and Enhanced by Dendritic Cell Transfer through DC-SIGN. *Journal of virology* **2004**, *78 (11)*, 5642-5650.

27. Lenza, M. P.; Oyenarte, I.; Diercks, T.; Quintana, J. I.; Gimeno, A.; Coelho, H.; Diniz, A.; Peccati, F.; Delgado, S.; Bosch, A.; Valle, M.; Millet, O.; Abrescia, N. G. A.; Palazón, A.; Marcelo, F.; Jiménez-Osés, G.; Jiménez-Barbero, J.; Ardá, A.; Ereño-Orbea, J., Structural Characterization of N-Linked Glycans in the Receptor Binding Domain of the SARS-CoV-2 Spike Protein and their Interactions with Human Lectins. *Angewandte Chemie (International ed. in English)* **2020**, *59 (52)*, 23763-23771.

28. Geijtenbeek, T. B.; Kwon, D. S.; Torensma, R.; van Vliet, S. J.; van Duijnhoven, G. C.; Middel, J.; Cornelissen, I. L.; Nottet, H. S.; KewalRamani, V. N.; Littman, D. R.; Figdor, C. G.; van Kooyk, Y., DC-SIGN, a dendritic cell-specific HIV-1-binding protein that enhances trans-infection of T cells. *Cell* **2000**, *100 (5)*, 587-97.

29. Clausen, B. E.; Kel, J. M., Langerhans cells: critical regulators of skin immunity? *Immunology and cell biology* **2010**, *88 (4)*, 351-60.

30. Fehres, C. M.; Duinkerken, S.; Bruijns, S. C.; Kalay, H.; van Vliet, S. J.; Ambrosini, M.; de Gruijl, T. D.; Unger, W. W.; Garcia-Vallejo, J. J.; van Kooyk, Y., Langerin-mediated internalization of a modified peptide routes antigens to early endosomes and enhances cross-presentation by human Langerhans cells. *Cellular & molecular immunology* **2017**, *14 (4)*, 360-370.

31. Singh, P. K.; Schaefer, A. L.; Parsek, M. R.; Moninger, T. O.; Welsh, M. J.; Greenberg, E. P., Quorum-sensing signals indicate that cystic fibrosis lungs are infected with bacterial biofilms. *Nature* **2000***, 407 (6805)*, 762-4.

32. Gagneux, P.; Varki, A., Evolutionary considerations in relating oligosaccharide diversity to biological function. *Glycobiology* **1999**, *9 (8)*, 747-755.

33. Kadam, R. U.; Bergmann, M.; Hurley, M.; Garg, D.; Cacciarini, M.; Swiderska, M. A.; Nativi, C.; Sattler, M.; Smyth, A. R.; Williams, P.; Camara, M.; Stocker, A.; Darbre, T.; Reymond, J. L., A glycopeptide dendrimer inhibitor of the galactose-specific lectin LecA and of Pseudomonas aeruginosa biofilms. *Angewandte Chemie (International ed. in English)* **2011**, *50 (45)*, 10631-5.

34. Sabin, C.; Mitchell, E. P.; Pokorná, M.; Gautier, C.; Utille, J.-P.; Wimmerová, M.; Imberty, A., Binding of different monosaccharides by lectin PA-IIL from Pseudomonas aeruginosa: Thermodynamics data correlated with X-ray structures. *FEBS letters* **2006**, *580 (3)*, 982-987.

35. Wagner, S.; Sommer, R.; Hinsberger, S.; Lu, C.; Hartmann, R. W.; Empting, M.; Titz, A., Novel Strategies for the Treatment of Pseudomonas aeruginosa Infections. *Journal of medicinal chemistry* **2016**, *59 (13)*, 5929-69.

36. Poole, J.; Day, C. J.; von Itzstein, M.; Paton, J. C.; Jennings, M. P., Glycointeractions in bacterial pathogenesis. *Nature Reviews Microbiology* **2018**, *16 (7)*, 440-452.

37. Van Breedam, W.; Pöhlmann, S.; Favoreel, H. W.; de Groot, R. J.; Nauwynck, H. J., Bitter-sweet symphony: glycan–lectin interactions in virus biology. *FEMS Microbiology Reviews* **2014**, *38 (4)*, 598-632.

38. Liu, Z.; Zhang, Q.; Peng, H.; Zhang, W.-z., Animal Lectins: Potential Antitumor Therapeutic Targets in Apoptosis. *Applied Biochemistry and Biotechnology* **2012**, *168 (3)*, 629-637.

39. Buzás, E. I.; György, B.; Pásztói, M.; Jelinek, I.; Falus, A.; Gabius, H. J., Carbohydrate recognition systems in autoimmunity. *Autoimmunity* **2006**, *39 (8)*, 691-704.

40. Hopkins, A. L.; Groom, C. R., The druggable genome. *Nature reviews. Drug discovery* **2002**, *1 (9)*, 727-30.

41. Imberty, A.; J, H. P., Structural Biology of Glycan Recognition. *In Essentials of Glycobiology*, Varki, A.; Cummings, R. D.; Esko, J. D.; Stanley, P.; Hart, G. W.; Aebi, M.; Darvill, A. G.; Kinoshita, T.; Packer, N. H.; Prestegard, J. H.; Schnaar, R. L.; Seeberger, P. H., Eds. Cold Spring Harbor Laboratory Press
Copyright 2015-2017 by The Consortium of Glycobiology Editors, La Jolla, California. All rights reserved.: Cold Spring Harbor (NY), 2015; pp 387-400.

42. Riccardi, L.; Genna, V.; De Vivo, M., Metal–ligand interactions in drug design. *Nature Reviews Chemistry* **2018**, *2 (7)*, 100-112.

43. Ernst, B.; Magnani, J. L., From carbohydrate leads to glycomimetic drugs. Nature reviews. *Drug discovery* **2009**, *8 (8)*, 661-677.

44. Wamhoff, E.-C.; Hanske, J.; Schnirch, L.; Aretz, J.; Grube, M.; Varón Silva, D.; Rademacher, C., 19F NMR-Guided Design of Glycomimetic Langerin Ligands. *ACS chemical biology* **2016**, *11 (9)*, 2407-2413.

45. Thépaut, M.; Guzzi, C.; Sutkeviciute, I.; Sattin, S.; Ribeiro-Viana, R.; Varga, N.; Chabrol, E.; Rojo, J.; Bernardi, A.; Angulo, J.; Nieto, P. M.; 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. *Journal of the American Chemical Society* **2013**, *135 (7)*, 2518-2529.

46. Garber, K. C.; Wangkanont, K.; Carlson, E. E.; Kiessling, L. L., A general glycomimetic strategy yields non-carbohydrate inhibitors of DC-SIGN. *Chemical communications* (Cambridge, England) **2010**, *46 (36)*, 6747-9.

47. Borrok, M. J.; Kiessling, L. L., Non-carbohydrate Inhibitors of the Lectin DC-SIGN. *Journal of the American Chemical Society* **2007**, *129 (42)*, 12780-12785.

48. Mangold, S. L.; Prost, L. R.; Kiessling, L. L., Quinoxalinone Inhibitors of the Lectin DC-SIGN. *Chemical science* **2012**, *3 (3)*, 772-777.

49. Sommer, R.; Wagner, S.; Rox, K.; Varrot, A.; Hauck, D.; Wamhoff, E.-C.; Schreiber, J.; Ryckmans, T.; Brunner, T.; Rademacher, C.; Hartmann, R. W.; Brönstrup, M.; Imberty, A.; Titz, A., Glycomimetic, Orally Bioavailable LecB Inhibitors Block Biofilm Formation of Pseudomonas aeruginosa. *Journal of the American Chemical Society* **2018**, *140 (7)*, 2537-2545.

50. Erlanson, D. A.; Fesik, S. W.; Hubbard, R. E.; Jahnke, W.; Jhoti, H., Twenty years on: the impact of fragments on drug discovery. N*ature Reviews Drug Discovery* **2016**, *15 (9)*, 605-619.

51. Congreve, M.; Carr, R.; Murray, C.; Jhoti, H., A 'rule of three' for fragment-based lead discovery? *Drug discovery today* **2003**, *8 (19)*, 876-7.

52. Hopkins, A. L.; Groom, C. R.; Alex, A., Ligand efficiency: a useful metric for lead selection. *Drug discovery today* **2004**, *9 (10)*, 430-1.

53. Baell, J. B.; Nissink, J. W. M., Seven Year Itch: Pan-Assay Interference Compounds (PAINS) in 2017—Utility and Limitations. *ACS chemical biology* **2018**, *13 (1)*, 36-44.

54. Hermann, J. C.; Chen, Y.; Wartchow, C.; Menke, J.; Gao, L.; Gleason, S. K.; Haynes, N. E.; Scott, N.; Petersen, A.; Gabriel, S.; Vu, B.; George, K. M.; Narayanan, A.; Li, S. H.; Qian, H.; Beatini, N.; Niu, L.; Gan, Q. F., Metal impurities cause false positives in high-throughput screening campaigns. *ACS Med Chem Lett* **2013**, *4 (2)*, 197-200.

55. Baell, J. B., Observations on screening-based research and some concerning trends in the literature. *Future Med Chem* **2010**, *2 (10)*, 1529-46.

56. Cohen, S. M., A Bioinorganic Approach to Fragment-Based Drug Discovery Targeting Metalloenzymes. *Accounts of Chemical Research* **2017**, *50 (8)*, 2007-2016.

57. Chen, A. Y.; Adamek, R. N.; Dick, B. L.; Credille, C. V.; Morrison, C. N.; Cohen, S. M., Targeting Metalloenzymes for Therapeutic Intervention. *Chemical reviews* **2019**, *119 (2)*, 1323-1455.

58. Marte, F.; Sankar, P.; Cassagnol, M., Captopril. In StatPearls, StatPearls Publishing Copyright © 2021, StatPearls Publishing LLC.: Treasure Island (FL), 2021.

59. Marks, P. A.; Breslow, R., Dimethyl sulfoxide to vorinostat: development of this histone deacetylase inhibitor as an anticancer drug. *Nature biotechnology* **2007**, *25 (1)*, 84-90.

60. Erlanson, D. A., Introduction to Fragment-Based Drug Discovery. *In Fragment-Based Drug Discovery and X-Ray Crystallography*, Davies, T. G.; Hyvönen, M., Eds. Springer Berlin Heidelberg: Berlin, Heidelberg, 2012; pp 1-32.

61. Joachim, I.; Rikker, S.; Hauck, D.; Ponader, D.; Boden, S.; Sommer, R.; Hartmann, L.; Titz, A., Development and optimization of a competitive binding assay for the galactophilic low affinity lectin LecA from Pseudomonas aeruginosa. *Organic & biomolecular chemistry* **2016**, *14 (33)*, 7933-48.

62. Schulze, J.; Baukmann, H.; Wawrzinek, R.; Fuchsberger, F. F.; Specker, E.; Aretz, J.; Nazaré, M.; Rademacher, C., CellFy: A Cell-Based Fragment Screen against C-Type Lectins. *ACS chemical biology* **2018**, *13 (12)*, 3229-3235.

63. Aretz, J.; Wamhoff, E.-C.; Hanske, J.; Heymann, D.; Rademacher, C., Computational and Experimental Prediction of Human C-Type Lectin Receptor Druggability. *Frontiers in Immunology* **2014**, *5* (323).

64. Troelsen, N. S.; Shanina, E.; Gonzalez-Romero, D.; Danková, D.; Jensen, I. S. A.; Śniady, K. J.; Nami, F.; Zhang, H.; Rademacher, C.; Cuenda, A.; Gotfredsen, C. H.; Clausen, M. H., The 3F Library: Fluorinated Fsp3-Rich Fragments for Expeditious 19F NMR Based Screening. *Angewandte Chemie International Edition* **2020**, *59 (6)*, 2204-2210.

65. Shanina E., K. S., Lal K., Seeberger P.H., Imberty A. and Rademacher C., Identification of druggable allosteric pockets in β-propeller lectins. *In submission* **2021**.

66. Hajduk, P. J.; Huth, J. R.; Fesik, S. W., Druggability indices for protein targets derived from NMR-based screening data. *Journal of medicinal chemistry* **2005**, 48 (7), 2518-25.

67. Shanina, E.; Siebs, E.; Zhang, H.; Varón Silva, D.; Joachim, I.; Titz, A.; Rademacher, C., Protein-observed 19F NMR of LecA from Pseudomonas aeruginosa. *Glycobiology* **2021**.

68. Kuhaudomlarp, S.; Siebs, E.; Shanina, E.; Topin, J.; Joachim, I.; da Silva Figueiredo Celestino Gomes, P.; Varrot, A.; Rognan, D.; Rademacher, C.; Imberty, A.; Titz, A., Non-Carbohydrate Glycomimetics as Inhibitors of Calcium(II)-binding Lectins. *Angewandte Chemie (International ed. in English)* **2020**.

5. Final conclusions and perspectives

Carbohydrate-lectin interactions mediate important biological processes including pathogenic infections and cellular communication. In case of pathogen recognition, bacteria and viruses developed strategies to invade the host by mimicking or exploiting the host carbohydrates (glycans).[174] Therefore, lectins emerged as auspicious targets to design inhibitors to combat viral and bacterial pathogens.[117, 149] However, drug-like inhibitors of lectins are still limited. Moreover, the interaction of lectins with drug-like fragments often demonstrates inherently low affinities, which requires highly sensitive methods for its detection. Since the work of Fesik et al. using ^1H-^{15}N HSQC NMR,[79] NMR methods became prevalent for small molecule discovery efforts in industry and academic biomedical research.[175] In particular, ^{19}F NMR offers the necessary 'sensitivity' to detect weak binding events. However, ^{19}F NMR-based techniques employing carbohydrates and lectins are still sparse.[176] Consequently, broadening the set of available methods for lectins would improve our understanding of carbohydrate-lectins interactions as well as accelerate the discovery of drug-like inhibitors for lectins. Cumulatively, this thesis focused on two aims: 1) establishing two broadly applicable protein- and ligand-observed ^{19}F NMR methods for characterization of carbohydrate- or fragment-lectin interaction, as well as 2) identifying novel chemotypes for mammalian and bacterial lectins using the fragment-based drug design (FBDD) approach.

In the first chapter, protein-observed ^{19}F (PrOF) NMR was demonstrated as a suitable method for assessing structural information and the binding affinities of small molecules to lectins. For this, PrOF NMR was explored for a bacterial lectin from *Pseudomonas aeruginosa* called PA-IL (LecA), which plays an important role in the formation of biofilms in chronic infections.[114] Therefore, development of inhibitors to disrupt LecA-mediated biofilms is desired, but discovery of drug-like molecules is difficult due to its weak affinities. Using this case study, it was demonstrated that tryptophanes in LecA could be ^{19}F-labeled using 5-fluoroindole (5FI) as a precursor of 5-fluorotryptophan (5FW). Unlike ligand-observed NMR techniques (e.g. STD NMR[177]), PrOF NMR showed an advantage by providing structural information on the binding site and affinities (K_d) of various ligands such as Ca^{2+}, D-galactose and D-GalNAc. Furthermore, ITC and competitive fluorescence polarization (FP) assay supported PrOF NMR data showing that 5FW LecA preserved its activity and the ligand preference similarly to LecA. In contrast to FP assay, PrOF NMR with 5FW LecA demonstrated a remarkable sensitivity for the identification of weak small molecule ligands as shown on example of D-GalNAc. Cumulatively, PrOF NMR could be applied

to LecA in order to assess the structural and binding information. Since tryptophan is by far the most frequently found amino acid in the carbohydrate-binding site of various lectins, this method can be employed on other lectins.[178] Moreover, PrOF NMR is fast in data acquisition and analysis unlike other protein-observed NMR methods such as HSQC NMR. Finally, PrOF NMR proved to be valuable for discovery of fragment-lectin interactions (**Chapters 4.3** and **4.4**), this method will support the future drug-discovery campaigns that aim to develop drug-like inhibitors for lectins.

The next aim was to expand the set of ligand-observed ^{19}F NMR methods for carbohydrate-lectin interactions. Therefore, ^{19}F NMR was employed in the second chapter to screen a small library of fluorinated glycans (F-glycans), i.e. Lewis type 2 antigens, which could serve as potential ^{19}F reporter molecules to discover new fragment-protein interactions. For this study, F-glycans were rapidly available using automated glycan assembly (AGA) and followed by its evaluation in ^{19}F NMR binding studies using mammalian and bacterial lectins as well as enzymes. The screening was performed in a high-throughput manner. Given that many reports have shown how to ^{19}F-label monosaccharides selectively,[179] new classes of F-glycans can be easily produced in AGA. Since ^{19}F NMR signals show the high dispersion, libraries of F-glycans with various chemical shifts can be designed to increase the throughput of this approach.[180] Furthermore, F-glycans can be used to derive its binding affinities (K_d) in ^{19}F NMR, as well as could serve as reporter molecules for competitive ^{19}F NMR.[62,101] Finally, F-glycans serve allowed to monitor the enzymatic reactions in real time allowing a fast check of the enzyme activity by NMR and the characterization of enzymatic processes.[181] Cumulatively, F-glycans are valuable probes for the discovery of the carbohydrate-lectin interactions and a potential application as ^{19}F reporters in drug discovery campaigns.

Both ^{19}F NMR methods were employed to discover drug-like molecules for targeting the allosteric and carbohydrate-binding sites of lectins. In the third chapter, a bacterial lectin BambL from the pathogen *Burkholderia ambifaria* was investigated for the druggable pockets. For this, ^{19}F NMR screening of a fluorinated fragment library was performed against BambL. As result of a high hit rate, the presence of druggable secondary sites in BambL was suggested. Notably, fragment screening has been reported previously as a successful approach for the identification of secondary sites in various proteins including a C-type lectin receptors namely DC-SIGN.[107,182] Interestingly, the fragment binding in PrOF NMR using 5FW BambL induced the perturbation of 5FW resonances in the carbohydrate-binding site region. Therefore, site-directed mutagenesis was used to introduce changes within the predicted site and the carbohydrate pocket to test the communication between the secondary and

orthosteric sites. Indeed, the perturbations of ^{15}N BambL backbone resonances in distal regions from the mutation sites were observed in TROSY NMR. This observation allowed us to assume the presence of a communication pathway between two spatially distant binding sites in BambL as such behavior is typical for allosteric proteins.[183-184] Finally, ^{19}F NMR revealed similar hit rates in other fungal and bacterial β-propeller lectins namely AFL and RSL from *A. fumigatus* and *R. solanacearum*, respectively. Given these lectins show similarities in structure and druggability hit rates to BambL, am allostery could also be present in other β-propeller lectins. Taken together, our findings demonstrated the druggability of β-propeller lectins suggesting that the secondary sites could be used to design allosteric inhibitors for bacterial and fungal β-propeller lectins.

In the forth chapter, the goal was to identify novel chemotypes for targeting the carbohydrate-binding site of the clinically relevant Ca^{2+}-dependent bacterial (LecA, LecB) and mammalian lectins (DC-SIGN, Langerin). Previously, identifying a suitable starting point for inhibitor development for the orthosteric site of lectins has been challenging. Therefore, many lectin inhibitors are based on carbohydrate scaffolds. Here, the NMR screening of three fragment libraries was applied to discover drug-like molecules for targeting the carbohydrate-binding site of Ca^{2+}-dependent lectins. As a result, metal-binding pharmacophores (MBP) were identified. Recently, A. Titz and coauthors reported the molecular basis of an MBP namely catechol binding to LecA, where it coordinated a Ca^{2+} ion in the carbohydrate-binding site of LecA and Langerin.[185] This discovery was encouraging to initiate a target-oriented MBP-FBDD campaign.

Experimental NMR-based screening of 142 commercial MBPs revealed several promising scaffolds for selective targeting of mammalian and bacterial lectins with one (LecA and Langerin) and multiple Ca^{2+}+ ions (DC-SIGN and LecB). In this thesis, a hydroxamic and malonic acids were identified as additional potential drug-like molecules to target Ca^{2+} ions in the carbohydrate-binding site of LecA. Since hydroxamates are the most widely used inhibitors of matrix metalloproteinases,[186] the SAR studies with synthesized, marketed and commercial hydroxamates, revealed a sterically optimal presentation of a hydroxamic acid group, which could contribute to design specific hydroxamate-based inhibitors for LecA. Such inhibitors have the potential to interfere with *P. aeruginosa* biofilm integrity and increase susceptibility to antibiotic treatment of patients with cystic fibrosis. Additionally, a malonate scaffold **58** was identified as a potent drug-like ligand for Ca^{2+}-dependent lectins such as LecA, LecB and DC-SIGN. Notably, malonates showed a low cytotoxicity (up to 10 mM) and selectivity for DC-SIGN$^+$, but not Langerin$^+$ Raji cells *in vitro* using cell-based assay

(cellFy). Aside from the Ca^{2+}-dependent binding of a common malonic acid group, additional interactions of the **58** electronegative CF_2 group with DC-SIGN (F313) and LecB were discovered by ^1H-^{15}N TROSY/HSQC NMR. Such a secondary interaction was not observed for LecA, which demonstrated a preference for a malonate with a different scaffold (**61**). Certainly, the potencies and selectivity of both compound series to lectins are still lower than those observed for metalloenzymes, and still have to be improved. Nevertheless, MBPs provide suitable starting points for developing the first drug-like inhibitors for the orthosteric sites of mammalian and bacterial Ca^{2+}-dependent lectins.

Taken together, ^{19}F NMR has continued to prove its versatility in several areas of chemical biology including carbohydrate-lectin interactions. Advances in the methods for ^{19}F labeling of proteins, as well as the synthesis of fluorinated monosaccharides and complex oligosaccharides have been the driver for the development of ^{19}F NMR methods for lectins reported in this thesis. Undoubtedly, PrOF NMR has proven its utility for discovery of fragment binding to the carbohydrate-binding site of lectins, whereas the combination of AGA with ^{19}F NMR is likely to amplify the application of F-glycans in the future. Notably, ^{19}F and TROSY NMR methods allowed us to assess the druggability of bacterial and fungal lectins identifying several druggable sites, which could be used for design of potential allosteric inhibitors. Combining NMR with other methods (X-ray, SPR and FP), metal-binding pharmacophores were identified as the first drug-like fragments targeting the orthosteric site of mammalian and bacterial Ca^{2+}-dependent lectins. These discoveries will assist many drug discovery campaigns aiming to develop lectin inhibitors in the future.

6. References

1.   Macarron, R.; Banks, M. N.; Bojanic, D.; Burns, D. J.; Cirovic, D. A.; Garyantes, T.; Green, D. V.; Hertzberg, R. P.; Janzen, W. P.; Paslay, J. W.; Schopfer, U.; Sittampalam, G. S., Impact of high-throughput screening in biomedical research. *Nature reviews. Drug discovery* **2011,** *10* (3), 188-95.

2.   Price, A. J.; Howard, S.; Cons, B. D., Fragment-based drug discovery and its application to challenging drug targets. *Essays in biochemistry* **2017,** *61* (5), 475-484.

3.   Hopkins, A. L.; Groom, C. R., The druggable genome. *Nature reviews. Drug discovery* **2002,** *1* (9), 727-30.

4.   Bollag, G.; Hirth, P.; Tsai, J.; Zhang, J.; Ibrahim, P. N.; Cho, H.; Spevak, W.; Zhang, C.; Zhang, Y.; Habets, G.; Burton, E. A.; Wong, B.; Tsang, G.; West, B. L.; Powell, B.; Shellooe, R.; Marimuthu, A.; Nguyen, H.; Zhang, K. Y. J.; Artis, D. R.; Schlessinger, J.; Su, F.; Higgins, B.; Iyer, R.; D'Andrea, K.; Koehler, A.; Stumm, M.; Lin, P. S.; Lee, R. J.; Grippo, J.; Puzanov, I.; Kim, K. B.; Ribas, A.; McArthur, G. A.; Sosman, J. A.; Chapman, P. B.; Flaherty, K. T.; Xu, X.; Nathanson, K. L.; Nolop, K., Clinical efficacy of a RAF inhibitor needs broad target blockade in BRAF-mutant melanoma. *Nature* **2010,** *467* (7315), 596-599.

5.   Hajduk, P. J.; Greer, J., A decade of fragment-based drug design: strategic advances and lessons learned. *Nature reviews. Drug discovery* **2007,** *6* (3), 211-9.

6.   Chessari, G.; Woodhead, A. J., From fragment to clinical candidate—a historical perspective. *Drug discovery today* **2009,** *14* (13), 668-675.

7.   Murray, C. W.; Blundell, T. L., Structural biology in fragment-based drug design. *Current opinion in structural biology* **2010,** *20* (4), 497-507.

8.   Erlanson, D. A.; Fesik, S. W.; Hubbard, R. E.; Jahnke, W.; Jhoti, H., Twenty years on: the impact of fragments on drug discovery. *Nature Reviews Drug Discovery* **2016,** *15* (9), 605-619.

9.   Roughley, S. D.; Hubbard, R. E., How well can fragments explore accessed chemical space? A case study from heat shock protein 90. *Journal of medicinal chemistry* **2011,** *54* (12), 3989-4005.

10.   Harner, M. J.; Frank, A. O.; Fesik, S. W., Fragment-based drug discovery using NMR spectroscopy. *Journal of biomolecular NMR* **2013,** *56* (2), 65-75.

11.   Hajduk, P. J.; Huth, J. R.; Tse, C., Predicting protein druggability. *Drug discovery today* **2005,** *10* (23-24), 1675-82.

12.   Hajduk, P. J.; Huth, J. R.; Fesik, S. W., Druggability indices for protein targets derived from NMR-based screening data. *Journal of medicinal chemistry* **2005,** *48* (7), 2518-25.

13.   Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J., Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Advanced Drug Delivery Reviews* **1997,** *23* (1), 3-25.

14.   Congreve, M.; Carr, R.; Murray, C.; Jhoti, H., A 'rule of three' for fragment-based lead discovery? *Drug discovery today* **2003,** *8* (19), 876-7.

15.   Lau, W. F.; Withka, J. M.; Hepworth, D.; Magee, T. V.; Du, Y. J.; Bakken, G. A.; Miller, M. D.; Hendsch, Z. S.; Thanabal, V.; Kolodziej, S. A.; Xing, L.; Hu, Q.; Narasimhan, L. S.; Love, R.; Charlton, M. E.; Hughes, S.; van Hoorn, W. P.; Mills, J. E., Design of a multi-purpose fragment screening library using molecular complexity and orthogonal diversity metrics. *Journal of Computer - Aided Molecular Design* **2011,** *25* (7), 621-36.

16.   Doak, B. C.; Morton, C. J.; Simpson, J. S.; Scanlon, M. J., Design and Evaluation of the Performance of an NMR Screening Fragment Library. *Australian Journal of Chemistry* **2013,** *66* (12), 1465-1472.

17.   Albert, J. S.; Blomberg, N.; Breeze, A. L.; Brown, A. J.; Burrows, J. N.; Edwards, P. D.; Folmer, R. H.; Geschwindner, S.; Griffen, E. J.; Kenny, P. W.; Nowak, T.; Olsson, L. L.; Sanganee, H.; Shapiro, A. B., An integrated approach to fragment-based lead generation: philosophy, strategy and case studies from AstraZeneca's drug discovery programmes. *Current topics in medicinal chemistry* **2007,** *7* (16), 1600-29.

18.   Jhoti, H.; Williams, G.; Rees, D. C.; Murray, C. W., The 'rule of three' for fragment-based drug discovery: where are we now? *Nature Reviews Drug Discovery* **2013,** *12* (8), 644-644.

19.   Leach, A. R.; Hann, M. M., Molecular complexity and fragment-based drug discovery: ten years on. *Current opinion in chemical biology* **2011,** *15* (4), 489-96.

20.   Hopkins, A. L.; Groom, C. R.; Alex, A., Ligand efficiency: a useful metric for lead selection. *Drug discovery today* **2004,** *9* (10), 430-1.

21.   Hajduk, P. J., Fragment-Based Drug Design: How Big Is Too Big? *Journal of medicinal chemistry* **2006,** *49* (24), 6972-6976.

22.   Kuntz, I. D.; Chen, K.; Sharp, K. A.; Kollman, P. A., The maximal affinity of ligands. *Proceedings of the National Academy of Sciences* **1999,** *96* (18), 9997-10002.

23.   Leeson, P. D.; St-Gallay, S. A., The influence of the 'organizational factor' on compound quality in drug discovery. *Nature reviews. Drug discovery* **2011,** *10* (10), 749-65.

24.   Hann, M., Molecular obesity, potency and other addictions in drug discovery. *MedChemComm* **2011,** *2*, 349-355.

25.   Hajduk, P. J.; Greer, J., A decade of fragment-based drug design: strategic advances and lessons learned. *Nature Reviews Drug Discovery* **2007,** *6* (3), 211-219.

26.   Baell, J.; Walters, M. A., Chemistry: Chemical con artists foil drug discovery. *Nature* **2014,** *513* (7519), 481-483.

27.   Baell, J. B.; Holloway, G. A., New Substructure Filters for Removal of Pan Assay Interference Compounds (PAINS) from Screening Libraries and for Their Exclusion in Bioassays. *Journal of medicinal chemistry* **2010,** *53* (7), 2719-2740.

28.   Hermann, J. C.; Chen, Y.; Wartchow, C.; Menke, J.; Gao, L.; Gleason, S. K.; Haynes, N. E.; Scott, N.; Petersen, A.; Gabriel, S.; Vu, B.; George, K. M.; Narayanan, A.; Li, S. H.; Qian, H.; Beatini, N.; Niu, L.; Gan, Q. F., Metal impurities cause false positives in high-throughput screening campaigns. *ACS Med Chem Lett* **2013,** *4* (2), 197-200.

29.   Baell, J. B.; Nissink, J. W. M., Seven Year Itch: Pan-Assay Interference Compounds (PAINS) in 2017—Utility and Limitations. *ACS chemical biology* **2018,** *13* (1), 36-44.

30.   Baell, J. B., Observations on screening-based research and some concerning trends in the literature. *Future Med Chem* **2010,** *2* (10), 1529-46.

31.   Friedberg, F., Effects of Metal Binding on Protein Structure. *Quarterly Reviews of Biophysics* **2009,** *7* (1), 1-33.

32.   Hajduk, P. J.; Sheppard, G.; Nettesheim, D. G.; Olejniczak, E. T.; Shuker, S. B.; Meadows, R. P.; Steinman, D. H.; Carrera, G. M.; Marcotte, P. A.; Severin, J.; Walter, K.; Smith, H.; Gubbins, E.; Simmer, R.; Holzman, T. F.; Morgan, D. W.; Davidsen, S. K.; Summers, J. B.; Fesik, S. W., Discovery of Potent Nonpeptide Inhibitors of Stromelysin Using SAR by NMR. *Journal of the American Chemical Society* **1997,** *119* (25), 5818-5827.

33.   Olejniczak, E.; Hajduk, P.; Marcotte, P.; Nettesheim, D.; Meadows, R.; Edalji, R.; Holzman, T.; Fesik, S., Stromelysin Inhibitors Designed from Weakly Bound Fragments: Effects of Linking and Cooperativity. *Journal of the American Chemical Society* **1997,** *119*, 5828-5832.

34.   Cohen, S. M., A Bioinorganic Approach to Fragment-Based Drug Discovery Targeting Metalloenzymes. *Accounts of Chemical Research* **2017,** *50* (8), 2007-2016.

35.   Jacobsen, J. A.; Fullagar, J. L.; Miller, M. T.; Cohen, S. M., Identifying chelators for metalloprotein inhibitors using a fragment-based approach. *Journal of medicinal chemistry* **2011,** *54* (2), 591-602.

36.   Jacobsen, F. E.; Lewis, J. A.; Cohen, S. M., The design of inhibitors for medicinally relevant metalloproteins. *ChemMedChem* **2007,** *2* (2), 152-71.

37.   White, R. J.; Margolis, P. S.; Trias, J.; Yuan, Z., Targeting metalloenzymes: a strategy that works. *Current opinion in pharmacology* **2003,** *3* (5), 502-7.

38.   Yang, Y.; Hu, X. Q.; Li, Q. S.; Zhang, X. X.; Ruan, B. F.; Xu, J.; Liao, C., Metalloprotein Inhibitors for the Treatment of Human Diseases. *Current topics in medicinal chemistry* **2016,** *16* (4), 384-96.

39.   Credille, C. V.; Chen, Y.; Cohen, S. M., Fragment-Based Identification of Influenza Endonuclease Inhibitors. *Journal of medicinal chemistry* **2016,** *59* (13), 6444-6454.

40.   Perez, C.; Li, J.; Parlati, F.; Rouffet, M.; Ma, Y.; Mackinnon, A. L.; Chou, T.-F.; Deshaies, R. J.; Cohen, S. M., Discovery of an Inhibitor of the Proteasome Subunit Rpn11. *Journal of medicinal chemistry* **2017,** *60* (4), 1343-1361.

41.   Akritopoulou-Zanze, I.; Hajduk, P. J., Kinase-targeted libraries: the design and synthesis of novel, potent, and selective kinase inhibitors. *Drug discovery today* **2009,** *14* (5-6), 291-7.

42.   Ostermann, N.; Ruedisser, S.; Ehrhardt, C.; Breitenstein, W.; Marzinzik, A.; Jacoby, E.; Vangrevelinghe, E.; Ottl, J.; Klumpp, M.; Hartwieg, J. C.; Cumin, F.; Hassiepen, U.; Trappe, J.; Sedrani, R.; Geisse, S.; Gerhartz, B.; Richert, P.; Francotte, E.; Wagner, T.; Krömer, M.; Kosaka, T.; Webb, R. L.; Rigel, D. F.; Maibaum, J.; Baeschlin, D. K., A novel class of oral direct renin inhibitors: highly potent 3,5-disubstituted piperidines bearing a tricyclic p3-p1 pharmacophore. *Journal of medicinal chemistry* **2013,** *56* (6), 2196-206.

43.   Agrawal, A.; Johnson, S. L.; Jacobsen, J. A.; Miller, M. T.; Chen, L.-H.; Pellecchia, M.; Cohen, S. M., Chelator fragment libraries for targeting metalloproteinases. *ChemMedChem* **2010,** *5* (2), 195-199.

44.   An, W. F., Fluorescence-based assays. *Methods in molecular biology (Clifton, N.J.)* **2009,** *486*, 97-107.

45.   Hartshorn, M. J.; Murray, C. W.; Cleasby, A.; Frederickson, M.; Tickle, I. J.; Jhoti, H., Fragment-Based Lead Discovery Using X-ray Crystallography. *Journal of medicinal chemistry* **2005,** *48* (2), 403-413.

46.   Navratilova, I.; Hopkins, A. L., Fragment Screening by Surface Plasmon Resonance. *ACS Medicinal Chemistry Letters* **2010,** *1* (1), 44-48.

47.   Hofstadler, S. A.; Sannes-Lowery, K. A., Applications of ESI-MS in drug discovery: interrogation of noncovalent complexes. *Nature Reviews Drug Discovery* **2006,** *5* (7), 585-595.

48.   Laurent, H.; Jean-Paul, R.; Dragos, H., Fragment-Based Drug Design: Computational and Experimental State of the Art. *Combinatorial Chemistry & High Throughput Screening* **2011,** *14* (6), 500-520.

49.   Carr, R. A. E.; Congreve, M.; Murray, C. W.; Rees, D. C., Fragment-based lead discovery: leads by design. *Drug discovery today* **2005,** *10* (14), 987-992.

50. Jhoti, H.; Cleasby, A.; Verdonk, M.; Williams, G., Fragment-based screening using X-ray crystallography and NMR spectroscopy. *Current opinion in chemical biology* **2007,** *11* (5), 485-493.

51. Wu, B.; Zhang, Z.; Noberini, R.; Barile, E.; Giulianotti, M.; Pinilla, C.; Houghten, Richard A.; Pasquale, Elena B.; Pellecchia, M., HTS by NMR of Combinatorial Libraries: A Fragment-Based Approach to Ligand Discovery. *Chemistry & Biology* **2013,** *20* (1), 19-33.

52. Śledź, P.; Abell, C.; Ciulli, A., Ligand-Observed NMR in Fragment-Based Approaches. In *NMR of Biomolecules*, 2012; pp 264-280.

53. Jordan, J. B.; Poppe, L.; Xia, X.; Cheng, A. C.; Sun, Y.; Michelsen, K.; Eastwood, H.; Schnier, P. D.; Nixey, T.; Zhong, W., Fragment based drug discovery: practical implementation based on ^{19}F NMR spectroscopy. *Journal of medicinal chemistry* **2012,** *55* (2), 678-87.

54. Gerig, J., Fluorine NMR. **2001**.

55. Spotswood, T. M.; Evans, J. M.; Richards, J. H., Enzyme-substrate interaction by nuclear magnetic resonance. *Journal of the American Chemical Society* **1967,** *89* (19), 5052-5054.

56. Torre, B. G. d. l.; Albericio, F., The Pharmaceutical Industry in 2020. An Analysis of FDA Drug Approvals from the Perspective of Molecules. *Molecules* **2021,** *26* (3), 627.

57. Schweizer, E.; Hoffmann-Röder, A.; Schärer, K.; Olsen, J. A.; Fäh, C.; Seiler, P.; Obst-Sander, U.; Wagner, B.; Kansy, M.; Diederich, F., A fluorine scan at the catalytic center of thrombin: C--F, C--OH, and C--OMe bioisosterism and fluorine effects on pKa and log D values. *ChemMedChem* **2006,** *1* (6), 611-21.

58. Gerebtzoff, G.; Li-Blatter, X.; Fischer, H.; Frentzel, A.; Seelig, A., Halogenation of drugs enhances membrane binding and permeation. *Chembiochem : a European journal of chemical biology* **2004,** *5* (5), 676-84.

59. Fielding, L., NMR methods for the determination of protein-ligand dissociation constants. *Current topics in medicinal chemistry* **2003,** *3* (1), 39-53.

60. Carr, H. Y.; Purcell, E. M., Effects of Diffusion on Free Precession in Nuclear Magnetic Resonance Experiments. *Physical Review* **1954,** *94* (3), 630-638.

61. Meiboom, S.; Gill, D., Modified Spin-Echo Method for Measuring Nuclear Relaxation Times. *Review of Scientific Instruments* **1958,** *29* (8), 688-691.

62. Dalvit, C.; Vulpetti, A., Ligand-Based Fluorine NMR Screening: Principles and Applications in Drug Discovery Projects. *Journal of medicinal chemistry* **2019,** *62* (5), 2218-2244.

63.   Ghitti, M.; Musco, G.; Spitaleri, A., NMR and computational methods in the structural and dynamic characterization of ligand-receptor interactions. *Advances in experimental medicine and biology* **2014,** *805*, 271-304.

64.   Stockman, B.; Dalvit, C., NMR screening techniques in drug discovery and drug design. *Progress in nuclear magnetic resonance spectroscopy* **2002,** *41*, 187-231.

65.   Fernández, C.; Jahnke, W., New approaches for NMR screening in drug discovery. *Drug discovery today. Technologies* **2004,** *1* (3), 277-83.

66.   Dalvit, C.; Fagerness, P. E.; Hadden, D. T.; Sarver, R. W.; Stockman, B. J., Fluorine-NMR experiments for high-throughput screening: theoretical aspects, practical considerations, and range of applicability. *Journal of the American Chemical Society* **2003,** *125* (25), 7696-703.

67.   Jahnke, W., Spin Labels as a Tool to Identify and Characterize Protein–Ligand Interactions by NMR Spectroscopy. *Chembiochem : a European journal of chemical biology* **2002,** *3* (2-3), 167-173.

68.   Hull, W. E.; Sykes, B. D., Fluorotyrosine alkaline phosphatase. Fluorine-19 nuclear magnetic resonance relaxation times and molecular motion of the individual fluorotyrosines. *Biochemistry* **1974,** *13* (17), 3431-3437.

69.   Liu, J. J.; Horst, R.; Katritch, V.; Stevens, R. C.; Wuthrich, K., Biased signaling pathways in beta2-adrenergic receptor characterized by 19F-NMR. *Science (New York, N.Y.)* **2012,** *335* (6072), 1106-10.

70.   Kitevski-LeBlanc, J. L.; Prosser, R. S., Current applications of 19F NMR to studies of protein structure and dynamics. *Progress in nuclear magnetic resonance spectroscopy* **2012,** *62*, 1-33.

71.   Arntson, K. E.; Pomerantz, W. C., Protein-Observed Fluorine NMR: A Bioorthogonal Approach for Small Molecule Discovery. *Journal of medicinal chemistry* **2016,** *59* (11), 5158-71.

72.   Sharaf, N. G.; Gronenborn, A. M., (19)F-modified proteins and (19)F-containing ligands as tools in solution NMR studies of protein interactions. *Methods in enzymology* **2015,** *565*, 67-95.

73.   Tobola, F.; Lelimousin, M.; Varrot, A.; Gillon, E.; Darnhofer, B.; Blixt, O.; Birner-Gruenberger, R.; Imberty, A.; Wiltschi, B., Effect of Noncanonical Amino Acids on Protein-Carbohydrate Interactions: Structure, Dynamics, and Carbohydrate Affinity of a Lectin Engineered with Fluorinated Tryptophan Analogs. *ACS chemical biology* **2018,** *13* (8), 2211-2219.

74.   Gee, C. T.; Arntson, K. E.; Urick, A. K.; Mishra, N. K.; Hawk, L. M.; Wisniewski, A. J.; Pomerantz, W. C., Protein-observed (19)F-NMR for fragment screening, affinity quantification and druggability assessment. *Nature protocols* **2016,** *11* (8), 1414-27.

75.   Lee, J. H.; Kim, Y. G.; Cho, M. H.; Kim, J. A.; Lee, J., 7-fluoroindole as an antivirulence compound against Pseudomonas aeruginosa. *FEMS Microbiol Lett* **2012,** *329* (1), 36-44.

76.   Gee, C. T.; Koleski, E. J.; Pomerantz, W. C. K., Fragment screening and druggability assessment for the CBP/p300 KIX domain through protein-observed 19F NMR spectroscopy. *Angewandte Chemie (International ed. in English)* **2015,** *54* (12), 3735-3739.

77.   Wider, G.; Wüthrich, K., NMR spectroscopy of large molecules and multimolecular assemblies in solution. *Current opinion in structural biology* **1999,** *9* (5), 594-601.

78.   Pervushin, K.; Riek, R.; Wider, G.; Wüthrich, 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. *Proceedings of the National Academy of Sciences of the United States of America* **1997,** *94* (23), 12366-71.

79.   Shuker, S. B.; Hajduk, P. J.; Meadows, R. P.; Fesik, S. W., Discovering high-affinity ligands for proteins: SAR by NMR. *Science (New York, N.Y.)* **1996,** *274* (5292), 1531-4.

80.   Shi, L.; Zhang, N., Applications of Solution NMR in Drug Discovery. *Molecules* **2021,** *26* (3), 576.

81.   Bishop, J. R.; Gagneux, P., Evolution of carbohydrate antigens--microbial forces shaping host glycomes? *Glycobiology* **2007,** *17* (5), 23r-34r.

82.   Varki, A., In *Essentials of Glycobiology*, Varki, A.; Cummings, R. D.; Esko, J. D.; Freeze, H. H.; Stanley, P.; Bertozzi, C. R.; Hart, G. W.; Etzler, M. E., Eds. Cold Spring Harbor Laboratory Press: Cold Spring Harbor (NY), 2009.

83.   Varki, A., Nothing in glycobiology makes sense, except in the light of evolution. *Cell* **2006,** *126* (5), 841-5.

84.   Poole, J.; Day, C. J.; von Itzstein, M.; Paton, J. C.; Jennings, M. P., Glycointeractions in bacterial pathogenesis. *Nature Reviews Microbiology* **2018,** *16* (7), 440-452.

85.   Van Breedam, W.; Pöhlmann, S.; Favoreel, H. W.; de Groot, R. J.; Nauwynck, H. J., Bitter-sweet symphony: glycan–lectin interactions in virus biology. *FEMS Microbiology Reviews* **2014,** *38* (4), 598-632.

86. Liu, Z.; Zhang, Q.; Peng, H.; Zhang, W.-z., Animal Lectins: Potential Antitumor Therapeutic Targets in Apoptosis. *Applied Biochemistry and Biotechnology* **2012,** *168* (3), 629-637.

87. Buzás, E. I.; György, B.; Pásztói, M.; Jelinek, I.; Falus, A.; Gabius, H. J., Carbohydrate recognition systems in autoimmunity. *Autoimmunity* **2006,** *39* (8), 691-704.

88. Becker, J. W.; Reeke, G. N., Jr.; Wang, J. L.; Cunningham, B. A.; Edelman, G. M., The covalent and three-dimensional structure of concanavalin A. III. Structure of the monomer and its interactions with metals and saccharides. *The Journal of biological chemistry* **1975,** *250* (4), 1513-24.

89. Imberty, A.; J, H. P., Structural Biology of Glycan Recognition. In *Essentials of Glycobiology*, Varki, A.; Cummings, R. D.; Esko, J. D.; Stanley, P.; Hart, G. W.; Aebi, M.; Darvill, A. G.; Kinoshita, T.; Packer, N. H.; Prestegard, J. H.; Schnaar, R. L.; Seeberger, P. H., Eds. Cold Spring Harbor Laboratory Press

Copyright 2015-2017 by The Consortium of Glycobiology Editors, La Jolla, California. All rights reserved.: Cold Spring Harbor (NY), 2015; pp 387-400.

90. Ernst, B.; Magnani, J. L., From carbohydrate leads to glycomimetic drugs. *Nature reviews. Drug discovery* **2009,** *8* (8), 661-677.

91. Taylor, M. E.; Drickamer, K.; Schnaar, R. L.; Etzler, M. E.; Varki, A., Discovery and Classification of Glycan-Binding Proteins. In *Essentials of Glycobiology*, Varki, A.; Cummings, R. D.; Esko, J. D.; Stanley, P.; Hart, G. W.; Aebi, M.; Darvill, A. G.; Kinoshita, T.; Packer, N. H.; Prestegard, J. H.; Schnaar, R. L.; Seeberger, P. H., Eds. Cold Spring Harbor Laboratory Press

Copyright 2015-2017 by The Consortium of Glycobiology Editors, La Jolla, California. All rights reserved.: Cold Spring Harbor (NY), 2015; pp 361-72.

92. Cecioni, S.; Imberty, A.; Vidal, S., Glycomimetics versus multivalent glycoconjugates for the design of high affinity lectin ligands. *Chemical reviews* **2015,** *115* (1), 525-61.

93. Touaibia, M.; Wellens, A.; Shiao, T. C.; Wang, Q.; Sirois, S.; Bouckaert, J.; Roy, R., Cover Picture: Mannosylated G(0) Dendrimers with Nanomolar Affinities to Escherichia coli FimH (ChemMedChem 8/2007). *ChemMedChem* **2007,** *2* (8), 1089-1089.

94. Zahorska, E.; Kuhaudomlarp, S.; Minervini, S.; Yousaf, S.; Lepsik, M.; Kinsinger, T.; Hirsch, A. K. H.; Imberty, A.; Titz, A., A rapid synthesis of low-nanomolar divalent LecA inhibitors in four linear steps from d-galactose pentaacetate. *Chemical Communications* **2020,** *56* (62), 8822-8825.

95.   Boukerb, A. M.; Rousset, A.; Galanos, N.; Méar, J.-B.; Thépaut, M.; Grandjean, T.; Gillon, E.; Cecioni, S.; Abderrahmen, C.; Faure, K.; Redelberger, D.; Kipnis, E.; Dessein, R.; Havet, S.; Darblade, B.; Matthews, S. E.; de Bentzmann, S.; Guéry, B.; Cournoyer, B.; Imberty, A.; Vidal, S., Antiadhesive Properties of Glycoclusters against Pseudomonas aeruginosa Lung Infection. *Journal of medicinal chemistry* **2014,** *57* (24), 10275-10289.

96.   Rodrigue, J.; Ganne, G.; Blanchard, B.; Saucier, C.; Giguere, D.; Shiao, T. C.; Varrot, A.; Imberty, A.; Roy, R., Aromatic thioglycoside inhibitors against the virulence factor LecA from Pseudomonas aeruginosa. *Organic & biomolecular chemistry* **2013,** *11* (40), 6906-18.

97.   Kadam, R. U.; Garg, D.; Schwartz, J.; Visini, R.; Sattler, M.; Stocker, A.; Darbre, T.; Reymond, J. L., CH-pi "T-shape" interaction with histidine explains binding of aromatic galactosides to Pseudomonas aeruginosa lectin LecA. *ACS Chem Biol* **2013,** *8* (9), 1925-30.

98.   Wagner, S.; Hauck, D.; Hoffmann, M.; Sommer, R.; Joachim, I.; Müller, R.; Imberty, A.; Varrot, A.; Titz, A., Covalent Lectin Inhibition and Application in Bacterial Biofilm Imaging. *Angewandte Chemie (International ed. in English)* **2017,** *56* (52), 16559-16564.

99.   Johansson, E. M.; 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.; Reymond, J. L., Inhibition and dispersion of Pseudomonas aeruginosa biofilms by glycopeptide dendrimers targeting the fucose-specific lectin LecB. *Chem Biol* **2008,** *15* (12), 1249-57.

100.   Richichi, B.; Imberty, A.; Gillon, E.; Bosco, R.; Sutkeviciute, I.; Fieschi, F.; Nativi, C., Synthesis of a selective inhibitor of a fucose binding bacterial lectin from Burkholderia ambifaria. *Organic & biomolecular chemistry* **2013,** *11* (24), 4086-4094.

101.   Wamhoff, E.-C.; Hanske, J.; Schnirch, L.; Aretz, J.; Grube, M.; Varón Silva, D.; Rademacher, C., 19F NMR-Guided Design of Glycomimetic Langerin Ligands. *ACS chemical biology* **2016,** *11* (9), 2407-2413.

102.   Medve, L.; Achilli, S.; Guzman-Caldentey, J.; Thépaut, M.; Senaldi, L.; Le Roy, A.; Sattin, S.; Ebel, C.; Vivès, C.; Martin-Santamaria, S.; Bernardi, A.; Fieschi, F., Enhancing Potency and Selectivity of a DC-SIGN Glycomimetic Ligand by Fragment-Based Design: Structural Basis. *Chemistry (Weinheim an der Bergstrasse, Germany)* **2019,** *25* (64), 14659-14668.

103.   Changeux, J. P., Allostery and the Monod-Wyman-Changeux model after 50 years. *Annual review of biophysics* **2012,** *41*, 103-33.

104.   Goodey, N. M.; Benkovic, S. J., Allosteric regulation and catalysis emerge via a common route. *Nature Chemical Biology* **2008,** *4* (8), 474-482.

105.   Aretz, J.; Anumala, U. R.; Fuchsberger, F. F.; Molavi, N.; Ziebart, N.; Zhang, H.; Nazaré, M.; Rademacher, C., Allosteric Inhibition of a Mammalian Lectin. *Journal of the American Chemical Society* **2018,** *140* (44), 14915-14925.

106.   Keller, B. G.; Rademacher, C., Allostery in C-type lectins. *Current opinion in structural biology* **2020,** *62*, 31-38.

107.   Aretz, J.; Baukmann, H.; Shanina, E.; Hanske, J.; Wawrzinek, R.; Zapol'skii, V. A.; Seeberger, P. H.; Kaufmann, D. E.; Rademacher, C., Identification of Multiple Druggable Secondary Sites by Fragment Screening against DC-SIGN. *Angewandte Chemie International Edition* **2017,** *56* (25), 7292-7296.

108.   Waldron, T. T.; Springer, T. A., Transmission of allostery through the lectin domain in selectin-mediated cell adhesion. *Proceedings of the National Academy of Sciences of the United States of America* **2009,** *106* (1), 85-90.

109.   Hanske, J.; Aleksić, S.; Ballaschk, M.; Jurk, M.; Shanina, E.; Beerbaum, M.; Schmieder, P.; Keller, B. G.; Rademacher, C., Intradomain Allosteric Network Modulates Calcium Affinity of the C-Type Lectin Receptor Langerin. *Journal of the American Chemical Society* **2016,** *138* (37), 12176-12186.

110.   Breitenbach Barroso Coelho, L. C.; Marcelino dos Santos Silva, P.; Felix de Oliveira, W.; de Moura, M. C.; Viana Pontual, E.; Soares Gomes, F.; Guedes Paiva, P. M.; Napoleão, T. H.; dos Santos Correia, M. T., Lectins as antimicrobial agents. *Journal of Applied Microbiology* **2018,** *125* (5), 1238-1252.

111.   Meiers, J.; Siebs, E.; Zahorska, E.; Titz, A., Lectin antagonists in infection, immunity, and inflammation. *Current opinion in chemical biology* **2019,** *53*, 51-67.

112.   Esko, J. D.; Sharon, N., Microbial Lectins: Hemagglutinins, Adhesins, and Toxins. In *Essentials of Glycobiology*, Varki, A.; Cummings, R. D.; Esko, J. D.; Freeze, H. H.; Stanley, P.; Bertozzi, C. R.; Hart, G. W.; Etzler, M. E., Eds. Cold Spring Harbor Laboratory Press

Copyright © 2009, The Consortium of Glycobiology Editors, La Jolla, California.: Cold Spring Harbor (NY), 2009.

113.   Sharon, N., Lectins. In *eLS*.

114.   Diggle, S. P.; Stacey, R. E.; Dodd, C.; Camara, M.; Williams, P.; Winzer, K., The galactophilic lectin, LecA, contributes to biofilm development in Pseudomonas aeruginosa. *Environmental microbiology* **2006,** *8* (6), 1095-104.

115.   Tielker, D.; Hacker, S.; Loris, R.; Strathmann, M.; Wingender, J.; Wilhelm, S.; Rosenau, F.; Jaeger, K. E., Pseudomonas aeruginosa lectin LecB is located in the

outer membrane and is involved in biofilm formation. *Microbiology (Reading, England)* **2005,** *151* (Pt 5), 1313-1323.

116. Santajit, S.; Indrawattana, N., Mechanisms of Antimicrobial Resistance in ESKAPE Pathogens. *BioMed research international* **2016,** *2016*, 2475067.

117. Meiers, J.; Siebs, E.; Zahorska, E.; Titz, A., Lectin antagonists in infection, immunity, and inflammation. *Current opinion in chemical biology* **2019,** *53*, 51-67.

118. Grishin, A. V.; Krivozubov, M. S.; Karyagina, A. S.; Gintsburg, A. L., Pseudomonas Aeruginosa Lectins As Targets for Novel Antibacterials. *Acta Naturae* **2015,** *7* (2), 29-41.

119. Singh, P. K.; Schaefer, A. L.; Parsek, M. R.; Moninger, T. O.; Welsh, M. J.; Greenberg, E. P., Quorum-sensing signals indicate that cystic fibrosis lungs are infected with bacterial biofilms. *Nature* **2000,** *407* (6805), 762-4.

120. Cioci, G.; Mitchell, E. P.; Gautier, C.; Wimmerova, M.; Sudakevitz, D.; Perez, S.; Gilboa-Garber, N.; Imberty, A., Structural basis of calcium and galactose recognition by the lectin PA-IL of Pseudomonas aeruginosa. *FEBS letters* **2003,** *555* (2), 297-301.

121. Mitchell, E.; Houles, C.; Sudakevitz, D.; Wimmerova, M.; Gautier, C.; Pérez, S.; Wu, A. M.; Gilboa-Garber, N.; Imberty, A., Structural basis for oligosaccharide-mediated adhesion of Pseudomonas aeruginosa in the lungs of cystic fibrosis patients. *Nature Structural Biology* **2002,** *9* (12), 918-921.

122. Kadam, R. U.; Bergmann, M.; Hurley, M.; Garg, D.; Cacciarini, M.; Swiderska, M. A.; Nativi, C.; Sattler, M.; Smyth, A. R.; Williams, P.; Camara, M.; Stocker, A.; Darbre, T.; Reymond, J. L., A glycopeptide dendrimer inhibitor of the galactose-specific lectin LecA and of Pseudomonas aeruginosa biofilms. *Angewandte Chemie (International ed. in English)* **2011,** *50* (45), 10631-5.

123. Perret, S.; Sabin, C.; Dumon, C.; Pokorná, M.; Gautier, C.; Galanina, O.; Ilia, S.; Bovin, N.; Nicaise, M.; Desmadril, M.; Gilboa-Garber, N.; Wimmerová, M.; Mitchell, E. P.; Imberty, A., Structural basis for the interaction between human milk oligosaccharides and the bacterial lectin PA-IIL of Pseudomonas aeruginosa. *The Biochemical journal* **2005,** *389* (Pt 2), 325-32.

124. Sabin, C.; Mitchell, E. P.; Pokorná, M.; Gautier, C.; Utille, J.-P.; Wimmerová, M.; Imberty, A., Binding of different monosaccharides by lectin PA-IIL from Pseudomonas aeruginosa: Thermodynamics data correlated with X-ray structures. *FEBS letters* **2006,** *580* (3), 982-987.

125. Imberty, A.; wimmerová, M.; Mitchell, E. P.; Gilboa-Garber, N., Structures of the lectins from Pseudomonas aeruginosa: insight into the molecular basis for host glycan recognition. *Microbes and infection* **2004,** *6* (2), 221-8.

126. Garber, N.; Guempel, U.; Gilboa-Garber, N.; Royle, R. J., Specificity of the fucose-binding lectin of Pseudomonas aeruginosa. *FEMS Microbiology Letters* **1987,** *48* (3), 331-334.

127. Gilboa-Garber, N.; Katcoff, D. J.; Garber, N. C., Identification and characterization of pseudomonas aeruginosa PA-IIL lectin gene and protein compared to PA-IL. *FEMS immunology and medical microbiology* **2000,** *29* (1), 53-7.

128. Wu, A. M.; Wu, J. H.; Singh, T.; Liu, J. H.; Tsai, M. S.; Gilboa-Garber, N., Interactions of the fucose-specific Pseudomonas aeruginosa lectin, PA-IIL, with mammalian glycoconjugates bearing polyvalent Lewis(a) and ABH blood group glycotopes. *Biochimie* **2006,** *88*, 1479–1492.

129. Rhim, A. D.; Kothari, V. A.; Park, P. J.; Mulberg, A. E.; Glick, M. C.; Scanlin, T. F., Terminal glycosylation of cystic fibrosis airway epithelial cells. *Glycoconj J* **2000,** *17* (6), 385-91.

130. Scanlin, T. F.; Glick, M. C., Glycosylation and the cystic fibrosis transmembrane conductance regulator. *Respir Res* **2001,** *2* (5), 276-279.

131. Glick, J.; Garber, N., The intracellular localization of Pseudomonas aeruginosa lectins. *Journal of general microbiology* **1983,** *129* (10), 3085-90.

132. Flemming, H.-C.; Wingender, J., The biofilm matrix. *Nature Reviews Microbiology* **2010,** *8* (9), 623-633.

133. Tielker, D.; Hacker, S.; Loris, R.; Strathmann, M.; Wingender, J.; Wilhelm, S.; Rosenau, F.; Jaeger, K. E., Pseudomonas aeruginosa lectin LecB is located in the outer membrane and is involved in biofilm formation. *Microbiology (Reading)* **2005,** *151* (Pt 5), 1313-1323.

134. Garber, N.; Guempel, U.; Belz, A.; Gilboa-Garber, N.; Doyle, R. J., On the specificity of the D-galactose-binding lectin (PA-I) of Pseudomonas aeruginosa and its strong binding to hydrophobic derivatives of D-galactose and thiogalactose. *Biochimica et biophysica acta* **1992,** *1116* (3), 331-3.

135. Bergmann, M.; Michaud, G.; Visini, R.; Jin, X.; Gillon, E.; Stocker, A.; Imberty, A.; Darbre, T.; Reymond, J. L., Multivalency effects on Pseudomonas aeruginosa biofilm inhibition and dispersal by glycopeptide dendrimers targeting lectin LecA. *Organic & biomolecular chemistry* **2016,** *14* (1), 138-48.

136. Wagner, S.; Sommer, R.; Hinsberger, S.; Lu, C.; Hartmann, R. W.; Empting, M.; Titz, A., Novel Strategies for the Treatment of Pseudomonas aeruginosa Infections. *Journal of medicinal chemistry* **2016,** *59* (13), 5929-69.

137. Nannini, E. C.; Ponessa, A.; Muratori, R.; Marchiaro, P.; Ballerini, V.; Flynn, L.; Limansky, A. S., Polyclonal outbreak of bacteremia caused by Burkholderia cepacia complex and the presumptive role of ultrasound gel. *The Brazilian journal of infectious*

diseases : an official publication of the Brazilian Society of Infectious Diseases **2015,** *19* (5), 543-5.

138.   Coenye, T.; Mahenthiralingam, E.; Henry, D.; LiPuma, J. J.; Laevens, S.; Gillis, M.; Speert, D. P.; Vandamme, P., Burkholderia ambifaria sp. nov., a novel member of the Burkholderia cepacia complex including biocontrol and cystic fibrosis-related isolates. *International journal of systematic and evolutionary microbiology* **2001,** *51* (Pt 4), 1481-1490.

139.   Glick, M. C.; Kothari, V. A.; Liu, A.; Stoykova, L. I.; Scanlin, T. F., Activity of fucosyltransferases and altered glycosylation in cystic fibrosis airway epithelial cells. *Biochimie* **2001,** *83* (8), 743-7.

140.   Venkatakrishnan, V.; Packer, N. H.; Thaysen-Andersen, M., Host mucin glycosylation plays a role in bacterial adhesion in lungs of individuals with cystic fibrosis. *Expert review of respiratory medicine* **2013,** *7* (5), 553-76.

141.   Audfray, A.; Claudinon, J.; Abounit, S.; Ruvoën-Clouet, N.; Larson, G.; Smith, D. F.; Wimmerová, M.; Le Pendu, J.; Römer, W.; Varrot, A.; Imberty, A., Fucose-binding lectin from opportunistic pathogen Burkholderia ambifaria binds to both plant and human oligosaccharidic epitopes. *The Journal of biological chemistry* **2012,** *287* (6), 4335-4347.

142.   Bonnardel, F.; Kumar, A.; Wimmerova, M.; Lahmann, M.; Perez, S.; Varrot, A.; Lisacek, F.; Imberty, A., Architecture and Evolution of Blade Assembly in β-propeller Lectins. *Structure* **2019,** *27* (5), 764-775.e3.

143.   Wilhelm, I.; Levit-Zerdoun, E.; Jakob, J.; Villringer, S.; Frensch, M.; Übelhart, R.; Landi, A.; Müller, P.; Imberty, A.; Thuenauer, R.; Claudinon, J.; Jumaa, H.; Reth, M.; Eibel, H.; Hobeika, E.; Römer, W., Carbohydrate-dependent B cell activation by fucose-binding bacterial lectins. *Science Signaling* **2019,** *12* (571), eaao7194.

144.   Goyard, D.; Baldoneschi, V.; Varrot, A.; Fiore, M.; Imberty, A.; Richichi, B.; Renaudet, O.; Nativi, C., Multivalent Glycomimetics with Affinity and Selectivity toward Fucose-Binding Receptors from Emerging Pathogens. *Bioconjugate Chemistry* **2018,** *29* (1), 83-88.

145.   Kuhaudomlarp, S.; Cerofolini, L.; Santarsia, S.; Gillon, E.; Fallarini, S.; Lombardi, G.; Denis, M.; Giuntini, S.; Valori, C.; Fragai, M.; Imberty, A.; Dondoni, A.; Nativi, C., Fucosylated ubiquitin and orthogonally glycosylated mutant A28C: conceptually new ligands for Burkholderia ambifaria lectin (BambL). *Chemical science* **2020,** *11* (47), 12662-12670.

146.   Galanos, N.; Chen, Y.; Michael, Z. P.; Gillon, E.; Dutasta, J.-P.; Star, A.; Imberty, A.; Martinez, A.; Vidal, S., Cyclotriveratrylene-Based Glycoclusters as High

Affinity Ligands of Bacterial Lectins from Pseudomonas aeruginosa and Burkholderia ambifaria. *ChemistrySelect* **2016,** *1* (18), 5863-5868.

147.   Ligeour, C.; Audfray, A.; Gillon, E.; Meyer, A.; Galanos, N.; Vidal, S.; Vasseur, J.-J.; Imberty, A.; Morvan, F., Synthesis of branched-phosphodiester and mannose-centered fucosylated glycoclusters and their binding studies with Burkholderia ambifaria lectin (BambL). *RSC Advances* **2013,** *3* (42), 19515-19524.

148.   Drickamer, K., C-type lectin-like domains. *Current opinion in structural biology* **1999,** *9* (5), 585-90.

149.   Mitchell, C. A.; Ramessar, K.; O'Keefe, B. R., Antiviral lectins: Selective inhibitors of viral entry. *Antiviral Research* **2017,** *142*, 37-54.

150.   Schulze, J.; Rentzsch, M.; Kim, D.; Bellmann, L.; Stoitzner, P.; Rademacher, C., A Liposomal Platform for Delivery of a Protein Antigen to Langerin-Expressing Cells. *Biochemistry* **2019,** *58* (21), 2576-2580.

151.   Osorio, F.; Reis e Sousa, C., Myeloid C-type Lectin Receptors in Pathogen Recognition and Host Defense. *Immunity* **2011,** *34* (5), 651-664.

152.   Garcia-Vallejo, J. J.; van Kooyk, Y., The physiological role of DC-SIGN: A tale of mice and men. *Trends in Immunology* **2013,** *34* (10), 482-486.

153.   Geijtenbeek, T. B.; Krooshoop, D. J.; Bleijs, D. A.; van Vliet, S. J.; van Duijnhoven, G. C.; Grabovsky, V.; Alon, R.; Figdor, C. G.; van Kooyk, Y., DC-SIGN-ICAM-2 interaction mediates dendritic cell trafficking. *Nature immunology* **2000,** *1* (4), 353-7.

154.   Geijtenbeek, T. B.; Torensma, R.; van Vliet, S. J.; van Duijnhoven, G. C.; Adema, G. J.; van Kooyk, Y.; Figdor, C. G., Identification of DC-SIGN, a novel dendritic cell-specific ICAM-3 receptor that supports primary immune responses. *Cell* **2000,** *100* (5), 575-85.

155.   Montoya, M. C.; Sancho, D.; Bonello, G.; Collette, Y.; Langlet, C.; He, H. T.; Aparicio, P.; Alcover, A.; Olive, D.; Sánchez-Madrid, F., Role of ICAM-3 in the initial interaction of T lymphocytes and APCs. *Nature immunology* **2002,** *3* (2), 159-168.

156.   Geijtenbeek , T. B. H.; van Vliet , S. J.; Koppel , E. A.; Sanchez-Hernandez , M.; Vandenbroucke-Grauls , C. M. J. E.; Appelmelk , B.; van Kooyk , Y., Mycobacteria Target DC-SIGN to Suppress Dendritic Cell Function. *Journal of Experimental Medicine* **2002,** *197* (1), 7-17.

157.   van den Berg, L. M.; Geijtenbeek, T. B., Antiviral immune responses by human langerhans cells and dendritic cells in HIV-1 infection. *Advances in experimental medicine and biology* **2013,** *762*, 45-70.

158.   Alvarez, C. P.; Lasala, F.; Carrillo, J.; Muñiz, O.; Corbí, A. L.; Delgado, R., C-type lectins DC-SIGN and L-SIGN mediate cellular entry by Ebola virus in cis and in trans. *Journal of virology* **2002,** *76* (13), 6841-4.

159.   Yang, Z.-Y.; Huang, Y.; Ganesh, L.; Leung, K.; Kong, W.-P.; Schwartz, O.; Subbarao, K.; Nabel, G. J., pH-Dependent Entry of Severe Acute Respiratory Syndrome Coronavirus Is Mediated by the Spike Glycoprotein and Enhanced by Dendritic Cell Transfer through DC-SIGN. *Journal of virology* **2004,** *78* (11), 5642-5650.

160.   Lenza, M. P.; Oyenarte, I.; Diercks, T.; Quintana, J. I.; Gimeno, A.; Coelho, H.; Diniz, A.; Peccati, F.; Delgado, S.; Bosch, A.; Valle, M.; Millet, O.; Abrescia, N. G. A.; Palazón, A.; Marcelo, F.; Jiménez-Osés, G.; Jiménez-Barbero, J.; Ardá, A.; Ereño-Orbea, J., Structural Characterization of N-Linked Glycans in the Receptor Binding Domain of the SARS-CoV-2 Spike Protein and their Interactions with Human Lectins. *Angewandte Chemie (International ed. in English)* **2020,** *59* (52), 23763-23771.

161.   Amraei, R.; Yin, W.; Napoleon, M. A.; Suder, E. L.; Berrigan, J.; Zhao, Q.; Olejnik, J.; Chandler, K. B.; Xia, C.; Feldman, J.; Hauser, B. M.; Caradonna, T. M.; Schmidt, A. G.; Gummuluru, S.; Mühlberger, E.; Chitalia, V.; Costello, C. E.; Rahimi, N., CD209L/L-SIGN and CD209/DC-SIGN act as receptors for SARS-CoV-2 and are differentially expressed in lung and kidney epithelial and endothelial cells. *bioRxiv* **2020**, 2020.06.22.165803.

162.   Feinberg, H.; Mitchell, D. A.; Drickamer, K.; Weis, W. I., Structural Basis for Selective Recognition of Oligosaccharides by DC-SIGN and DC-SIGNR. *Science (New York, N.Y.)* **2001,** *294* (5549), 2163-2166.

163.   Feinberg, H.; Taylor, M. E.; Razi, N.; McBride, R.; Knirel, Y. A.; Graham, S. A.; Drickamer, K.; Weis, W. I., Structural basis for langerin recognition of diverse pathogen and mammalian glycans through a single binding site. *Journal of molecular biology* **2011,** *405* (4), 1027-39.

164.   Holla, A.; Skerra, A., Comparative analysis reveals selective recognition of glycans by the dendritic cell receptors DC-SIGN and Langerin. *Protein engineering, design & selection : PEDS* **2011,** *24* (9), 659-69.

165.   Curtis, B. M.; Scharnowske, S.; Watson, A. J., Sequence and expression of a membrane-associated C-type lectin that exhibits CD4-independent binding of human immunodeficiency virus envelope glycoprotein gp120. *Proceedings of the National Academy of Sciences of the United States of America* **1992,** *89* (17), 8356-60.

166.   Pederson, K.; Mitchell, D. A.; Prestegard, J. H., Structural Characterization of the DC-SIGN–LewisX Complex. *Biochemistry* **2014,** *53* (35), 5700-5709.

167.   van Die, I.; van Vliet, S. J.; Nyame, A. K.; Cummings, R. D.; Bank, C. M. C.; Appelmelk, B.; Geijtenbeek, T. B. H.; van Kooyk, Y., The dendritic cell–specific C-type lectin DC-SIGN is a receptor for Schistosoma mansoni egg antigens and recognizes the glycan antigen Lewis x. *Glycobiology* **2003,** *13* (6), 471-478.

168.   Appelmelk, B. J.; van Die, I.; van Vliet, S. J.; Vandenbroucke-Grauls, C. M. J. E.; Geijtenbeek, T. B. H.; van Kooyk, Y., Cutting Edge: Carbohydrate Profiling Identifies New Pathogens That Interact with Dendritic Cell-Specific ICAM-3-Grabbing Nonintegrin on Dendritic Cells. *The Journal of Immunology* **2003,** *170* (4), 1635-1639.

169.   Thépaut, M.; Guzzi, C.; Sutkeviciute, I.; Sattin, S.; Ribeiro-Viana, R.; Varga, N.; Chabrol, E.; Rojo, J.; Bernardi, A.; Angulo, J.; Nieto, P. M.; 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. *Journal of the American Chemical Society* **2013,** *135* (7), 2518-2529.

170.   Tomašić, T.; Hajšek, D.; Švajger, U.; Luzar, J.; Obermajer, N.; Petit-Haertlein, I.; Fieschi, F.; Anderluh, M., Monovalent mannose-based DC-SIGN antagonists: targeting the hydrophobic groove of the receptor. *European journal of medicinal chemistry* **2014,** *75*, 308-26.

171.   Garber, K. C.; Wangkanont, K.; Carlson, E. E.; Kiessling, L. L., A general glycomimetic strategy yields non-carbohydrate inhibitors of DC-SIGN. *Chemical communications (Cambridge, England)* **2010,** *46* (36), 6747-9.

172.   Borrok, M. J.; Kiessling, L. L., Non-carbohydrate Inhibitors of the Lectin DC-SIGN. *Journal of the American Chemical Society* **2007,** *129* (42), 12780-12785.

173.   Mangold, S. L.; Prost, L. R.; Kiessling, L. L., Quinoxalinone Inhibitors of the Lectin DC-SIGN. *Chemical science* **2012,** *3* (3), 772-777.

174.   Gagneux, P.; Varki, A., Evolutionary considerations in relating oligosaccharide diversity to biological function. *Glycobiology* **1999,** *9* (8), 747-755.

175.   Bentley, M.; Doak, B. C.; Mohanty, B.; Scanlon, M. J., Applications of NMR Spectroscopy in FBDD. In *Modern Magnetic Resonance*, Webb, G. A., Ed. Springer International Publishing: Cham, 2018; pp 2211-2231.

176.   Linclau, B.; Ardá, A.; Reichardt, N.-C.; Sollogoub, M.; Unione, L.; Vincent, S. P.; Jiménez-Barbero, J., Fluorinated carbohydrates as chemical probes for molecular recognition studies. Current status and perspectives. *Chemical Society Reviews* **2020,** *49* (12), 3863-3888.

177.   Mayer, M.; Meyer, B., Group epitope mapping by saturation transfer difference NMR to identify segments of a ligand in direct contact with a protein receptor. *Journal of the American Chemical Society* **2001,** *123* (25), 6108-17.

178.   Taroni, C.; Jones, S.; Thornton, J. M., Analysis and prediction of carbohydrate binding sites. *Protein engineering* **2000,** *13* (2), 89-98.

179.   Dax, K.; Albert, M.; Ortner, J.; Paul, B. J., Synthesis of deoxyfluoro sugars from carbohydrate precursors. *Carbohydrate research* **2000,** *327* (1-2), 47-86.

180.   Martínez, J. D.; Manzano, A. I.; Calviño, E.; Diego, A. d.; Rodriguez de Francisco, B.; Romanò, C.; Oscarson, S.; Millet, O.; Gabius, H.-J.; Jiménez-Barbero, J.; Cañada, F. J., Fluorinated Carbohydrates as Lectin Ligands: Simultaneous Screening of a Monosaccharide Library and Chemical Mapping by 19F NMR Spectroscopy. *The Journal of Organic Chemistry* **2020,** *85* (24), 16072-16081.

181.   Keenan, T.; Parmeggiani, F.; Malassis, J.; Fontenelle, C. Q.; Vendeville, J.-B.; Offen, W.; Both, P.; Huang, K.; Marchesi, A.; Heyam, A.; Young, C.; Charnock, S. J.; Davies, G. J.; Linclau, B.; Flitsch, S. L.; Fascione, M. A., Profiling Substrate Promiscuity of Wild-Type Sugar Kinases for Multi-fluorinated Monosaccharides. *Cell Chemical Biology* **2020,** *27* (9), 1199-1206.e5.

182.   Ludlow, R. F.; Verdonk, M. L.; Saini, H. K.; Tickle, I. J.; Jhoti, H., Detection of secondary binding sites in proteins using fragment screening. *Proceedings of the National Academy of Sciences* **2015,** *112* (52), 15910-15915.

183.   Grutsch, S.; Brüschweiler, S.; Tollinger, M., NMR Methods to Study Dynamic Allostery. *PLOS Computational Biology* **2016,** *12* (3), e1004620.

184.   Manley, G.; Loria, J. P., NMR insights into protein allostery. *Arch Biochem Biophys* **2012,** *519* (2), 223-231.

185.   Kuhaudomlarp, S.; Siebs, E.; Shanina, E.; Topin, J.; Joachim, I.; da Silva Figueiredo Celestino Gomes, P.; Varrot, A.; Rognan, D.; Rademacher, C.; Imberty, A.; Titz, A., Non-Carbohydrate Glycomimetics as Inhibitors of Calcium(II)-binding Lectins. *Angewandte Chemie (International ed. in English)* **2020**.

186.   Verma, R. P., Hydroxamic Acids as Matrix Metalloproteinase Inhibitors. In *Matrix Metalloproteinase Inhibitors: Specificity of Binding and Structure-Activity Relationships*, Gupta, S. P., Ed. Springer Basel: Basel, 2012; pp 137-176.

7. Supporting information

7.1. Supporting Information for Subchapter 4.1.

Supplementary data

Protein-observed 19F NMR of LecA from *Pseudomonas aeruginosa*

Elena V. Shanina[1,2], Eike Siebs[3,4,5], Hengxi Zhang[1,2], Daniel Varón Silva[1,2], Ines Joachim[3,4,5], Alexander Titz[3,4,5] & Christoph Rademacher[1,2,*]

[1] Max Planck Institute of Colloids and Interfaces, Department of Biomolecular Systems, Am Mühlenberg 14424, Potsdam, Germany;

[2] Free University of Berlin, Department of Biochemistry and Chemistry, 14195 Berlin, Germany;

[3] Chemical Biology of Carbohydrates (CBCH), Helmholtz Institute for Pharmaceutical Research Saarland (HIPS), Helmholtz Centre for Infection Research, 66123 Saarbrücken, Germany;

[4] Saarland University, Department of Pharmacy, 66123 Saarbrücken, Germany;

[5] German Center for Infection Research (DZIF), Hannover-Braunschweig, Germany.

Table of contents

Supplementary materials and methods

Site−directed mutagenesis

Liquid chromatography−mass spectrometry (LC−MS)

Materials

Supplementary tables

1

Table SIV. PrOF NMR and FP assay affinity values

Supplementary figures

Fig. S1. Temperature optimization for 5FW LecA PrOF NMR**.**

Fig. S2. Optimization of 5FW LecA concentration for PrOF NMR**.**

Fig. S3. Buffer optimization of 5FW LecA for PrOF NMR**.**

Fig. S4. PrOF NMR titration of Ca^{2+} to apo 5FW LecA.

Fig. S5. PrOF NMR titration of D-Gal to holo 5FW LecA.

Fig. S6. ITC titration of Ca^{2+} to apo LecA.

Fig. S7. Competitive binding assay based on fluorescence polarization with LecA and 5FW LecA.

Fig. S8. PrOF NMR titration of Ph−β−D−Gal to holo 5FW LecA.

2

Supplementary materials and methods

Materials

All chemicals and buffers used within this work were purchased from Sigma Aldrich (St. Louis, MO, USA) or Carl Roth (Karlsruhe, Germany) unless otherwise indicated.

The LecA−encoding gene from *P. aeruginosa* PAO1 cloned into pET25(b+) vector resulted in pET25pa1l plasmid and LecA was produced recombinantly as reported previously (Blanchard, et al. 2008).

Site−directed mutagenesis

To distinguish fluorine resonances, we mutated tryptophan (W) against phenylalanine (F) using Quick−Change Lightning Mutagenesis Kit (Agilent, Santa Clara, USA). We used primers shown in (**Table SII**) following the manufacturer's protocol to generate mutants W33F, W42F and W84F. The plasmids were sequenced to confirm the mutagenesis using primers for T7 promoter. The plasmid for W2F mutant was obtained from GenScript Biotech (USA).

Liquid chromatography−mass spectrometry (LC−MS)

LC−MS analysis of LecA was performed using an ESI Waters Xevo G2−XS mass spectrometer coupled to an Acquity H−class UPLC system. LC separation was done on an Acquity UPLC Protein BEH C4 column (21 x 50 mm, 1.7 μm) using a gradient from 10 to 70% of acetonitrile in water with 0.5% formic acid in 12 min. The MS analysis was performed in positive mode using a 1.5 kV capillary voltage. The data were analyzed using MassLynx 4.2 and Biopharmalynx 1.3 software.

3

Supplementary tables

Table SI. Line width values of four tryptophan residues measured in PrOF NMR upon variation of buffer, temperature and protein concentration (*ND* = line width could not be determined due to low signal-to-noise quality of spectra).

Peak	Line width [Hz]							
	Buffer		**Temperature**			**Protein concentration**		
	TBS pH 7.8	MES pH 6.0	285 K	298 K	310 K	50 µM	100 µM	200 µM
W42	180	148	*ND*	*ND*	153	*ND*	170	178
W2	172	172	*ND*	161	125	*ND*	152	159
W33	161	144	*ND*	150	145	*ND*	167	165
W85	115	113	*ND*	127	116	*ND*	109	109

Table SII. Primer sequences used for single-point mutagenesis (F = forward, R = reverse).

Primer	Sequence
LecA W33F F	CGTAACTGGCGAAACCGGCGGCGACGA
LecA W33F R	TCGTCGCCGCCGGTTTCGCCAGTTACGGA
LecA W42F F	CCTGCGGCCCGAATTTCTGGGTAGGTCCGTA
LecA W42F R	TACGGACCTACCCAGAAATTCGGGCCGCAGG
LecA W84F F	ATTGGGTGCAACGAAACGGAACAACC
LecA W84F R	GGTTGTTCCGTTTCGTTGCACCCAAT
T7 PROMOTOR	TAATACGACTCACTATATAGG
T7 TERMINATOR	GCTAGTTATTGCTCAGCGG

4

Table SIII. List of components used to prepare minimal medium (M9) supplemented with an amino acid cocktail and 5−fluoroindole (5FI).

Component	Amount per 1 L expression
Autoclaved MiliQ water	500 mL
5x Amino acid cocktail	250 mL
10x M9 salts	100 mL
20% D-glucose	20 mL
100x Trace elements	10 mL
100 mg mL^{-1} Ampicillin	1 mL
1 mg mL^{-1} Vitamins	1 mL
1 M CaCl$_2$	300 µL
1 M MgSO$_4$	4 mL
Prepare 5−fluoroindole (60mg) in 250 µL DMSO	

Amino acids cocktail (5x)

* For 1 L of stock solution, dissolve one by one in 50 mL of distilled MiliQ water and combine in following order:

Component	
Alanine	500 mg
Arginine	400 mg
Asparagine Monohydrate	400 mg
Aspartic acid	400 mg
Cysteine-HCl-H$_2$O	50 mg
Glutamine	400 mg
Glutamic acid	650 mg
Glycine	550 mg
Histidine	100 mg
Isoleucine	230 mg
Leucine	230 mg
Lysine-HCl	420 mg
Methionine	250 mg
Phenylalanine	130 mg
Proline	100 mg
Serine	2.1 g
Threonine	230 mg
Tyrosine	170 mg
Valine	230 mg
Sodium acetate	1.5 g
Succinic acid	1.5 g
Potassium phosphate (dibasic)	10.5 g
Adenine	500 mg
Guanosine	650 mg
Thymine	200 mg
Uracil	500 mg
Cytosine	200 mg

5

Table SIII. List of components used to prepare minimal medium (M9) supplemented with an

amino acid cocktail and 5−fluoroindole (5FI).

Component	Amount per 1 L expression
Autoclaved MiliQ water	500 mL
5x Amino acid cocktail	250 mL
10x M9 salts	100 mL
20% D-glucose	20 mL
100x Trace elements	10 mL
100 mg mL^{-1} Ampicillin	1 mL
1 mg mL^{-1} Vitamins	1 mL
1 M CaCl$_2$	300 µL
1 M MgSO$_4$	4 mL
Prepare 5−fluoroindole (60mg) in 250 µL DMSO	

Amino acids cocktail (5x)

* For 1 L of stock solution, dissolve one by one in 50 mL of distilled MiliQ water and combine
in following order:

Component	
Alanine	500 mg
Arginine	400 mg
Asparagine Monohydrate	400 mg
Aspartic acid	400 mg
Cysteine-HCl-H$_2$O	50 mg
Glutamine	400 mg
Glutamic acid	650 mg
Glycine	550 mg
Histidine	100 mg
Isoleucine	230 mg
Leucine	230 mg
Lysine-HCl	420 mg
Methionine	250 mg
Phenylalanine	130 mg
Proline	100 mg
Serine	2.1 g
Threonine	230 mg
Tyrosine	170 mg
Valine	230 mg
Sodium acetate	1.5 g
Succinic acid	1.5 g
Potassium phosphate (dibasic)	10.5 g
Adenine	500 mg
Guanosine	650 mg
Thymine	200 mg
Uracil	500 mg
Cytosine	200 mg

5

M9 salt solution (10x)	
* For 1 L stock solution, dissolve in 800 mL, adjust the pH to 7.2 with NaOH, autoclave for 15 min at 121°C. Use non-^{15}N labeled NH_4Cl.	
Na_2HPO_4-$2H_2O$	75.2 g
KH_2PO_4	30 g
NaCl	5 g
NH_4Cl	5 g
Trace elements (10x)	
* For 1 L stock solution, sterilize the solution over a 0.22 µm filter. Store in dark.	
EDTA	5 g
In 800 mL MQ, adjust pH to 7.5 with NaOH	
$FeCl_3$ (anhydrous)	498 mg
$ZnCl_2$	84 mg
0.1 M $CuCl_2$-$2H_2O$	765 µL
0.2 M $CoCl_2$-$2H_2O$	210 µL
0.1 M H_3BO_3	1.6 mL
1 M $MnCl_2$-$4H_2O$	8.1 µL
Vitamin stock (1000x)	
* For 50 mL stock solution, sterilize the solution over a 0.22 µm filter. Prepare 1 mL aliquots and store at -20°C.	
Biotin	50 mg
Thiamin-HCl	50 mg

6

Table SIV. PrOF NMR and FP assay affinity values. Competitive binding assay of Me−α−D−Gal, D−GalNAc and D−Gal with LecA and 5FW LecA based on fluorescence polarization are shown. IC_{50} values are given as averages of three independent (LecA) or three technical replicates (5FW LecA) and is a relative affinity value depending on the concentration of ligand (8 mM) used in this assay, which was required to reduce binding of Me−α−D−Gal to half of the uninhibited value. K_d values for Me−α−D−Gal, D−GalNAc, D−Gal, Ph−β−D−Gal and pNPGal are defined as half of binding sites of 100 µM 5FW LecA being bound and are given as average of three independent replicates (*ND* = not determined).

	Competitive binding assay		PrOF NMR
	LecA	5FW LecA	5FW LecA
	IC_{50} [µM]	IC_{50} [µM]	K_d [µM]
Me−α−D−Gal	140 ± 30	195	*ND*
D−Gal	330 ± 40	317	360 ± 50
D−GalNAc	1230 ± 200	1991	780 ± 97
Ph−β−D−Gal	*ND*	*ND*	166 ± 42
pNPGal	*ND*	*ND*	54 ± 6

7

Supplementary figures
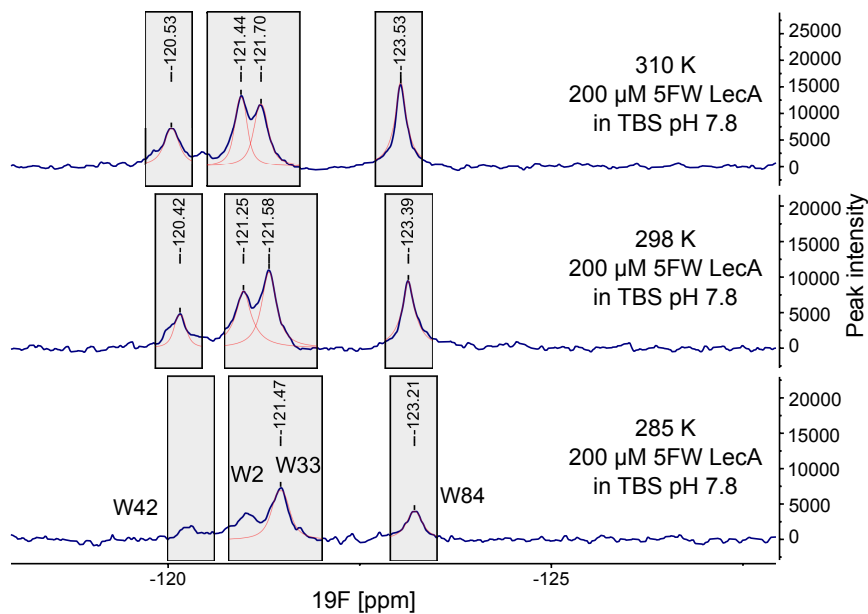
Fig. S1. Temperature optimization for 5FW LecA PrOF NMR. Shown is PrOF NMR spectrum of 200 µM holo 5FW LecA in TBS pH 7.8 at temperatures 285 K, 298 K and 310 K. Peak line widths of tryptophan resonances were measured in MestReNova using Line Fitting function for manual fitting of peaks (*red line*). The line width values for W42 and W84 at different temperatures are shown in **Table SI**, PrOF NMR spectrum of 5FW LecA at 298 K and 310 K resulted in similarly well-resolved spectra. The line width values of W2 and W33 deviated between measurements due to the overlap of both signals. As result, we used 310 K for PrOF NMR experiments with 5FW LecA**.**

8
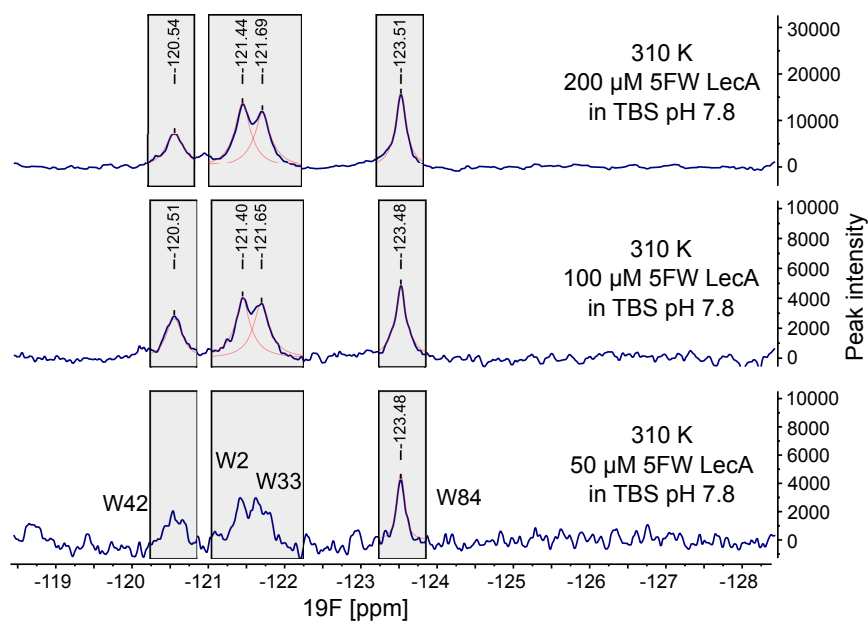
Fig. S2. Optimization of 5FW LecA concentration for PrOF NMR. Shown are PrOF NMR spectra of 50 μM (*bottom*), 100 μM (*middle*) and 200 μM (*top*) holo 5FW LecA in TBS pH 7.8 buffer at 310 K. Peak line widths of tryptophan resonances were measured in MestReNova x64 using Line Fitting function for manual fitting of peaks (*red line*). The line width values for W42 and W84 at concentrations 100 μM and 200 μM of 5FW LecA are shown in **Table SI**. Since PrOF NMR spectrum containing 50 μM 5FW LecA resulted in a low signal-to-noise spectrum, we did not measure line widths of tryptophan resonances. As result, 100 – 200 μM 5FW LecA concentrations can be used in future experiments.

9
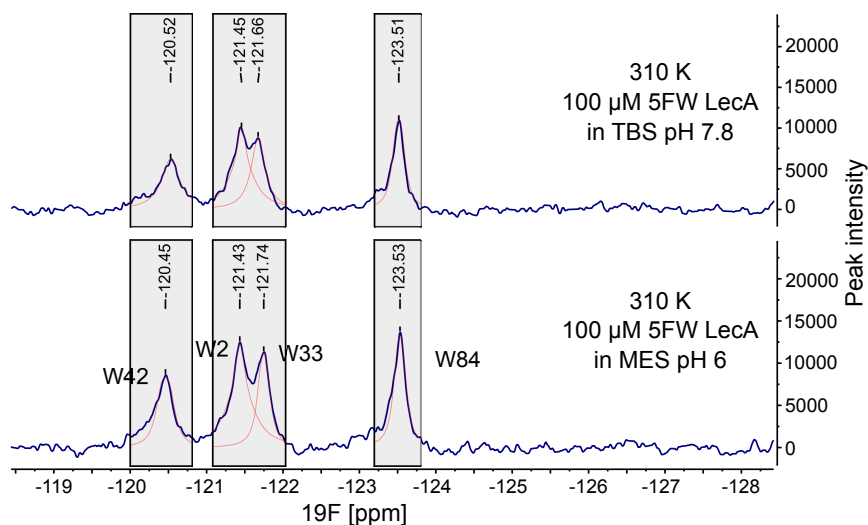
Fig. S3. Buffer optimization of 5FW LecA PrOF NMR. Shown is PrOF NMR spectrum of 100 µM holo 5FW LecA in a buffer with low (MES pH 6.0, *bottom*) *vs* medium salt (TBS pH 7.8, *top*) concentration. Both buffer systems show a comparable result on resolution of fluorine resonances. Peak line width values of tryptophan resonances were measured in MestReNova x64 using Line Fitting function for manual fitting of peaks (*red line*). The line width values for W42 and W84 are shown in **Table SI**, whereas line widths of W2 and W33 deviated between measurements due to the overlap of both signals.

10
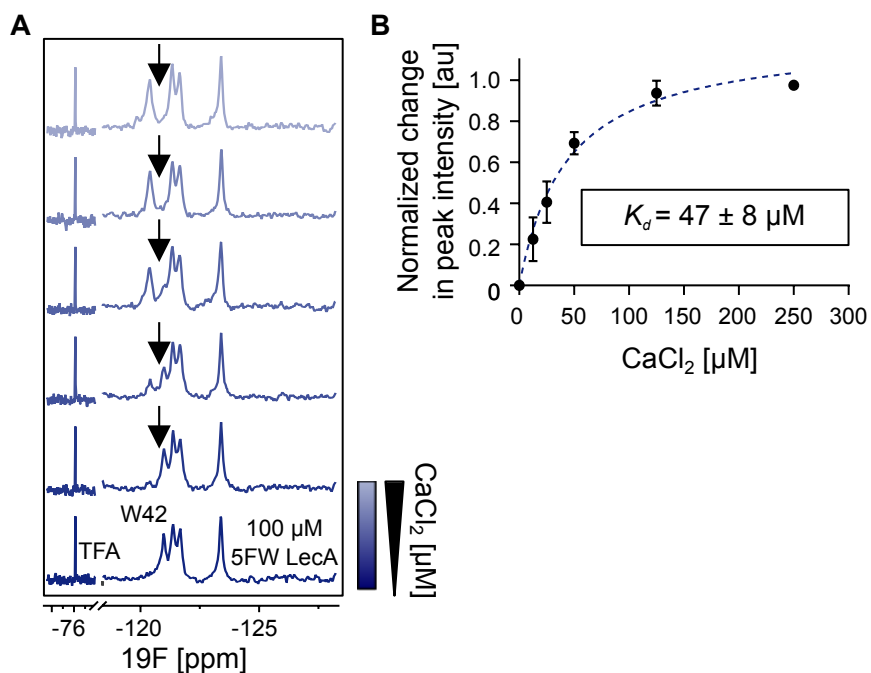
Fig. S4. PrOF NMR titration of Ca^{2+} to apo 5FW LecA. **A.** PrOF NMR titration of Ca^{2+} to Ca^{2+}−free (apo) 5FW LecA in MES pH 6 at 310 K. The W42 resonance undergoes a slow exchange on the chemical shift timescale and the signal intensity of Ca^{2+}−free W42 peak (*arrow*) can be followed to determine K_d. **B.** Shown is normalized change in unbound W42 peak intensity upon addition of Ca^{2+}. For this, the change in W42 peak intensity in presence of the Ca^{2+} compared to the reference was dived by W42 peak intensity of the reference spectrum delivering the normalized change in peak intensity values plotted on Y-axis. Titration data was fitted to one−site−binding model to obtain K_d value of 47±8 µM. The error bars represent the standard deviation of three independent titrations.

11
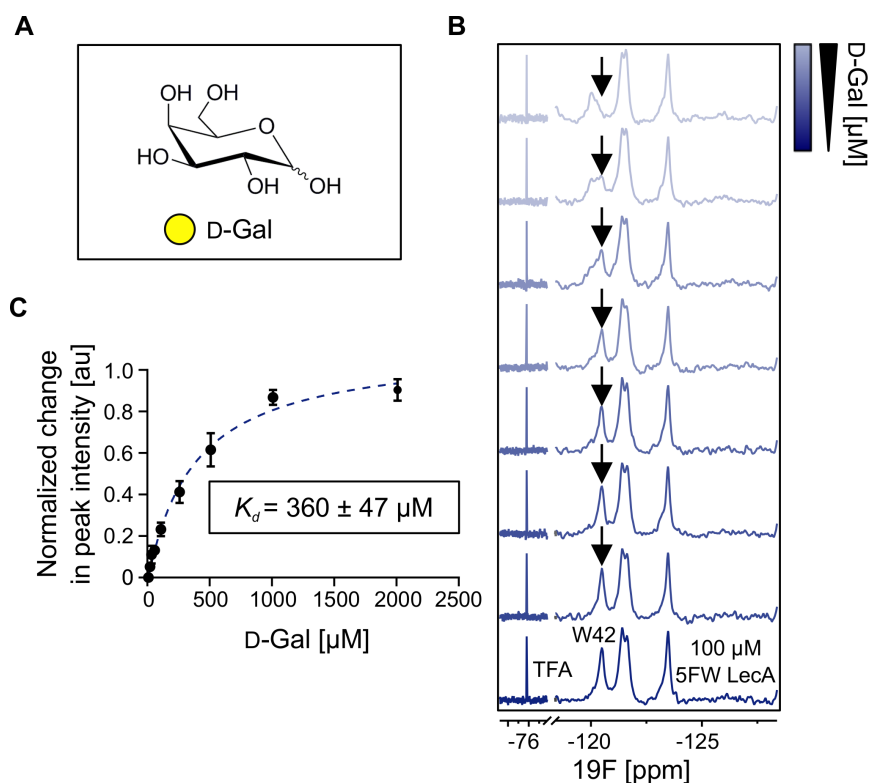
Fig. S5. PrOF NMR titration of D-Gal to holo 5FW LecA. **A.** The structure of D-Gal. **B.** PrOF NMR titration of D-Gal to holo 5FW LecA in TBS pH 7.8 at 310 K. The W42 resonance from free protein appeared with decreasing intensity (*arrow*) upon D-Gal addition and can be followed to determine K_d. **C.** Shown is normalized change in W42 peak intensity upon addition of D-Gal. For this, the change in W42 peak intensity in presence of the D-Gal was dived by W42 peak intensity of the reference spectrum delivering the normalized change in peak intensity values plotted on Y-axis. Titration data was fitted to one−site−binding model to obtain K_d value of 360 ± 47 µM. The error bars represent the standard deviation of three independent titrations.

12
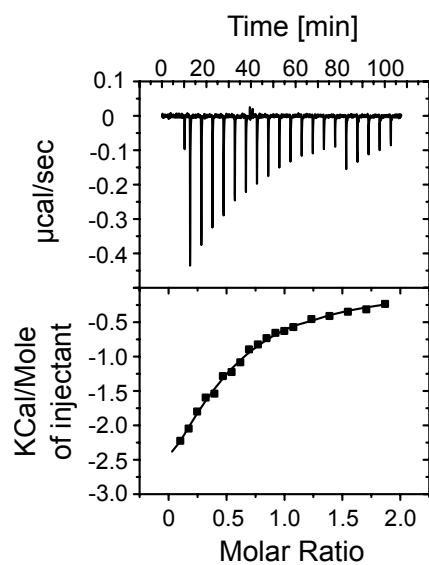
Fig. S6. ITC titration of Ca^{2+} to apo LecA. One representative graph of an isothermal titration microcalorimetry (ITC) of unlabeled LecA with Ca^{2+} is depicted. The K_d of 60±20 µM was determined from a minimum of three independent titrations.
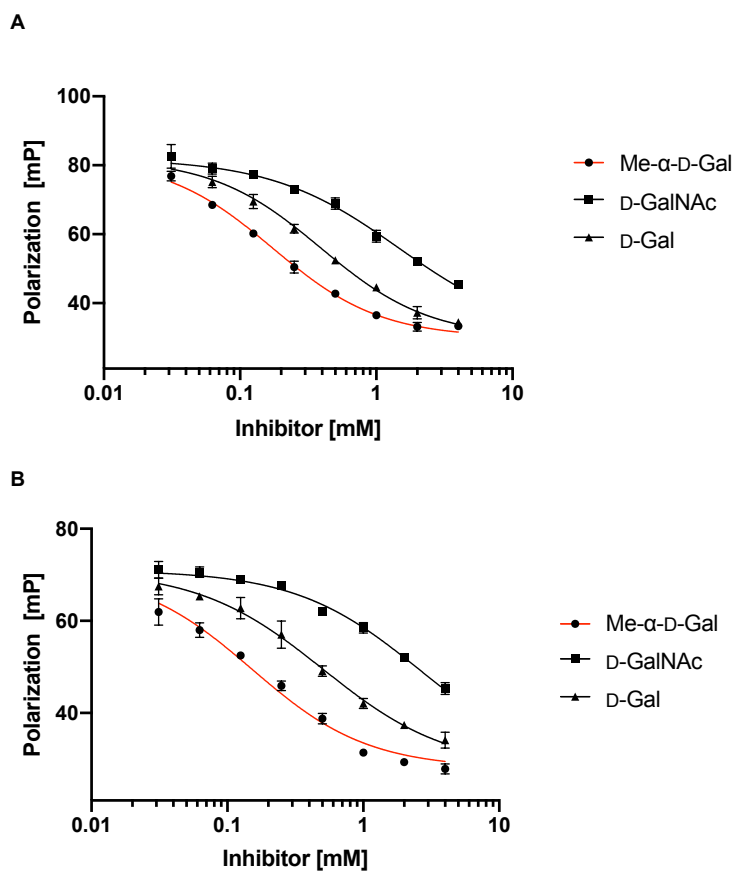
13

A

B

Fig. S7. Competitive binding assay based on fluorescence polarization with LecA and 5FW LecA. One representative titration of Me−α−D−Gal (positive control), D−GalNAc and D−Gal to **A.** LecA and **B.** 5FW LecA is depicted. IC_{50} values are given in **Table I** as average of three independent (LecA) or three technical replicates (5FW LecA).

14
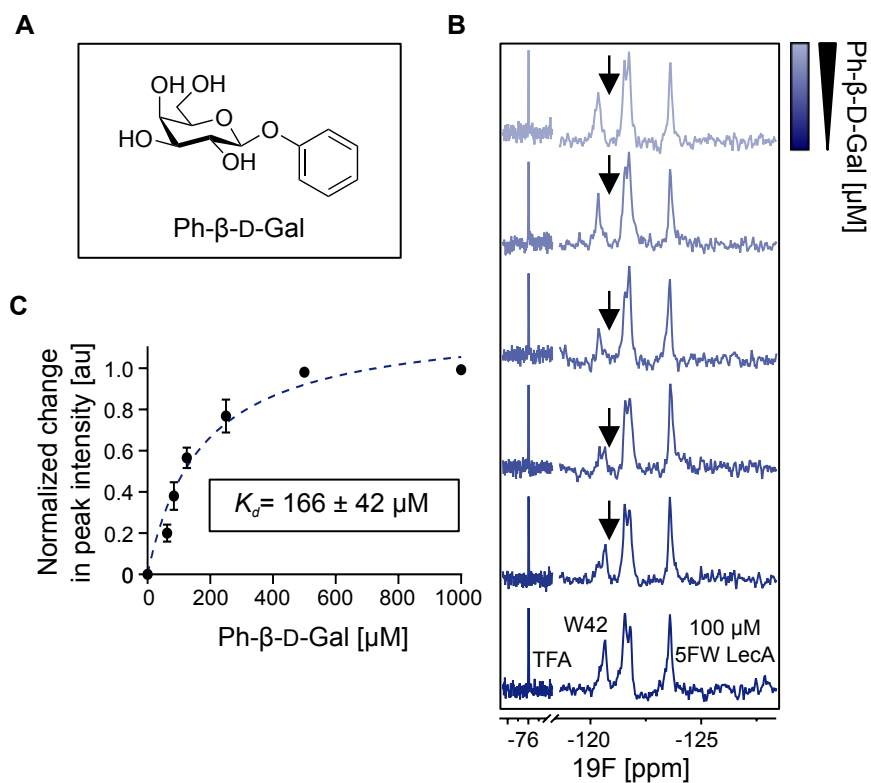
Fig. S8. PrOF NMR titration of Ph−β−D−Gal to holo 5FW LecA. **A.** The structure of Ph−β−D−Gal. **B.** The PrOF NMR titration of Ph−β−D−Gal to holo 5FW LecA in TBS pH 7.8 at 310 K. The W42 resonance from free protein appeared with decreasing intensity (*arrow*) upon Ph−β−D−Gal addition. The change in signal intensity of free W42 peak can be followed to determine K_d of 5FW LecA for Ph−β−D−Gal. **B.** Shown is normalized change in W42 peak intensity upon addition of Ph−β−D−Gal. For this, peak intensity in presence of Ph−β−D−Gal was dived by W42 peak intensity of the reference spectrum delivering the normalized change in peak intensity values plotted on Y-axis. Titration data was fitted to one−site−binding model to obtain K_d value of 166±42 µM. The error bars represent the standard deviation of three independent titrations.

15

7.2. Supporting Information for Subchapter 4.2.

Angewandte
Chemie
Eine Zeitschrift der Gesellschaft Deutscher Chemiker

Supporting Information

Automated Glycan Assembly of ^{19}F-labeled Glycan Probes Enables High-Throughput NMR Studies of Protein–Glycan Interactions

Giulio Fittolani[+], Elena Shanina[+], Mónica Guberman, Peter H. Seeberger,
Christoph Rademacher,* and Martina Delbianco*

anie_202102690_sm_miscellaneous_information.pdf

Table of Contents

1

A.   General materials and methods

All chemicals used were reagent grade and used as supplied unless otherwise noted. The automated syntheses were performed on a home-built synthesizer developed at the Max Planck Institute of Colloids and Interfaces. Analytical thin-layer chromatography (TLC) was performed on Merck silica gel 60 F254 plates (0.25 mm). Compounds were visualized by UV irradiation or dipping the plate in a staining solution (sugar stain: 10% H_2SO_4 in EtOH; CAM: 48 g/L ammonium molybdate, 60 g/L ceric ammonium molybdate in 6% H_2SO_4 aqueous solution). Flash column chromatography was carried out by using forced flow of the indicated solvent on Fluka Kieselgel 60 M (0.04 – 0.063 mm). Analysis and purification by normal and reverse phase HPLC was performed by using an Agilent 1200 series. Products were lyophilized using a Christ Alpha 2-4 LD plus freeze dryer. ^1H, ^{13}C and HSQC NMR spectra were recorded on a Varian 400-MR (400 MHz), Varian 600-MR (600 MHz), or Bruker Biospin AVANCE700 (700 MHz) spectrometer. Spectra were recorded in $CDCl_3$ by using the solvent residual peak chemical shift as the internal standard ($CDCl_3$: 7.26 ppm ^1H, 77.0 ppm ^{13}C) or in D_2O using the solvent as the internal standard in ^1H NMR (D_2O: 4.79 ppm ^1H). High resolution mass spectra were obtained using a 6210 ESI-TOF mass spectrometer (Agilent) and a MALDI-TOF autoflex^{TM} (Bruker). MALDI and ESI mass spectra were run on IonSpec Ultima instruments. IR spectra were recorded on a Perkin-Elmer 1600 FTIR spectrometer. Optical rotations were measured by using a Perkin-Elmer 241 and Unipol L1000 polarimeter.

2
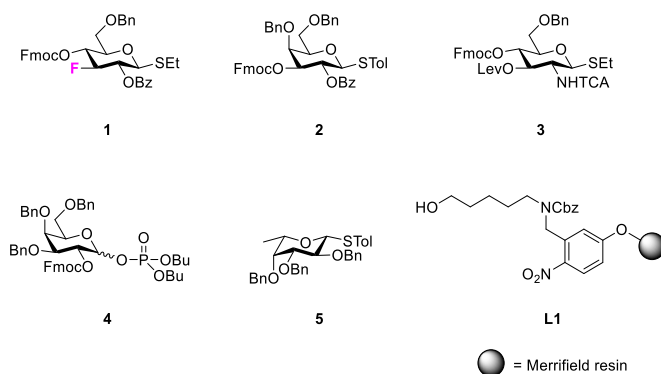
B.   Building blocks

Building blocks **1** and **4** were synthesized according to previous literature procedures.[1,2] Building blocks **2**, **3**, and **5** were purchased from GlycoUniverse (Germany, product codes Gal32.11140202, GlcN30.15131402, Fuc32.020202 respectively). Merrifield resin equipped with a photocleavable linker (**L1**, loading 0.30 mmol/g) was prepared according to previous literature.[3]

3

C.   Automated glycan assembly

General materials and method

The automated syntheses were performed on a home-built synthesizer developed at the Max Planck Institute of Colloids and Interfaces. All solvents used were HPLC-grade. The solvents used for the building block, activator, TMSOTf and capping solutions were taken from an anhydrous solvent system (J.C. Meyer) and further dried with molecular sieves (4 Å) for moisture-sensitive solutions. The building blocks were co-evaporated three times with toluene and dried for 1 h on high vacuum before use. Oven-heated, argon-flushed flasks were used to prepare all moisture-sensitive solutions. Activator, capping, deprotection, acidic wash and building block solutions were freshly prepared and kept under argon during the automation run. All yields of products obtained by AGA were calculated on the basis of resin loading. Resin loading was determined following previously established procedures.[4]

Preparation of stock solutions

- **Building block solution**: Between 0.06 and 0.10 mmol of building block (depending on the BB, see Module C1 and C2) was dissolved in DCM (1 mL).
- **NIS/TfOH activator solution**: 1.35 g (6.0 mmol) of recrystallized NIS was dissolved in 40 mL of a 2:1 v/v mixture of anhydrous DCM and anhydrous dioxane. Then triflic acid (55 µL, 0.6 mmol) was added. The solution is kept at 0°C for the duration of the automation run.
- **Fmoc deprotection solution**: A solution of 20% piperidine in DMF (v/v) was prepared.
- **Lev deprotection solution**: Hydrazine acetate (550 mg, 5.97 mmol) was dissolved in pyridine/AcOH/H2O (40mL, v/v, 32:8:2) and sonicated for 10 min.
- **TMSOTf solution**: TMSOTf (0.45 mL, 2.49 mmol) was added to DCM (40 mL).
- **Capping solution**: A solution of 10% acetic anhydride and 2% methanesulfunic acid in DCM (v/v) was prepared.

Modules for automated synthesis

Module A: Resin preparation for synthesis (20 min)

All automated syntheses were performed on 0.0135 mmol scale. Resin (**L1**, 45 mg) was placed in the reaction vessel and swollen in DCM for 20 min at room temperature prior to synthesis. During this time, all reagent lines needed for the synthesis were washed and primed. After the swelling, the resin was washed with DMF, THF, and DCM (three times each with 2 mL for 25 s).

4

Module B: Acidic wash with TMSOTf solution (20 min)

The resin was swollen in 2 mL DCM and the temperature of the reaction vessel was adjusted to -20 °C. Upon reaching the low temperature, TMSOTf solution (1 mL) was added drop wise to the reaction vessel. After bubbling for 3 min, the acidic solution was drained and the resin was washed with 2 mL DCM for 25 s.

Action	Cycles	Solution	Amount	T (°C)	Incubation time
Cooling	-	-	-	-20	(15 min)*
Deliver	1	DCM	2 mL	-20	-
Deliver	1	TMSOTf solution	1 mL	-20	3 min
Wash	1	DCM	2 mL	-20	25 sec

*Time required to reach the desired temperature.

Module C1: Thioglycoside glycosylation (35 min-55 min)

The building block solution (0.10 mmol of BB in 1 mL of DCM per glycosylation) was delivered to the reaction vessel. After the set temperature was reached, the reaction was started by dropwise addition of the NIS/TfOH activator solution (1.0 mL, excess). The glycosylation conditions (T_1, T_2, t_1, and t_2) are building block dependent and are reported in a table below. After completion of the reaction, the solution was drained and the resin was washed with DCM, DCM:dioxane (1:2, 3 mL for 20 s) and DCM (two times, each with 2 mL for 25 s). The temperature of the reaction vessel was increased to 25 °C for the next module.

Action	Cycles	Solution	Amount	T (°C)	Incubation time
Cooling	-	-	-	T_1	-
Deliver	1	BB solution	1 mL	T_1	-
Deliver	1	NIS/TfOH activator solution	1 mL	T_1	-
Reaction time (BB dependent)	1			T_1 to T_2	t_1 t_2
Wash	1	DCM	2 mL	T_2	5 sec
Wash	1	DCM : Dioxane (1:2)	2 mL	T_2	20 sec
Heating	-	-	-	25	-
Wash	2	DCM	2 mL	> 0	25 sec

The AGA glycosylation conditions employed for thioglycoside BBs were previously reported.[1,2]

5

BB	Equiv.	t_1 (min)	T_1 (°C)	t_2 (min)	T_2 (°C)
1	8	5	-20	20	0
2	8	5	-20	20	0
3	8	5	-20	40	0
5	8	5	-40	20	-20

Module C2: Glycosyl phosphate glycosylation (45 min)

The building block solution (0.06 mmol of BB in 1 mL of DCM per glycosylation) was delivered to the reaction vessel. After the set temperature was reached, the reaction was started by drop wise addition of the TMSOTf solution (1.0 mL, stoichiometric). After completion of the reaction, the solution was drained and the resin washed with DCM (six times, each with 2 mL for 25 s). The temperature of the reaction vessel was increased to 25 °C for the next module.

Action	Cycles	Solution	Amount	T (°C)	Incubation time
Cooling	-	-	-	-35	-
Deliver	1	BB solution	1 mL	-35	-
Deliver	1	TMSOTf solution	1 mL	-35	-
Reaction time (BB dependent)	1			-35 to -15	5 min 30 min
Wash	1	DCM	2 mL	-15	5 sec
Heating	-	-	-	25	-
Wash	6	DCM	2 mL	> 0	25 sec

The AGA glycosylation conditions employed for the glycosyl phosphate BB were previously reported.[2]

BB	Equiv.	t_1 (min)	T_1 (°C)	t_2 (min)	T_2 (°C)
4	5	5	-35	30	-15

Module D: Capping (30 min)

The resin was washed with DMF (two times with 2 mL for 25 s) and the temperature of the reaction vessel was adjusted to 25 °C. 2 mL of Pyridine solution (10% in DMF) was delivered into the reaction vessel. After 1 min, the reaction solution was drained and the resin washed with DCM (three times with 3 mL for 25 s). 4 mL of capping solution was delivered into the

6

117

reaction vessel. After 20 min, the reaction solution was drained and the resin washed with DCM (three times with 3 mL for 25 s).

Action	Cycles	Solution	Amount	T (°C)	Incubation time
Heating	-	-	-	25	(5 min)*
Wash	2	DMF	2 mL	25	25 sec
Deliver	1	10% Pyridine in DMF	2 mL	25	1 min
Wash	3	DCM	2 mL	25	25 sec
Deliver	1	Capping Solution	4 mL	25	20 min
Wash	3	DCM	2 mL	25	25 sec

*Time required to reach the desired temperature.

Module E1: Fmoc deprotection (9 min)

The resin was washed with DMF (three times with 2 mL for 25 s) and the temperature of the reaction vessel was adjusted to 25 °C. 2 mL of Fmoc deprotection solution was delivered to the reaction vessel and kept under Ar bubbling. After 5 min, the reaction solution was drained and the resin washed with DMF (three times with 3 mL for 25 s) and DCM (five times each with 2 mL for 25 s). The temperature of the reaction vessel was decreased to -20 °C for the next module.

Action	Cycles	Solution	Amount	T (°C)	Incubation time
Wash	3	DMF	2 mL	25	25 sec
Deliver	1	Fmoc depr. solution	2 mL	25	5 min
Wash	1	DMF	2 mL		
Cooling	-	-	-	-20	-
Wash	3	DMF	2 mL	< 25	25 sec
Wash	5	DCM	2 mL	< 25	25 sec

Module E2: Lev deprotection (65 min)

The resin was washed with DCM (three times with 2 mL for 25 s). DCM (1.3 mL) was delivered to the reaction vessel and the temperature of the reaction vessel was adjusted to 25 °C. 2 mL of Lev deprotection solution was delivered to the reaction vessel that was kept under pulsed Ar bubbling for 30 min. This procedure was repeated twice. The reaction solution was drained and the resin washed with DMF (three times with 3 mL for 25 s) and DCM (five times each with 2 mL for 25 s).

Action	Cycles	Solution	Amount	T (°C)	Incubation

7

118

					time
Wash	3	DMF	2 mL	25	25 sec
Deliver	2	Lev depr. solution	2 mL	25	30 min
Wash	1	DMF	2 mL		
Cooling	-	-	-	-20	-
Wash	3	DMF	2 mL	< 25	25 sec
Wash	5	DCM	2 mL	< 25	25 sec

Note:

With the current setup the automated synthesizer has four BB lines. Therefore, for AGA syntheses requiring the use of five BBs (*i.e.* AGA of **F-Le^x**, **F-H type 2**, and **F-Le^y**) a first cycle with BB **1** was performed and, upon completion, BB **1** was replaced by BB **5** solution to continue the AGA.

Post-synthesizer manipulations (Post-AGA)

Module F: On-resin methanolysis

The resin was suspended THF (4 mL). 0.4 mL of NaOMe in MeOH (0.5 M) was added and the suspension was gently shaken at room temperature. After micro-cleavage (see Module G1) indicated the complete removal of benzoyl groups, the resin was repeatedly washed with MeOH (2mL x 3) and DCM (2mL x 3).

Module G: Cleavage from solid support

The oligosaccharides were cleaved from the solid support using a continuous-flow photoreactor as described previously.[5]

Module G1: Micro-cleavage from solid support

Trace amount of resin (around 20 beads) was dispersed in DCM (0.1 mL) and irradiated with a UV lamp (6 watt, 356 nm) for 10 minutes. ACN (10 μL) was then added to the resin and the resulting solution analyzed by MALDI.

Module H1: Hydrogenolysis

The crude compound obtained from *Module G* was dissolved in 2 mL of EA:*t*BuOH:H_2O (2:1:1). 100% by weight Pd-C (10%) was added and the reaction was stirred in H_2 bomb with 60 psi pressure. The reaction progress was monitored to avoid undesired side products formation.

8

Upon completion, the reaction was filtered and washed with EA, *t*BuOH and H_2O. The filtrates were concentrated *in vacuo.*

Module H2: Hydrogenolysis at ambient pressure

The crude compound obtained from *Module G* was dissolved in 2 mL of EA:*t*BuOH:H_2O (2:1:1). 100% by weight Pd-C (10%) was added and the reaction was stirred in a flask equipped with a H_2 balloon. The reaction progress was monitored to avoid undesired side products formation. Upon completion, the reaction was filtered and washed with EA, *t*BuOH and H_2O. The filtrates were concentrated *in vacuo.*

Module I: Purification

The purification of the crudes was conducted using a C_{18} silica column or reverse phase HPLC (Agilent 1200 Series, Method B and Method C). The pure compound was analyzed using analytical HPLC (Agilent 1200 Series, Method A).

- **Method A:** (Hypercarb column, 150 x 4.6 mm, 3 µm) flow rate of 0.7 mL/min with H_2O (0.1% formic acid) as eluents [isocratic (5 min), linear gradient to 30% ACN (30 min), linear gradient to 100% ACN (5 min)].
- **Method B:** (Hypercarb column, 150 x 10 mm, 5 µm), flow rate of 3 mL /min with H_2O (0.1% formic acid) as eluents [isocratic (5 min), linear gradient to 30% ACN (30 min), linear gradient to 100% ACN (5 min)].
- **Method C:** (Manual reverse phase C_{18} silica gel column chromatography): H_2O (0.1% formic acid, 10 mL), 3% MeOH (10 mL), 6% MeOH (10 mL), 9% MeOH (10 mL), 15% MeOH (10 mL).

Following final purification, all deprotected products were lyophilized on a Christ Alpha 2-4 LD plus freeze dryer prior to characterization.

9

Oligosaccharides synthesis
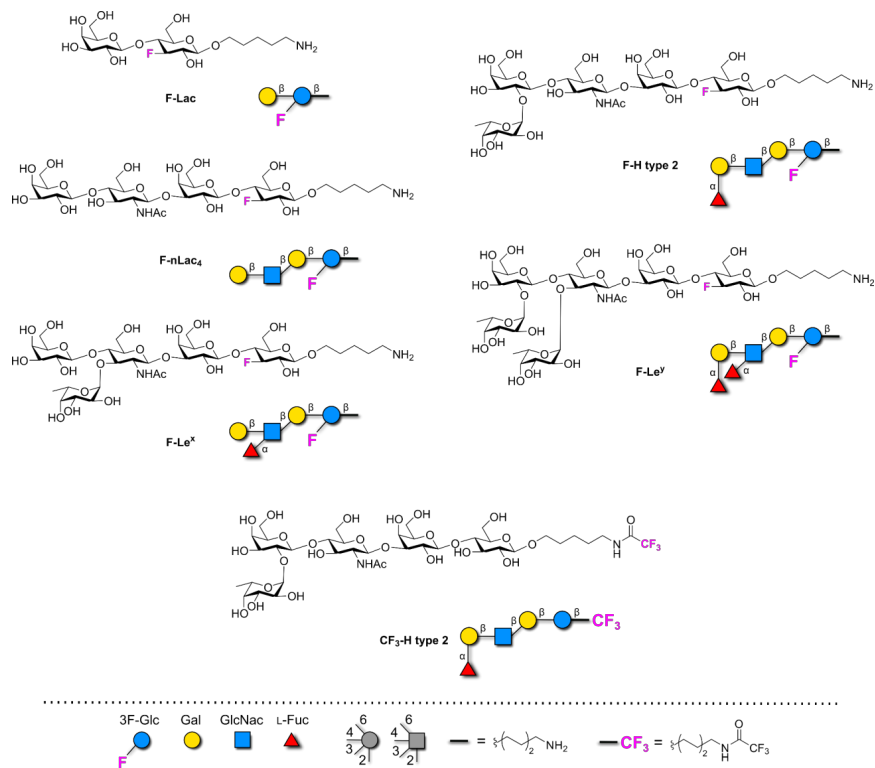
Figure S1 Collection of ^{19}F labelled Lewis type-II antigens synthesized by AGA.

10

Synthesis of **F-Lac**

F-Lac

Step	BB	Modules	Notes
AGA		**A**	**L1** swelling
	1	**B**, **C1**, **D**, **E1**	**C1**: (**1**, -20°C for 5 min, 0°C for 20 min)
	2	**B**, **C1**, **D**, **E1**	**C1**: (**2**, -20°C for 5 min, 0°C for 20 min)
Post-AGA		**F**, **G**, **H2**, **I**	**F**: (16 h)
			H2: (3 h)
			I: (Method B, t_R = 19.6 min)

Automated synthesis, global deprotection, and purification afforded **F-Lac** as a white solid (1.17 mg, 20% overall yield).

Analytical data for **F-Lac**:

^{1}H NMR (600 MHz, Deuterium Oxide) δ 4.67 – 4.55 (m, 1H, **H-1 Glc**, H-3 Glc), 4.54 (d, J = 8.1 Hz, 1H, **H-1 Gal**), 4.50 (d, J = 7.9 Hz, 1H), 4.05 – 3.92 (m, 4H), 3.87 (dd, J = 12.3, 5.1 Hz, 1H), 3.81 (dd, J = 11.8, 7.7 Hz, 1H), 3.78 – 3.65 (m, 4H), 3.64 – 3.53 (m, 3H), 3.01 (t, J = 7.5 Hz, 2H, CH_2-NH$_3^+$ linker), 1.70 (h, J = 7.4, 6.8 Hz, 4H, 2x CH_2 linker), 1.47 (p, J = 7.7 Hz, 2H, CH_2 linker). ^{13}C NMR (151 MHz, Deuterium Oxide) δ 102.81 (s, **C-1 Gal**), 101.11 (d, J = 12.1 Hz, **C-1 Glc**), 94.77 (d, J = 183.6 Hz, C-3 Glc), 75.46 (d, J = 16.9 Hz), 75.18, 73.80 (d, J = 8.1 Hz), 72.57, 71.72 (d, J = 18.1 Hz), 70.92, 70.17 (s, CH_2-O linker), 68.43, 60.79, 59.77, 39.31 (s, CH_2-

11

NH_3^+ linker), 28.07 (s, CH_2 linker), 26.49 (s, CH_2 linker), 22.01 (s, CH_2 linker). ^{19}F NMR (564 MHz, Deuterium Oxide) δ -192.23 (dt, J = 52.1, 14.0 Hz). (ESI-HRMS) m/z 430.208 [M+H]$^+$ ($C_{17}H_{33}FNO_{10}$ requires 430.208).

RP-HPLC of F-Lac (ELSD trace, Method A, t_R = 21.4 min)

12

123
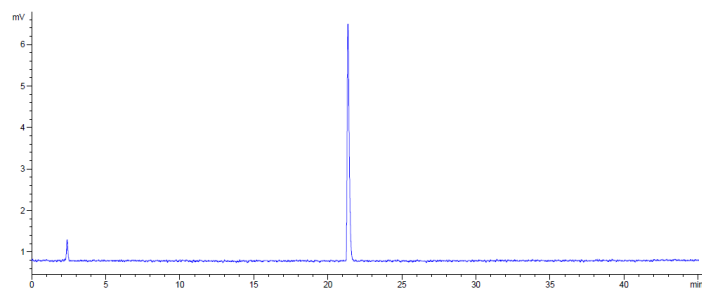
1H NMR of F-Lac (600 MHz, D2O)
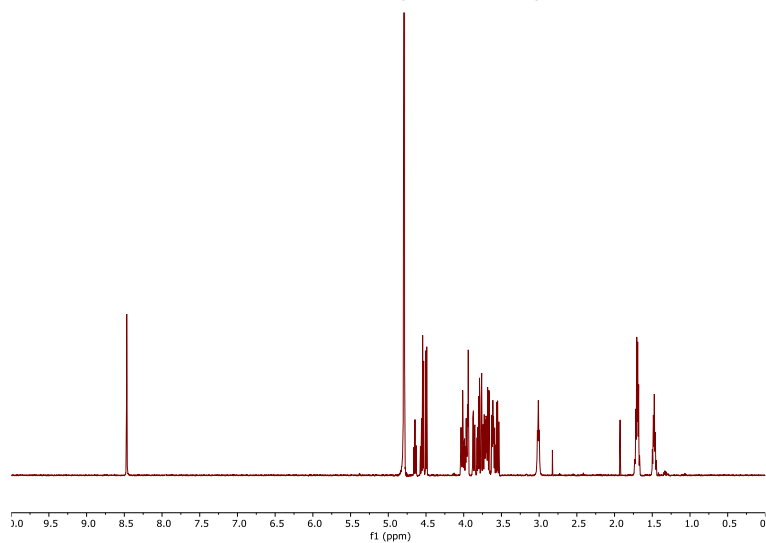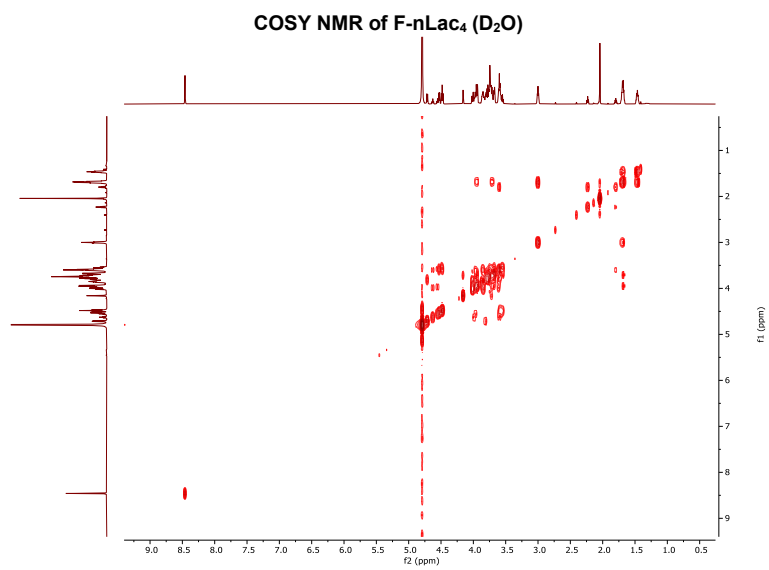
13C NMR of F-Lac (151 MHz, D2O)

13
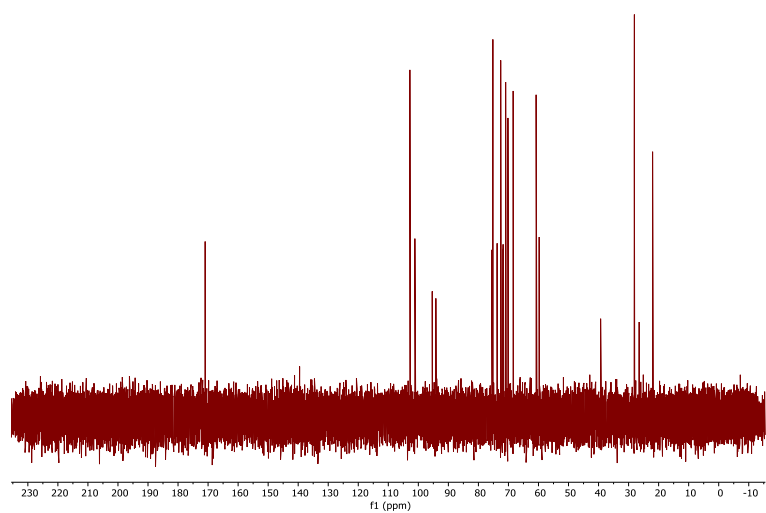
¹⁹F NMR of F-Lac (564 MHz, D₂O)

HSQC NMR of F-Lac (D₂O)

14

COSY NMR of F-Lac (D$_2$O)

15
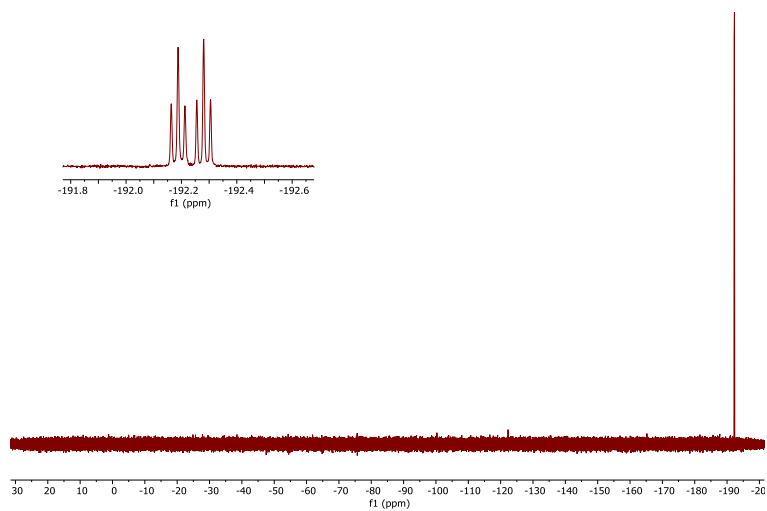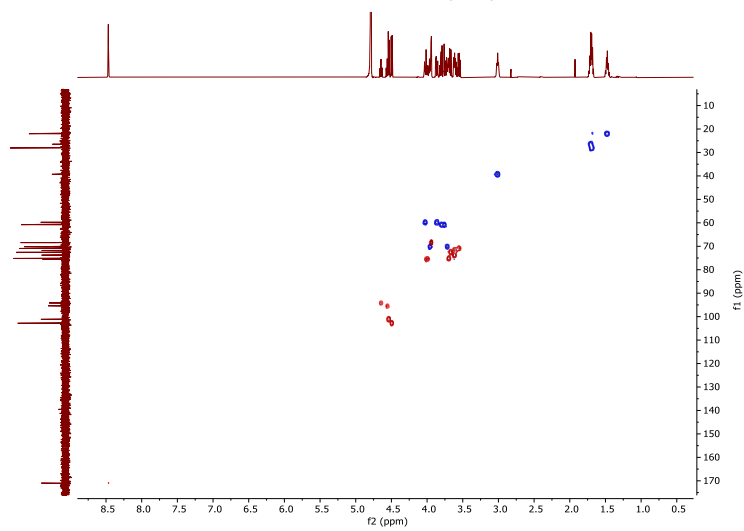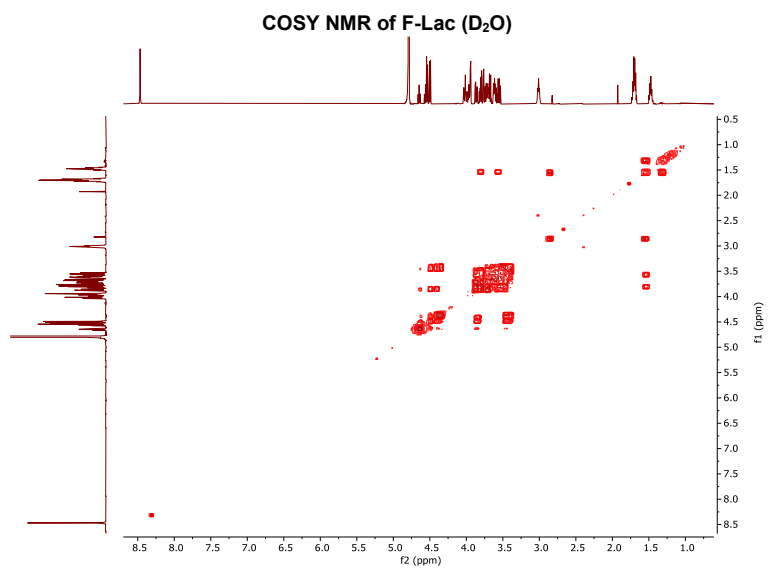
Synthesis of **F-nLac₄**
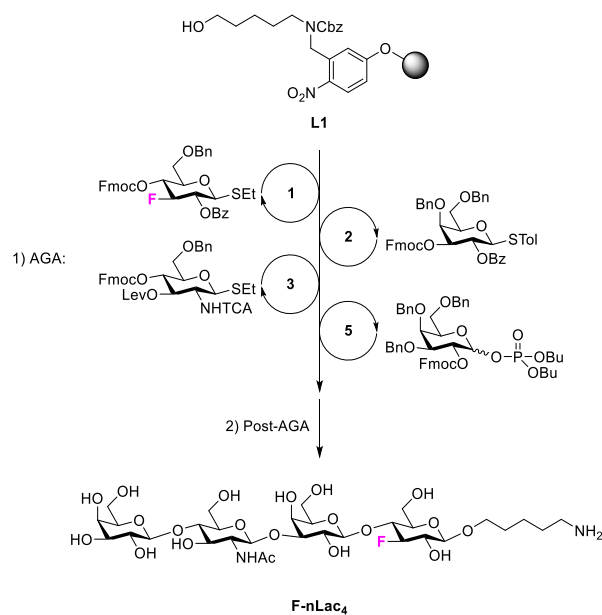
L1

1) AGA:

2) Post-AGA

F-nLac₄

Step	BB	Modules	Notes
AGA		**A**	**L1** swelling
	1	**B**, **C1**, **D**, **E1**	**C1**: (**1**, -20°C for 5 min, 0°C for 20 min)
	2	**B**, **C1**, **D**, **E1**	**C1**: (**2**, -20°C for 5 min, 0°C for 20 min)
	3	**B**, **C1**, **D**, **E1**	**C1**: (**3**, -20°C for 5 min, 0°C for 40 min)
	4	**B**, **C2**, **D**, **E1**, **E2**	**C2**: (**4**, -35°C for 5 min, -15°C for 30 min)
Post-AGA	**F**, **G**, **H1**, **I**		**F:** (3 d)
			H1: (3 d)
			I: (Method C)

Automated synthesis, global deprotection, and purification afforded **F-nLac₄** as white solid (1.64 mg, 16% overall yield).

Analytical data for **F-nLac₄**:

16

^{1}H NMR (700 MHz, Deuterium oxide) 4.71 (d, *J* = 8.4 Hz, 1H, **H-1**), 4.59 (dt, *J* = 52.0, 7.8 Hz, 1H, H-3 Glc), 4.53 (d, *J* = 8.0 Hz, 1H, **H-1**), 4.48 (t, *J* = 8.8 Hz, 2H, 2x **H-1**), 4.16 (s, 1H), 4.07 – 3.89 (m, 5H), 3.89 – 3.65 (m, 14H), 3.64 – 3.50 (m, 5H), 3.00 (t, *J* = 7.7 Hz, 2H, C*H$_2$*-NH$_3^+$ linker), 2.04 (s, 3H, C*H$_3$* NHAc), 1.73 – 1.64 (m, 4H, 2x C*H$_2$* linker), 1.53 – 1.43 (m, 2H, *CH$_2$* linker). ^{13}C NMR (151 MHz, Deuterium oxide) δ 174.84, 102.83 (s, **C-1**), 102.80 (s, **C-1**), 102.66 (s, **C-1**), 101.13 (s, **C-1**) 94.72 (d, *J* = 184.3 Hz, C-3 Glc), 82.13, 78.14, 75.57, 75.29, 74.72, 74.50, 73.81, 72.44, 72.12, 71.66, 70.90, 70.15 (s, *CH$_2$*-O linker), 69.92, 68.48, 68.18, 60.96, 60.72, 59.78, 55.14, 39.28 (s, *CH$_2$* linker), 28.06 (s, *CH$_2$* linker), 26.33 (s, *CH$_2$* linker), 22.11 (s, *CH$_3$* NHAc), 21.99 (s, *CH$_2$* linker). ^{19}F NMR (564 MHz, Deuterium oxide) δ -192.16 (dt, *J* = 52.1, 14.0 Hz). (ESI-HRMS) m/z 795.344 [M+H]$^+$ (C$_{31}$H$_{56}$FN$_2$O$_{20}$ requires 795.341).

RP-HPLC of F-nLac$_4$ (ELSD trace, Method A, t$_R$ = 28.1 min)

17
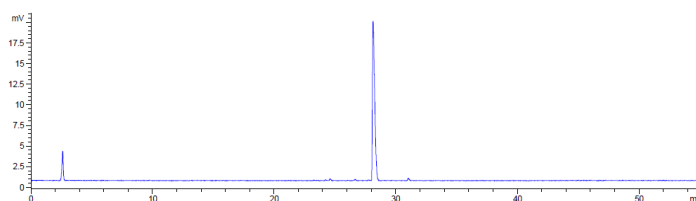
^{1}H NMR of F-nLac$_4$ (700 MHz, D$_2$O)
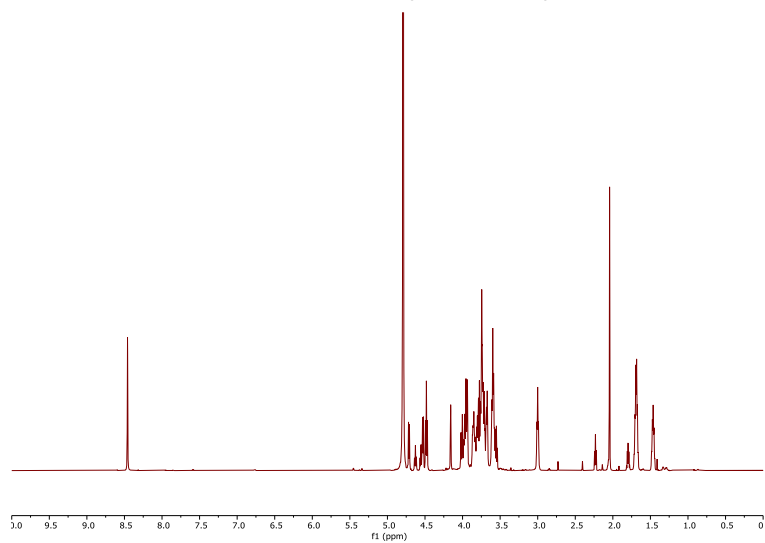
^{13}C NMR of F-nLac$_4$ (151 MHz, D$_2$O)

18
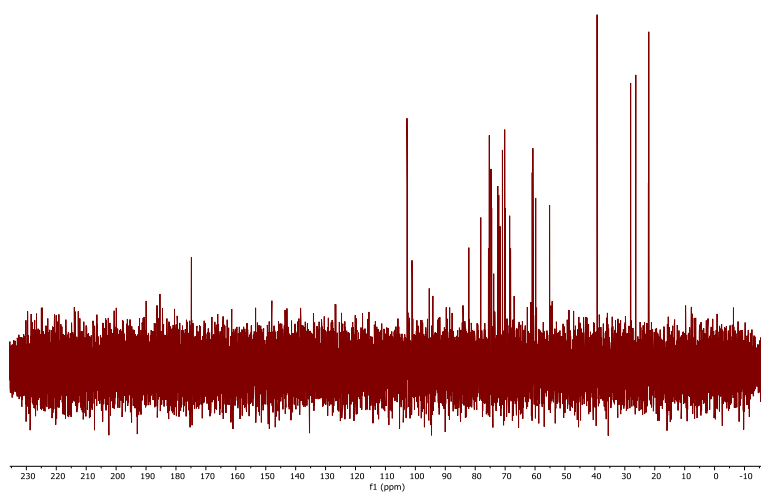
^{19}F NMR of F-nLac$_4$ (376 MHz, D$_2$O)

HSQC NMR of F-nLac$_4$ (D$_2$O)

19

COSY NMR of F-nLac$_4$ (D$_2$O)

20

Synthesis of **F-Le^x**

L1

1) AGA:

2) Post-AGA
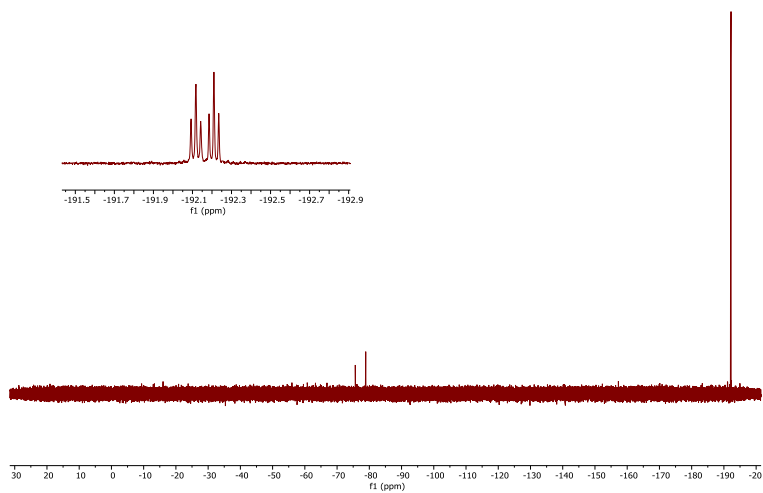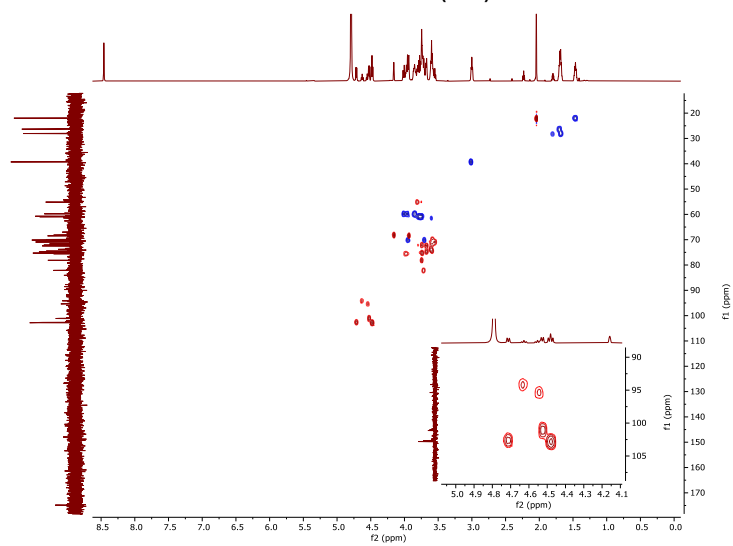
F-Le^x

Step	BB	Modules	Notes
AGA		**A**	**L1** swelling
	1	**B**, **C1**, **D**, **E1**	**C1**: (**1**, -20°C for 5 min, 0°C for 20 min)
	2	**B**, **C1**, **D**, **E1**	**C1**: (**2**, -20°C for 5 min, 0°C for 20 min)
	3	**B**, **C1**, **D**, **E2**	**C1**: (**3**, -20°C for 5 min, 0°C for 40 min)
	5	**B**, **C1**, **D**, **E1**	**C1**: (**5**, -40°C for 5 min, -20°C for 20 min)
	4	**B**, **C2**, **D**, **E1**	**C2**: (**4**, -35°C for 5 min, -15°C for 30 min)
Post-AGA		**F**, **G**, **H1**, **I**	**F:** (3 d)
			H1: (3 d)
			I: (Method C)

Automated synthesis, global deprotection, and purification afforded **F-Le^x** as a white solid (1.1 mg, 8% overall yield).

21

Analytical data for **F-Le^x**:

^{1}H NMR (700 MHz, Deuterium Oxide) δ 5.14 (s, 1H, **H-1 Fuc α-1,3**), 4.89 – 4.83 (m, 1H, H-5 Fuc α-1,3), 4.72 (d, J = 8.4 Hz, 1H, **H-1**), 4.67 – 4.51 (m, 2H, H-3 Glc, **H-1**), 4.50 – 4.44 (m, 2H, 2x **H-1**), 4.21 – 4.13 (m, 1H), 4.06 – 3.82 (m, 11H), 3.82 – 3.64 (m, 10H), 3.59 (ddt, J = 17.3, 10.2, 3.8 Hz, 5H), 3.51 (t, J = 8.9 Hz, 1H), 3.00 (t, J = 7.6 Hz, 2H, CH_2-NH$_3^+$ linker), 2.03 (s, 3H, CH_3 NHAc), 1.68 (p, J = 7.8, 7.1 Hz, 4H, 2x CH_2 linker), 1.46 (p, J = 7.5 Hz, 2H, CH_2 linker), 1.18 (d, J = 6.7 Hz, 3H, CH_3 Fuc). ^{13}C NMR (151 MHz, Deuterium oxide) δ 170.97 (s, C=O NHAc), 102.82 (s, **C-1**), 102.46 (s, **C-1**), 101.71 (s, **C-1**), 101.13 (s, **C-1**), 98.54 (s, **C-1 Fuc α-1,3**), 94.73 (d, J = 178.3 Hz, C-3 Glc), 82.16, 75.07, 74.85, 74.71, 73.81, 73.01, 72.42, 71.85, 70.99, 70.15 (s, CH_2-O linker), 69.92, 69.14, 68.29, 68.16, 67.65, 66.63 (s, C-5 Fuc α-1,3), 61.44, 60.73, 59.58, 39.29 (s, CH_2-NH$_3^+$ linker), 28.06 (s, CH_2 linker), 26.35 (s, CH_2 linker), 22.19 (s, CH_3 NHAc), 21.99 (s, CH_2 linker), 15.24 (s, CH_3 Fuc). ^{19}F NMR (564 MHz, Deuterium oxide) δ -192.15 (dt, J = 52.2, 14.0 Hz). (ESI-HRMS) m/z 941.403 [M+H]$^+$ ($C_{37}H_{66}FN_2O_{24}$ requires 941.398).

RP-HPLC of F-Le^x (ELSD trace, Method A, t_R = 24.3 min)

22

133
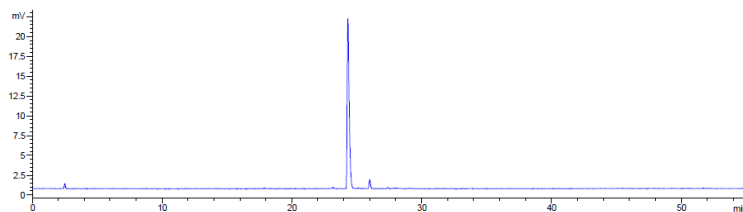
¹H NMR of F-Leˣ (700 MHz, D₂O)
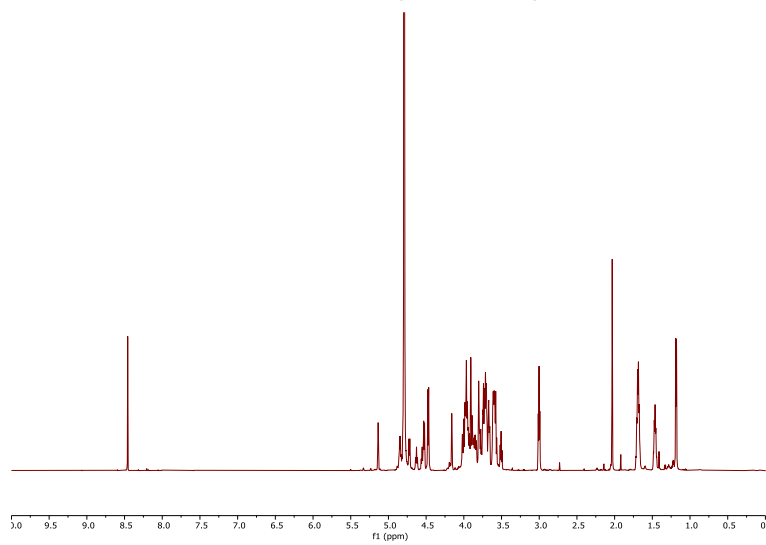

¹³C NMR of F-Leˣ (151 MHz, D₂O)

23
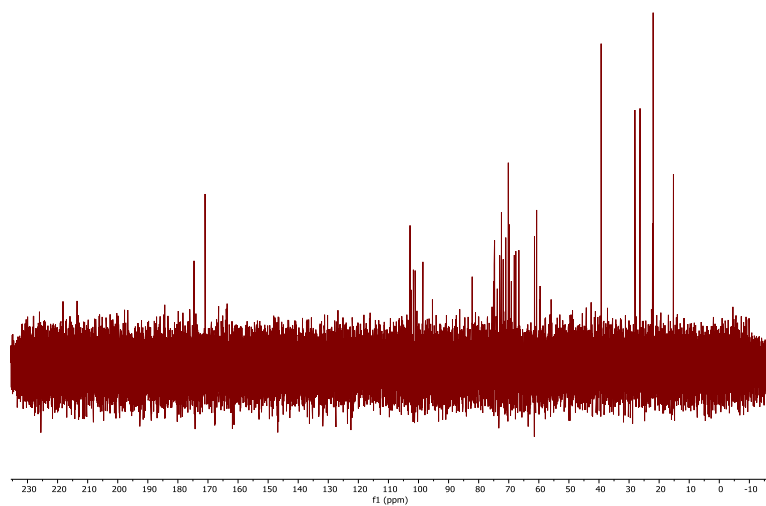
¹⁹F NMR of F-Leˣ (376 MHz, D₂O)
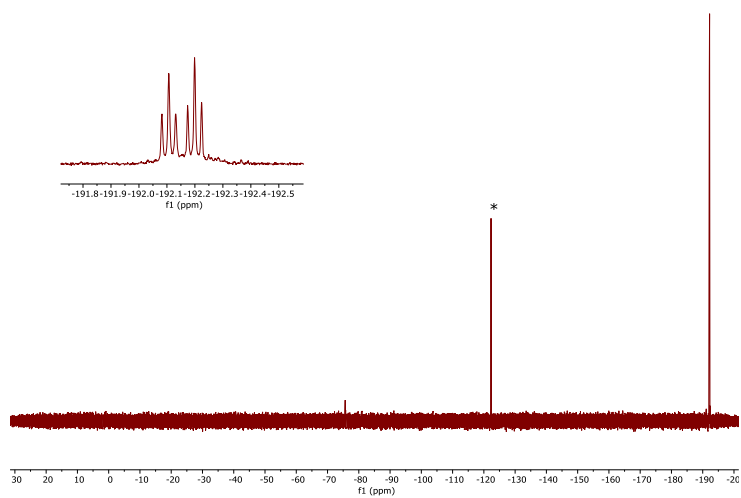
* The peak at -122.29 ppm is a fluoride impurity

HSQC NMR of F-Leˣ (D₂O)

24
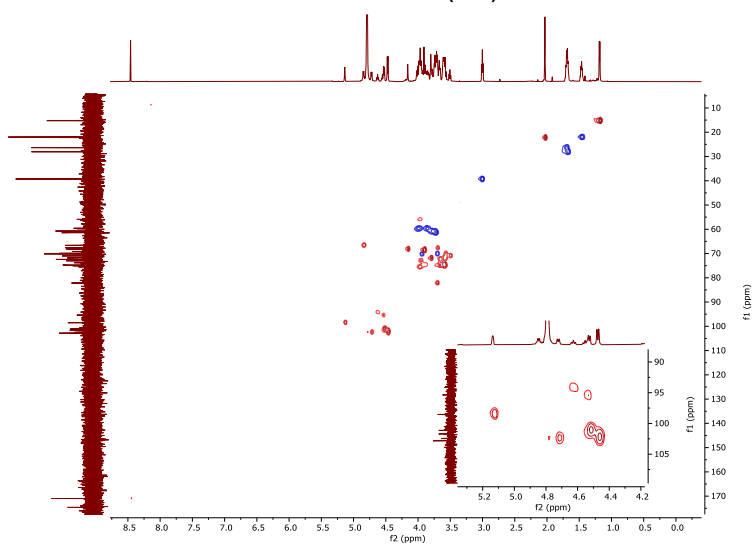
COSY NMR of F-Le^x (D$_2$O)

25
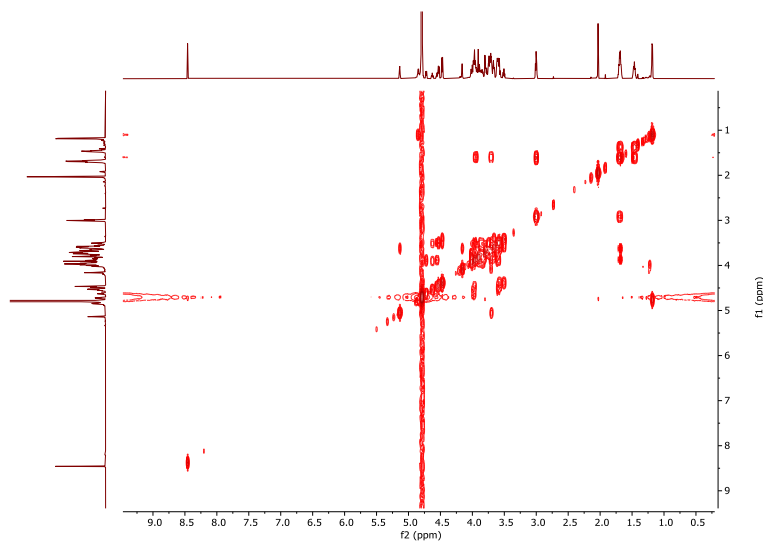
Synthesis of **F-H type 2**

F-H type 2

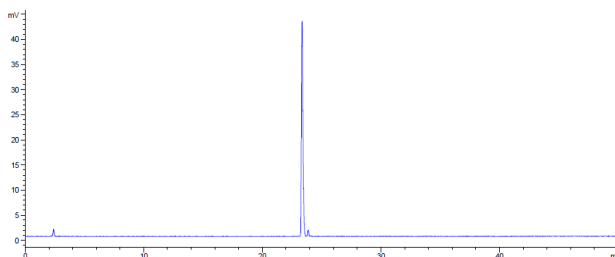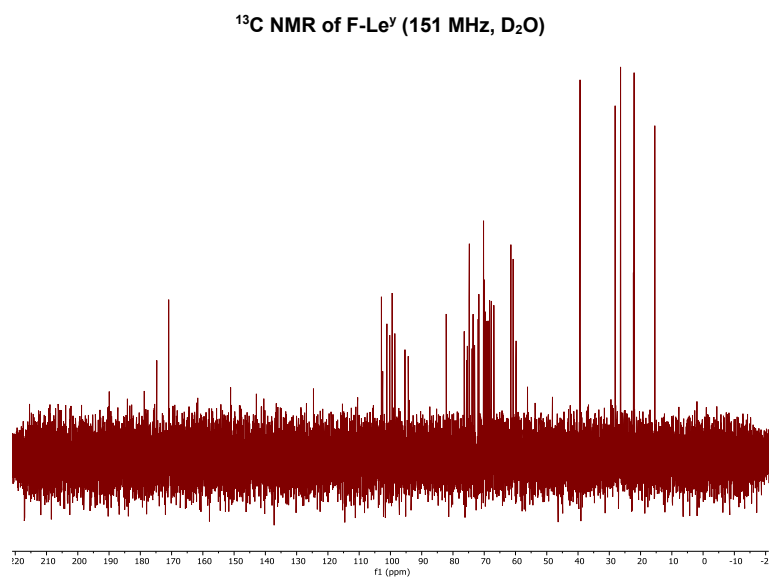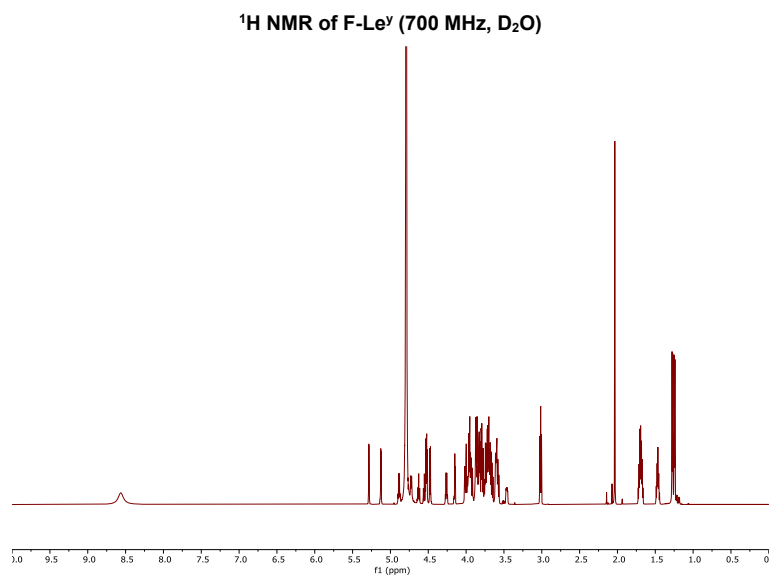| Step	BB	Modules	Notes
		A	**L1** swelling
	1	**B**, **C1**, **D**, **E1**	**C1**: (**1**, -20°C for 5 min, 0°C for 20 min)
	2	**B**, **C1**, **D**, **E1**	**C1**: (**2**, -20°C for 5 min, 0°C for 20 min)
AGA	**3**	**B**, **C1**, **D**, **E1**	**C1**: (**3**, -20°C for 5 min, 0°C for 40 min)
	4	**B**, **C2**, **D**, **E1**	**C2**: (**4**, -35°C for 5 min, -15°C for 30 min)
	5	**B**, **C1**, **D**, **E2**	**C1**: (**5**, -40°C for 5 min, -20°C for 20 min)
Post-AGA		**F**, **G**, **H1**, **I**	**F**: (2 d)
			H1: (3 d)
			I: (Method C)

26

Automated synthesis, global deprotection, and purification afforded **F-H type 2** as a white solid (1.3 mg, 10% overall yield).

Analytical data for **F-H type 2**:

^{1}H NMR (700 MHz, Deuterium Oxide) δ 5.32 (s, 1H, **H-1 Fuc α-1,2**), 4.71 (d, J = 8.5 Hz, 1H, **H-1**), 4.68 – 4.44 (m, 4H, 4x **H-1**, H-3 Glc), 4.23 (q, J = 6.5 Hz, 1H, H-5 Fuc α-1,2), 4.20 – 3.64 (m, 25H), 3.64 – 3.50 (m, 2H), 3.47 (m, 1H), 3.01 (t, J = 7.6 Hz, 2H, CH_2-NH$_3^+$ linker), 2.05 (s, 3H, CH_3 NHAc), 1.80 – 1.62 (m, 4H, 2x CH_2 linker), 1.46 (p, J = 8.0 Hz, 2H, CH_2 linker), 1.27 – 1.20 (d, J = 6.1 Hz, 3H, CH_3 Fuc). ^{13}C NMR (151 MHz, Deuterium oxide) δ 170.79 (s, C=O NHAc), 102.82 (s, **C-1**), 102.68 (s, **C-1**), 101.13 (s, **C-1**), 100.19 (s, **C-1**), 99.35 (s, **C-1 Fuc α-1,2**), 94.75 (d, J = 196.3 Hz, C-3 Glc), 82.07, 76.39, 75.85, 75.58, 75.19, 75.04, 74.70, 73.76, 73.46, 72.00, 71.61, 70.15 (s, CH_2-O linker), 69.94, 69.55, 69.05, 68.14, 66.87 (s, C-5 Fuc α-1,2), 61.05, 60.68, 59.78, 55.33, 39.28 (s, CH_2-NH$_3^+$ linker), 28.06 (s, CH_2 linker), 26.33 (s, CH_2 linker), 22.13 (s, CH_3 NHAc), 21.99 (s, CH_2 linker), 15.24 (s, CH_3 Fuc). ^{19}F NMR (564 MHz, Deuterium oxide) δ -192.17 (dt, J = 52.1, 14.1 Hz). (ESI-HRMS) m/z 941.404 [M+H]$^+$ (C$_{37}$H$_{66}$FN$_2$O$_{24}$ requires 941.398).

RP-HPLC of F-H type 2 (ELSD trace, Method A, t$_R$ = 27.4 min)

27
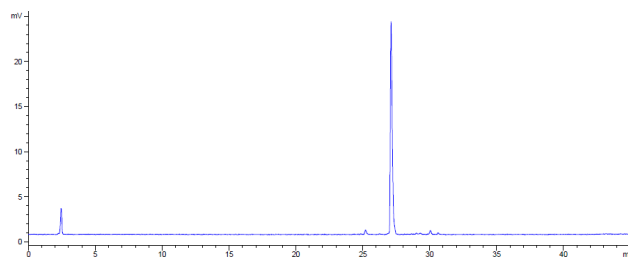
^{1}H NMR of F-H type 2 (700 MHz, D$_2$O)
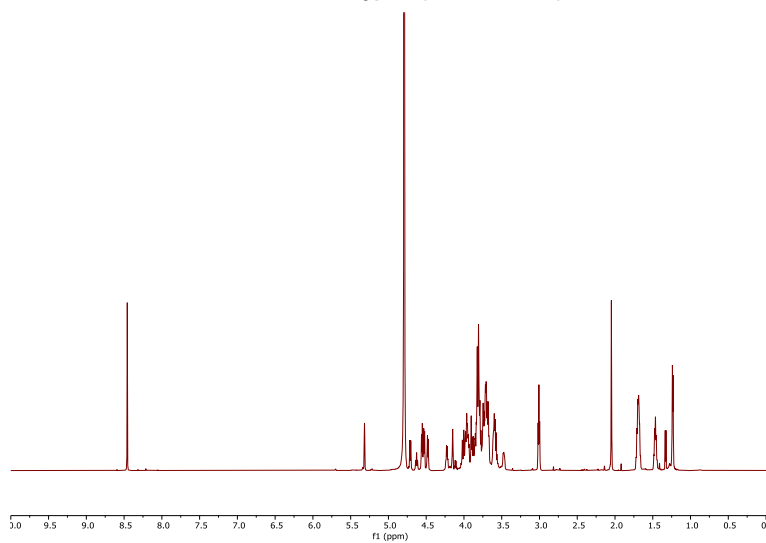
^{13}C NMR of F-H type 2 (151 MHz, D$_2$O)

28
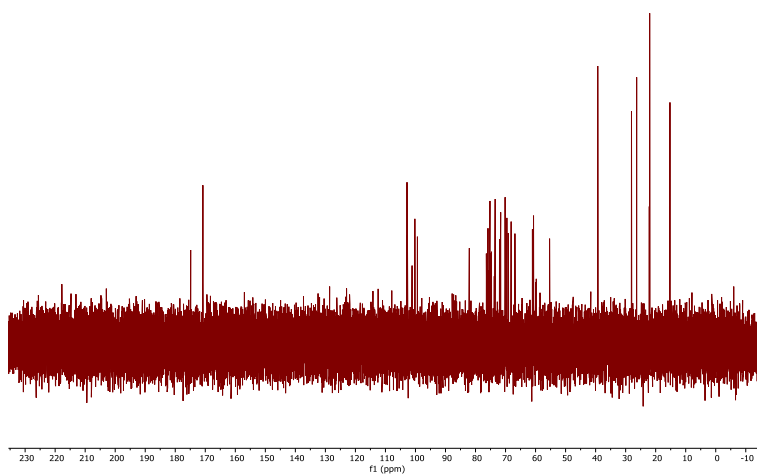
^{19}F NMR of F-H type 2 (564 MHz, D$_2$O)

HSQC NMR of F-H type 2 (D$_2$O)

29

COSY NMR of F-H type 2 (D₂O)

30
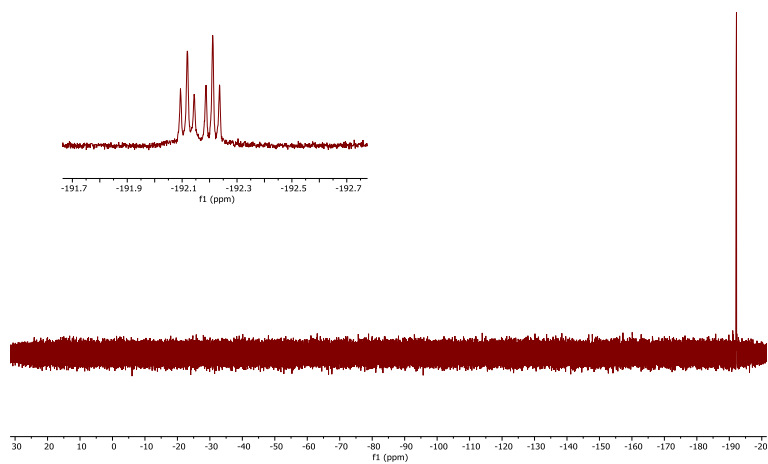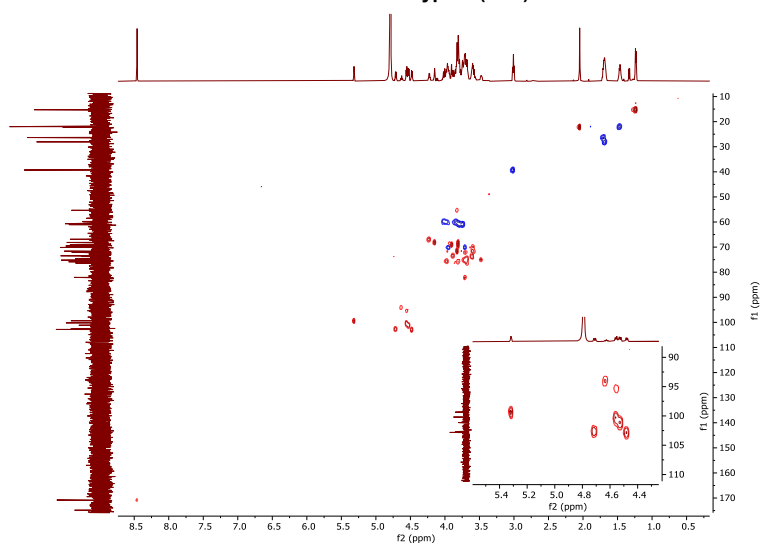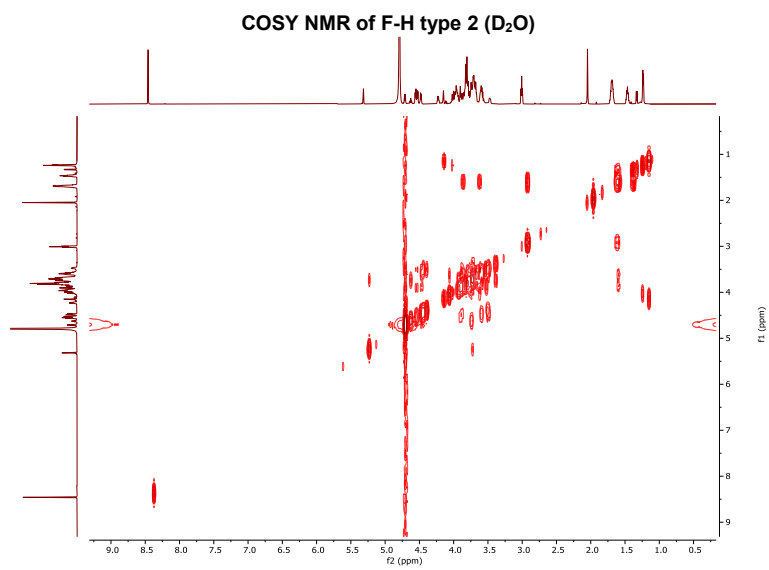
Synthesis of **F-Le^y**

Step	BB	Modules	Notes
		A	**L1** swelling
	1	**B**, **C1**, **D**, **E1**	**C1**: (**1**, -20°C for 5 min, 0°C for 20 min)
	2	**B**, **C1**, **D**, **E1**	**C1**: (**2**, -20°C for 5 min, 0°C for 20 min)
AGA	3	**B**, **C1**, **D**, **E2**	**C1**: (**3**, -20°C for 5 min, 0°C for 40 min)
	5	**B**, **C1**, **D**, **E1**	**C1**: (**5**, -40°C for 5 min, -20°C for 20 min)
	4	**B**, **C2**, **D**, **E1**	**C2**: (**4**, -35°C for 5 min, -15°C for 30 min)
	5	**B**, **C1**, **D**	**C1**: (**5**, -40°C for 5 min, -20°C for 20 min)
Post-AGA		**F**, **G**, **H1**, **I**	**F:** (3 d)
			H1: (3 d)

31

142

I: (Method B, t_R = 22.0 min)

Automated synthesis, global deprotection, and purification afforded **F-Le^y** as a white solid (0.8 mg, 5% overall yield).

Analytical data for **F-Le^y**:

^{1}H NMR (700 MHz, Deuterium Oxide) δ 5.28 (d, *J* = 3.6 Hz, 1H, **H-1 Fuc α-1,2**), 5.12 (d, *J* = 4.0 Hz, 1H, **H-1 Fuc α-1,3**), 4.89 (q, *J* = 6.8 Hz, 1H, H-5 Fuc **α-1,3**), 4.73 (d, *J* = 8.4 Hz, 1H, **H-1**), 4.59 (dt, *J* = 52.3, 8.8 Hz, 1H, H-3 Glc), 4.54 – 4.50 (m, 2H, 2x **H-1**), 4.48 (d, *J* = 7.9 Hz, 1H, **H-1**), 4.26 (q, *J* = 6.6 Hz, 1H, H-5 Fuc α-1,2), 4.16 – 4.13 (m, 1H), 4.05 – 3.90 (m, 7H), 3.90 – 3.64 (m, 17H), 3.63 – 3.55 (m, 5H), 3.46 (ddd, *J* = 10.1, 5.2, 2.4 Hz, 1H), 3.05 – 2.99 (m, 2H, CH_2-NH$_3^+$ linker), 2.03 (s, 3H, CH_3 NHAc), 1.69 (dp, *J* = 13.8, 7.1 Hz, 4H, 2x CH_2 linker), 1.47 (qd, *J* = 9.6, 8.9, 6.5 Hz, 2H, CH_2 linker), 1.27 (d, *J* = 6.6 Hz, 3H, CH_3 Fuc), 1.24 (d, *J* = 6.6 Hz, 3H, CH_3 Fuc). ^{13}C NMR (176 MHz, Deuterium oxide) δ 170.92 (s, *C*=O NHAc), 102.88 (s, **C-1**), 101.20 (s, **C-1**), 101.13 (s, **C-1**), 100.22 (s, **C-1**), 99.44 (s, **C-1 Fuc α-1,2**), 98.60 (s, **C-1 Fuc α-1,3**), 94.83 (d, *J* = 182.8 Hz, C-3 Glc), 82.16, 76.40, 75.63, 75.39, 74.87, 74.76, 73.88, 73.83, 73.56, 73.09, 71.95, 71.85, 71.72, 70.22 (s, CH_2-O linker), 70.02, 69.74, 69.19, 68.75, 68.29, 68.22, 67.72, 66.92 (s, C-5 Fuc α-1,2), 66.80 (s, C-5 Fuc α-1,3), 61.48, 60.75, 59.86, 59.80, 56.15, 39.35 (s, CH_2-NH$_3^+$ linker), 28.13 (s, CH_2 linker), 26.40 (s, CH_2 linker), 22.28 (s, CH_3 NHAc), 22.06 (s, CH_2 linker), 15.46 (s, CH_3 Fuc), 15,44 (s, CH_3 Fuc). ^{19}F NMR (564 MHz, Deuterium oxide) δ -192.16 (dt, *J* = 52.4, 14.1 Hz). (ESI-HRMS) m/z 1087.462 [M+H]$^+$ (C$_{43}$H$_{76}$FN$_2$O$_{28}$ requires 1087.456).

RP-HPLC of F-Le^y (ELSD trace, Method A, t_R = 23.0 min)

^{1}H NMR of F-Le^y (700 MHz, D$_2$O)

^{13}C NMR of F-Le^y (151 MHz, D$_2$O)

33

19F NMR of F-Le^y (564 MHz, D$_2$O)

HSQC NMR of F-Le^y (D$_2$O)

34

COSY NMR of F-Le^y (D$_2$O)

35
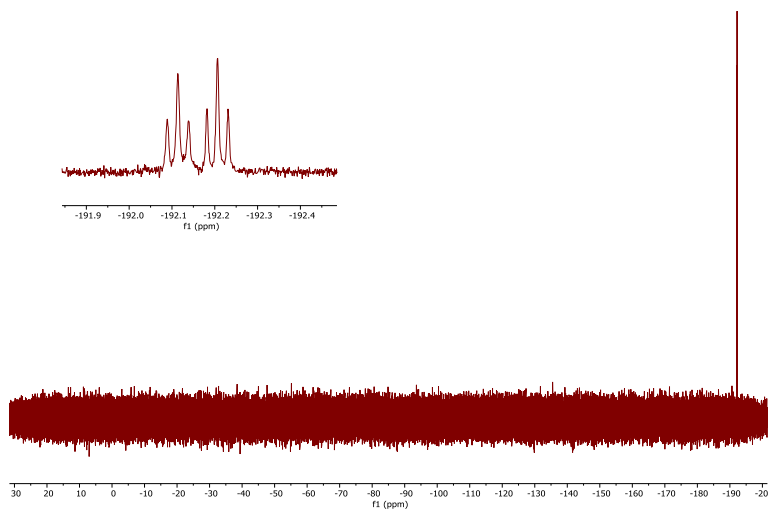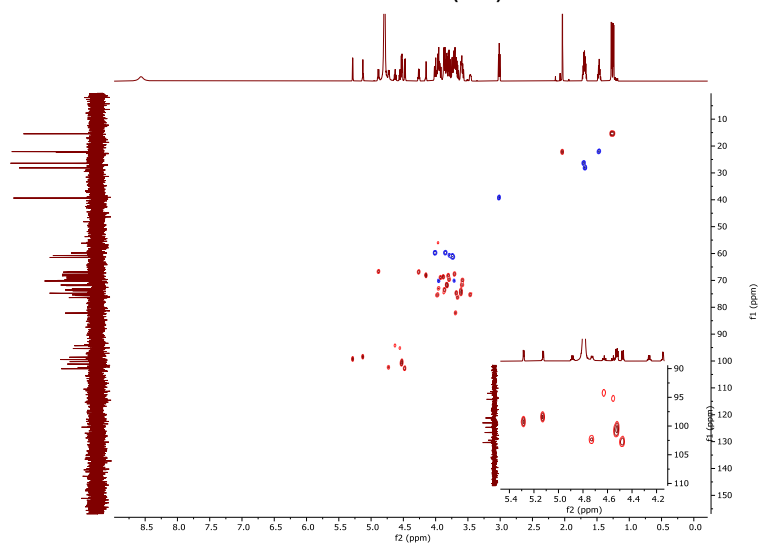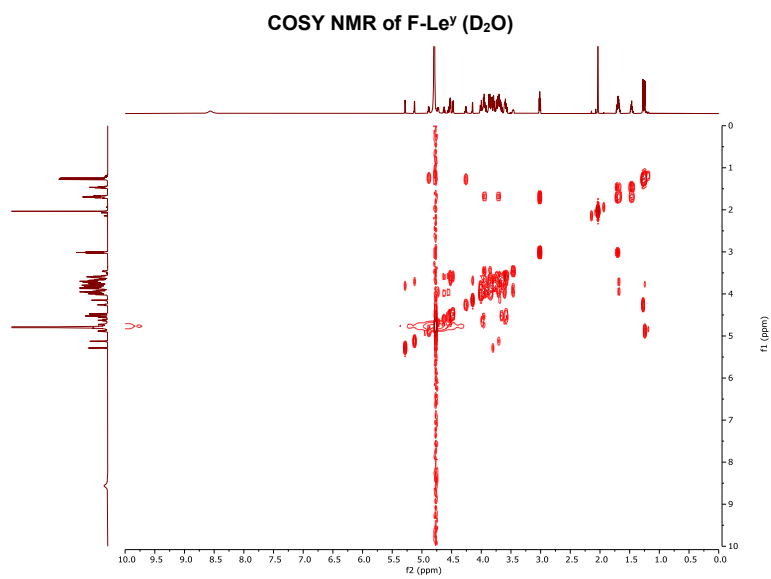
Synthesis of **CF₃-H type 2**
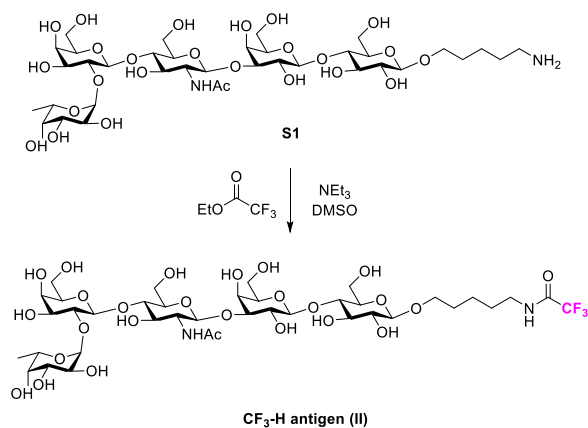
S1 was prepared by automated synthesis according to previously established protocols.[2]

S1 (250 μg, 0.27 μmol) was dissolved in DMSO (150 μL) and a triethylamine solution (10 μL, 0.1 M in DMSO, 1 μmol) was added. A solution of ethyltrifluoroacetate (20 μL, 0.1 M in DMSO, 2 μmol) was added and the mixture stirred at RT overnight. The reaction was then diluted with water, lyophilized, and purified by RP HPLC (Method B, t_R = 38.6 min). **CF₃-type 2** was obtained as a white solid (140 μg, 51% yield).

Analytical data for **CF₃-H type 2**:

¹H NMR (600 MHz, Deuterium Oxide) δ 5.32 (d, J = 3.1 Hz, 1H, **H-1 Fuc α-1,2**), 4.72 (d, J = 8.4 Hz, 1H, **H-1**), 4.56 (d, J = 7.7 Hz, 1H, **H-1**), 4.49 (d, J = 8.0 Hz, 1H, **H-1**), 4.46 (d, J = 7.9 Hz, 1H, **H-1**), 4.23 (q, J = 6.6 Hz, 1H, H-5 Fuc α-1,2), 4.16 (d, J = 3.4 Hz, 1H), 4.02 – 3.87 (m, 5H), 3.87 – 3.57 (m, 22H), 3.51 – 3.45 (m, 1H), 3.35 (t, J = 7.0 Hz, 2H, CH_2-NH₃⁺ linker), 3.33 – 3.29 (m, 1H), 2.06 (s, 3H, CH_3 NHAc), 1.70 – 1.59 (m, 4H, 2x CH_2 linker), 1.42 (p, J = 7.7 Hz, 2H, CH_2 linker), 1.24 (d, J = 6.6 Hz, 3H, CH_3 Fuc). ¹⁹F NMR (564 MHz, Deuterium oxide) δ -75.92. (ESI-HRMS) m/z 1057.372 [M+Na]⁺ ($C_{39}H_{65}F_3N_2O_{26}Na$ required 1057.367).
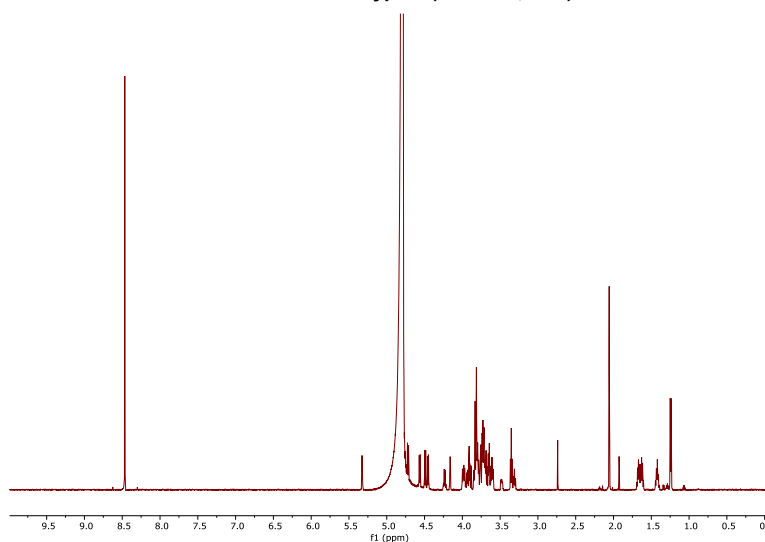
36

147

RP-HPLC of the crude reaction mixture (ELSD trace, Method A)
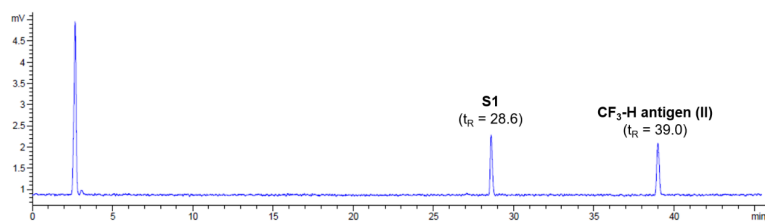
RP-HPLC of pure CF$_3$-H type 2 (ELSD trace, Method A, t$_R$ = 39.1 min)
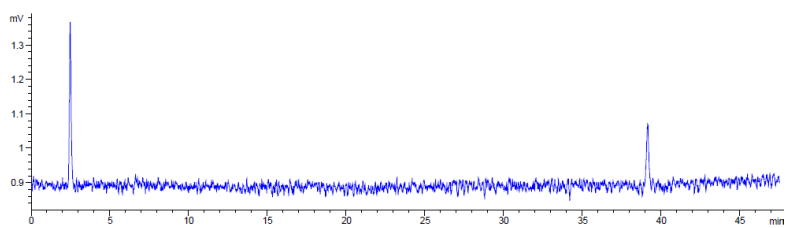
^{1}H NMR of CF$_3$-H type 2 (600 MHz, D$_2$O)

37

¹⁹F NMR of CF₃-H type 2 (564 MHz, D₂O)

* The peak at -122.29 ppm is a fluoride impurity

38
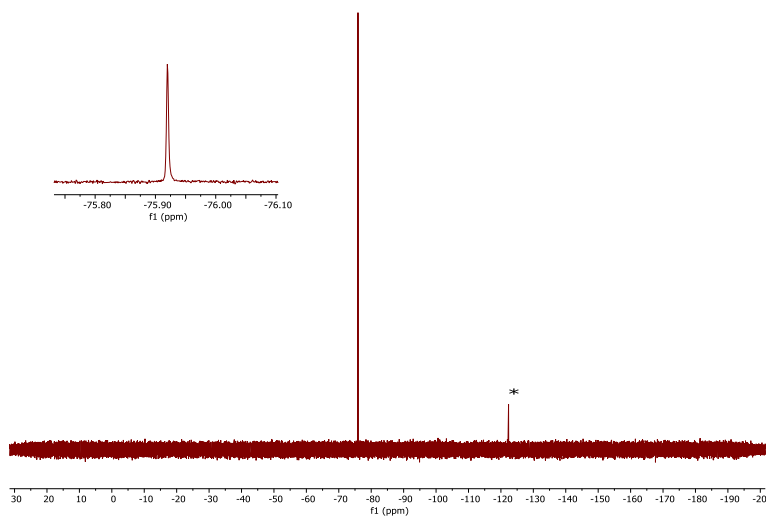
D.   NMR studies of glycan-protein interactions

General materials and methods

Chemicals

All commercial carbohydrates and building blocks were purchased from Carbosynth (UK) or TCI (Germany).

Recombinant proteins

The non-labelled or ^{15}N-labeled LecA[6], LecB[7] and BambL[8] were purified in soluble form as reported previously. LecA labelled with 5-fluorotryptophan (5FW) was expressed and purified as described previously.[9] Human Langerin ECD, DC-SIGN CRD and ECD constructs were expressed in inclusion bodies and prepared as described previously.[10,11]

^{19}F NMR

^{19}F NMR was used to screen F-glycans for bacterial (LecA, LecB, BambL), mammalian lectins (Langerin, DC-SIGN), and sialyltransferases (α(2,3)-sialyltransferase from Pasteurella multocida (Pmα23ST) and α(2,6)-sialyltransferase from Photobacterium damsela (Pdα26ST)). The enzymes α(2,3)-sialyltransferase from *Pasteurella multocida* (Pmα23ST, *E. coli* EC no. 2.4.99.4), and α(2,6)-sialyltransferase from *Photobacterium damsela* (Pdα26ST, *E. coli* EC no. 2.4.99.1) were purchased from Sigma Aldrich. All experiments were performed on a Bruker Ascend™700 (AvanceIII HD) spectrometer equipped with a 5 mm TCI700 CryoProbe™ in 3 mm tubes (Norell S-3-800-7) at 298 K. The proteins were screened at 10 μM (or 67 mU/mL for enzymes) against F-glycans: **F-Lac**, **F-nLac$_4$**, **F-Le^x**, **F-H type 2** and **F-Le^y** in 25 mM Tris-HCl pH 7.8, 150 mM NaCl, 10% D$_2$O, 25 μM TFA and 5 mM CaCl$_2$ at 298 K. For this, proteins were mixed at 1:1 in volume ratio with the F-glycans resulting in the concentrations of 10 μM (or 67 mU/mL for enzymes) and 50 μM, respectively. The enzymes were screened in the absence of donor (*i.e.* CMP-Neu5Ac). ^{19}F spectra were recorded with 512 scans, a spectral width of 10 ppm, a transmitter offset at -190 ppm, acquisition time of 2 s and 1 s relaxation time. T2-filtered spectra were recorded using a CPMG pulse sequence with a 180° pulse repetition rate of 0.38 s using same acquisition and relaxation times.[12,13] ^{19}F R$_2$-filtered experiments for 0.6 mM **F-H type 2** and 0.15 mM LecA were performed with a CPMG pulse sequence with a 180° pulse repetition rate of 0, 0.24, 0.48, 0.72, 0.96 and 1.2 s using the same buffer and acquisition conditions as described above. ^{19}F and CPMG NMR spectra of 50 μM **CF$_3$-H type 2** were recorded in presence of 10 μM DC-SIGN ECD and 5 μM BambL with 16 scans, a spectral width of 10 ppm, a transmitter offset at -75 ppm, acquisition time of 2 s and 1 s relaxation time. Data

39

were recorded without proton decoupling. All spectra were analyzed in MestReNova 11.0.0 (Mestrelab Research SL). The binding strength (**Figure 1B**) was defined depending on the changes observed in the NMR after addition of the protein. A decrease in peak intensity higher than -25% or a CSP higher than 0.01 ppm in the normal ^{19}F NMR was interpreted as strong binding (blue), a decrease in peak intensity higher than -25% in the CPMG filtered ^{19}F NMR as weak/medium binding (light blue), and a decrease in peak lower than -25% in normal or CPMG filtered ^{19}F NMR as no binding (white).

K_d determination with F-glycans in ^{19}F NMR

To derive the affinities of BambL to **F-H type 2** and **F-Le^y**, we recorded ^{19}F NMR of 50 µM compound alone and in presence of BambL (6.25 µM to 100 µM) in 25 mM Tris-HCl pH 7.8, 150 mM NaCl (TBS) with 25 µM TFA and 10% D_2O at 298 K. ^{19}F spectra were recorded with 512 scans, a spectral width of 5 ppm, a transmitter offset at -191 ppm, acquisition time of 2 s and 1 s relaxation time. All spectra were normalized to internal reference trifluoroacetic acid (TFA) at -75.6 ppm and analyzed for the changes in the peak intensities. The decreasing intensity of F-glycan peak in the free state was followed to determine the K_d values of **F-H type 2** and **F-Le^y**. Next, we normalized the changes in the fluorine peak intensities ($I_{normalized}$) following the equation (1) resulting in values plotted on Y-axis.

$$I_{normalized} = \frac{I_0 - I_{measured}}{I_0} \quad (1),$$

where I_0 was F-glycan in the reference spectrum without protein, $I_{measured}$ was F-glycan with protein. The K_d values were calculated according to the one- and two−site binding models in Origin(Pro) 2020b (OriginLab Corp., USA) from three independent titrations.

^1H-^{15}N HSQC and TROSY NMR

To validate binding F-glycans to DC-SIGN CRD and BambL, we ^{15}N-labeled proteins for ^1H-^{15}N HSQC and TROSY NMR, respectively. All ^1H-^{15}N HSQC and TROSY experiments were measured on a Bruker Ascend™700 (AvanceIII HD) spectrometer equipped with a 5 mm TCI700 CryoProbe™ in 3 mm tubes (Norell S-3-800-7) at 298 K and 310 K, respectively.

Briefly, ^{15}N DC-SIGN HSQC and ^{15}N BambL TROSY experiments were recorded with 100 µM protein in 20 mM HEPES pH 7.4, 150 mM NaCl with 5 mM $CaCl_2$, 10% D_2O and 100 µM 4,4-dimethyl-4-silapentane-1-sulfonic acid (DSS) as internal reference. A ^1H-^{15}N HSQC pulse sequence *hsqcf3gpph19* with 128 increments and 8 scans per increment was applied for ^{15}N DC-SIGN CRD. A ^1H-^{15}N TROSY pulse sequence *trosyf3gpphsi19* with 128 increments and 32 scans per increment was applied for ^{15}N BambL. We recorded ^{15}N HSQC NMR of DC-SIGN

40

CRD in presence of 1 mM **F-H type 2** and 250 µM **F-Le^y** or **F-Le^x**. ^{15}N TROSY NMR titration experiments of [15]N BambL were recorded with 30 to 600 µM **F-H type 2** or 90 to 880 µM **F-Le^y**. Data were processed with NMRpipe[14] and further analyzed with CcpNmr analysis.[15]

^{1}H-^{15}N HSQC DC-SIGN CRD resonances were assigned as reported previously.[11] [1]H-^{15}N TROSY BambL resonances were indexed with IDs from 1 to 72 due to a lack of the protein backbone resonance assignment. Next, resonance IDs from BambL and DC-SIGN CRD spectra were transferred to the spectra obtained in the presence of F-glycan for a comparison of the changes in the peak intensities and chemical shift perturbations (CSPs) indicating a slow and fast exchange regimes on the chemical shift timescale, respectively. The changes in peak intensities upon addition of F-glycans were calculated in CcpNmr analysis followed by normalization according to Equation (1). The K_d values were calculated according to the one−site binding model in Origin(Pro) 2020b (OriginLab Corp., USA).

The changes in chemical shift perturbations (CSPs) were calculated according to Equation (2):

$$\Delta\delta = \sqrt{\frac{1}{2}[\Delta\delta_H^2 + (\alpha\Delta\delta_N)^2]} \ (2)$$

in which δ is the difference in chemical shift (in ppm) and α is an empirical weighting factor of 0.14 for all amino acid backbone resonances.[16] The threshold value was set based on three independent measurements of reference spectra to 0.01 ppm and 0.005 ppm for LecA, BambL and DC-SIGN CRD, respectively.

Protein-observed ^{19}F (PrOF) NMR

To validate binding of **F-H type 2** to LecA, we recorded protein-observed ^{19}F (PrOF) NMR with recombinant 5FW LecA as reported previously.[9] Briefly, 150 µM 5FW LecA was recorded alone and in presence of 1 mM **F-H type 2** in 25 mM Tris-HCl pH 7.8, 150 mM NaCl, 5 mM CaCl$_2$ with 10% D$_2$O and 50 µM TFA as internal reference at 310 K. Data analysis was performed in MestReNova 11.0.0 (Mestrelab Research SL) after applying the exponential function (30 Hz) and baseline correction.

Real-time ^{19}F NMR kinetic measurement

The enzymes β-galactosidase (*E. coli*, CAS: 9031-11-2), and α(2,3)-sialyltransferase from *Pasteurella multocida* (Pmα23ST, *E. coli* EC no. 2.4.99.4) were purchased from Sigma Aldrich. The enzymatic reactions with β-galactosidase (0.15 µM), and Pmα23ST (100 mU/mL) were performed at 310 K in 25 mM Tris-HCl pH 7.8, 150 mM NaCl with 2 mM MgCl$_2$, 100 µM TFA and 10% D$_2$O. The concentration of **F-Lac** was 250 µM. The reactions with β-galactosidase were monitored every 149 s. All data were recorded without proton decoupling with 32 scans,

41

acquisition time of 2 s and relaxation delay of 1 s. All spectra were referenced to TFA at -75.6 ppm. To derive the kinetics we used the peak integral of product at -194.8 ppm to derive its concentration. Data were analyzed with nonlinear least-squares methods in the GraphPad Prism 8 software. The best fit of the experimental data provides the value of K_M the Henri−Michaelis−Menten equation. To monitor the sialylation of 250 µM **F-Lac** and **F-nLac₄** with Pmα23ST, we used the buffer conditions reported above with 275 µM CMP-Neu5Ac and recorded data at time points: 0 min, 13 min, 32 min, and 246 min. All data were recorded without proton decoupling with 128 scans, acquisition time of 2 s and relaxation delay of 1 s. All spectra were referenced to TFA at -75.6 ppm. To determine the yield of this reaction, we used the ratio of peak integrals of substrate and product at -192.1 ppm and -192.3 ppm, respectively.

Note to the ^{19}F NMR experiments

All the experiments described in this study were performed without [1]H-^{19}F decoupling due to the lack of the specific hardware component in the spectrometer used. The ^{19}F NMR analysis would benefit in sensitivity and resolution when recording ^{19}F{^{1}H } NMR spectra.
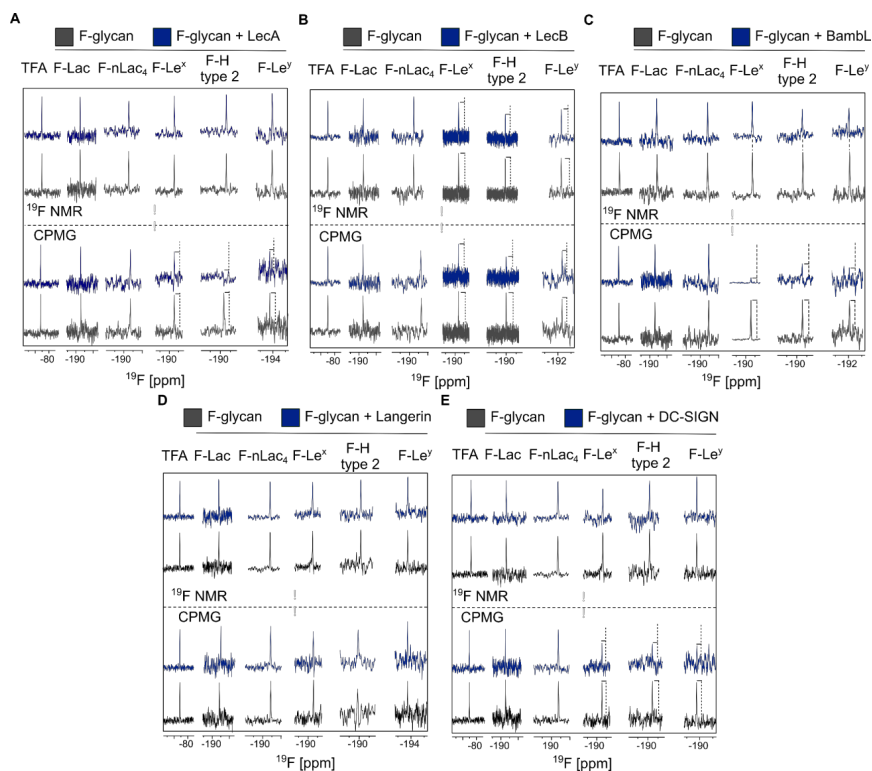
42

Figure S2 *Binding of F-glycans to mammalian and bacterial lectins.* ^{19}F and CPMG NMR spectra of F-glycans alone (*gray*) and in presence of DC-SIGN ECD (*blue*). A chemical shift perturbation of fucosylated Lewis antigens in presence of BambL (**C**, *upper panel*) in ^{19}F NMR is shown (*dashed line*). Decrease in peak intensity (*dashed line*) in CPMG NMR indicates binding of fucosylated Lewis antigens to bacterial: LecA (**A**), LecB (**B**) and BambL (**C**, *lower panel*) and mammalian: Langerin ECD (**D**) DC-SIGN ECD (**E**) lectins. No binding to F-glycans has been observed in presence of Langerin ECD.

Differences in signal-to-noise ratio were observed for the spectra acquired with different quality of shimming.

43
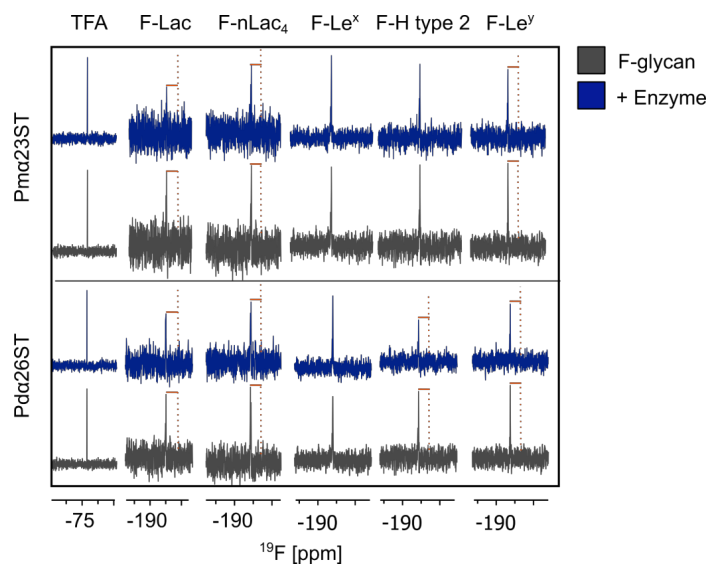
Figure S3 *Binding of sialyltransferases to F-glycans.* CPMG NMR of F-glycans alone and in presence of Pmα23ST or Pdα26ST in the absence of CMP-Neu5Ac donor. F-glycan binding to both enzymes is shown as a reduction in peak intensity in presence of enzymes (*orange*).
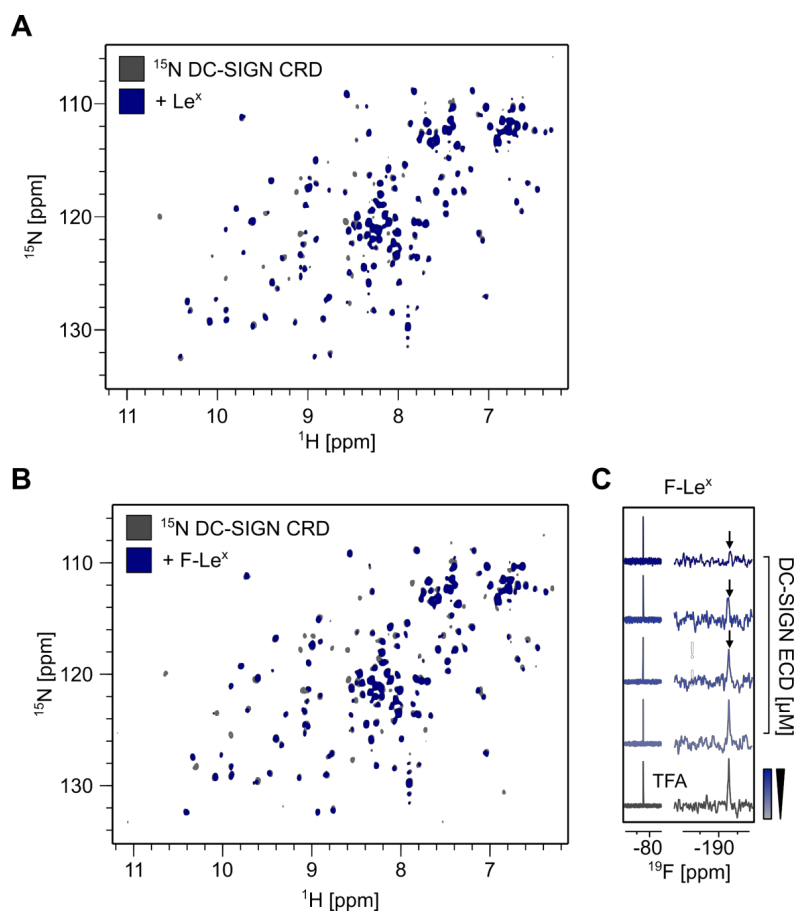
44

Figure S4 *Interaction of **F-Le**^{x} *and **Le**^{x} *with a mammalian lectin DC-SIGN*. **A** HSQC NMR of ^{15}N-labeled DC-SIGN CRD binding to **Le**^{x}. **B** HSQC NMR of ^{15}N-labeled DC-SIGN CRD binding to **F-Le**^{x}. **C** CPMG NMR spectra of 25 µM **F-Le**^{x} alone (*gray*) and in presence of increasing concentration of DC-SIGN ECD: 2.5, 5, 10, 25 and 50 µM (*blue*). The largest difference in a bound *vs* free state was observed at 50 µM DC-SIGN ECD.

45
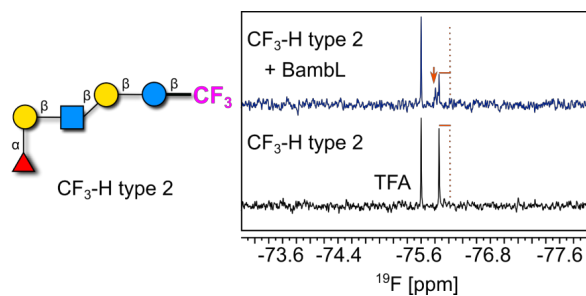
Figure S5 *Binding of CF₃-H type 2 to BambL in ¹⁹F NMR.* ¹⁹F NMR spectrum of **CF₃-H type 2** alone and in presence of BambL (6 µM). Given that **CF₃-H type 2** undergoes a slow exchange in presence of BambL a new peak for the bound ligand arises (*arrow*). This proves that the CF_3 reporter does not have to be near the carbohydrate-binding site of a lectin and can be placed even in remote position (*i.e.* aminopentyl linker attached to the reducing end) of the glycan in order to be detected in ¹⁹F NMR.

46
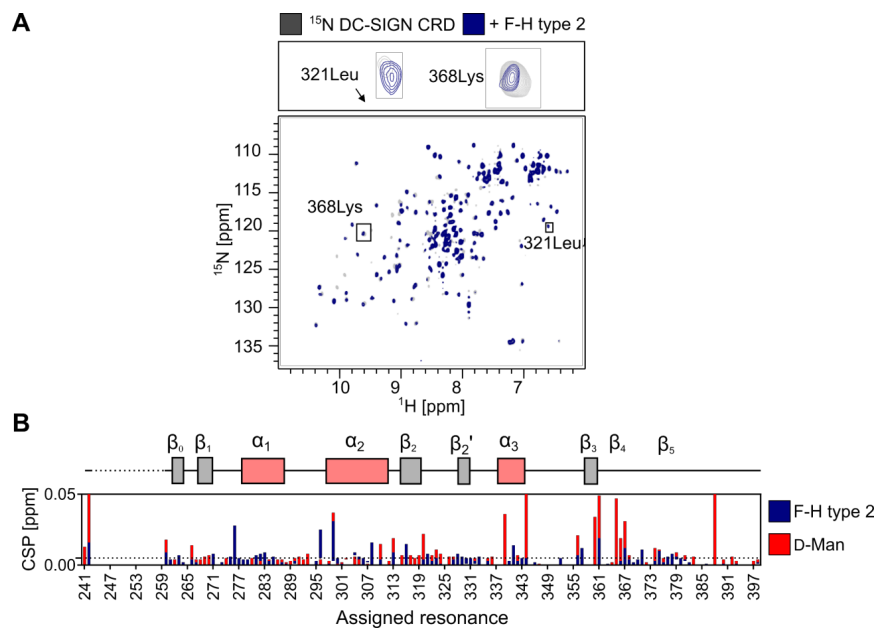
Figure S6 *Weak interaction of **F-H type 2** with DC-SIGN CRD*. **A** HSQC NMR of ^{15}N-labeled DC-SIGN CRD in complex with **F-H type 2**. **F-H type 2** promoted CSPs in the carbohydrate-binding site of DC-SIGN CRD (321Leu and Lys368) and remote parts of DC-SIGN CRD (*not shown*) similarly to D-mannose. **B** The magnitude of **F-H type 2** interaction with DC-SIGN CRD was comparable to D-mannose showing a weaker interaction of **F-H type 2** compared to **F-Le[x]** or **F-Le[y]**.
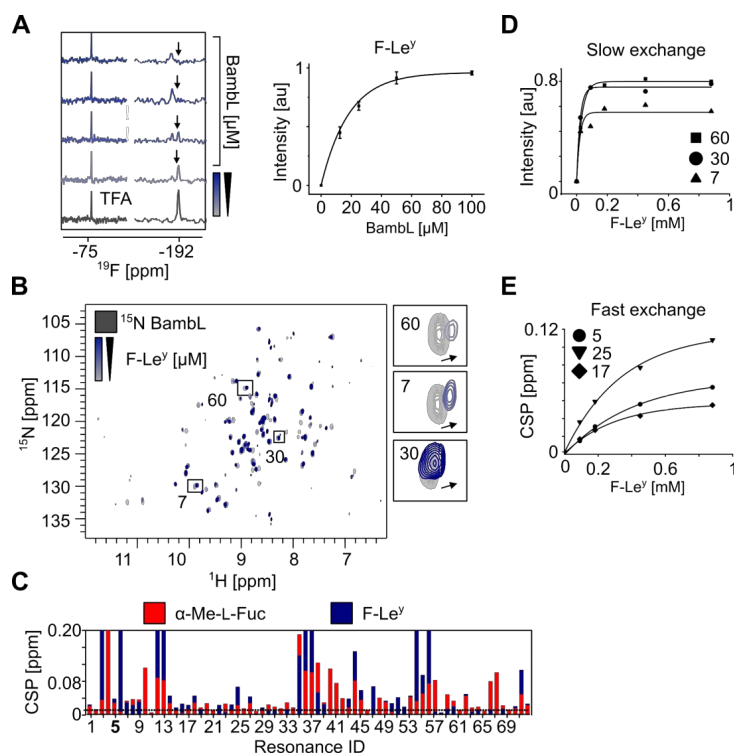
47

Figure S7 Interaction of *F-Le^y with a bacterial lectin BambL*. **A** ^{19}F and CPMG NMR screening of F-glycan (25 μM) alone and in presence of BambL (10 μM). DC-SIGN interacts with **F-Le^y** as shown in ^{19}F NMR titration spectra (*left panel*) and one-site-binding model fit used to derive the dissociation constant (K_d) of **F-Le^y** to BambL (*right panel*). The apparent K_d was calculated using changes in peak intensity (*arrow*) and estimated to be at P:L ratio of 1:0.38 which (referenced to the concentration of 25 μM **F-Le^y** in this assay) corresponds to K_d of 9.2 ± 1.5 μM. **B** TROSY NMR verified **F-Le^y** binding to ^{15}N-labeled BambL. Given that BambL has two binding sites, peaks showing a slow (60, 7 and 30), intermediate and fast exchange (5, 17 and 25) on the chemical shift timescale have been observed upon titration of **F-Le^y**. **C** CSP plot showing the resonances perturbed in presence of α-Me-L-fucose and **F-Le^y**. This verifies that **F-Le^y** targets the carbohydrate-binding site of BambL. Notably, the magnitude of **F-Le^y** promoted effects is similarly to L-fucose suggesting that the avidity effect does not contribute to **F-Le^y**

48

interaction with BambL, but is rather due to binding of L-fucose alone. One-site model for slow (**D**) and fast exchange (**E**) peaks was applied to derive the K_d values of 17±3 µM and 245±29 µM, respectively.
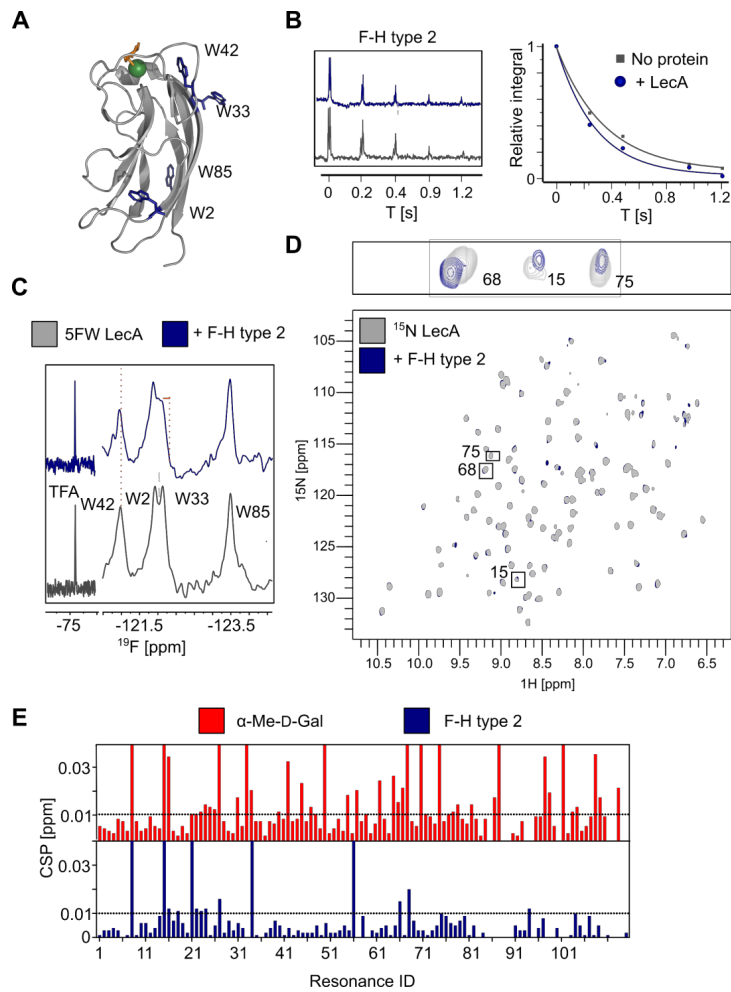
49

Figure S8 *F-H type 2 binding to a bacterial lectin LecA*. **A** A ribbon diagram of the crystal structure of a LecA monomer (PDB: 1OKO) bound to galactose (*orange*) in a Ca^{2+}-dependent manner (*green sphere*). The four tryptophanes (*blue*) are labelled with 5-fluorotryptophanes for PrOF NMR. **B** ^{19}F R_2-filtered NMR assay performed on **F-H type 2** and the decay curves of 0.6 mM **F-H type 2** alone (*grey*) and in presence of 0.15 mM LecA (*blue*). **C** PrOF NMR spectrum of 5-fluorotryptophan labelled (5FW) LecA alone (*grey*) and in presence of **F-H type 2** (*blue*). **F-H**

50

type 2 binds weakly the carbohydrate-binding site of 5FW LecA as shown by CSP and peak intensity decrease of W42 and W33 resonances, respectively. **D** Shown is a TROSY NMR spectrum of ^{15}N-labeled LecA bound to **F-H type 2**, which perturbs resonances similarly to α-Me-D-galactose (*e.g.* 68, 15 and 75). **D** The magnitude of **F-H type 2** promoted perturbations is lower compared to α-Me-D-galactose as shown in a CSP plot.

51

Table S1 List of residue IDs in a CSP plot indicating an assigned resonance in ^{15}N-labeled DC-SIGN CRD (n.a. = not assigned)

Residue ID	Assigned resonance
1	241Leu
2	242Val
3	n.a.
4	n.a.
5	n.a.
6	n.a.
7	n.a.
8	n.a.
9	n.a.
10	n.a.
11	n.a.
12	n.a.
13	n.a.
14	n.a.
15	n.a.
16	n.a.
17	n.a.
18	n.a.
19	n.a.
20	260Trp
21	261Thr
22	262Phe
23	263Phe
24	264Gln
25	n.a.
26	266Asn
27	267Cys
28	268Tyr
29	269Phe
30	270Met
31	271Ser
32	n.a.
33	273Ser
34	274Gln
35	275Arg
36	276Asn
37	277Trp
38	278His
39	279Asp
40	280Ser
41	281Ile
42	282Thr
43	283Ala
44	284Cys
45	285Lys
46	286Glu
47	287Val

52

48	288Gly
49	289Ala
50	290Gln
51	291Leu
52	292Val
53	293Val
54	n.a.
55	295Lys
56	296Ser
57	n.a.
58	298Glu
59	299Glu
60	300Gln
61	301Asn
62	302Phe
63	n.a.
64	304Gln
65	305Leu
66	306Gln
67	307Ser
68	308Ser
69	n.a.
70	310Ser
71	311Asn
72	312Arg
73	313Phe
74	n.a.
75	315Trp
76	316Met
77	317Gly
78	318Leu
79	319Ser
80	320Asp
81	321Leu
82	322Asn
83	323Gln
84	324Glu
85	n.a.
86	326Thr
87	327Trp
88	328Gln
89	329Trp
90	330Val
91	331Asp
92	332Gly
93	333Ser
94	n.a.
95	335LeuH
96	336LeuH
97	n.a.
98	n.a.

53

99	339Phe
100	340Lys
101	341Gln
102	342Tyr
103	343Trp
104	344Asn
105	n.a.
106	346Gly
107	347Glu
108	n.a.
109	n.a.
110	n.a.
111	n.a.
112	352Gly
113	n.a.
114	n.a.
115	n.a.
116	356Cys
117	357Ala
118	358Glu
119	n.a.
120	360Ser
121	361Gly
122	n.a.
123	363Gly
124	364Trp
125	365Asn
126	366Asp
127	367Asp
128	368Lys
129	369Cys
130	370Asn
131	371Leu
132	372Ala
133	373Lys
134	374Phe
135	375Trp
136	376Ile
137	377Cys
138	378Lys
139	379Lys
140	380Ser
141	381Ala
142	382Ala
143	383Ser
144	n.a.
145	n.a.
146	386Arg
147	n.a.
148	388Glu
149	n.a.

54

150	390Gln
151	n.a.
152	392Leu
153	393Ser
154	n.a.
155	n.a.
156	n.a.
157	397Ala
158	398Thr

55
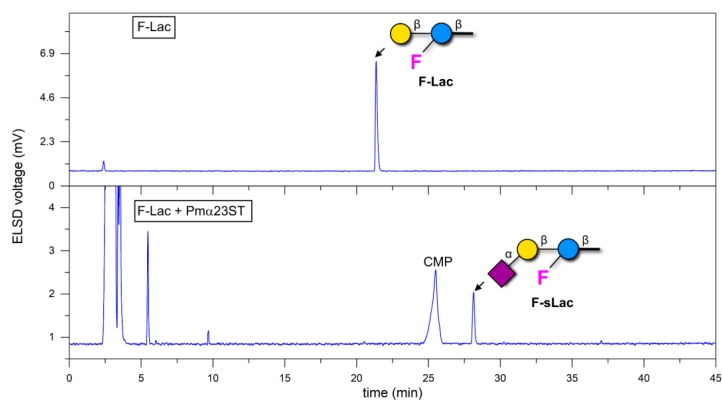
Figure S9 *Enzymatic sialylation of* ***F-Lac*** *with Pmα23ST confirmed by HPLC.* RP-HPLC trace (ELSD trace, Method A) shows consumption of the starting material (**F-Lac**, t_R = 21.4 min, *top*) and formation of a new peak in the crude reaction mixture after 246 min (**F-sLac**, t_R = 28.1 min, *bottom*). HPLC retention times match previously reported data.[17]

56

Figure S10 *Enzymatic sialylation of **F-nLac** with Pmα23ST.* **A** Real-time ^{19}F NMR monitoring of **F-nLac$_4$** incubated with Pmα23ST in the presence of CMP-Neu5Ac. No chemical shift perturbation was observed after 6 h incubation time. **B** RP-HPLC trace (ELSD trace, Method A) shows incomplete consumption of the starting material (**F-nLac$_4$**, t_R = 28.1 min, *top*) and formation of the sialylated product (**F-snLac$_4$**, t_R = 33.3 min, *bottom*) in the crude reaction mixture.

57

E.   References

(1)   Yu, Y.; Tyrikos-Ergas, T.; Zhu, Y.; Fittolani, G.; Bordoni, V.; Singhal, A.; Fair, R. J.; Grafmüller, A.; Seeberger, P. H.; Delbianco, M. Systematic Hydrogen-Bond Manipulations To Establish Polysaccharide Structure-Property Correlations. *Angew. Chem., Int. Ed.* **2019**, *58* (37), 13127– 13132. https://doi.org/10.1002/anie.201906577.

(2)   Guberman, M.; Bräutigam, M.; Seeberger, P. H. Automated Glycan Assembly of Lewis Type I and II Oligosaccharide Antigens. *Chem. Sci.* **2019**, *10* (21), 5634–5640. https://doi.org/10.1039/C9SC00768G.

(3)   Le Mai Hoang, K.; Pardo-Vargas, A.; Zhu, Y.; Yu, Y.; Loria, M.; Delbianco, M.; Seeberger, P. H. Traceless Photolabile Linker Expedites the Chemical Synthesis of Complex Oligosaccharides by Automated Glycan Assembly. *J. Am. Chem. Soc.* **2019**, *141* (22), 9079–9086. https://doi.org/10.1021/jacs.9b03769.

(4)   Gude, M.; Ryf, J.; White, P. D. An Accurate Method for the Quantitation of Fmoc-Derivatized Solid Phase Supports. *Lett. Pept. Sci.* **2002**, *9* (4), 203–206. https://doi.org/doi/10.1023/A:1024148619149.

(5)   Hurevich, M.; Kandasamy, J.; Ponnappa, B. M.; Collot, M.; Kopetzki, D.; McQuade, D. T.; Seeberger, P. H. Continuous Photochemical Cleavage of Linkers for Solid-Phase Synthesis. *Org. Lett.* **2014**, *16* (6), 1794–1797. https://doi.org/10.1021/ol500530q.

(6)   Denavit, V.; Lainé, D.; Bouzriba, C.; Shanina, E.; Gillon, É.; Fortin, S.; Rademacher, C.; Imberty, A.; Giguère, D. Stereoselective Synthesis of Fluorinated Galactopyranosides as Potential Molecular Probes for Galactophilic Proteins: Assessment of Monofluorogalactoside–LecA Interactions. *Chem. – A Eur. J.* **2019**, *25* (17), 4478–4490. https://doi.org/10.1002/chem.201806197.

(7)   Sommer, R.; Hauck, D.; Varrot, A.; Wagner, S.; Audfray, A.; Prestel, A.; Möller, H. M.; Imberty, A.; Titz, A. Cinnamide Derivatives of <scp>d</Scp> -Mannose as Inhibitors of the Bacterial Virulence Factor LecB from *Pseudomonas Aeruginosa*. *ChemistryOpen* **2015**, *4* (6), 756–767. https://doi.org/10.1002/open.201500162.

(8)   Audfray, A.; Claudinon, J.; Abounit, S.; Ruvoën-Clouet, N.; Larson, G.; Smith, D. F.; Wimmerová, M.; Le Pendu, J.; Römer, W.; Varrot, A.; Imberty, A. Fucose-Binding Lectin from Opportunistic Pathogen Burkholderia Ambifaria Binds to Both Plant and Human Oligosaccharidic Epitopes. *J. Biol. Chem.* **2012**, *287* (6), 4335–4347. https://doi.org/10.1074/jbc.M111.314831.

(9)   Shanina, E.; Siebs, E.; Zhang, H.; Varón Silva, D.; Joachim, I.; Titz, A.; Rademacher, C. Protein-Observed 19F NMR of LecA from Pseudomonas Aeruginosa. *Glycobiology* **2020**, *2020*, 1–7. https://doi.org/10.1093/glycob/cwaa057.

(10)   Hanske, J.; Aleksić, S.; Ballaschk, M.; Jurk, M.; Shanina, E.; Beerbaum, M.; Schmieder, P.; Keller, B. G.; Rademacher, C. Intradomain Allosteric Network Modulates Calcium Affinity of the C-Type Lectin Receptor Langerin. *J. Am. Chem. Soc.* **2016**, *138* (37), 12176–12186. https://doi.org/10.1021/jacs.6b05458.

(11)   Aretz, J.; Baukmann, H.; Shanina, E.; Hanske, J.; Wawrzinek, R.; Zapol'skii, V. A.; Seeberger, P. H.; Kaufmann, D. E.; Rademacher, C. Identification of Multiple Druggable Secondary Sites by Fragment Screening against DC-SIGN. *Angew. Chem., Int. Ed.* **2017**, *56* (25), 7292–7296. https://doi.org/10.1002/anie.201701943.

58

(12)   Carr, H. Y. Steady-State Free Precession in Nuclear Magnetic Resonance. *Phys. Rev.* **1958**, *112* (5), 1693–1701. https://doi.org/10.1103/PhysRev.112.1693.

(13)   Meiboom, S.; Gill, D. Modified Spin-Echo Method for Measuring Nuclear Relaxation Times. *Rev. Sci. Instrum.* **1958**, *29* (8), 688–691. https://doi.org/10.1063/1.1716296.

(14)   Delaglio, F.; Grzesiek, S.; Vuister, G.; Zhu, G.; Pfeifer, J.; Bax, A. NMRPipe: A Multidimensional Spectral Processing System Based on UNIX Pipes. *J. Biomol. NMR* **1995**, *6* (3), 277–293. https://doi.org/10.1007/BF00197809.

(15)   Vranken, W. F.; Boucher, W.; Stevens, T. J.; Fogh, R. H.; Pajon, A.; Llinas, M.; Ulrich, E. L.; Markley, J. L.; Ionides, J.; Laue, E. D. The CCPN Data Model for NMR Spectroscopy: Development of a Software Pipeline. *Proteins Struct. Funct. Genet.* **2005**, *59* (4), 687–696. https://doi.org/10.1002/prot.20449.

(16)   Williamson, M. P. Using Chemical Shift Perturbation to Characterise Ligand Binding. *Prog. Nucl. Magn. Reson. Spectrosc.* **2013**, *73*, 1–16. https://doi.org/10.1016/j.pnmrs.2013.02.001.

(17)   Fair, R. J.; Hahm, H. S.; Seeberger, P. H. Combination of Automated Solid-Phase and Enzymatic Oligosaccharide Synthesis Provides Access to α(2,3)-Sialylated Glycans. *Chem. Commun.* **2015**, *51* (28), 6183–6185. https://doi.org/10.1039/C5CC01368B.

59

7.3. Supporting Information for Subchapter 4.3.

7.3.1. Materials and Methods

Chemicals

The carbohydrates methyl-α-L-fucose (MeFuc, CAS: 14687-15-1) and 2-deoxy-2-fluoro-L-fucose (2FF, CAS: 70763-62-1) were purchased from Biosynth-Carbosynth (UK) or TCI (Germany). F-H type 2 was synthesized using AGA at MPICI Potsdam as reported previously. [1]

Recombinant proteins

Non-labeled proteins BambL[2], RSL[3] and AFL[4] were purified in soluble form as reported previously. Recombinant ^{15}N-labeled BambL WT used pET15b, whereas mutants (T18S, L87R, T25S, W51F, W8F, W72F and W74F) used pProEx as expression plasmids. Mutagenesis was performed following QuikChange II site-directed mutagenesis method (Agilent). The plasmids were individually transformed into *E. coli* BL21 (DE3). The transformed cells were grown in M9 medium (100 µg mL^{-1} ampicillin) at 37°C with agitation (120 rpm) until OD$_{600}$ reached 0.6. Protein production was induced with 200 µM IPTG at 20°C and harvested in 4 h.

For production of recombinant 5-fluorotryptophan (5FW)-labeled BambL WT and the carbohydrate-binding site mutants (W51F, W8F, W72F and W74F), we transformed *E. coli* with the plasmids and grew it in LB medium (100 µg mL^{-1} ampicillin) at 37°C with agitation (120 rpm) until OD$_{600}$ reached 0.6. 1 L of culture was harvested by centrifugation at 1000 g, 10 min and resuspended in modified minimal M9 medium prepared as reported previously.[5] It was shaken at 37°C for 60 min as a recovery time for bacteria and followed by addition of 250 µL of 5−fluoroindole (Santa Cruz, USA; 240 mg/mL in DMSO). Protein production was induced with 200 µM IPTG at 20°C and harvested in 4 h.

Cell pellets containing either 5FW- or ^{15}N-labeled BambL were resuspended in buffer A (20 mM Tris-HCl pH 7.4, 137 mM NaCl, 2.6 mM KCl) supplemented with 1 mM PMSF and DNaseI (Applichem, Darmstadt, Germany). The cells were lysed by cell disruption (Branson Digital Sonifier) at 50% power 10 s on and 40 s off pulses following removal of cell debris by centrifugation (10 000 g, 30 min, 4°C). The supernatant was loaded onto a 2 mL D-mannose agarose column (Sigma-Aldrich, Germany) that was equilibrated with 3−fold column volume of buffer A. Bound BambL was eluted with buffer B (20 mM Tris-HCl pH 7.4, 137 mM NaCl, 2.6 mM KCl, 50 mM D-mannose). Protein was dialyzed in MilliQ water and TBS buffer (20 mM Tris−HCl pH 7.8, 100 mM NaCl) or HBS buffer (20 mM HEPES pH 7.4, 100

mM NaCl) for 5FW- and ^{15}N-labaled BambL three times for 4 h and once overnight at 4°C, respectively. The protein solution was flash frozen and stored at -80°C.

^{19}F NMR screening

Fluorinated (^{19}F) fragment library was prepared as reported previously.[6] Briefly, 20 µM BambL, 20 µM AFL and 40 µM RSL were screened against 350 fluorinated fragments in mixtures at 50 µM each using ^{19}F NMR and ^{19}F T_2-filtered (CPMG) spectra on a Bruker Ascend™700 (AvanceIII HD) spectrometer equipped with a 5 mm TCI700 CryoProbe™ in 3 mm tubes (Norell S-3-800-7) at 298 K.

^{19}F NMR screening was performed in the presence and absence of protein, subsequently adding methyl-α-L-fucose (MeFuc) to a final concentration of 10 mM. Two separate ^{19}F spectra were recorded for CF_3 and CF groups with 32 and 64 scans, a spectral width of 100 ppm, a transmitter offset at -50 and -150 ppm, acquisition time of 2 s and 1 s relaxation time, respectively. T_2-filtered spectra were recorded using a CPMG pulse sequence with a 180° pulse repetition rate of 384 ms using same acquisition and relaxation times with 64 and 256 scans for CF_3 and CF compounds, respectively. Data was recorded without proton decoupling.

All spectra were analyzed in MestReNova 11.0.0 (Mestrelab Research SL) for the changes in peak intensity and a chemical shift. Compounds binding in presence of protein were used to derive a total hit rate, whereas compounds competed with MeFuc were defined as targeting the carbohydrate-binding site. Intensity changes in the T_2-filtered spectra of 20-50% or more than 50% change were defined as 'high' and 'low' confidence hits, respectively. For ^{19}F spectra, chemical shift changes of 0.01 ppm or intensity changes between 25-50% were defined as 'high' and 'low' confidence hits, respectively. Afterwards, only compounds fulfilling three criteria: 1) ^{19}F only, 2) ^{19}F and CPMG and 3) 'high confidence' CPMG only, were followed up in the counter-screening.

^{1}H-^{15}N TROSY NMR

To validate fragment binding to BambL, we used ^{15}N-labeled BambL in TROSY NMR. All TROSY NMR experiments were measured on the machine described above in 3 mm tubes at 298 K or 310 K for fragment-binding validation or recording spectra of BambL mutants, respectively.

Briefly, ^{15}N BambL TROSY experiments were recorded with 70-100 µM protein in 20 mM HEPES pH 7.4, 150 mM NaCl, 10% D_2O and 100 µM 4,4-dimethyl-4-silapentane-1-sulfonic acid (DSS) as internal reference. A TROSY pulse sequence *trosyf3gpphsi19* with 128 increments and 20 scans per increment was applied. For fragment-binding validation experiments, ^{15}N TROSY spectra of ^{15}N BambL were recorded with DMSO as reference, 75-1000 µM MeFuc or 2 mM fragments.

Titration experiments with fragments were recorded by stepwise addition of fragments to ^{15}N BambL (0.25, 0.5, 0.75, 1 and 2 mM) at 298 K. Binding of the compound **24** and its SAR to ^{15}N BambL were validated at 2 mM at 298 K. To improve the resolution of TROSY spectra for mutants, spectra were recorded at 310 K without (W8F, W51F, W72F, W74F, T18S, L87R and T25S) and with 1 mM **24** and **83** (W51F, T18S, L87R and T25S). Data were processed with NMRpipe[7] and further analyzed with CcpNmr analysis.[8]

BambL resonances were indexed with IDs from 1 to 75 due to a lack of protein backbone resonance assignment. Next, resonance IDs from BambL WT spectra were transferred to the spectra of mutants and in the presence of **24** and **83** in order to compare the changes in chemical shift perturbations (CSPs) for intermediate and fast exchange peaks. The CSPs were calculated according to Equation (1):

$$\Delta\delta = \sqrt{\frac{1}{2}[\Delta\delta_H^2 + (\alpha\Delta\delta_N)^2]} \; (1)$$

in which δ is the difference in chemical shift (in ppm) and α is an empirical weighting factor of 0.14 for all amino acid backbone resonances.[9] The threshold value was set based on three independent measurements of reference spectra to 0.01 ppm.

^{19}F R_2-filtered NMR

To derive the affinity of **24** for BambL, experiments using 0.1 mM BambL were performed in 25 mM Tris-HCl pH 7.8, 150 mM NaCl (TBS) with 50 µM TFA and 10% D_2O at 298 K. All spectra were referenced to internal reference trifluoroacetic acid (TFA) at -75.6 ppm and analyzed for changes in peak intensity. For each spectrum 64 scans were recorded in 3 mm tubes at sample volume of 150 µL. Relaxation rates $R_{2,obs}$ were determined with the CPMG pulse sequence by fitting Equation (2) to the integrals of the ^{19}F resonance of **24**. [10] T is the relaxation time and I_0 is the integral at $T = 0$ s. The relaxation delay and acquisition time were set to 2 s and 1 s, respectively.

$$I = I_0 e^{-R_{2,obs}T} \; (2)$$

The K_d and $R_{2,b}$ value of **24** were derived from Equation (3) by detection of ^{19}F relaxation rates $R_{2,obs}$. $R_{2,b}$ is the relaxation rate of the reporter in the protein bound form, whereas p_b is the bound fraction of the ligand with the concentrations of ligand and protein $[L]_T$ and $[P]_T$, respectively. The relaxation rate of the free ligand $R_{2,f}$ was measured at 0.035 mM and 0.1 mM **24** in absence of BambL. The K_d values were fitted in a two parameter fit Origin(Pro) 2020b (OriginLab Corp., USA) from two independent titrations.

$$R_{2,obs} = R_{2,f} + (R_{2,b} - R_{2,f})p_b \quad (3)$$

with

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

SPR

All experiments were performed on a BIACORE X100 instrument (GE Healthcare) at 25 °C BamBL was immobilized onto a CM7 chip (BIACORE) following standard amine coupling procedures using phosphate buffer saline as a running buffer (10 mM phosphate buffer pH 7.4, 2.7 mM KCl, 137 mM NaCl, 0.05% Tween 20): the CM7 chip was activated by three injections of a NHS/EDC mixture with a contact time of 540 s at a flow rate of 10 µL/min until the response exceeded 800 RU, followed by multiple injections of BamBL dissolved in 10 mM sodium acetate pH 4.5 (100 µg/mL) onto channel 2 (contact time of 540 s at a flow rate of 10 µL·min^{-1}). A minimum of 7,000 RU of BamBL was captured onto the chip. Binding experiments were performed with the injections of 0.2 and 1 mM of each compound diluted in phosphate buffer saline supplemented with 5% DMSO (10 mM phosphate buffer pH 7.4, 2.7 mM KCl, 137 mM NaCl, 0.05% Tween 20, 5% DMSO). Each compound was analyzed using the following parameters: the association 30 s, dissociation 60 s, 30 µL·min^{-1} flow rate. Compounds eliciting dose-response behaviors (more than twice binding response at equilibrium) are identified as binders. All data evaluation was performed using BIACORE X100 evaluation software (version 2.0).

In silico prediction of binding sites

The crystal structures of BambL in complex with α-L-fucose (PDB ID: 3ZW0) and H-type 2 tetrasaccharide (PDB ID: 3ZZV) were used for prediction of the possible secondary binding sites using SiteMap[11] tool. This tool creates a grid of points on the protein surface based on depth, size, van der Waals interaction energy, hydrophilicity, hydrophobicity and assign a single scoring function (SiteScore) to the potential druggable regions. The score helps to assess a site's propensity for ligand binding and prioritize the pharmaceutically relevant regions in the target protein. For BambL, the calculations identified three regions at the interface in the trimer as potential druggable sites. The same approach was applied to the crystal structure of RSL in complex with L-fucose (PDB ID: 3ZI8) and Lewis x trisaccharide (PDB ID: 5AJB), which identified three pockets structurally similar to BambL. The studies further extended to the crystal structure of AFL in complex with fucoside (PDB ID: 4AGI) also indicated the presence of a secondary binding site.

Docking of fragments into binding site

Preparation of protein model for docking

All the calculations were performed using the Schrödinger Suite through Maestro (version 2018-1) graphical interface.[12] Atomic coordinates from the high resolution crystal structure of BambL (PDB ID: 3ZW0) was taken from the Protein Data Bank.[13] The asymmetric unit contains three peptide chains and a carbohydrate ligand (α-methyl-L-fucoside), around a 3-fold pseudo axis of symmetry. The water molecules were removed and hydrogen atoms were added. pKa was predicted for protein residues using the PROPKA[14] method at pH 7.4. Protonation state (δ-nitrogen protonated) was assigned to the histidine (His58) residue. Finally, the complex was subjected to restrained minimization with convergence of heavy atoms to an RMSD of 0.3 Å using the OPLS3 force field.[15]

Preparation of ligand models for docking

The ligands were prepared for docking using the LigPrep[16] tool and generated tautomers, stereoisomers and protonation states at pH 7.4. The calculations yield 25 structures.

Models for docking study

For docking grid generation, the centroids of residues from chain B (Gly67, Thr69, Gly86, Leu87) and chain C (Thr18, Asn20, Lys23, Thr25) were selected to define a cubic grid box with dimensions 32×32×32 Å. The grid was used for docking studies using extra precision (XP) and standard precision (SP) scoring functions. All the calculations were accomplished by Glide (version 7.8)[17] using the flexible docking approach.

Competition ^{19}F T_2-filtered NMR using 24 as a reporter

Competition ^{19}F T_2-filtered NMR experiments with 2FF and **24** as a reporter molecule were performed in 25 mM Tris-HCl pH 7.8, 150 mM NaCl (TBS) with 50 μM TFA and 10% D_2O at 298 K using 1 mM **24** and 0.1 mM BambL in presence of eight 2FF concentrations. ^{19}F T_2-filtered spectra were recorded with 64 scans, a spectral width of 50 ppm, a transmitter offset at -75 ppm, acquisition time of 0.8 s and 1 s relaxation time. T_2-filtered spectra were recorded using a CPMG pulse sequence with a 180° pulse repetition rate of d_{20}=3 ms and L_4=128. All data was recorded without proton decoupling. The integrals of **24** in presence of BambL were followed to fit and to derive the IC_{50} value in Origin(Pro) 2020b (OriginLab Corp., USA). Next, we normalized the changes of the fluorine integrals ($I_{normalized}$) following the Equation (4) resulting in values plotted on Y-axis.

$$I_{normalized} = \frac{I_0 - I_{measured}}{I_0} \quad (4),$$

where I_0 is the integral **24** in the reference spectrum without protein, $I_{measured}$ is **24** with protein.

One-point binding ^{19}F T$_2$-filtered NMR experiments with 0.01 mM **24** were performed with 25 µM BambL WT, T18S or L87R in 25 mM Tris-HCl pH 7.8, 150 mM NaCl (TBS) with 50 µM TFA and 10% D$_2$O at 298 K. To check for a better derivative of the compound **24**, we added 0.01 mM **83**. ^{19}F T$_2$-filtered spectra were recorded with 512 scans, a spectral width of 50 ppm, a transmitter offset at -75 ppm, acquisition time of 0.8 s, 2 s relaxation time and CPMG filter of 384 ms.

K_d determination in ^{19}F T$_2$-filtered NMR

To derive the affinities of BambL to 2FF and F-H type 2, we recorded ^{19}F NMR spectra of 50 µM 2FF or F-H type 2 alone and in presence of BambL (2.5 µM to 50-100 µM) in 25 mM Tris-HCl pH 7.8, 150 mM NaCl (TBS) with 25 µM TFA and 10% D$_2$O at 298 K. ^{19}F spectra were recorded with 512 scans, a spectral width of 10 ppm, a transmitter offset at -205 ppm, acquisition time of 0.8 s and 2 s relaxation time. For F-H type 2, the transmitter offset (tof) was set at -191 ppm. All spectra were analyzed in MestReNova 11.0.0 (Mestrelab Research SL) and referenced to the internal reference trifluoroacetic acid (TFA) at -75.6 ppm and analyzed for changes in peak intensity. The decreasing intensity of 2FF or F-H type in the free state was followed to determine K_d values. Next, we normalized the changes in the fluorine peak intensities *Peak(I$_{normalized}$)* following the equation (5) resulting in values plotted on Y-axis. The K_d values were calculated according to the one−site−binding model in Origin(Pro) 2020b (OriginLab Corp., USA) from three independent titrations.

$$Peak(I)_{normalized} = \frac{Peak(I_0) - Peak(I_{measured})}{Peak(I_0)} \quad (5),$$

where *Peak(I$_0$)* is the peak intensity of 2FF/F-H type 2 in the reference spectrum without protein, *Peak(I$_{measured}$)* is 2FF/F-H type 2 with protein.

Protein−observed fluorine (PrOF) NMR

All experiments were conducted on Bruker Ascend^{TM}700 (AvanceIII HD) spectrometer equipped with a 5 mm TCI700 CryoProbe^{TM} *cpmg1d* sequence in 3 mm tubes (Norell S−3−800−7) with following parameters: time domain of 3946, relaxation delay 1 s, acquisition time of 0.3 s, tof of -124 ppm, d$_{20}$ of 1 ms, spectral width of 10 ppm and 1024 scans.

PrOF NMR of 100 µM 5FW BambL was recorded in TBS pH 7.8 with 10% D$_2$O and 100 µM TFA at 310 K. Single-point PrOF NMR titration experiments used 1 mM **24**, 1 mM **83**, 0.5 mM Me−α−L−fucose (MeFuc). For 5FW resonance assignment, we recorded spectra using 5FW BambL WT and mutants (W8F, W51F, W72F and W74F) were measured at 100 µM alone and in presence of 500 µM 2FF. Titration experiments were performed at 0.1 mM 5FW BambL WT with 1) 2FF: 0.05, 0.1,

0.25, 0.5 and 1 mM, 2) **24**: 0.25, 0.5, 1, 1.5 mM following changes in chemical shift perturbations (CSPs) of 5FW resonances.
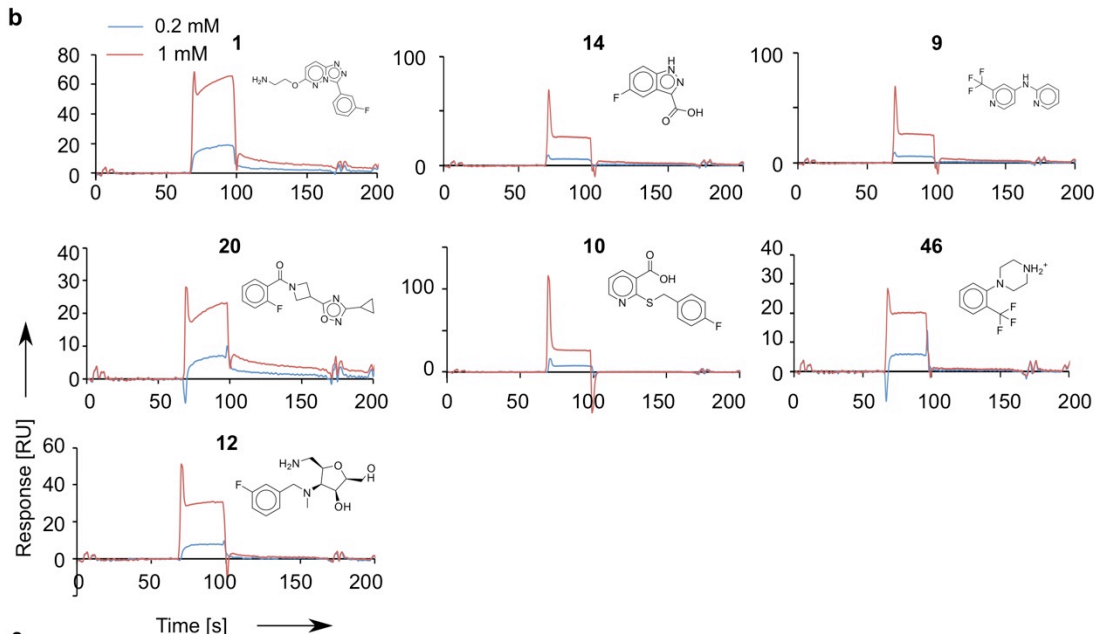
The data analysis was performed with MestReNova 11.0.0 (Mestrelab Research SL, Santiago de Compostela, Spain). All spectra were referenced to trifluoroacetic acid (TFA) as internal reference at -75.6 ppm after applying the Exponential function (30 Hz) and baseline correction.
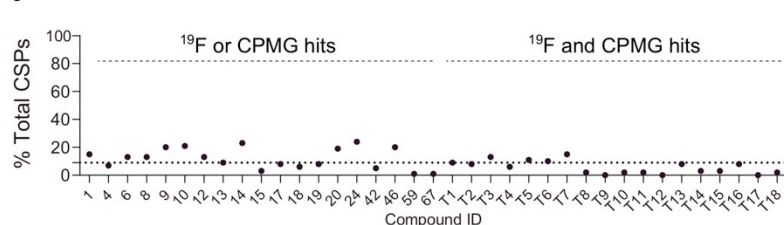
7.3.2. Supplementary figures
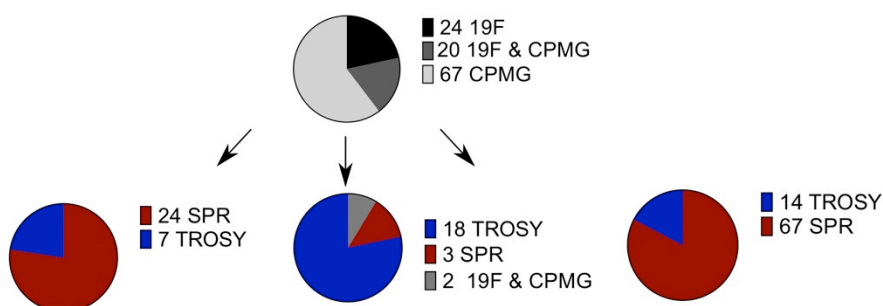
Figure 4.3-S1 Validation of ^{19}F NMR hits by SPR and ^{1}H-^{15}N TROSY NMR.

a The chart overview of validated ^{19}F NMR screening hits shows numbers of compounds validated in SPR and TROSY NMR. TROSY NMR titration data was used to rank compounds resulting in **24**, **10** and **12** hits. **b** Dose-response SPR sensograms of screening hits. **C** Total % of perturbed resonances in [15]N BambL WT in presence of 1 mM fragments identified in ^{19}F NMR screening. We followed up only fragments above average of 9% in ^{1}H-^{15}N TROSY NMR, whereas structures of the best hits are shown in **b**.
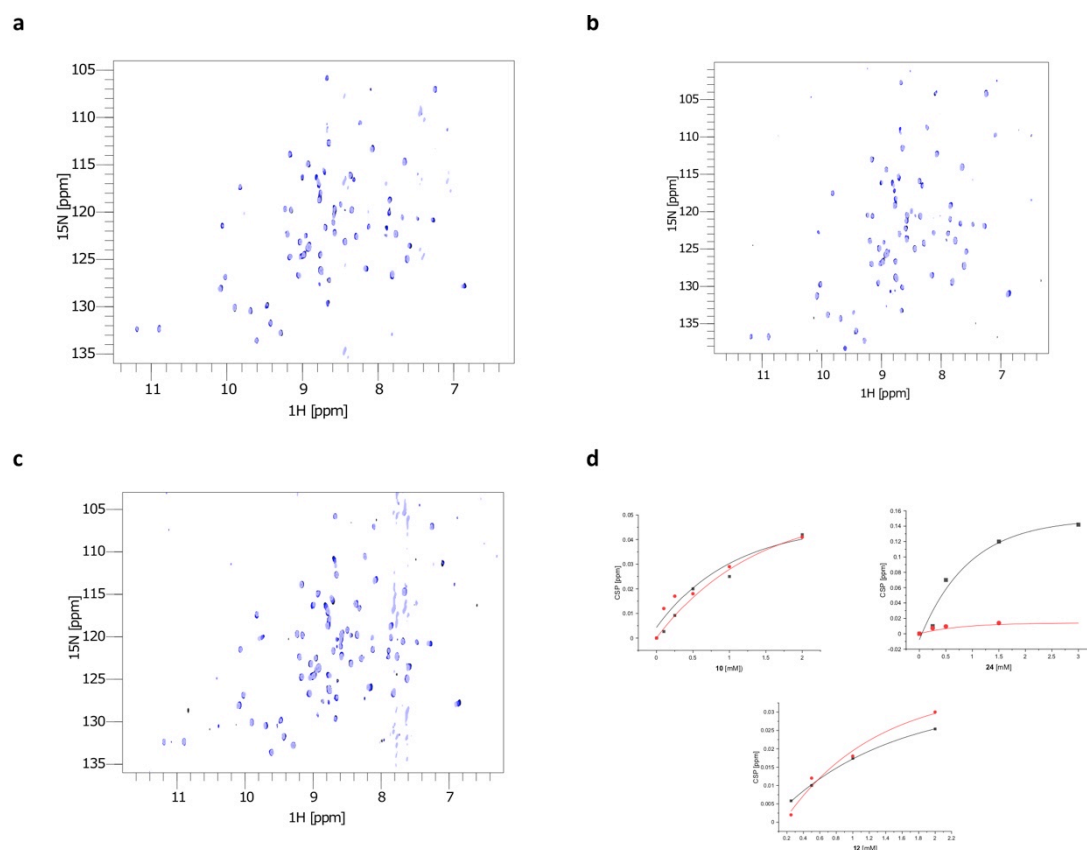
Figure 4.3-S2 Titration ^{1}H-^{15}N TROSY NMR experiments with 10, 12 and 24.

a Fingerprint of 0.05 mM [15]N BambL in presence of the compound **12** at the concentrations 0.25 to 2 mM (*dark* to *light blue*). **b** Fingerprint of 0.05 mM [15]N BambL in presence of 0.25 to 2 mM **10** (*dark* to *light blue*). **c** Fingerprint of 0.05 mM [15]N BambL in presence of 0.25 to 2 mM **24** (*dark* to *light blue*). **d** One-site binding model was applied to derive the binding affinities (K_d) of compound **10**, **12** and **24**.
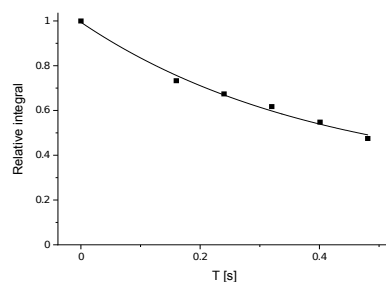
Figure 4.3-S3 Supporting information for the ^{19}F R$_2$-filtered NMR assay with 24.

The decay curve at 0.1 mM **24** in absence of BambL to determine R$_{2,free.}$
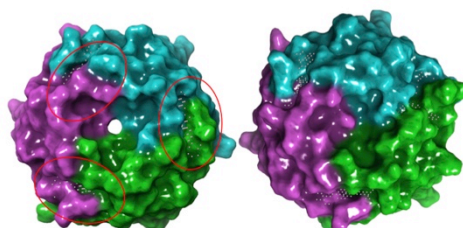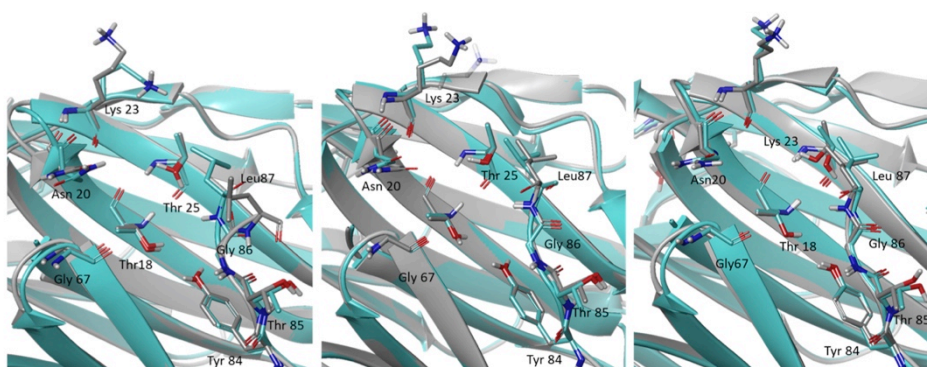
Figure 4.3-S4 The computational analysis of potential druggable binding sites in BambL.

a Three binding pockets (*red*) were identified in PDB structures 3ZZV (*left*) and 3ZW0 (*right*) using the SiteMap tool. **b** Superimposition of the predicted three binding sites at the interface of chains A, B (A) A, C (B) and B, C (C) in the crystal structures of BambL (PDB 3ZW0, *gray* and 3ZVV, *cyan*). Lys23 (in all the sites) and Leu87 (in one site) show significant difference in the side chain orientation, which slightly changes the shape and size of the predicted sites.

Figure 4.3-S5 Docking poses of 24, 12 and 10.

Binding poses of the ligands **24** (A), **12** (B) and **10** (C) predicted by docking SP studies. The key residues identified in the binding site are shown in the binding pose of **24**.

Figure 4.3-S6 Interaction maps of 24, 12 and 10.

Key residues involved in the interaction with the molecules **24** (A), **12** (B) and **10** (C) predicted by **a** XP and **b** SP docking.
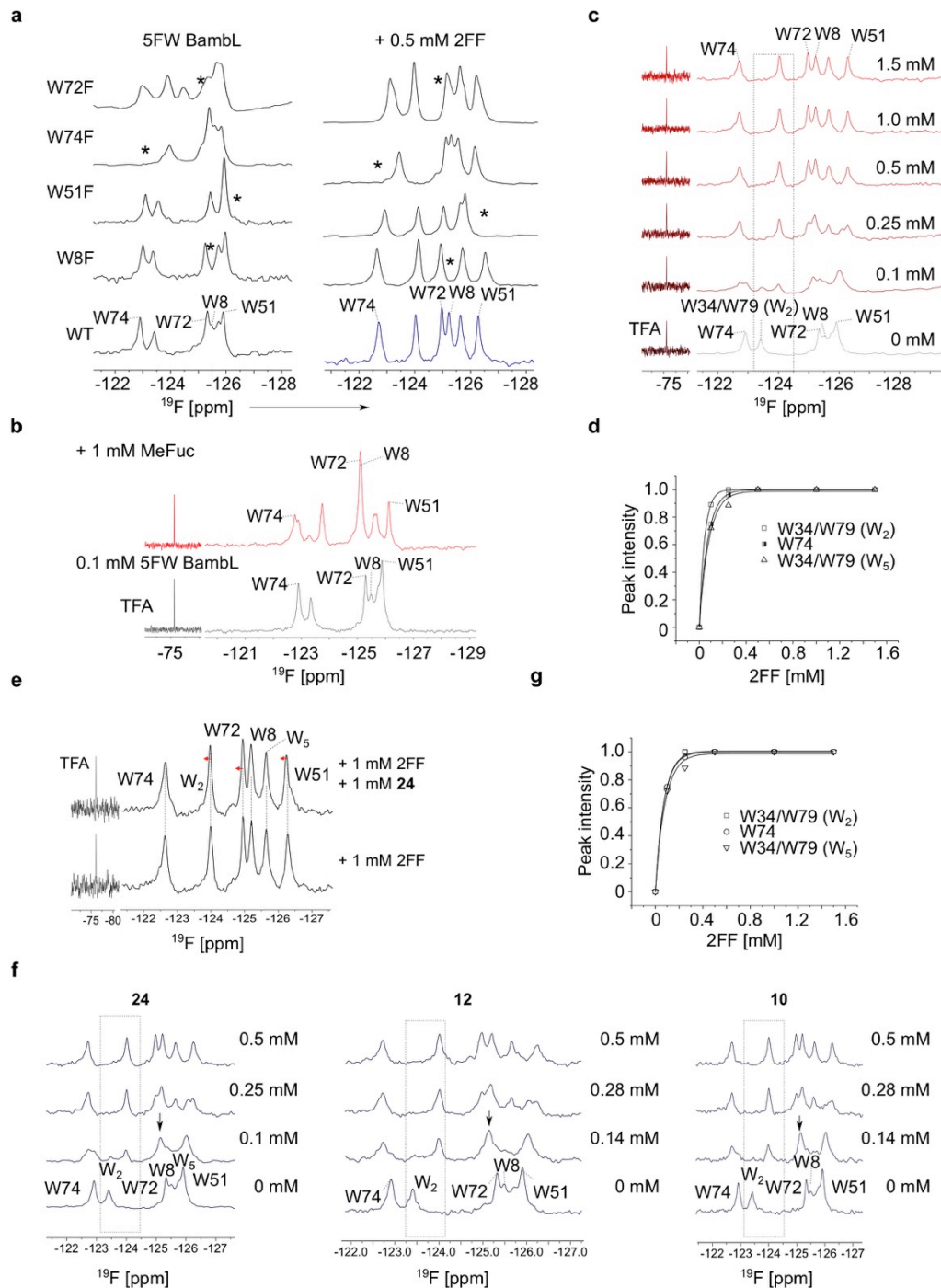
Figure 4.3-S7 PrOF NMR with 5FW BambL.

a Assignment of 5FW BambL resonances in PrOF NMR: * - indicates a missing 5FW resonance, which corresponds to a mutation. Additionally, PrOF NMR of all four mutants has been recorded in presence of 2FF to ensure the protein activity. **b** PrOF NMR spectra of 5FW BambL WT alone and in presence of methyl-α-L-fucose (MeFuc). Chemical shift resonance perturbations of all six 5FW show all resonances undergo a slow/intermediate exchange on NMR time scale. **c** Titration PrOF NMR spectra of 5FW BambL WT in presence of 2FF to derive the binding affinity. The single 5FW resonances were better resolved in presence of

2FF compared to MeFuc. Therefore, we used 2FF as a positive control for binding in this assay. **d** One-site binding model was applied to derive the affinity of 5FW BambL for 2FF. Hereby, the changes in peak intensities of indicated 5FW resonances were fitted to one-site binding model to derive the binding affinity. In this assay, fitting to two-site model did not deliver reliable results due to a large deviation in K_d *(not shown).* **e** PrOF NMR spectra of 5FW BambL with 1 mM 2FF without (*bottom*) or with 1 mM **24** (*top*). *Dashed lines* show chemical shift perturbations of 5FW resonances that remained perturbed in presence of **24** (*arrow*). W_2 and W_5 correspond to W79/W34 mutants, which were not assigned. **f** Titration PrOF NMR spectra of 5FW BambL WT in presence of 1 mM **24**, **12** or **10** and upon addition of 2FF at various concentrations. **g** One-site binding model to derive the binding affinity of 2FF to 5FW BambL in presence of 1 mM **24**.
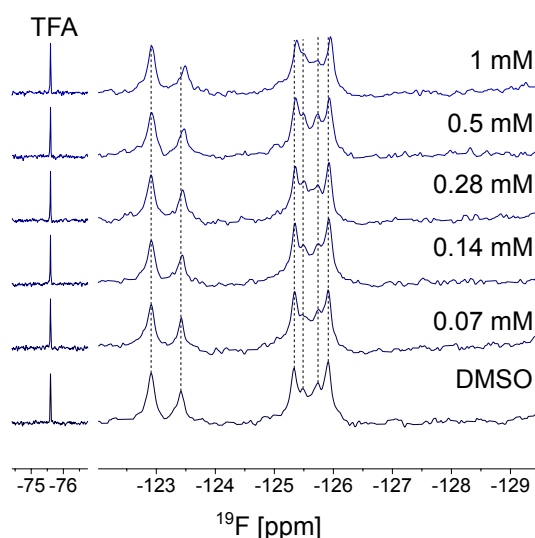
Figure 4.3-S8 Titration PrOF NMR spectra of 5FW BambL WT with 24.

Shown are PrOF NMR spectra of 5FW BambL WT in presence of the negative control DMSO or increasing concentration of the compound **24**. *Dashed lines* are shown to visualize the chemical shift perturbations of 5FW resonances upon addition of the compound **24**.
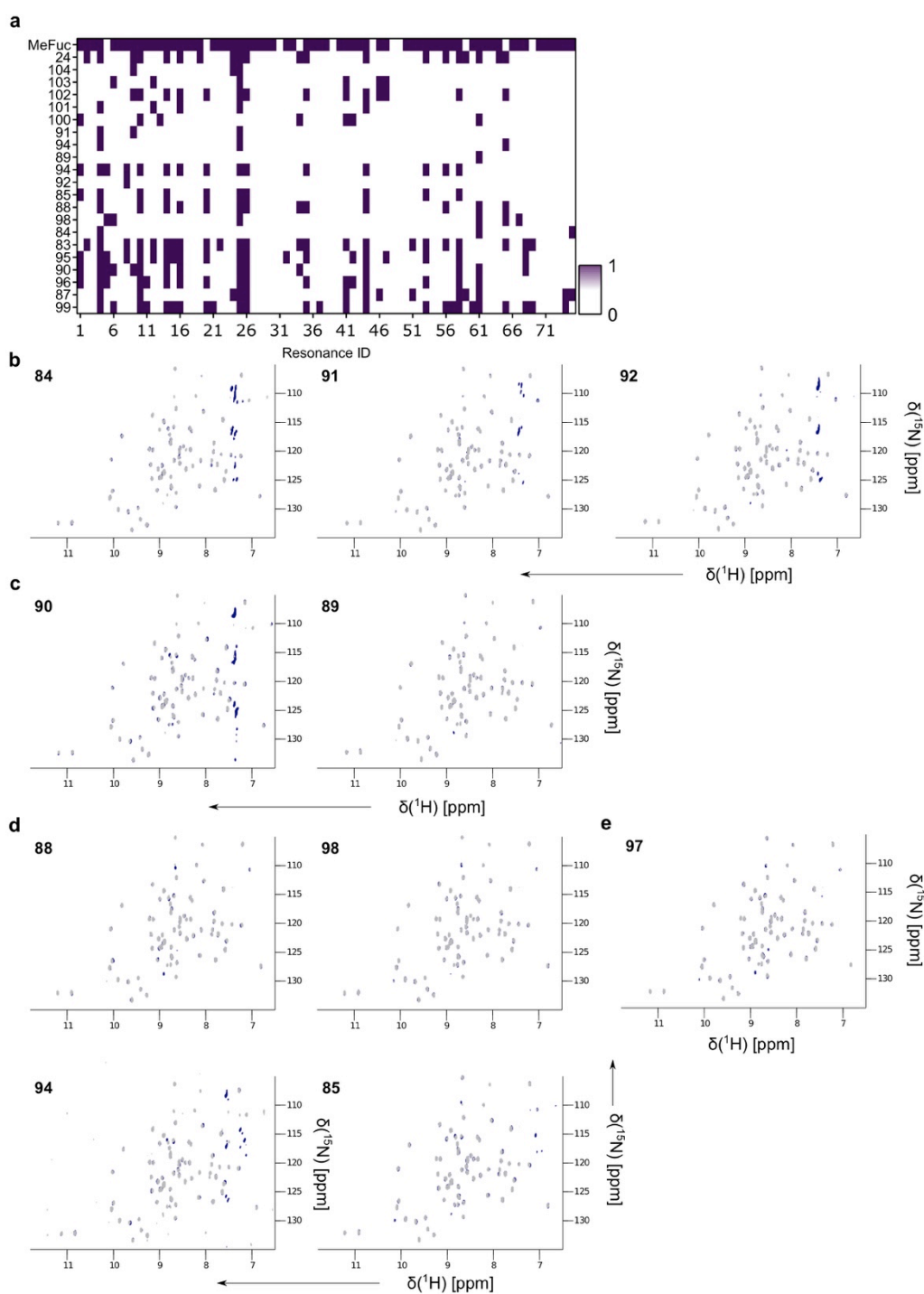
Figure 4.3-S9 Validation of the analogues of the hit 24.

a The 1:0 plot shows the conformational perturbations in [15]N BambL resonances in presence of **24** derivatives: 1 is CSP>0.01 ppm, whereas 0 is CSP<0.01 ppm. MeFuc was used as a positive control. **b-e** show the ^{1}H-^{15}N TROSY NMR spectra of [15]N BambL alone (*gray*) and in presence of 1 mM fragments.

Figure 4.3-S10 Validation of the derivatives of the compound 83.

Shown are ^{1}H-^{15}N TROSY NMR spectra of ^{15}N BambL in presence of derivatives of the compound **83**. The compound **99** improved binding to ^{15}N BambL compared to **83**. The compound **101** bound ^{15}N BambL better compared to **100**, **102**, **103** and **104**.

Figure 4.3-S11 The computational analysis of the derivatives of the compound 24.

a and **b** show the docking poses of four structural derivatives of the compound **24** (**84**, **87**, **90** and **94**) and its interaction maps with the predicted binding pocket, respectively. No electrostatic interactions have been observed for **84**, whereas modifications on the amino group (**90**) and a change of benzyl group (**94**) can be tolerated. Moreover, change of the benzyl group position (**87**) can be tolerated as well.

Figure 4.3-S12 ^{1}H-^{15}N TROSY NMR of ^{15}N BambL mutants.

Shown are ^{1}H-^{15}N TROSY NMR spectra of ^{15}N BambL mutants: 1) carbohydrate binding site region (W72F, W51F, W74F and W8F) and 2) in the predicted pocket (T18S, L87R and T25S). All ^{15}N BambL mutants were folded and active.

185
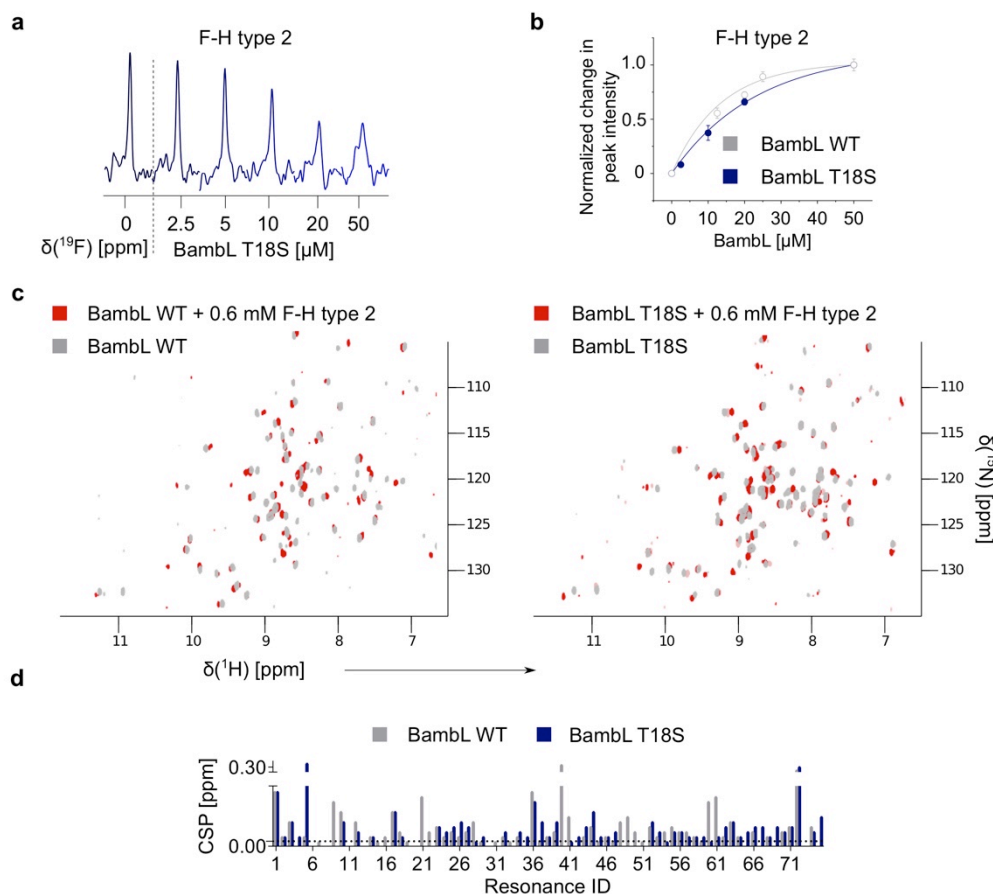
Figure 4.3-S13 F-H type 2 interaction with BambL WT and T18S.

a Shown are ^{19}F NMR spectra of 0.1 mM F-H type 2 alone and in presence of various protein BambL T18S concentrations. The changes in the peak intensities were used to derive the affinity of F-H type 2 for BambL T18S (n=3). **b** One-site binding model was applied to determine the K_d value of F-H type 2 for BambL T18S. BambL T18S revealed a two-fold affinity decrease compared to BambL WT, which was reported previously.[1] **c** ^{1}H-^{15}N TROSY NMR analysis of F-H type 2 interaction with ^{15}N BambL WT (*left*) and T18S mutant (*right*). **d** Quantitative analysis of NMR chemical shift perturbations (CSPs) in ^{15}N BambL WT *vs* T18S upon F-H type 2 binding. Overall, T18S mutation reduced the magnitude of CSPs in ^{15}N BambL suggesting a negative modulatory role of the pocket on the carbohydrate-binding site.
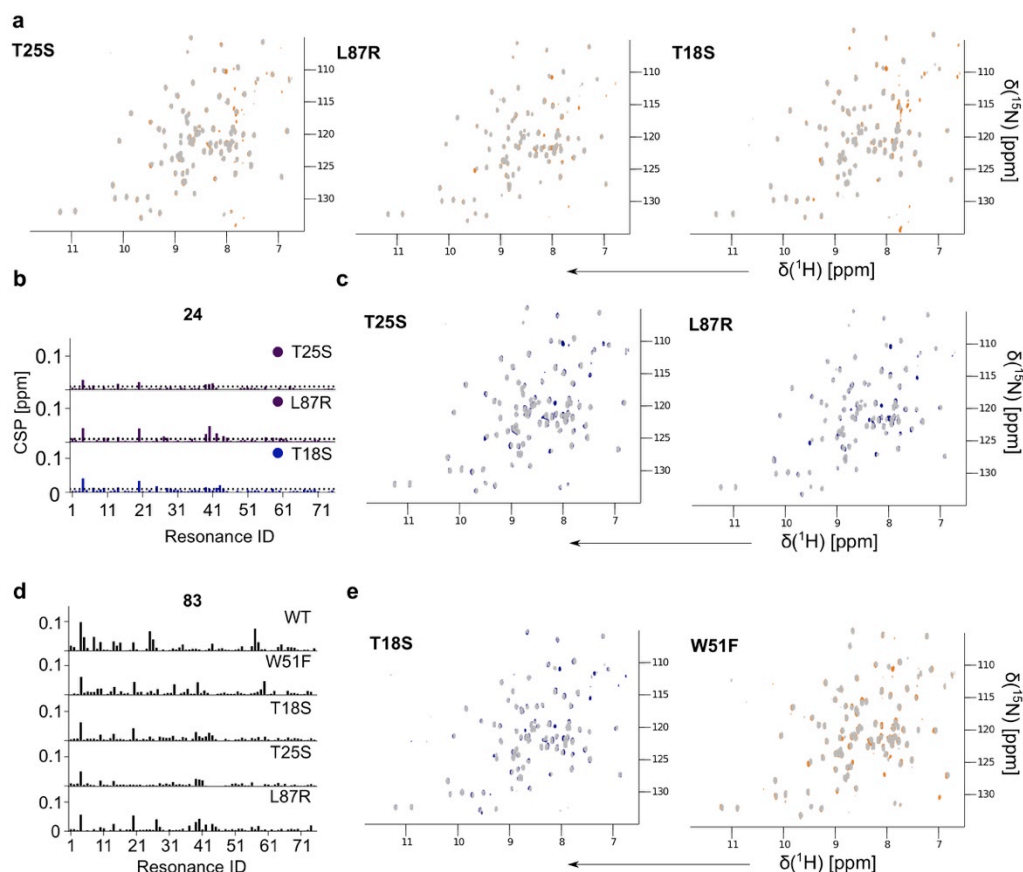
Figure 4.3-S14 Characterization of the interactions between [15]N BambL mutants and the compounds 24 and 83 in ^{1}H-^{15}N TROSY NMR.

a Shown are ^{1}H-^{15}N TROSY NMR spectra of [15]N BambL mutants (T25S, L87R and T18S) alone (*gray*) and in complex with 1 mM **24** (*orange*). **b** The plots show chemicals shift perturbations (CSPs) of the backbone resonances in presence of **24**. Compared to BambL WT, we observed a reduced magnitude of CSPs in presence of **24** in all three mutants. This suggests that the predicted pocket was partially closed for **24** binding. **c** ^{1}H-^{15}N TROSY NMR spectra of [15]N BambL pocket mutants: L87R and T25S (*gray*) in complex with 1 mM **83** (*blue*). **d** The CSP plots show the changes in the backbone resonances in presence of **83** demonstrating a reduced binding of **83** to [15]N BambL mutants (T25S, L87R and T18S) similar to **24**. **e** ^{1}H-^{15}N TROSY NMR spectra of [15]N BambL the predicted site (T18S) and the carbohydrate-binding (W51F) mutants with 1 mM **83** (*blue* and *orange*). As shown in **d** the carbohydrate-binding site mutant W51F demonstrated a reduced binding to **83** compared to [15]N BambL WT similar to the predicted site mutants. This supports the existence of a communication between both sites.
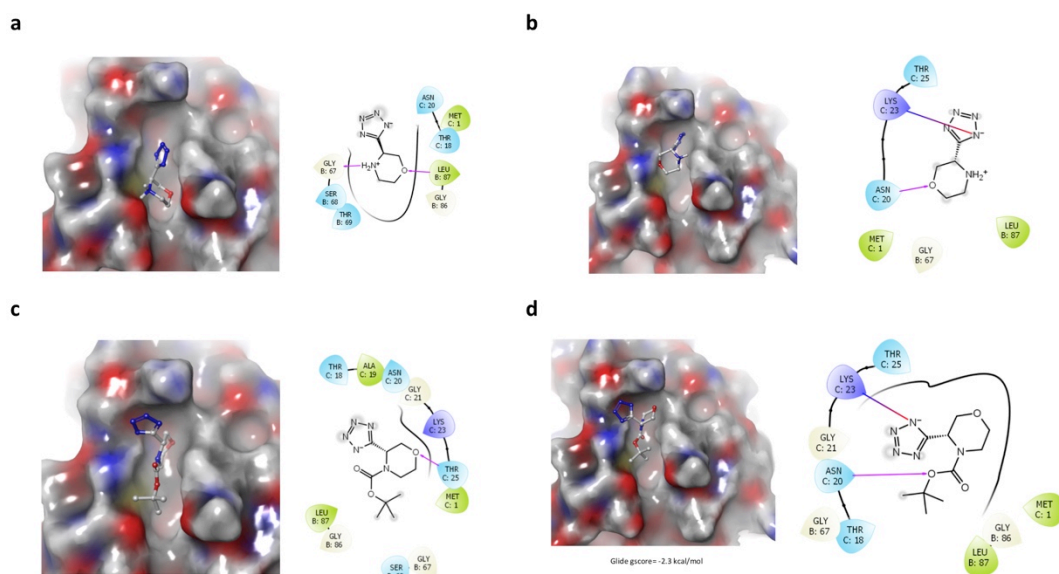
Figure 4.3-S15 Computational docking analysis of the compound 83.

Shown are docking pose orientations 1 (**a**) and 2 (**b**) of the compound **83**, whereas **c** and **d** are both docking poses of its derivative **99**.
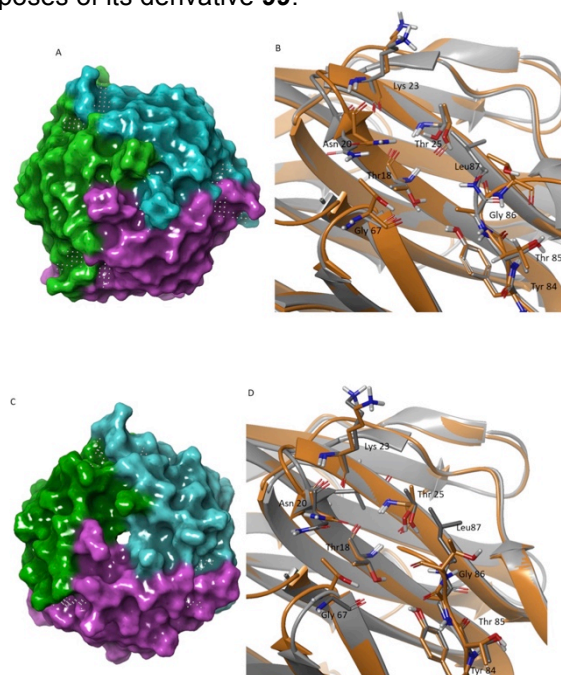
Figure 4.3-S16 Computational analysis of RSL.

Computational analysis of potential druggable binding sites in apo (A, B) and holo (C, D) forms of RSL**.** Three binding pockets comparable to BambL were identified in PDB structures 3Z8I (apo) and 5AJB (holo) using SiteMap tool. Show is the superimposition (B, D) of one of the predicted binding sites at the interface of chains B, C of apo (B) and holo (D) forms of BambL (*gray*) and RSL (*orange*). This demonstrates the partial similarity of the binding site residues. However, terminal residues in the flexible loop region have significant differences.

AFL (PDB 4AGI) top view   AFL (PDB 4AGI) side view
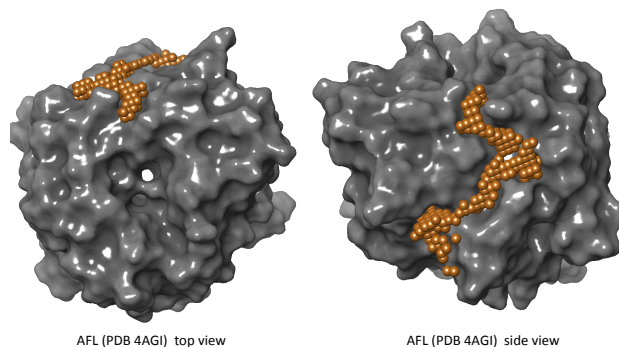
Figure 4.3-S17 Computational analysis of AFL.

Computational analysis of potential druggable binding sites in AFL reveals one druggable site (*dots*) as shown in the top (*left*) and side (*right*) views.

7.3.3. Supplementary tables

Table 4.3-S1 List of chemical shift perturbations (CSP) of 5FW resonances derived in PrOF NMR.
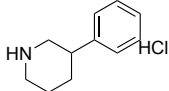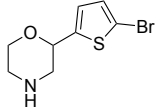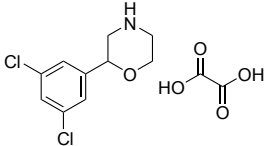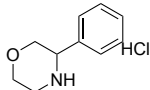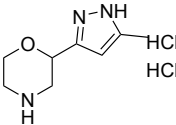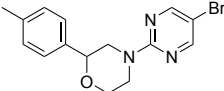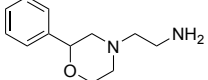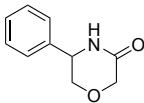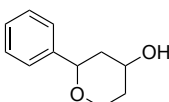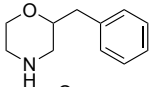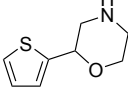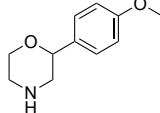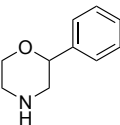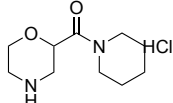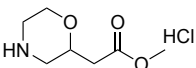
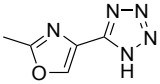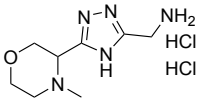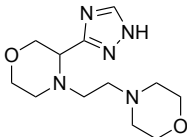| 5FW	CSP [ppm]				
	DMSO*	24	12	10	83
W74	0.01	0	0.01	0.01	0.01
W34/W79	0.01	0.04	0.03	0.03	0.08
W72	0.01	0.02	0.02	0.02	0.03
W8	0.01	0	0	0.01	0.04
W34/W79	0.01	0.03	0.01	0	0.08
W51	0	0.02	0.03	0.02	0.05
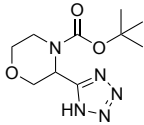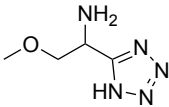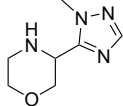
* Average value of 3 independent measurements

Table 4.3-S2 Commercial analogues of the compound 24.

Compound ID	MW [Da]	Structure
24	321.26	
83	191.62	

84	197.71	
85	248.14	
86	322.15	
87	199.68	
88	240.13	
89	334.22	
90	206.29	
91	177.20	
92	178.23	
93	267.28	
94	169.25	
95	193.25	
96	163.22	
97	234.73	
98	195.65	

99	255	
100	143.15	
101	151.13	
102	270.16	
103	267.32	
104	168.19	

7.3.4. Supplementary references

[1] G. Fittolani, E. Shanina, M. Guberman, P. H. Seeberger, C. Rademacher, M. Delbianco, *Angewandte Chemie International Edition* **2021**, 60, 13302-13309.

[2] A. Audfray, J. Claudinon, S. Abounit, N. Ruvoën-Clouet, G. Larson, D. F. Smith, M. Wimmerová, J. Le Pendu, W. Römer, A. Varrot, A. Imberty, *J Biol Chem* **2012**, 287, 4335-4347.

[3] N. Kostlánová, E. P. Mitchell, H. Lortat-Jacob, S. Oscarson, M. Lahmann, N. Gilboa-Garber, G. Chambat, M. Wimmerová, A. Imberty, *Journal of Biological Chemistry* **2005**, 280, 27839-27849.

[4] J. Houser, J. Komarek, G. Cioci, A. Varrot, A. Imberty, M. Wimmerova, *Acta Crystallographica Section* D **2015**, 71, 442-453.

[5] E. Shanina, E. Siebs, H. Zhang, D. Varón Silva, I. Joachim, A. Titz, C. Rademacher, *Glycobiology* **2021**, 31, 159-165.

[6] J. Aretz, E.-C. Wamhoff, J. Hanske, D. Heymann, C. Rademacher, *Frontiers in Immunology* **2014**, 5.

[7] F. Delaglio, S. Grzesiek, G. W. Vuister, G. Zhu, J. Pfeifer, A. Bax, *J Biomol NMR* **1995**, 6, 277-293.

[8] W. F. Vranken, W. Boucher, T. J. Stevens, R. H. Fogh, A. Pajon, M. Llinas, E. L. Ulrich, J. L. Markley, J. Ionides, E. D. Laue, *Proteins: Structure, Function, and Bioinformatics* **2005**, 59, 687-696.

[9] M. P. Williamson, *Prog Nucl Magn Reson Spectrosc* **2013**, 73, 1-16.

[10] H. Y. Carr, E. M. Purcell, *Physical Review* **1954**, 94, 630-638.

[11] T. A. Halgren, J Chem Inf Model **2009**, 49, 377-389.

[12] S. Schrödinger Release 2018-1: Maestro, LLC, New York, NY, **2018**.

[13] H. M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov, P. E. Bourne, *Nucleic Acids Research* **2000**, 28, 235-242.

[14] aM. H. Olsson, C. R. Sondergaard, M. Rostkowski, J. H. Jensen*, J. Chem. Theory Comput.* **2011**, 7, 525-537; bH. Li, A. D. Robertson, J. H. Jensen, *Proteins* **2005**, 61, 704-721; cD. C. Bas, D. M. Rogers, J. H. Jensen, *Proteins* **2008**, 73, 765-783.

[15] E. Harder, W. Damm, J. Maple, C. Wu, M. Reboul, J. Y. Xiang, L. Wang, D. Lupyan, M. K. Dahlgren, J. L. Knight, J. W. Kaus, D. S. Cerutti, G. Krilov, W. L. Jorgensen, R. Abel, R. A. Friesner, *J. Chem. Theory Comput.* **2016**, 12, 281-296.

[16] S. Schrödinger Release 2018-1: LigPrep, LLC, New York, NY, 2018.

[17] R. A. Friesner, J. L. Banks, R. B. Murphy, T. A. Halgren, J. J. Klicic, D. T. Mainz, M. P. Repasky, E. H. Knoll, M. Shelley, J. K. Perry, D. E. Shaw, P. Francis, P. S. Shenkin, *J. Med. Chem.* **2004**, 47, 1739-1749.

7.3.5. Author contribution

C.R. and A.I. conceived and supervised the experiments. The manuscript was written by E.S. with contributions of C.R., A.I., S.K. and K.L.; all NMR experiments were carried out by E.S.; S.K. and K.L. supported this work by performing SPR studies and computational analysis, respectively. P.H.S. provided the institutional support for this work. All authors have read and agreed to the published version of the manuscript.

7.4. Supporting Information for Subchapter 4.4.

7.4.1. Supplementary Results and Discussion

Virtual screening and evaluation of LecA hits

For the virtual screening, we used the Bioinfo Database (http://bioinfo-pharma.u-strasbg.fr/bioinfo/), a curated database of drug-like compounds, filtered for molecules with at least two potential hydrogen bond donors and acceptors to match the polar nature of the carbohydrates. Approximately 3 million molecules were docked into the crystallographic structure of LecA. We carefully inspected all the available crystallographic LecA structures (**Table 4.4-S5**) complexed with small molecules to see if there was considerable side-chain variation that could impact the virtual screening process. The carbohydrate-binding site showed to be conserved through all structures. The key aspect of post-processing of docking poses was based on the interaction similarity pattern between the docked compounds and the available ligands co-crystallized with LecA. The similarity was calculated using GRIM, a knowledge-based approach to convert protein-ligand complexes in interaction pattern graphs and score docking solutions by similarity of predicted interaction patterns to the already visited in the PDB. The similarity was quantified using the GRIM score (GrSc). Since most of the co-crystallized ligands were carbohydrate-like, we expected to have drug-like hits with similar interactions with the carbohydrate-binding site residues. Other additional filters were used to ensure the elimination of the carbohydrate-like molecules, but still maintaining the correct interaction pattern, such as presence of interaction with the Ca^{2+} ion. In addition, the molecules with the tetrahydrofuran or tetrahydropyran scaffolds and number of polar interactions with binding site residues superior to 3 were discarded. The remaining ligands were clustered by the maximum common substructure (MCS) to select a chemically diverse set of compounds and after visual inspection for the presence of key hydrogen bonds shared by galactose and the binding site residues, a total of 46 hits were selected as hits and 37 were purchased for testing.

Virtual screening and evaluation of LecB hits

A similar protocol was followed for the virtual screening of LecB. The available crystal structures were analyzed (**Table 4.4-S6**), but no significant variation was found for the fucose binding site residues. We screened the same database in our virtual screening protocol, the commercially available drug-like compounds from Bioinfo and the same filtering was applied for LecB. Post-processing of docking poses followed the same principle applied for LecA, which mainly focused on comparing the interaction patterns between the docked compounds and the available co-crystallized ligands with LecB

using GRIM. Pose filtering followed the same criteria described for LecA and an additional filter was added for LecB, in which shape similarity with fucose was considered to improve the hit number. After selecting a diverse set through clustering and visually inspecting the molecules for the presence of key hydrogen bonds, a total of 42 molecules were selected as hits.

Evaluation of virtual screening hits for LecA

To evaluate the 37 commercial fragments identified in virtual screening, we performed experimental analysis using SPR and ^{1}H-^{15}N TROSY NMR. First, we tested compound binding in SPR, where 19 compounds demonstrated a dose-dependent response. However, compounds perturbed the resonances in ^{15}N-labeled LecA in a similar manner to methyl-α-D-galactose (hereafter, MeGal) in ^{1}H-^{15}N TROSY NMR only slightly resulting in 8 compounds (**Figures 4.4-S1c-S1e**). Altogether, we identified two hits being confirmed in both SPR and TROSY NMR (**Figure 4.4-S1f**), whereas the compound **5c** demonstrating the strongest effect in TROSY NMR was not confirmed in SPR (**Figure 4.4-S1e**). Therefore, we performed a biochemical study based on fluorescence polarization, which did not confirm competitive properties of VS hits (*data not shown*). This suggests that hits observed in TROSY and SPR NMR were very weak ligands of LecA. Cumulatively, virtual screening identified 37 compounds that could bind to the carbohydrate-binding site of LecA, but the experimental validation did not deliver hits for future fragment evaluation studies.

Evaluation of virtual screening hits for LecB

In virtual screening against LecB, we identified 42 commercially available fragments. To validate these compounds, we performed ^{1}H-^{15}N TROSY NMR and FP assay only. Due to the low hit rates for LecB, we combined fragments in mixtures of 10 compounds and validated its binding to LecB in ^{1}H-^{15}N TROSY NMR. The changes in spectra in presence of compounds were compared to the positive control methyl-α-L-fucose (hereafter, MeFuc) delivering no hits for the carbohydrate-binding site of LecB (*not shown*). Finally, the analysis of compounds in a fluorescence polarization assay did not confirm hits from virtual screening (*not shown*). Taken together, virtual screening identified 42 compounds that could potentially bind to the carbohydrate-binding site of LecB, but these were not confirmed experimentally.

Chemical derivatization of the hydroxamic acid hit 1

The expansion of the hydroxamic acid library was guided by an SAR study by TROSY-NMR experiments and later extended after receiving the first co-crystal structure of the hydroxamic acid **35** with LecA (see below). The initial derivatization focused on the core structure of hit **1**. The hydroxamic acid functional group was methylated on either the oxygen or the nitrogen atoms, starting from 2-phenylacetyl chloride (**S2**) and

performing an amidation with methoxyamine or *N*-methylhydroxylamine, respectively, resulting in **7** and **8** (**Scheme 4.4-1A**). Removing the hydroxy or the amine group of **1** led to commercial compounds **3** and **7** (Scheme 4.4-2). Next, we modified the linker between the hydroxamic acid moiety and the phenyl ring by either removing the ring (**18**) or varying its length **35**, **43** and **44**. The synthesis was performed with hydroxylamine and the corresponding acyl chlorides after synthesis from their acids using oxalyl chloride (**Scheme 4.4-1B**). Additionally, the phenyl ring was replaced with a cyclohexyl ring in the compound **6**. To cover a broad spectrum of hydroxamic acids, compounds with a substituted phenyl ring (**9**, **10**, **15** and **11**), cyclic linkers such as a thiazole **27**, furane **34**, and 1-hydroxypyridin-2(1*H*)-ones were purchased. Later, derivatives of **35** were synthesized using the corresponding acid and oxalyl chloride, followed by the reaction with hydroxylamine. Linker and hydroxamic acid functional group were kept unchanged and electron donating/ withdrawing substituents **36**, **40**, **41** and **39** were introduced on the phenyl ring, the ring itself was replaced by a thiophen residue **37**, and compound **42** containing a double bond as a spacer was designed to increase rigidity.

^{1}H-^{15}N TROSY NMR of hydroxamate derivatives

Protein-observed NMR technique is a valuable method in detection of weak protein-ligand interaction in initial FBDD campaigns.[1] Therefore, we used the previously established ^{1}H-^{15}N TROSY NMR with ^{15}N-labeled LecA to rank binding of hydroxamates to LecA. Hereby, we derived the total number of promoted chemical shift perturbations (CSPs) in ^{15}N LecA and compared it to the positive control MeGal (**Figure 4.4-S4a**). Notably, compounds with a terminal benzyl elucidated a better binding compared to **1**, but worse than MeGal. Briefly, we observed that the changes on the hydroxamic acid group (**3**, **7** and **8**) are not tolerated. To demonstrate the importance of the benzyl group, we replaced it with a methylcyclohexane group (**6**), which preserved binding to LecA compared to a free hydroxamic acid (**18**). Moreover, the linker connecting both groups cannot be too flexible (**47**) or rigid (**42, 43**) resulting in two compounds (**2**, **35**) preserving the binding (**Figures 4.4-S4b-c**). Modifications on the linker position 2 (**29**) and benzyl group (**5**, **36**) are suitable for further fragment expansion. Interestingly, the structural rescaffolding of hydroxamic acid to a cyclic form (**20**, **21**) preserves LecA binding, which has not been reported for metalloenzymes previously.[2] This is not surprising given the shallow binding site of LecA compared to rather deep active site pockets in metalloenzymes, which on the other hand require long and linear scaffolds as shown on example of marketed drugs for MMPs (**Table 4.4-S1**, Group 4).

SPR analysis of hydroxamate derivatives

In order to establish structure-activity relationship (SAR) study of the hydroxamic acid derivatives by SPR, we assessed LecA binding capability of 27 commercial and 7 in-house synthetic hydroxamic acid derivatives by surface plasmon resonance (SPR). Binding responses of each compound were recorded at 0.2 and 1 mM to establish dose-dependent responses. The binding responses were normalized as binding efficiency (%) to account for the molecular weight differences of the compounds and the amount of active immobilized LecA during the analyses (see **Supplementary Materials and Methods** for normalization calculation). In agreement with ^{1}H-^{15}N TROSY NMR data, no binding response was observed when hydroxamic acid was replaced by amide (**3**), confirming that the binding to LecA was dependent on the presence of hydroxamic acid functional group.

Notably, compounds with a single terminal benzyl are predominant in our screen and elicit good dose-response and normalized binding responses. None of the in-house synthetic compounds showed positive dose-dependent binding responses, contradictory to the findings from NMR analyses (**Figure 4.4-S6a-b**), whereas 8 of the commercial hydroxamic acid compounds (**4**, **9**, **11**, **15**, **26**, **27**, **29**, **35**) exhibited more than twice dose responses with the normalized binding responses greater than 10% (**Figure 4.4-S6c**). This discrepancy emphasizes the importance of performing several orthogonal assays in the analysis of protein-fragment interaction. Especially, commercial compounds may contain impurities, such as metals, causing false-positive responses in SPR, which were taken care of in our in-house synthesized compounds.[3-4] Moreover, a rather weak affinity of hydroxamate derivatives was a limiting factor for employing SPR and thus, explaining the discrepancy with ^{1}H-^{15}N TROSY NMR results. Taken together, we concluded that SPR was not reliable to prioritize the hydroxamate derivatives and thus, other orthogonal methods were employed.

Competitive binding assay of hydroxamate derivatives with LecA

The hydroxamic acid derivatives were tested in a competitive binding assay (**Table 4.4-S1**). The parent compound *N*-hydroxy-2-phenylacetamide (**2**) showed a low millimolar binding affinity to LecA (K_d=6.1±0.9 mM). We observed that the hydroxamic acid scaffold was essential for binding the carbohydrate-binding site of LecA. The compounds, with methylated **3** or amide **7** lost their binding and inhibitory effect. A modification at the nitrogen atom was tolerated but did not improve the binding affinity as shown by the methylated compound **8** (inh.=19.9±2.7%) compared to **2** (inh.=18.3±1.2%, K_d=6.1±0.9 mM). An interaction of the hydroxamic acid functional group alone, without the phenyl ring **18**, and LecA could not be detected in any assay. None of the tested modifications on the phenyl ring of *N*-hydroxy-2-phenylacetamide led to improved potencies **9**, **10**, **15**, **11**, inh.=14–16%). On the other hand, substitution

of the phenyl ring with a cyclohexyl ring slightly improved the affinity (**6**, inh.=21.4±0.6%, K_d =4.4±0.6 mM). The optimal linker length between the phenyl ring and the hydroxamic acid moiety consist of 3 methylene groups (**35**, inh.=26±0.7%, K_d=4.6±0.9 mM). Longer spacers led to a binding decrease (**43**, inh.=7.6±3.0% and **44**, inh.=16.7±2.4%), similarly, diminished activity was observed for the rigid olefin (**42**, inh.=12.7±2.7%). Finally, substituted *N*-hydroxy-4-phenylbutanamide revealed that electron donating substituents in *para* position increases binding affinity (-Me **36**, -OMe **40**, -OH **41**, inh.=27–39%), and the electron withdrawing groups dramatically decrease the potency (-NO$_2$ **39** inh.=4.1±2.0) hinting that a CH-π-stacking arises with LecA and the phenyl ring. Interestingly, the 1-hydroxypyridin-2(1H)-one derivatives were as potent as the best *N*-hydroxy-4-phenylbutanamides (**19** inh.=37.2±1.8 and **20** inh.=37.2±0.7%), possibly due to the difference in predicted pK_A values for the linear (**35** pK_A≈9.41±0.20) and cyclic hydroxamic acids with lower pK_As (**19** pK_A≈6.0±0.1) leading to the stronger chelating effects on the calcium ion and hence increased affinity.

Competitive ^{19}F T_2-filtered NMR with hydroxamate derivative 5

We evaluated the selectivity and Ca^{2+}-dependency of the hydroxamate-LecA interaction and compared it to other Ca^{2+}-dependent lectins LecB, DC-SIGN and Langerin. For this, we used a hydroxamate derivative of **1** (**5**) as a fluorinated reporter molecule in a competitive ^{19}F NMR. Here, we tested 100 µM **5** binding in absence (5 mM EDTA) and presence of 10 mM CaCl$_2$ alone and with 10 µM lectins. Both Langerin and DC-SIGN demonstrated Ca^{2+}-independent binding to **5**. For Langerin, the interaction persisted with and without CaCl$_2$, demonstrating **5** targeting a secondary site. Notably, this interaction was weak, as it did not bind to ^{15}N Langerin CRD in ^{1}H-^{15}N HSQC NMR (**Figures 4.4-S5b-c**). For DC-SIGN, we observed a partial recovery of the fluorine peak **5** in the presence of CaCl$_2$. Such behavior in ^{19}F NMR has been previously observed for fragments hits identified in ^{19}F NMR targeting the secondary sites in DC-SIGN.[5-6] Therefore, we validated **5** binding to DC-SIGN in ^{1}H-^{15}N HSQC and STD NMR. Both assays revealed a very weak binding of **5** being partially competed with 30 mM D-mannose in STD NMR and fully competed with 5 mM EDTA, suggesting **5** bound to multiple sites in DC-SIGN (**Figures 4.4-S5a-d**). This is not surprising given the presence of an aryl ring in **5** besides the hydroxamic acid group, which has potentially interacted with a secondary site of DC-SIGN. To rank the hydroxamate derivatives, we expected a stronger binding ligand to compete the reporter molecule from the orthosteric site (**Figure 4.4-S7**). In this study, the reporter molecule **5** interacted with LecA as shown by a decrease of the fluorine peak intensity. Notably, **5** recovered upon addition of **35** and **2**, whereas the initial hit **1** did not, as

well as the negative control **3**. Given other 7 compounds show rather a comparable competition, we concluded that **35** was the best binder in this assay.

Docking study of malonate 58 with lectins

PA-IL (LecA)

Docking simulation of compound **58** with LecA (PDB: 4CP9) revealed that both carboxylates of the malonic acid fit in the binding-pocket of the lectin (**Figure 4.4-S8a-b**). Moreover, the data suggested that only one carboxylate interacts with the calcium ion, whereas the other carboxylate can form a hydrogen bond with the water molecule buried in the pocket. This is in line with the SAR study using PrOF NMR, where compounds with only one carboxylate (**63**, **70**) failed to bind to the protein. Taken together, this indicated that the carbohydrate-binding site of LecA can accommodate two adjacent carboxylates (**58**) and that both of them are required: one to interact with calcium ion, and one with the water molecule.

PA-IIL (LecB)

The compound **58** was docked with LecB (PDB: 1OXC) and two binding poses gained our attention. The highest-ranking showed the pose where one carboxylate interacted with both calcium ions, whereas the other carboxylate interacted with the protein surface namely S22 and S23. Additionally, it revealed a potential interaction of the CF_2-group on the cyclopentyl group and T98 (**Figure 4.4-3a** and **4.4-S11b**). The second docking pose displayed both carboxylates interacting with the calcium atoms, as well as G97 and S22 (**Figure 4.4-11a-b**). Together, both poses indicated that both carboxylates are required for binding, what is also in line with data from the SAR study, where compounds with one carboxylate moiety (**63**, **70**) failed to bind to the protein.

DC-SIGN CRD (CD209)

The compound **58** was docked with DC-SIGN (PDB: 2XR5). The data showed that in the carbohydrate binding-site, **58** was able to interact with the calcium with one carboxylate, while the other carboxylate made an H-bond with N344. Interestingly, a hydrogen atom of the cyclopentyl ring, alpha to the CF_2, is able to form a CH-π interaction with the aromatic ring of F313.

It is interesting to note that **58** appeared to favor binding to the carbohydrate-binding site rather than in the secondary binding site, although the latter bears two calcium atoms, indicating that the interaction of **58** with the protein was not driven by sheer electrostatic forces alone. Indeed the secondary binding site, which is more solvent exposed, probably offers to the fragment less possibilities for the interactions with DC-SIGN. Therefore, this further indicated that the malonate moiety alone could not explain the selectivity in fragment binding. On the other hand, interactions with the

amino acids neighboring the calcium ions can drive selective binding for malonate-containing fragments.

Competitive ^{19}F T_2-filtered NMR with malonate derivative 61

To investigate the Ca^{2+}-dependency and selectivity of the malonates-lectin interaction, we subjected metal-dependent lectins to binding studies in ^{19}F T_2-filtered NMR, where BambL was expected not to interaction with MBPs such as malonates. As result, **61** bound to $CaCl_2$ alone as indicated by a chemical shift perturbation (CSP) of the fluorine peak in presence of 20 mM $CaCl_2$, but not in its absence (5 mM EDTA). In presence of $CaCl_2$ and 10 µM LecA, the fluorine peak decreased in the peak intensity demonstrating that this interaction was Ca^{2+}-dependent. We observed a similar pattern for LecB and DC-SIGN. For LecB, the competition experiment with 5 mM MeFuc showed a full recovery of the fluorine peak demonstrating that **58** targeted the carbohydrate-binding site of LecB. Moreover, ^{15}N TROSY NMR revealed that **58** was fully displaced from its binding site upon addition of 5 mM EDTA supporting our ^{19}F NMR data. However, in the competition experiment with 30 mM D-mannose and DC-SIGN, the fluorine peak of **61** recovered only partially in presence of the competitor. This observation was likely due to a weak affinity of 30 mM D-mannose for DC-SIGN (K_d=3 mM compared to **61** (K_d = 1.9 mM) given **58** was displaced from DC-SIGN CRD upon addition of 10 mM EDTA in ^{1}H-^{15}N HSQC NMR (**Figure 4.4-S17b**). Further, the off-target effect of **58** to Langerin was observed as the fluorine peak decreased in the peak intensity in presence of 5 mM EDTA suggesting that interaction was Ca^{2+}-independent. However, the **61** peak showed a stronger decrease in presence of $CaCl_2$ and binding to BambL. Given the presence of the secondary sites in both lectins,[5-7] we aimed to confirm that **61** targeted the secondary sites in both lectins. For this, we added a competitor (30 mM D-mannose and 10 mM MeFuc) expecting both ^{19}F peaks to remain unchanged. Indeed, competitors did not influence interactions with **61**, verifying its binding to a remote site in BambL and Langerin.

SAR study of malonate 58 with LecA

PrOF NMR with 5FW-labeled lectin was used to gain information on the binding site of derivatives of the malonate **58**. Similar to hydroxamates, **58** and its analogues (**61-63, 66-69**) perturbed W42 (**Figure 4.4-8c**). To rank the impact of 13 analogues of **58** on W42 to prioritize scaffolds for future fragment evolution, we quantified the magnitude of W42 chemical shift perturbations (**Table 4.4-S4**, **Figure 4.4-S8d**). Interestingly, the scaffold **61** had the largest effect on W42, whereas an effect also on W33, but not any of the other tryptophanes, was observed (**Figure 4.4-S8**). The compounds with an acetic acid group (**63** and **70**) did not bind to LecA, which supports our docking result

proposing the different interactions of both carboxylates with the LecA surface. Altogether, this discovery demonstrates that malonates bind to the carbohydrate-binding site of LecA bearing one Ca^{2+} ion, similar to hydroxamates.

SAR study of malonate 58 with LecB

Docking analysis predicted that **58** targeted the carbohydrate-binding site of LecB. Therefore, we investigated the derivatives of **58** for binding to ^{15}N LecB by ^{15}N TROSY NMR (**Figures 4.4-S11b** and **4.4-S12**). We observed some resonances being perturbed in ^{15}N LecB with 2 mM **58** and its analogues (**59**, **62**, **64**, **66** and **67**) similar to 1 mM MeFuc (**Figures 4.4-3e**, **4.4-S12a** and **4.4-S13**). Next, we determined the affinities (K_d) and LE values of the malonic acid derivatives using ^1H-^{15}N TROSY NMR (**Figure 4.4-S14**). Interestingly, all structural derivatives showed comparable affinities and LE values for ^{15}N LecB. Notably, the compounds **58**, **62** and **67** showed higher affinities than **64**, indicating the role of an electronegative group in binding to LecB. This supports the docking pose 1 of the compound **58**, where LecB was predicted to interact with the CF_2 group of **58** through T98. Moreover, the methyl group introduced in **59** decreased its K_d and LE compared to **58**, suggesting that malonic moiety directly interacted with the protein surface and thus, a substituent in this position can disrupt this interaction. Therefore, this position is not suitable for fragment evolution. Notably, three compounds with a different scaffold (**66**, **67** and **69**) showed K_d values in a similar range. Their lower molecular weight and thus, better LE compared to **58**, could render them superior starting points for fragment growing. Thus, two main scaffold series have been identified for LecB: 1) **58**, **62** and **67**, and 2) **66**, **67** and **69**. Given the lack of ^{15}N LecB protein assignment, co-crystallization studies are currently ongoing to define the most potent scaffold for future fragment evolution.

SAR study of malonate 58 with DC-SIGN

Similar to LecA and LecB, acetylated compounds (**63**, **65** and **70**) did not bind ^{15}N DC-SIGN CRD (**Figure 4.4-S14**). However, malonates **58**, **62** and **67** perturbed the resonances in the EPN motif coordinating Ca_1^{2+} and D367 the strongest, whereas L321 and E324 near Ca_2^{2+} and Ca_3^{2+} showed weaker effects. Quantitative analyses of the conformational changes caused by **58** and D-mannose revealed similar CSPs (**Figure 4.4-S14a**). Encouraged by these results, we derived the affinities of **58** analogues in ^1H-^{15}N HSQC NMR (**Figure 4.4-S16**). Interestingly, all compounds showed a similar affinity and thus, three scaffold groups were defined as interchangeable (**58**, **62** and **69**). Similar to LecB, the compounds with an electronegative group on the ring being (**58**, **62, 63** and **67**) were predominant and

thus, in agreement with the predicted F313 interaction of CF_2 group in **58**. In contrast to LecB, a methyl group in **59** was well tolerated in DC-SIGN CRD. Since **59** did not interact with LecA either, this position is potentially suitable for future fragment growing to gain malonates specificity towards DC-SIGN. Together, both computational and experimental data demonstrated malonates' ability to target the Ca_1^{2+} binding site of DC-SIGN similarly to D-mannose.

7.4.2. Supplementary Materials and Methods

Chemicals

The carbohydrates D-mannose (CAS: 3458-28-4) and methyl-α-D-Galactose (MeGal, CAS: 3396-99-4) were purchased from Sigma-Aldrich Chemie GmbH, whereas methyl-α-L-Fucose (MeFuc, CAS: 14687-15-1) was from Biosynth-Carbosynth (UK). Commercial analogues (**Tables S1**, **3**-**5**, **9**-**35**, **46**-**50**) of hydroxamic and malonic acids (**Table S3**) were purchased from Life Chemicals Europe Gmbh (Germany), Otava Chemicals (Lithuania), KeyOrganics (UK).

Recombinant proteins

The non-labeled or ^{15}N-labeled LecA[8], LecB[9] and BambL[10] were purified in soluble form as reported previously. Human ^{15}N-labeled Langerin CRD and non-labeled ECD or DC-SIGN CRD and non-labeled ECD constructs were expressed in inclusion bodies and prepared as described before.[5,11]

Virtual screening of drug-like molecules targeting LecA

Virtual screening (VS) for commercially available drug-like compounds was performed using the Bioinfo database (http://bioinfo-pharma.u-strasbg.fr/bioinfo/), v.18.1, applying the following filters: hydrogen bond acceptors (HBA) >=2 and hydrogen bond donors (HBD) >=2. Filter (OpenEye) was used to remove compounds with poor pharmacokinetic profile and undesirable functional groups. Approximately 3 million molecules composed the VS library. LecA protein structure was retrieved from the Protein Data Bank (PDB): ID 1oko. One water molecule was conserved and kept on the active site. Other water molecules and heteroatoms were stripped from the structure leaving only the Ca^{2+} metal. Hydrogen atoms were added using PROTOSS.[12] Ligands were prepared for docking using Surflex[13] v. 3066, following the *pgeom* protocol (pose accuracy parameter set), 20 poses were generated for each ligand. All docked poses were rescored using the GRIM method[14], a knowledge based approach to score docking solutions by similarity of prediction interaction patterns to that already visited in the PDB. The similarity was quantified using a score (GRIM score or GrSc) set to 0.7 or higher. The interaction patterns found on the virtual screening were compared to selected complexes available for the co-crystallized molecules with LecA

(IDs:1oko, 2vxj, 2wyf, 3zyb, 3zyf, 3zyh, 4a6s, 4ljh, 4lk6, 4lk7, 4yw6, 5d21, 5mih, 4al9, 4cp9). The following additional filters were used: presence of metal interactions; absence of "carbohydrate-like scaffolds", in this case the tetrahydrofuran or tetrahydropyran cores; Surflex score >= 5; number of polar interactions (with binding site residues) >=3; number of rings >0. The remaining ligands were clustered based on the maximum common substructure (MCS) using Chemaxon (https://www.chemaxon.com) available at Pipeline Pilot[15] with a cutoff set to 8. Later, 46 chemically diverse molecules were selected as hits and 37 were purchased for testing.

Virtual screening of drug-like molecules targeting LecB

Virtual screening (VS) for commercially available drug-like compounds was performed with an updated version of Bioinfo database, v.18.2, following the same filtering criteria as described for LecA. Ligands were prepared as described in the above section. LecB protein structure was retrieved from the Protein Data Bank (PDB): ID 1gzt. In addition to the two conserved Ca^{2+} ions, a conserved water molecule, bridging interactions between α-L-fucose O1 and O2 atoms and residues Thr98 and Asp99, was preserved for docking calculations. In addition to 1gzt, 13 PDB structures were selected for calculation of the similarity interaction patterns using GRIM: 1gzt, 1ovp, 1ovs, 2boj, 2bp6, 2vuc, 2vud, 3dcq, 3zdv, 5a6x, 5may, 5maz, 5mb. All the co-crystallized ligands and other heteroatoms, with exception of Ca^{2+} and the conserved water were striped from the protein structures. Hydrogens were added using PROTOSS. Docking was performed using Surflex, with the same parameters described for LecA. The post-processing of the docked poses followed a workflow similar to the carbohydrate-binding site of LecA, with GrSc cutoff set to >=0.7, and clustering by MCS with same parameters described above. Other filters used included: number of polar interactions >=3; SurflexScore >=5; number of rings >0; number of rotatable bonds <=10; number of aromatic rings <5. We additionally used ROCS (DOI: 10.1021/jm0603365), with a Tanimoto shape-similarity cutoff of 0.6 in respect to fucose co-crystallized with LecB. A total of 42 molecules were selected as hits.

Preparation of fragment libraries

Here, we prepared three FBDD libraries (3F Fsp[3]-rich, general and MBP) composed of fluorinated (^{19}F) and non-fluorinated fragments. The 3F Fsp[3]-rich and general ^{19}F FBDD libraries were prepared as reported previously.[16-17] In addition to ^{19}F fragments, we used the non-fluorinated fragments from general library of 650 non-fluorinated fragments, which were prepared in mixtures of 10 fragments at 1 mM for screening. Following, we designed an MBP library, which is composed of 142 commercially

available fragments purchased from Otava Chemicals (Lithuania). For this, fragments were chosen from the 'Chelator Fragment Library' based on the presence of a fluorine atom and various chelating groups represented by picolinic acids, pyrimidines, hydroxypyrones, hydroxypyridinones, salicylic acids, hydroxamic acids, sulfonamides and beta-diketones. All compounds were subjected to a quality control for solubility and purity in ^{1}H and ^{19}F NMR, yielding 98 fluorinated and 9 non-fluorinated fragments. Following, we combined ^{19}F MBP fragments in mixtures of 32 fragments at 100 μM in a two-fold concentrated TBS buffer (25 mM in Tris-HCl pH 7.8, 150 mM NaCl) with 100 μM TFA and 20% D_2O) and stored at -20°C upon use.

^{1}H-^{15}N HSQC and TROSY NMR

To validate a fragment binding to Langerin CRD, DC-SIGN CRD, LecA and LecB, we ^{15}N-labeled proteins for ^{1}H-^{15}N HSQC and TROSY NMR. Both experiments deliver information on the binding site of compounds, whereas HSQC sequence was applied for Langerin CRD and DC-SIGN CRD. ^{1}H-^{15}N TROSY NMR sequence was used for ^{15}N LecA and LecB. Briefly, ^{1}H-^{15}N HSQC and TROSY NMR experiments were measured on a Bruker Ascend™700 (AvanceIII HD) spectrometer equipped with a 5 mm TCI700 CryoProbe™ in 3 mm tubes (Norell S-3-800-7) at 298 K and 310 K, respectively. A ^{1}H-^{15}N HSQC pulse sequence hsqcf3gpph19 with 128 increments and 8 scans per increment was applied for ^{15}N DC-SIGN/Langerin CRD. A ^{1}H-^{15}N TROSY pulse sequence trosyf3gpphsi19 with 128 increments and 32 scans per increment was applied for ^{15}N LecA/LecB. For ^{1}H-^{15}N HSQC NMR, ^{15}N-labeled Langerin CRD and DC-SIGN CRD were prepared at 0.1 mM in HBS buffer (20 mM HEPES pH 7.4, 150 mM NaCl) with 10 mM $CaCl_2$, 10% D_2O and 100 μM sodium trimethylsilylpropanesulfonate (DSS) as internal standard. Similarly, ^{1}H-^{15}N TROSY NMR spectra of ^{15}N-labeled LecA and LecB were acquired at 0.15 and 0.07 mM, respectively. For the titration experiments, we increased the concentration of $CaCl_2$ to 30 mM. To validate the fragment binding to proteins, we recorded ^{1}H-^{15}N HSQC and TROSY NMR of proteins in presence of DMSO as negative control and 2-4 mM fragments. Natural ligands were added to lectins as positive controls at 1 mM MeGal, MeFuc and 30 mM D-mannose to LecA/LecB and DC-SIGN CRD/Langerin CRD, respectively.

All data were processed with NMRpipe[18] and analyzed with CcpNmr analysis.[19] For data analysis, the protein fingerprint of ^{15}N DC-SIGN and Langerin CRD was assigned as reported previously (**Table 4.4-S7**).[5] The ^{1}H-^{15}N TROSY resonances of ^{15}N LecA and ^{15}N LecB were indexed with IDs due to a lack of protein backbone resonance assignment (**Tables 4.4-S8-S9**). Next, resonance IDs from protein spectra were transferred to the spectra obtained in the presence of compounds in order to compare

the changes in chemical shift perturbations for fast exchange peaks. Titration experiments with ^{15}N-labeled DC-SIGN CRD and LecB were recorded with 0.25 to 4-8 mM malonic acid derivatives. The K_d values were calculated according to the one−site−binding model in Origin(Pro) 2020b (OriginLab Corp., USA).

The changes in chemical shift perturbations (CSPs) were calculated according to Equation (1):

$$\Delta\delta = \sqrt{\frac{1}{2}[\Delta\delta_H^2 + (\alpha\Delta\delta_N)^2]} \ (1)$$

in which δ is the difference in chemical shift (in ppm) and α is an empirical weighting factor of 0.14 for all amino acid backbone resonances.[20] The threshold value was set based on three independent measurements of reference spectra to 0.01 ppm for LecA, LecB, DC-SIGN CRD and Langerin CRD.

Experimental fragment screening

Binding of fluorinated and non-fluorinated fragments was performed in ^{19}F NMR and ^{1}H-^{15}N HSQC/TROSY NMR, respectively. For ^{19}F NMR, we prepared two samples containing a fragment mixture alone and with 10 μM protein in TBS buffer. To detect the fragment binding to lectins in a Ca^{2+}-dependent manner, we recorded two samples in the presence of 5 mM EDTA subsequently adding 10 mM $CaCl_2$. The samples were recorded using ^{19}F NMR and [19]F T_2-filtered (CPMG) sequences on a Bruker Ascend™700 (AvanceIII HD) spectrometer equipped with a 5 mm TCI700 CryoProbe™ in 3 mm tubes (Norell S-3-800-7) at 298 K. Two separate ^{19}F spectra were recorded for CF_3 and CF groups with 32 and 64 scans, a spectral width of 100 ppm, a transmitter offset at -50 and -150 ppm, acquisition time of 2 s and 1 s relaxation time, respectively. T_2-filtered spectra were recorded using a CPMG pulse sequence with a 180° pulse repetition rate of 384 ms using same acquisition and relaxation times with 64 and 256 scans for CF_3 and CF compounds, respectively. Data was acquired without proton decoupling. These parameters were applied for all three FBDD libraries (3F Fsp[3]-rich, general and MBP). All ^{19}F NMR spectra were referenced to the internal standard trifluoroacetic acid (TFA) at -75.6 ppm and analyzed in MestReNova 11.0.0 (Mestrelab Research SL) for changes in peak intensity and chemical shift perturbations. Fragments binding in presence of protein and 10 mM $CaCl_2$ were used to derive a total hit rate. For ^{19}F spectra, chemical shift changes of 0.01 ppm or intensity changes between 25-50% were defined as 'high' and 'low' confidence hits, respectively. Intensity changes in the T_2-filtered spectra of 20-50% or more than 50% change were defined as 'high' and 'low' confidence hits, respectively. Afterwards, the fragments competed with the carbohydrates: 1 mM MeGal (LecA), 1

mM MeFuc (LecB/BambL) and 30 mM D-mannose (DC-SIGN/Langerin) were defined as targeting the carbohydrate-binding site and were followed up.

To screen the general and MBP-FBDD libraries of 650 and 9 non-fluorinated fragments, we performed ^{1}H-^{15}N HSQC or TROSY NMR. For screening of general FBDD library against LecA, a mix of 10 compounds at 1 mM each was combined with 0.15 mM ^{15}N LecA in HBS buffer with 10 mM CaCl$_2$, 10% D$_2$O and 100 µM DSS, whereas the fragments with the MBP scaffolds were validated separately at 1 mM. Similarly, 9 non-fluorinated fragments from the MBP-FBDD library were screened against ^{15}N-labeled DC-SIGN, LecA and LecB.

^{19}F NMR with reporters 5 and 61

To derive the Ca^{2+}-dependency and selectivity of hydroxamates and malonates, we used the analogues of **1** (**5**) and **58** (**61**) as fluorinated reporter molecules. For ^{19}F T$_2$-filtered NMR, we prepared two samples containing 0.1 mM **5** or **61** alone and in presence of a 10 µM protein in TBS buffer with 50 µM TFA and 10% D$_2$O. To demonstrate the Ca^{2+}-dependency of the protein—reporter interactions, we recorded the ^{19}F spectra in presence of 5 mM EDTA followed by addition of 10 mM CaCl$_2$. For the competitive ^{19}F T$_2$-filtered NMR, we added the carbohydrates: 1 mM MeGal (LecA), 1 mM MeFuc (LecB/BambL) and 30 mM D-mannose (DC-SIGN/Langerin) to compete the reporter molecules from the lectin orthosteric-binding site. Similarly, we used 0.1 mM **5** to prioritize hydroxamate derivatives. Here, we added 3 mM analogues of **1** in presence of 10 mM CaCl$_2$ instead of the carbohydrates.

Briefly, the ^{19}F T$_2$-filtered spectra were recorded with 64 scans, a spectral width of 5 ppm, a transmitter offset at -155.5 ppm (5) and -155.5 ppm (61), acquisition time of 0.8 s, 2 s relaxation time and a T$_2$-filter of 384 ms. All spectra were recorded at 298 K, referenced to TFA and analyzed in MestReNova 11.0.0 (Mestrelab Research SL). We considered reporter molecule binding to protein or competed if reduced signal intensity or a chemical shift change of 0.01 ppm in ^{19}F or T$_2$-filtered spectrum has been observed in presence of protein.

SPR

All experiments were performed on a BIACORE X100 instrument (GE Healthcare) at 25°C in phosphate-buffered saline (10 mM phosphate buffer pH 7.4, 2.7 mM KCl, 137 mM NaCl, 0.05% Tween 20, 100 µM CaCl$_2$, 5% DMSO). LecA was immobilized onto a CM7 chip (BIACORE) following standard amine coupling procedures: the CM7 chip was activated by three injections of a NHS/EDC mixture with a contact time of 540 s at a flow rate of 10 µL min^{-1} until the response exceeded 800 RU, followed by multiple injections of LecA dissolved in 10 mM sodium acetate pH 4.5 (100 µg/mL) onto channel 2 (contact time of 540 s at a flow rate of 10 µL min^{-1}). A minimum of 10 000

RU of LecA was captured onto the chip. Initial binding screens were performed with the injections of 0.2 and 1 mM of each hydroxamate (association 30 s, dissociation 60 s, 30 µL min[-1] flow rate), to identify the initial binders eliciting dose-response behaviors. Injection of the positive control (0.1 mM 4-nitrophenyl-β-D-galactoside) was included after every forth injection cycle to monitor the activity of immobilized LecA throughout the binding screen experiments. All data evaluation was performed using BIACORE X100 evaluation software (version 2.0).

Competitive binding fluorescence polarization (FP) assay

The FP assay was performed in a black 384-well plate (Greiner Bio-One, Germany, 781900) with the final volume of 20 µL as described previously.[21] Briefly, 10 µL of a sample solution series (10–0.078 mM) in TBS buffer containing 10 mM Ca^{2+} (TBS/Ca^{2+}-buffer) and 20% DMSO were added in technical triplicates to 10 µL LecA (40 µM) pre-incubated with a galactose-based Cy5 conjugate[22] (20 nM) in TBS/Ca^{2+}-buffer. Two positive controls were included: MeGal, IC_{50}=140±30 µM and GalNAc, IC_{50}=1230±200[23], and two negative control: TBS/Ca^{2+}-buffer containing 10% DMSO and LecA (20 µM) with the Cy5 dye (10 nM) in TBS/Ca^{2+}-buffer. The plate was sealed (EASYseal, Greiner Bio-One, 676001), centrifuged (1 min, 1500 rpm, 25°C) and incubated in a wet chamber for 16 h. The fluorescence was measured with an excitation 590 nm and emission 675 nm filter[22] on a PheraStar FS plate reader (BMG Labtech GmbH, Germany). The signal of TBS/Ca^{2+}-buffer containing 10% DMSO was subtracted and the compounds were analyzed with the MARS Data Analysis Software (BMG Labtech GmbH Germany) using the four-parameter variable slope model. The top and bottom plateaus were defined according to the positive controls. The graphs were visualized using Graphpad Prism 5. The percentage (%) of inhibition was calculated at the highest concentration compared to 10 mM MeGal. The pK_A values were derived using Advanced Chemistry Development (ACD/Labs) Software V11.02 (© 1994-2021 ACD/Labs), **2005**.

Protein-observed ^{19}F (PrOF) NMR

To validate binding of hydroxamic and malonic acids to the orthosteric site of LecA, we used protein-observed ^{19}F (PrOF) NMR with recombinant 5-fluorotryptophane (5FW)-labeled LecA as reported previously.[24] Briefly, 150 µM 5FW LecA was recorded alone and in presence of 2-4 mM compounds in TBS/Ca^{2+} buffer with 10% D_2O and 50 µM TFA as internal reference at 310 K. All experiments were conducted on Bruker Ascend[TM]700 (AvanceIII HD) spectrometer equipped with a 5 mm TCI700 CryoProbe[TM] in 3 mm tubes (Norell S−3−800−7) with following parameters: time domain of 1972, relaxation delay 1 s, acquisition time of 0.15 s, spectral width of 10 ppm and 1024 scans resulting in measurement time of 20 minutes pro spectrum. Data

analysis was performed in MestReNova 11.0.0 (Mestrelab Research SL) after applying the exponential function (50 Hz) and baseline correction.

The difference in chemical shift perturbations of W42 in 5FW LecA free vs compound-bound forms was followed to determine K_d values of compounds. The K_d values were calculated according to the one−site−binding model in Origin(Pro) 2020b (OriginLab Corp., USA) from two or three independent titrations.

Crystallography study

Lyophilized powder of recombinant LecA was dissolved in MilliQ water containing 1 mM $CaCl_2$ to the final protein concentration of 11.7 mg mL^{-1}. 100 mM hydroxamate 1 stock solution in DMSO was diluted in the LecA solution to the final concentration of 20 mM and incubated at 25°C to allow hydroxamate 1 to interact with LecA. 1.2 µL of the protein solution containing hydroxamate 1 was then mixed with 0.3 µL of LecA seed solution containing LecA microcrystals. 1.5 µL of reservoir solution (20% PEG6K, 1 M LiCl, 0.1 M sodium acetate pH 4) was added to the mixture. The entire mixture (total volume of 3 µL) was deposited on a siliconized glass slide. Crystallization was performed by the hanging drop vapor diffusion method on a 24-well plate with sealant (Hampton Research) at 19 °C. Crystals were cryo-protected in 30% PEG6000, 1 M LiCl, 100 mM sodium acetate pH 4, supplemented with 10 mM hydroxamate 1 and flash cooled in liquid nitrogen. X-ray diffraction data were collected at SOLEIL-PROXIMA2 (Saint Aubin, France) using ADSC Quantum 315r CCD detector. The recorded data were indexed, integrated, and scaled using XDS[25] and merged using AIMLESS[26]. The structures were solved by molecular replacement using 1OKO as a searching template in PHASER[27], followed by further iterations of manual rebuilding in COOT[28] and restrained refinement in REFMAC5[29]. Hydroxamic acid ligand was manually built in ACEDRG[30] in CCP4i2 suite[31]. The final model was validated with MOLPROBITY[32], PDB-redo (https://PDB-redo.eu/) and wwPDB validation service (http://validate-rcsb-1.wwPDB.org/) prior to submission to the Protein Data Bank. All structural figures were prepared using CCP4MG[33]. Data processing, refinement statistics and PDB ID of the deposited structure are provided in the Supporting Information (**Table 4.4-S2**).

CellFy

The cellFy experiments using Langerin$^+$ and DC-SIGN$^+$ Raji cell lines were performed as described before.[34] Briefly, 50k cells were plated in a 96-well plate (clear, round bottom; Greiner Bio-One and mixed with varying concentrations of malonates **58**, D-mannose and 0.025 mg mL^{-1} FITC-conjugated dextran (500 kDa, Sigma Aldrich) in a final volume of 50 µl following incubation for 30 min on ice. After centrifugation at 500 g for 3 min at 4°C, supernatant was discarded. After washing cells were treated with

50 µl 4% paraformaldehyde (Roti-Histofix, Carl Roth) for 20 min on ice and resuspended in 100 µl fresh culture medium. Fluorescence of cells was measured by flow cytometry (MACSQuant Analyzer 16). Data was analyzed in FlowJo.

Docking of malonic acid derivatives

Docking poses were obtained by docking compound 58 in MOE (Molecular Operating Environment (MOE), 2019.01; Chemical Computing Group ULC, 1010 Sherbooke St. West, Suite #910, Montreal, QC, Canada, H3A 2R7, 2021) using the Triangle Matcher and the Rigid Receptor as placement and refinement methods, respectively. For LecA and DC-SIGN CRD, only the highest-ranking pose was taken into account, whereas for LecB the best two ones were considered. Interaction maps were also retrieved from MOE.

Chemical synthesis of hydroxamate derivatives

All reactions were performed under inert gas (N_2) by using the Schlenk technique. The chemicals and solvents were bought from TCI, Merck or Roth and used without further purifications. The reactions were followed by either TLC (aluminum plates coated with silica gel 60, Merck KGaA, Damstadt, Germany) by using molybdenum-stain (0.02 M solution of $Ce(NH_4)_4(SO_4)_4 \cdot 2H_2O$ and $(NH_4)_6Mo_7O_{24} \cdot 4H_2O$ in aqueous 10% H_2SO_4) or by HPLC-MS (Thermo Dionex Ultimate 3000 HPLC coupled to a Bruker amaZon SL mass spectrometer, with UV detection at 254 nm, using a C18 column (100/2 Nucleoshell RP18plus, 2.7 µM from Macherey-Nagel, Germany). The crude products were purified by MPLC (Teledyne Isco Combiflash Rf200) by using self-packed silica gel 60 columns (60Å, 400 mesh particle size, Fluka) as a stationary phase or by HPLC (Waters 2545 Binary Gradient Module with Waters 2489 UV/Visible detector) using RP-18 column (250/21 Nucleodur C18 Gravity SB, 5 µM from Macherey-Nagel, Germany). The synthesized compounds were analyzed by NMR spectroscopy (Bruker Avance III 500 ultra shield spectrometer) at 500 MHz (^{1}H) or 126 MHz (^{13}C) using deuterated solvents (Eurisotop, Saarbrücken, Germany) and analyzed with MestReNova (Version 12.0.2). Chemical shifts are given in parts per million compared to an internal solvent peak (MeOH-d_4 = 3.31, 49.00 ppm, DMSO-d_6 = 2.50, 39.52 ppm)[35] and the multiplicities as s (singlet), d (doublet), t (triplet), q (quartet) and m (multiplet).

N,3-dihydroxy-3-(3-methoxyphenyl)propenamide (31)

The previously reported procedure of *N*,3-dihydroxy-3-(3-methoxyphenyl)-propenamide (**31**) was slightly modified and optimized.[36] A: Zinc was activated by washing in a 1N HCl bath for 30 min, then filtrated washed with H_2O, EtOH, Et_2O and dried on high vacuum for 20 min. B: The activated zinc powder (653.8 mg, 10 mmol,

2.0 eq.) was dissolved in dry THF (8 mL) and refluxed for 5 min. The heating was stopped and a solution of ethyl bromoacetate (1 g, 6 mmol, 1.2 eq.) and *m*-anisaldehyde (680.75 mg, 5 mmol, 1.0 eq.) in dry THF (5 mL) was carefully added over 30 min (syringe pump). Then, the mixture was refluxed for 16 h. The reaction was stopped by adding NH_4Cl (20 mL) and stirred for additional 15 min. The phases were separated and the aq. phase extracted with EtOAc (3x). The combined org. phase was dried over Na_2SO_4, concentrated, purified by flash chromatography (PE:EtOAc - 8:2) and re-purified with Tol:EtOAc - 9:1. The ester intermediate was obtained as an clear oil (660 mg, 2.94 mmol, 59%, R_f = 0.35, PE: EtOAc – 8:2). C: Hydroxylamine hydrochloride (216.65 mg, 1.66 eq) was dissolved in MeOH (2.76 mL) and a NaOMe solution (5.66 M, 0.98 mL, 3.33 eq.) was added. The milky mixture was stirred for 15 min and the previously synthesized ester (421.2 mg, 1.878 mmol, 1.0 eq) in MeOH (1 mL, c_{final} = 0.5M) was added carefully. After 5 h, the reaction was stopped by adjusting the pH to 7 with 1 M HCl (ca. 20 drops) and H_2O was added. The mixture was diluted with H_2O and the aq. phase was extracted with EtOAc (3x). The combined org. phases were dried over Na_2SO_4 and the product was purified by flash chromatography (PE: EE, R_f = 0.4 pure EtOAc). The product was obtained as a clear oil which solidifies on high vacuum (128.1 mg, 0.50 mmol, 30%, Keto E/Z = 89:10). ^{1}H NMR (500 MHz, DMSO-d_6) δ 10.33 (s, 1H, Keto-E O<u>H</u>), 9.86 (s, 0H, Keto-Z O<u>H</u>), 9.07 (s, 0H, Keto-Z N<u>H</u>), 8.73 (s, 1H, Keto-E N<u>H</u>), 7.22 (t, *J* = 8.0 Hz, 1H, Ar<u>H</u>), 6.92 – 6.86 (m, 2H), Ar<u>H</u> 2x), 6.79 (ddd, *J* = 8.2, 2.6, 1.0 Hz, 1H, Ar<u>H</u>), 5.37 (d, *J* = 4.6 Hz, 1H, CHO<u>H</u>), 4.91 (dt, *J* = 8.6, 4.4 Hz, 1H, C<u>H</u>OH), 3.74 (s, 3H, OC<u>H</u>$_3$), 2.33 – 2.20 (m, 2H, C<u>H</u>$_2$). ^{13}C NMR (126 MHz, DMSO) δ 166.85 (<u>C</u>=O), 159.13 (Ar<u>C</u>), 147.07 (Ar<u>C</u>), 129.12 (Ar<u>C</u>H), 117.91 (Ar<u>C</u>H), 112.29 (Ar<u>C</u>H), 111.21 (Ar<u>C</u>H), 69.29 (<u>C</u>H), 54.93 (O<u>C</u>H$_3$), 42.76 (<u>C</u>H$_2$). LR-MS calcd $[C_{10}H_{14}NO_4]^+$: 212.09, found 212.17. The spectroscopic data differing slightly from the literature.[36]

*2-Cyclohexyl-N-hydroxyacetamide (**6**)*

The synthesis of 2-cyclohexyl-*N*-hydroxyacetamide (6) was performed as described before by Ohtsuka et al.[37] Solutions of $H_2NOH \cdot HCl$ (444.74 mg, 6.4 mmol) in dry MeOH (3.2 mL, *c* = 2 M) and KOH (71.82 mg, 0.4 eq.) in MeOH (3.2 mL, c = 0.4) were prepared. Both solutions were cooled to 0 °C, then, the alkali solution was added to the stirred hydroxylamine and the resulting suspension was left without stirring for 5 min. The white precipitate (KCl) was removed by suction filtration and the clear filtrate was added to methyl cyclohexylacetate (500 mg, 3.20 mmol). KOH (2 pellets) was added until pH = 10 was reached and the mixture was stirred overnight. The solvent was removed and the crude was diluted in H_2O leading to white precipitate. The

solution was acidified pH < 4 and the product was filtrated, resulting in a white solid (144 mg, 0.917 mmol, 29%, E/Z = 88:8, CH_2Cl_2/ MeOH – 19:1, R_f = 0.57). ^{1}H NMR (500 MHz, DMSO-d_6) δ 10.30 (br s, 1H, Keto-E O$\underline{H}$), 9.72 (s, 0H, Keto-Z OH), 8.96 (s, 0H, Keto-Z N$\underline{H}$) 8.65 (d, J = 1.7 Hz, 1H, Keto-E N$\underline{H}$), 1.81 (d, J = 6.8 Hz, 2H, C$\underline{H}_2$), 1.72 – 1.54 (m, 6H, C$\underline{H}_2$ 2.5x, C$\underline{H}$), 1.26 – 1.04 (m, 3H, C$\underline{H}_2$ 1.5x), 0.95 – 0.82 (m, 2H, C$\underline{H}_2$). ^{13}C NMR (126 MHz, DMSO) δ 168.19 ($\underline{C}$=O), 39.78 ($\underline{C}H_2$), 34.45 ($\underline{C}H$), 32.47 ($\underline{C}H_2$ 2x), 25.83 ($\underline{C}H_2$), 25.57 ($\underline{C}H_2$ 2x). HR-MS calcd $[C_8H_{16}NO_2]^+$: 158.1176, found 158.1174. The spectroscopic data matching the literature.[38]

*N-methoxy-2-phenylacetamide (**7**)*

N-methoxy-2-phenylacetamide (**7**) was synthesized according to the previously reported procedure by Kawase.[39] Methoxyamine hydrochloride (445.7 mg, 5.3 mmol, 1.1 eq.) and Na_2CO_3 (1028 mg, 9.7 mmol, 2.0 eq.) were dissolved in a mixture of toluene and H_2O (1:1, c ≈ 0.25 M) and cooled to 0 °C. Phenylacetyl chloride (641 µL, 4.8 mmol, 1.0 eq.) was added and the reaction was stirred for 20 h. The reaction was taken up in EtOAc, separated, and the org. phase was washed with brine and dried over Na_2SO_4. The crude product was purified by flash chromatography (PE:EE – 1:2, R_f = 0.3, $KMnO_4$) and the compound **7** was obtained as a white solid (487 mg, 2.95 mmol, 61%). ^{1}H NMR (500 MHz, DMSO-d_6) δ 11.25 (s, 1H, N$\underline{H}$), 7.34 – 7.27 (m, 2H, Ar$\underline{H}$ 2x), 7.27 – 7.20 (m, 3H, Ar$\underline{H}$ 3x), 3.57 (s, 3H, OC$\underline{H}_3$), 3.28 (s, 2H, C$\underline{H}_2$). ^{13}C NMR (126 MHz, DMSO) δ 166.91 ($\underline{C}$=O), 135.52 (Ar$\underline{C}$), 128.93 (Ar$\underline{C}H$ 2x), 128.30 (Ar$\underline{C}H$ 2x), 126.56 (Ar$\underline{C}H$), 63.20 (O$\underline{C}H_3$), 39.25 ($\underline{C}H_2$). HR-MS calcd $[C_9H_{12}NO_2]^+$: 166.0863, found 166.0860. The spectroscopic data are in accordance with the literature.[40]

*N-hydroxy-N-methyl-2-phenylacetamide (**8**)*

N-hydroxy-*N*-methyl-2-phenylacetamide (**8**) was synthesized by following the produce of Clark et al.[41] *N*-Methylhydroxylamine hydrochloride (500 mg, 5.98 mmol, 1.0 eq.) was dissolved in CH_2Cl_2 (*c* = 0.5 M) and cooled to 0 °C. Et_3N (1.66 mL, 2.0 eq.) was added and the mixture was stirred for 10 min, followed by phenylacetyl chloride drop-wise (≈ 797 µL, 5.98 mmol, 1.0 eq.). The mixture was stirred at room temperature for 16 h, then washed with 1 N HCl and brine. The org. phase was dried over Na_2SO_4, concentrated under vaccum and the yellow crude was purified by flash chromatography (PE: EtOAc – 1:2, R_f = 0.43, Mo-stain). The product was obtained as a colorless oil (314.7 mg, 1.90 mmol, 32%). ^{1}H NMR (500 MHz, DMSO-d_6) δ 9.98 (s, 1H, N$\underline{H}$), 7.34 – 7.25 (m, 2H, Ar$\underline{H}$ 2x), 7.25 – 7.18 (m, 3H, Ar$\underline{H}$ 3x), 3.69 (s, 2H C$\underline{H}_2$), 3.10 (s, 3H, C$\underline{H}_3$). ^{13}C NMR (126 MHz, DMSO) δ 170.77 ($\underline{C}$=O), 135.92 (Ar$\underline{C}$), 129.43 (Ar$\underline{C}H$ 2x), 128.11 (Ar$\underline{C}H$ 2x), 126.24 (Ar$\underline{C}H$), 38.32 ($\underline{C}H_2$), 35.80 ($\underline{C}H_3$). HR-MS calcd [C9H12NO2]+: 166.0863, found 166.0861.

General procedure for 35 to 44

The different hydroxamic acids **35** to **44** were synthesized according to a modified procedure of Trabulsi *et al.* starting from an acid moiety.[42] The corresponding acid (1eq.) was dissolved in dry CH_2Cl_2 (*c* = 0.25 M) and cooled to 0 °C. Oxalyl chloride (1.75 eq.) was added dropwise and the mixture was allowed to warm to room temperature and stirred for 16 h (yellowish). The solvent was removed und reduced pressure and an equimolar solution of hydroxyl amine (7 eq.) and NaOH (7 eq.) in H_2O (*c* = 2.5 M) was added. The mixture was left without stirring for 15 min (until the formed fog vanished) and the mixture was diluted with EtOAc, separated and the aq. phase was extracted with EtOAc (3x). The combined org. phases were dried over Na_2SO_4, concentrated under reduced pressure and purified by flash chromatography (CH_2Cl_2 MeOH – 19:1) resulting in white solids.

N-hydroxy-4-phenylbut-3-enamide (42)

N-hydroxy-4-phenylbut-3-enamide (**42**) was synthesized according to the general procedure starting from *trans*-styrylacetic acid (200 mg, 1.23 mmol, 1.0 eq.). A white solid was obtained (80 mg, 0.49 mmol, 40%, CH_2Cl_2/ MeOH – 19:1, R_f = 0.49, Keto-E/Z: 89:11). ^1H NMR (500 MHz, DMSO-d_6) δ 10.51 (s, 1H, Keto-E O$\underline{H}$), 9.94 (s, 0H, Keto-Z O$\underline{H}$), 9.19 (s, 0H, Keto-Z N$\underline{H}$), 8.79 (s, 1H, Keto-E N$\underline{H}$), 7.43 – 7.37 (m, 2H, ArC$\underline{H}$ 2x), 7.32 (dd, *J* = 8.5, 6.9 Hz, 2H, ArC$\underline{H}$ 2x), 7.26 – 7.20 (m, 1H, ArC$\underline{H}$), 6.51 – 6.43 (m, 1H, C$\underline{H}$CHCH$_2$), 6.29 (dt, *J* = 15.9, 7.1 Hz, 1H, CHC$\underline{H}$CH$_2$), 2.92 (dd, *J* = 7.1, 1.5 Hz, 2H, CHCHC$\underline{H}_2$). ^{13}C NMR (126 MHz, DMSO) δ 166.93 ($\underline{C}$=O), 136.77 (Ar$\underline{C}$), 132.09 (C$\underline{H}$CHCH$_2$), 128.62 (Ar$\underline{C}$H 2x), 127.33 (Ar$\underline{C}$H), 125.98 (Ar$\underline{C}$H 2x), 124.07 (CHC$\underline{H}$CH$_2$), 36.84 (CHCHC$\underline{H}_2$). HR-MS calcd $[C_{10}H_{12}NO_2]^+$: 178.0863, found 178.0860.

N-hydroxy-2-phenethoxyacetamide (43)

The product was obtained from 2-phenethoxyacetic acid (100 mg, 0.55 mmol, 1.0 eq.) as a white solid (38.7 mg, 0.198 mmol, 36%, CH_2Cl_2/MeOH – 19:1, R_f = 0.575, Keto-E/Z = 92:7). ^1H NMR (500 MHz, DMSO-d_6) δ 10.48 (s, 1H, Keto-E O$\underline{H}$), 9.96 (s, 0H, Keto-Z O$\underline{H}$), 9.00 (s, 0H, Keto-Z N$\underline{H}$), 8.83 (s, 1H, Keto-E N$\underline{H}$), 7.32 – 7.22 (m, 4H, Ar$\underline{H}$ 4x), 7.22 – 7.17 (m, 1H, Ar$\underline{H}$), 3.84 (s, 2H, C$\underline{H}_2$), 3.62 (t, *J* = 7.0 Hz, PhCH$_2$C$\underline{H}_2$O), 2.83 (t, *J* = 7.0 Hz, 1H, PhC$\underline{H}_2$CH$_2$O). ^{13}C NMR (126 MHz, DMSO) δ 165.59 ($\underline{C}$=O), 138.77 (Ar$\underline{C}$), 128.86 (Ar$\underline{C}$H 2x), 128.25 (Ar$\underline{C}$H 2x), 126.09 (Ar$\underline{C}$H), 71.73 (PhCH$_2$$\underline{C}H_2$O), 68.68 ($\underline{C}H_2$), 35.24 (PhC$\underline{H}_2CH_2$O). HR-MS calcd $[C_{10}H_{14}NO_3]^+$: 196.0968, found 196.0967.

*Nhydroxy-2-(3-phenylpropoxy)acetamide (**44**)*

The desired compound was obtained from 2-(3-phenylpropoxy)acetic acid (100 mg, 0.51 mmol, 1.0 eq.) as a white solid (35 mg, 0.18 mmol, 35%, CH_2Cl_2/ MeOH - 19:1, R_f = 0.85, Keto-E/Z = 93:9). ^{1}H NMR (500 MHz, DMSO-d_6) δ 10.50 (s, 1H, Keto-E O$\underline{H}$), 9.95 (s, 0H, Keto-Z O$\underline{H}$), 9.00 (s, 0H, Keto-Z N$\underline{H}$), 8.82 (s, 1H, Keto-E N$\underline{H}$), 7.27 (t, J = 7.5 Hz, 2H, ArC$\underline{H}$), 7.24 – 7.14 (m, 3H, ArC$\underline{H}$), 3.81 (s, 2H, C$\underline{H}_2$), 3.41 (t, J = 6.4 Hz, 2H, PhCH$_2$CH$_2$C$\underline{H}_2$O), 2.69 – 2.57 (m, 2H, PhC$\underline{H}_2$CH$_2$CH$_2$O), 1.89 – 1.76 (m, 2H, PhCH$_2$C$\underline{H}_2$CH$_2$O). ^{13}C NMR (126 MHz, DMSO) δ 165.65 (C=O), 141.79 (Ar$\underline{C}$), 128.33 (Ar$\underline{C}$H 2x), 128.29 (Ar$\underline{C}$H 2x), 125.72 (Ar$\underline{C}$H), 70.12 (PhCH$_2$CH$_2$$\underline{C}H_2$O), 68.77 ($\underline{C}H_2$), 31.56 (Ph$\underline{C}H_2CH_2CH_2$O), 30.80 (PhCH$_2$$\underline{C}H_2CH_2$O). HR-MS calcd [$C_{11}H_{16}NO_3$]$^+$: 210.1125, found 210.1121.

*N-hydroxy-4-phenylbutanamide (**35**)*

This product was obtained from 4-phenylbutric acid (200 mg, 1.22 mmol) as a white solid (131 mg, 0.73 mmol, 60%, CH_2Cl_2/MeOH – 19:1, R_f = 0.45, Keto-E/Z 85:14). ^{1}H NMR (500 MHz, DMSO-d_6) δ 10.35 (s, 1H, NHOH), 9.76 (s, 0H, Keto-Z O$\underline{H}$, 8.99 (s, 0H, Keto-Z N$\underline{H}$) 8.68 (s, 1H, NHOH), 7.28 (dd, J = 8.2, 6.9 Hz, 2H, ArH (3x)), 7.23 – 7.15 (m, 3H, ArH (3x)), 2.54 (t, J = 7.7 Hz, 2H, PhC$\underline{H}_2$CH$_2$CH$_2$CO), 1.96 (t, J = 7.4 Hz, 2H, PhCH$_2$CH$_2$C$\underline{H}_2$CO), 1.82 – 1.73 (m, 2H, PhCH$_2$C$\underline{H}_2$CH$_2$CO). ^{13}C NMR (126 MHz, DMSO) δ 168.86 ($\underline{C}$=O), 141.66 (ArC), 128.31 (Ar$\underline{C}$H, 4x), 125.79 (Ar$\underline{C}$H 1x), 34.60 (Ph$\underline{C}$H$_2$CH$_2$CH$_2$CO), 31.78 (PhCH$_2$CH$_2$$\underline{C}H_2$CO), 26.99 (PhCH$_2$$\underline{C}H_2CH_2$CO). HR-MS calcd [$C_{10}H_{14}NO_2$]$^+$: 180.1019, found 180.1017. The spectroscopic data are in accordance with the literature.[38]

*N-hydroxy-4-(p-tolyl)butanamide (**36**)*

The desired compound was obtained from 4-(*p*-tolyl)butanoic acid (200 mg, 1.12 mmol) after purification with CH_2Cl_2/MeOH – 19:1 (R_f = 0.22) as a white solid (177.5 mg, 0.92 mmol, 82%, E/Z= 85:11). The compound was further purified by prep-HPLC (CH_3CN/H_2O) for biological tests. ^{1}H NMR (500 MHz, DMSO-d_6) δ 10.34 (s, 1H, Keto-E O$\underline{H}$), 9.75 (s, 0H, Keto-Z-OH), 8.68 (s, 1H, Keto-E N$\underline{H}$), 7.11 – 7.04 (m, 4H, ArC$\underline{H}$(4x)), 2.52 – 2.50 (m, 2H, PhC$\underline{H}_2$CH$_2$CH$_2$CO), 2.26 (s, 1H, PhCH$_2$CH$_2$C$\underline{H}_2$CO), 1.95 (t, J = 7.5 Hz, 2H), 1.79 – 1.70 (m, 2H, PhCH$_2$C$\underline{H}_2$CH$_2$CO). ^{13}C NMR (126 MHz, DMSO) δ 169.36 (C=O), 138.98 (ArC), 135.10 (ArC), 129.34 ArCH (2x)), 128.65 (ArCH (2x)), 34.64 (Ph$\underline{C}$H$_2$CH$_2$CH$_2$CO), 32.23 (PhCH$_2$CH$_2$$\underline{C}H_2$CO), 27.53 (PhCH$_2$$\underline{C}H_2CH_2$CO), 21.09 (ArC$\underline{H}_3$). HR-MS calcd [$C_{11}H_{16}NO_2$]$^+$: 194.1176, found 194.1174.

N-hydroxy-4-(thiophen-2-yl)butanamide (37)

This compound was obtained from 4-(thiophen-2-yl)butanoic acid (200 mg, 1.18 mmol) after purification with CH_2Cl_2/MeOH – 19:1 (R_f = 0.50) as a slightly orange solid (62.2 mg, 0.33 mmol, 29%, E/Z = 84:16). ^{1}H NMR (500 MHz, DMSO-d_6) δ 10.37 (s, 1H, Keto-E O$\underline{H}$), ~~9.79~~ (s, 0H, Keto-Z OH), 9.01 (s, 0H, Keto-Z NH), 8.69 (s, 1H, Keto-E N$\underline{H}$), 7.32 (d, J = 1.2 Hz, 1H, thiophene-H), 6.93 (d, J = 5.2 Hz, 1H, thiophene-H), 6.84 (d, J = 1.1 Hz, 1H, thiophene-H), 2.77 (s, 2H, thiopheneC$\underline{H}_2$CH$_2$CH$_2$CO), 2.00 (d, J = 7.4 Hz, 2H, thiopheneCH$_2$CH$_2$C$\underline{H}_2$CO), 1.82 (s, 2H, thiopheneCH$_2$C$\underline{H}_2$CH$_2$CO). ^{13}C NMR (126 MHz, DMSO) δ 169.13 ($\underline{C}$=O), 144.65 (thiophene-$\underline{C}$), 127.40 thiophene-$\underline{C}$H, 125.01 thiophene-$\underline{C}$H, 124.03 thiophene-$\underline{C}$H, 32.00 (thiophene$\underline{C}$H$_2$CH$_2$CH$_2$CO), 29.10 (thiopheneCH$_2$CH$_2$$\underline{C}H_2$CO), 27.83 (thiopheneCH$_2$$\underline{C}H_2CH_2$CO). HR-MS calcd [$C_8H_{12}NO_2S$]$^+$: 186.0583, found 186.0580.

4-(4-bromophenyl)-N-hydroxybutanamide (38)

This compounds was obtained from 4-(4-bromophenyl)butanoic (200 mg, 0.82 mmol) after purification with CH_2Cl_2/MeOH – 19:1 (R_f = 0.19) as a white solid (100.8 mg, 0.39 mmol, 48%, E/Z= 87:13) and was further purified by prep-HPLC (CH_3CN/H_2O) for biological tests. ^{1}H NMR (500 MHz, DMSO-d_6) δ 10.35 (s, 1H, Keto-E O$\underline{H}$), 9.77 (s, 0H, Keto-Z OH), 8.98 (s, 0H, Keto-Z NH), 8.69 (s, 1H, Keto-E N$\underline{H}$), 7.57 – 7.29 (m, 2H (ArC$\underline{H}$ (2x)), 7.29 – 6.90 (m, 2H, ArC$\underline{H}$ (2x)), 2.53 (d, J = 7.6 Hz, 2H, PhC$\underline{H}_2$CH$_2$CH$_2$CO), 1.94 (t, J = 7.4 Hz, 2H, PhCH$_2$CH$_2$C$\underline{H}_2$CO), 1.86 – 1.53 (m, 2H, PhCH$_2$C$\underline{H}_2$CH$_2$CO). ^{13}C NMR (126 MHz, DMSO) δ 168.74 (C=O, 141.09 (Ar$\underline{C}$), 131.14 (Ar$\underline{C}$H), 130.63 (Ar$\underline{C}$H), 118.82 (Ar$\underline{C}$), 33.84 (Ph$\underline{C}$H$_2$CH$_2$CH$_2$CO), 31.59 (PhCH$_2$CH$_2$$\underline{C}H_2$CO), 26.71 (PhCH$_2$$\underline{C}H_2CH_2$CO). HR-MS calcd [$C_{10}H_{13}BrNO_2$]$^+$: 258.0124, found 258.0119.

N-hydroxy-4-(4-nitrophenyl)butanamide (39)

The product was obtained from 4-(4-nitrophenyl)butanoic acid (200 mg, 0.96) after purification with CH_2Cl_2/MeOH – 19:1 (R_f = 0.19) as a white solid (166 mg, 0.74 mmol, 77%, E/Z= 87:13). This compound was further purified for biological tests by prep-HPLC (CH_3CN/ H_2O). ^{1}H NMR (500 MHz, DMSO-d_6) δ 10.55 – 10.18 (m, 1H, Keto-E O$\underline{H}$), 9.81 (s, 0H, Keto-Z OH), 9.01 (s, 0H, Keto-Z NH), 8.71 (s, 1H, Keto-E N$\underline{H}$), 8.15 (d, J = 8.5 Hz, 2H, ArC$\underline{H}$ (2x)), 7.48 (d, J = 8.3 Hz, 2H, (ArC$\underline{H}$ (2x)), 2.70 (t, J = 7.7 Hz, 2H, PhC$\underline{H}_2$CH$_2$CH$_2$CO), 1.98 (t, J = 7.4 Hz, 2H, PhCH$_2$CH$_2$C$\underline{H}_2$CO), 1.83 (t, J = 7.3 Hz, 2H, PhCH$_2$C$\underline{H}_2$CH$_2$CO). ^{13}C NMR (126 MHz, DMSO) δ 168.61 ($\underline{C}$=O), 150.18 (Ar$\underline{C}$), 145.89 (Ar$\underline{C}$), 129.65 (Ar$\underline{C}$H (2x)), 123.48 (Ar$\underline{C}$H (2x)), 34.29 (Ph$\underline{C}$H$_2$CH$_2$CH$_2$CO), 31.55 (PhCH$_2$CH$_2$$\underline{C}H_2$CO), 26.39 (PhCH$_2$$\underline{C}H_2CH_2$CO). HR-MS calcd [$C_{10}H_{13}N_2O_4$]$^+$: 225.0870, found 225.0865.

N-hydroxy-4-(4-methoxyphenyl)butanamide (40)

The product was obtained from 4-(4-methoxyphenyl)butanoic (200 mg, 1.030 mmol) as a white solid (81 mg, 0.387 mmol, 38%, E/Z= 83:17, R_f = 0.50, CH_2Cl_2/MeOH – 19:1). ^{1}H NMR (500 MHz, DMSO-d_6) δ 10.34 (s, 1H, Keto-E O$\underline{H}$), 9.75 (s, 0H Keto-Z OH), 8.98 (s, 0H, Keto-Z NH), 8.67 (s, 1H, Keto-E N$\underline{H}$), 7.09 (d, *J* = 8.5 Hz, 2H, ArCH (2x)), 6.84 (d, *J* = 8.6 Hz, 2H, ArCH (2x)), 3.71 (s, 3H, OC$\underline{H}_3$), 2.47 (d, *J* = 7.8 Hz, 2H, PhC$\underline{H}_2$CH$_2$CH$_2$CO), 1.94 (t, *J* = 7.5 Hz, 2H, PhCH$_2$CH$_2$C$\underline{H}_2$CO), 1.86 – 1.61 (m, 2H, PhCH$_2$C$\underline{H}_2$CH$_2$CO). ^{13}C NMR (126 MHz, DMSO) δ 168.91 ($\underline{C}$=O), 157.42 (Ar$\underline{C}$), 133.49 (Ar$\underline{C}$), 129.21 (Ar$\underline{C}$H (2x)), 113.71 (Ar$\underline{C}$H (2x)), 54.96 (O$\underline{C}$H$_3$), 33.70 (Ph$\underline{C}$H$_2$CH$_2$CH$_2$CO), 31.73 (PhCH$_2$CH$_2$$\underline{C}H_2$CO), 27.23 (PhCH$_2$$\underline{C}H_2CH_2$CO). HR-MS calcd [$C_{11}H_16NO_3$]$^+$: 210.1125, found 210.1121.

N-hydroxy-4-(4-hydroxyphenyl)butanamide (41)

The reaction was performed in analogy to a previously reported procedure.[43] *N*-hydroxy-4-(4-methoxyphenyl)butanamide (15.8 mg, 0.076 mmol, 1 eq.) was dissolved in dry CH_2Cl_2 and cooled to -78 °C. BBr$_3$ (380 µL, 0.38 mmol, 5 eq.) was added and the reaction was warmed to room temperature and stirred for 3 days (CH_2Cl_2/MeOH – 19:1 R_f = 0.28). The reaction was taken up in KHSO$_4$ and extracted with EtOAc (3x). The combined org. phase was dried over Na$_2$SO$_4$ and concentrated under reduced pressure. The crude product was purified by prep-HPLC (CH$_3$CN/ H$_2$O) to obtain the product as a white solid (3.2 mg, 0.016 mmol, 22%, E/Z= 84:16). ^{1}H NMR (500 MHz, DMSO-d_6) δ 10.33 (s, 1H, Keto-E O$\underline{H}$), 9.74 (s, 0H, Keto-Z OH), 9.12 (s, 1H, ArO$\underline{H}$), 8.97 (s, 0H, Keto-Z-NH), 8.66 (s, 1H, Keto-E N$\underline{H}$), 6.95 (d, *J* = 8.4 Hz, 2H, ArC$\underline{H}$ (2x)), 6.66 (d, *J* = 8.4 Hz, 2H, ArC$\underline{H}$ (2x)), 2.42 (t, *J* = 7.6 Hz, 2H, PhC$\underline{H}_2$CH$_2$CH$_2$CO), 1.93 (t, *J* = 7.5 Hz, 2H, PhCH$_2$CH$_2$C$\underline{H}_2$CO), 1.76 – 1.67 (m, 2H, PhCH$_2$C$\underline{H}_2$CH$_2$CO). ^{13}C NMR (126 MHz, DMSO) δ 168.97 (C=O), 155.35 (Ar$\underline{C}$), 131.67 (Ar$\underline{C}$), 129.12 (Ar$\underline{C}$H (2x)), 115.04 (Ar$\underline{C}$H (2x)), 33.78 (Ph$\underline{C}$H$_2$CH$_2$CH$_2$CO), 31.77 (PhCH$_2$CH$_2$$\underline{C}H_2$CO), 27.31 (PhCH$_2$$\underline{C}H_2CH_2$CO). HR-MS calcd [$C_{10}H_{14}NO_3$]$^+$: 196.0968, found 196.0965.

7.4.3. Supplementary Figures
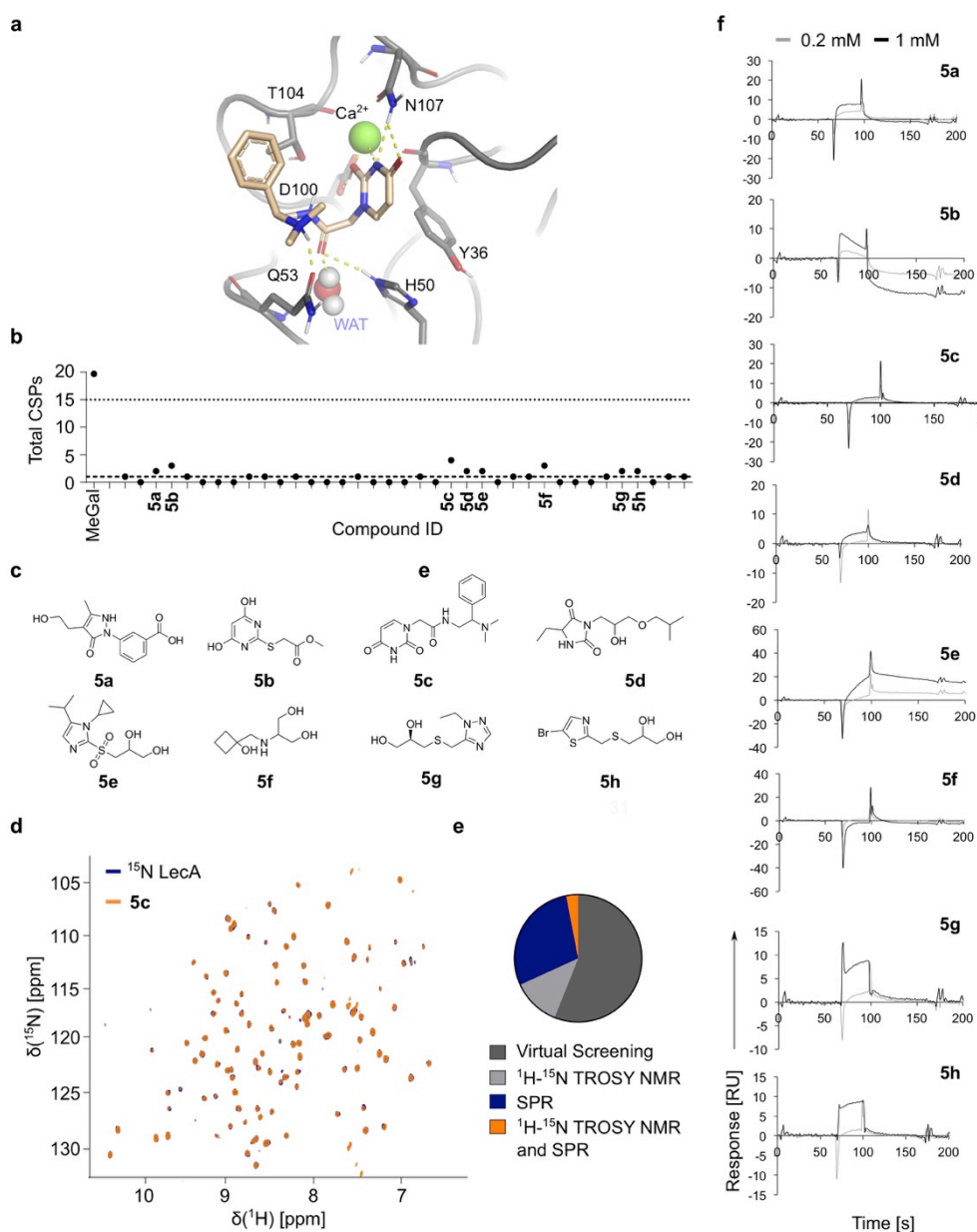
Figure 4.4-S1 Virtual screening of fragment and drug-like libraries for PA-IL (LecA).

a Shown is the binding mode of compound **5c** to the carbohydrate-binding pocket of LecA. **b** Virtual screening hits were validated in ^{1}H-^{15}N TROSY NMR. The plot shows the total number of chemical shift perturbations (CSPs) that were promoted in ^{15}N LecA in presence of 2 mM hits, whereas the structures of fragments, which exhibited binding similarly to MeGal above the threshold (*dashed line*) are shown in **c**. **d** ^{1}H-^{15}N TROSY NMR spectrum of 0.15 mM ^{15}N LecA in presence of 2 mM **5c** shows only small changes in protein upon fragment addition. **e** Overview of 36 LecA virtual screening hits, whereas 8 were tested positive in ^{1}H-^{15}N TROSY NMR, 19 in SPR and 2 in both assays. **f** SPR analysis of the most promising hits identified in ^{1}H-^{15}N TROSY NMR.
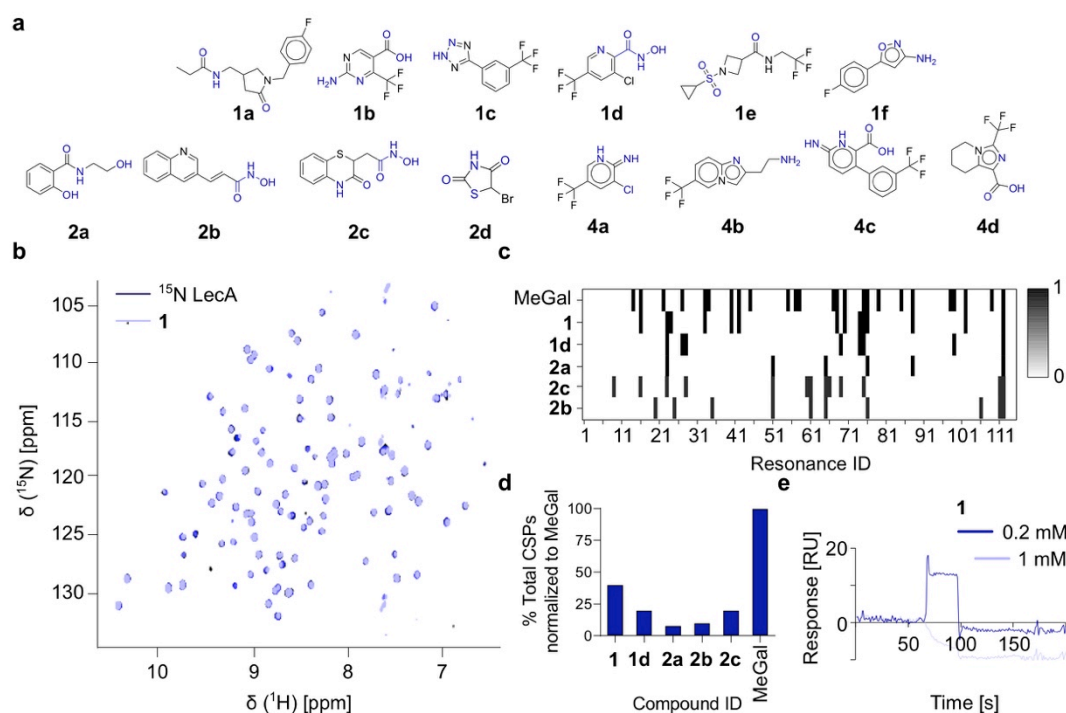
Figure 4.4-S2 MBP-like fragments identified for LecA in general FBDD screening.

a Example structures of best hits identified for LecA in [19]F (**1a**-**1f,** 4a**-**4d) and ^{15}N TROSY NMR (**2a**-**2c**) screenings of the general and MBP FBDD libraries. The metal-binding scaffold is highlighted in *blue*. Hydroxamates (**2a**-**2c**) are the most potent binders compared to ^{19}F hits (**1a**-**1f**). **b** Shown is the fingerprint of ^{15}N LecA in ^{1}H-^{15}N TROSY NMR in presence of a hydroxamate 0.25-2 mM **1**. **c** Shown is a quantitative analysis of perturbed [15]N LecA resonances upon addition of 2 mM hydroxamates **1**, **1d**, **2a-2c** and 1 mM MeGal in form of a 1:0 plot. The lack and presence of a chemical shift perturbation (CSP) or a change in peak intensity are set to 1 and 0, respectively. **1** promoted the strongest effect on [15]N LecA perturbing resonances similarly to MeGal. **d** The plot shows a total % of CSPs normalized to MeGal in [15]N LecA upon addition of hydroxamates, where **1** promoted 40% of CSPs in [15]N LecA similarly to MeGal. **e** Shown is a dose-dependent binding of **1** to LecA in SPR.

216
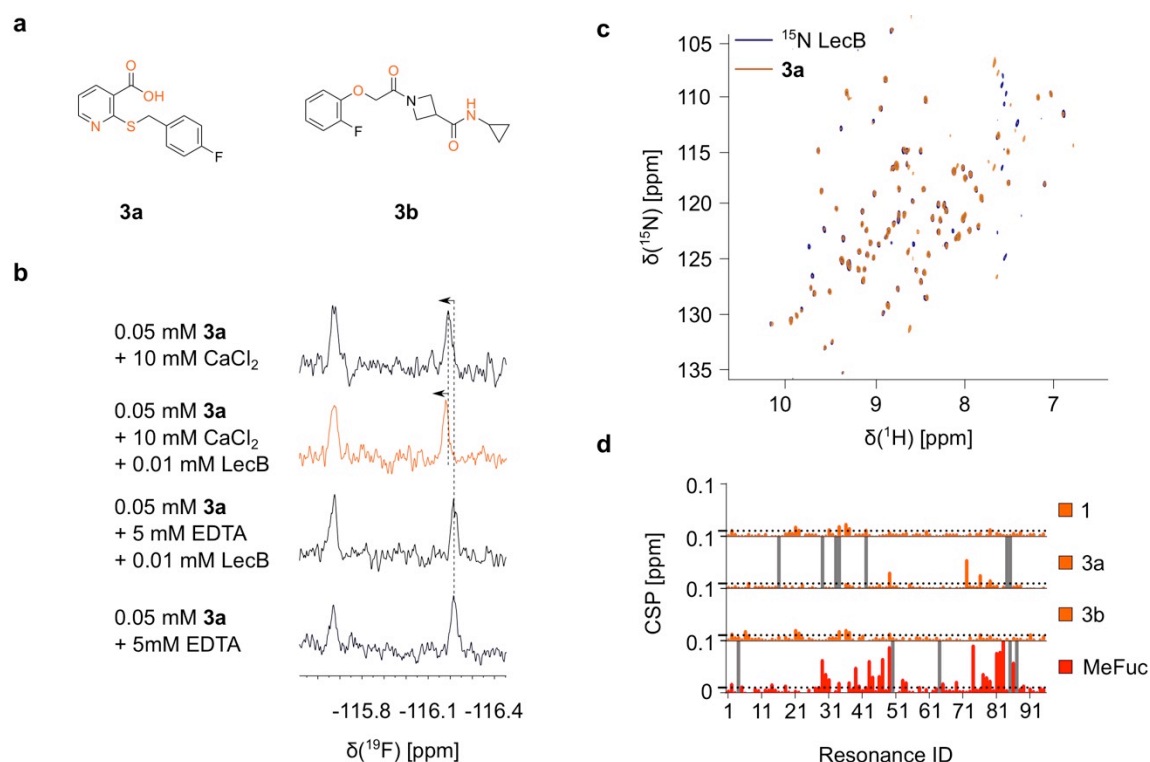
Figure 4.4-S3 MBP-like fragments identified for LecB in general FBDD screening.

a Shown are example structures of 2 out of 24 hits identified for LecB in ^{19}F NMR (**3a**-**3b**) screening of the general FBDD library. The metal-binding scaffold is highlighted in *orange*. **b** ^{19}F NMR spectra of a fragment mixture containing 0.05 mM **3a** and EDTA, where **3a** bound LecB only in presence of 10 mM CaCl$_2$ given the chemical shift perturbation of ^{19}F resonance (*orange*). **c** Fingerprint of ^{15}N LecB in ^{1}H-^{15}N TROSY NMR alone (*blue*) and in presence of 2 mM **3a**. **d** Shown is a quantitative analysis of perturbed ^{15}N LecB resonances upon addition of 2 mM **1**, **3a**, **3b** and 1 mM MeFuc. The significant CSPs are above 0.01 ppm (*dashed line*) and the decreased peak intensities were highlighted (*gray*). Notably, **3a** perturbed some resonances similarly to MeFuc. Notably, the hydroxamate **1** did not bind to ^{15}N LecB.
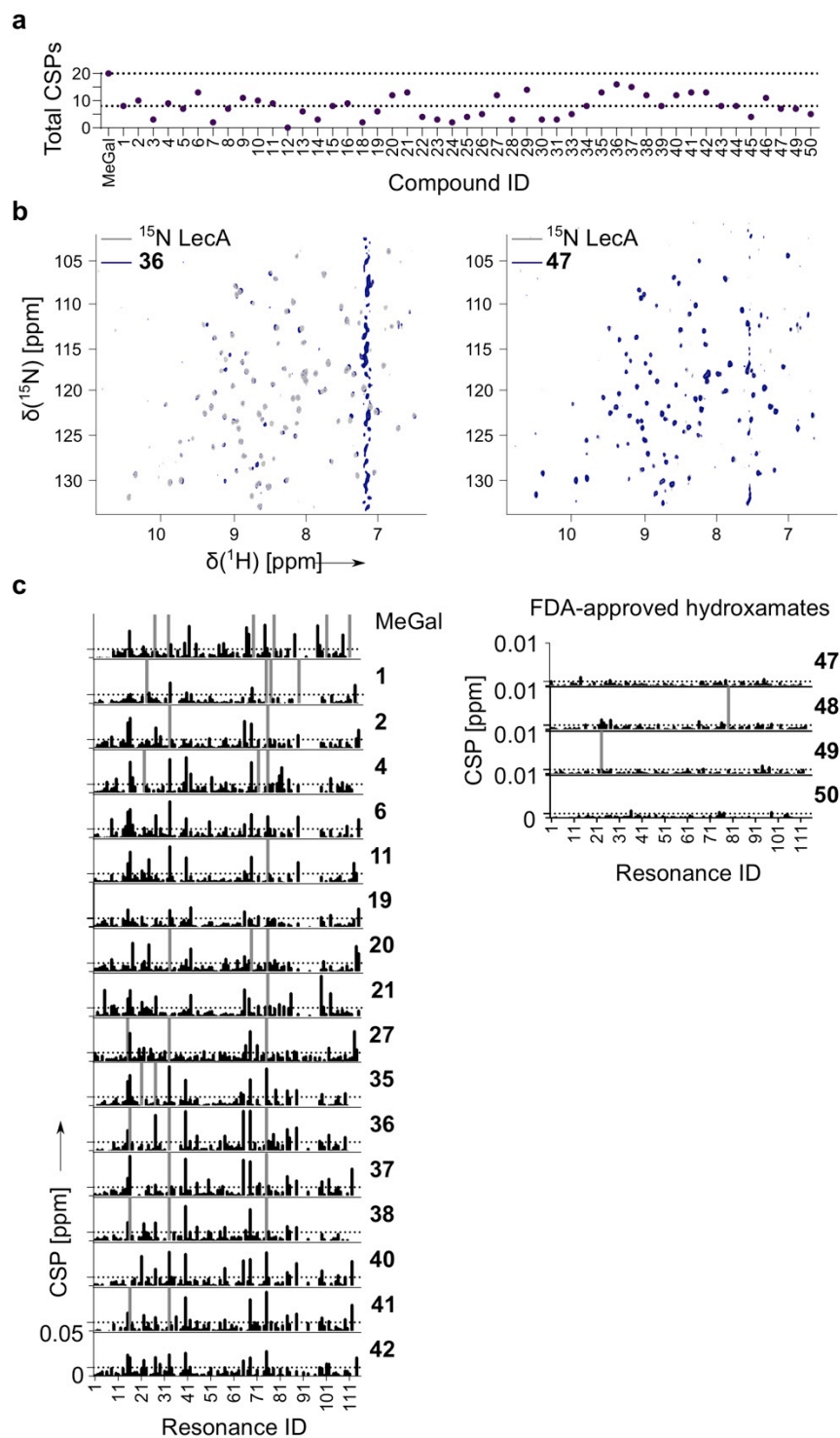
Figure 4.4-S4 Ranking of hydroxamate derivatives in ^{15}N TROSY NMR.

a The plot shows a total number of chemical shifts perturbed (CSPs) in ^{15}N LecA upon binding to 4 mM hydroxamate derivatives. To rank **1** analogues, we set the total number of CSPs upon addition of MeGal as the upper limit. The initial hit **1** was set as the lower limit. Further, compounds that demonstrated binding in this range were subjected to FP assay and PrOF NMR. **b** Shown are ^{1}H-^{15}N TROSY NMR spectra of ^{15}N LecA in presence of 4 mM **36**

and **47**. The compound **36** is a derivative of **35** demonstrated a slightly improved binding to ^{15}N LecA compared to **35**. Notably, no binding of the marketed drugs such as **47** has been observed. **c** Quantitative analysis of chemical shifts perturbed upon addition of the positive control 1 mM MeGal, 4 mM analogues of **1** or 1 mM marketed hydroxamates (**47**-**50**). A lower concentration of the marketed drugs has been used due to its high MW and thus, poor solubility at the concentrations above 1 mM. Notably, **35** and its derivatives (**36**-**40**) promoted the largest changes compared to the initial hit **1**, whereas the marketed hydroxamates did not interact with ^{15}N LecA.
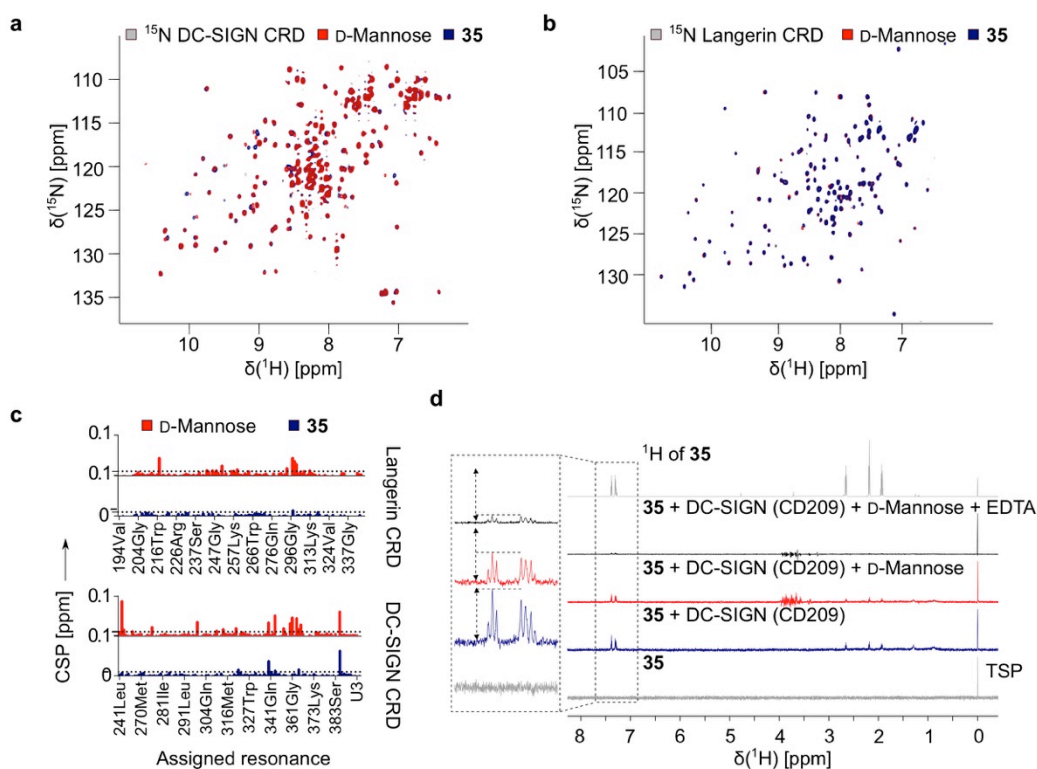
Figure 4.4-S5 Investigation of interactions between 35 and DC-SIGN or Langerin.

a and **b** are ^{1}H-^{15}N HSQC NMR spectra of DC-SIGN CRD and Langerin CRD (*gray*), respectively. Next, are shown ^{15}N fingerprints of both lectins with 3 mM **35** (*blue*) and 5 mM D-mannose. **c** Quantitative analysis of CSPs in presence of both ligands. The small CSPs above 0.01 ppm were observed only for DC-SIGN CRD. **d** ^{1}H and STD NMR spectra of 0.5 mM **35** (*gray*) shown on top and bottom, respectively. The STD NMR spectrum of **35** with 0.02 mM DC-SIGN ECD demonstrated **35** binding given the recovery of ^{1}H resonances in the STD spectrum. Following addition of 100 mM D-mannose a partial competition of the ligand has been observed. This supports hydroxamate binding to DC-SIGN ECD to a secondary binding site rather than a Ca^{2+}-binding site. The addition of 10 mM EDTA almost displaced **35**, verifying its binding to DC-SIGN ECD.
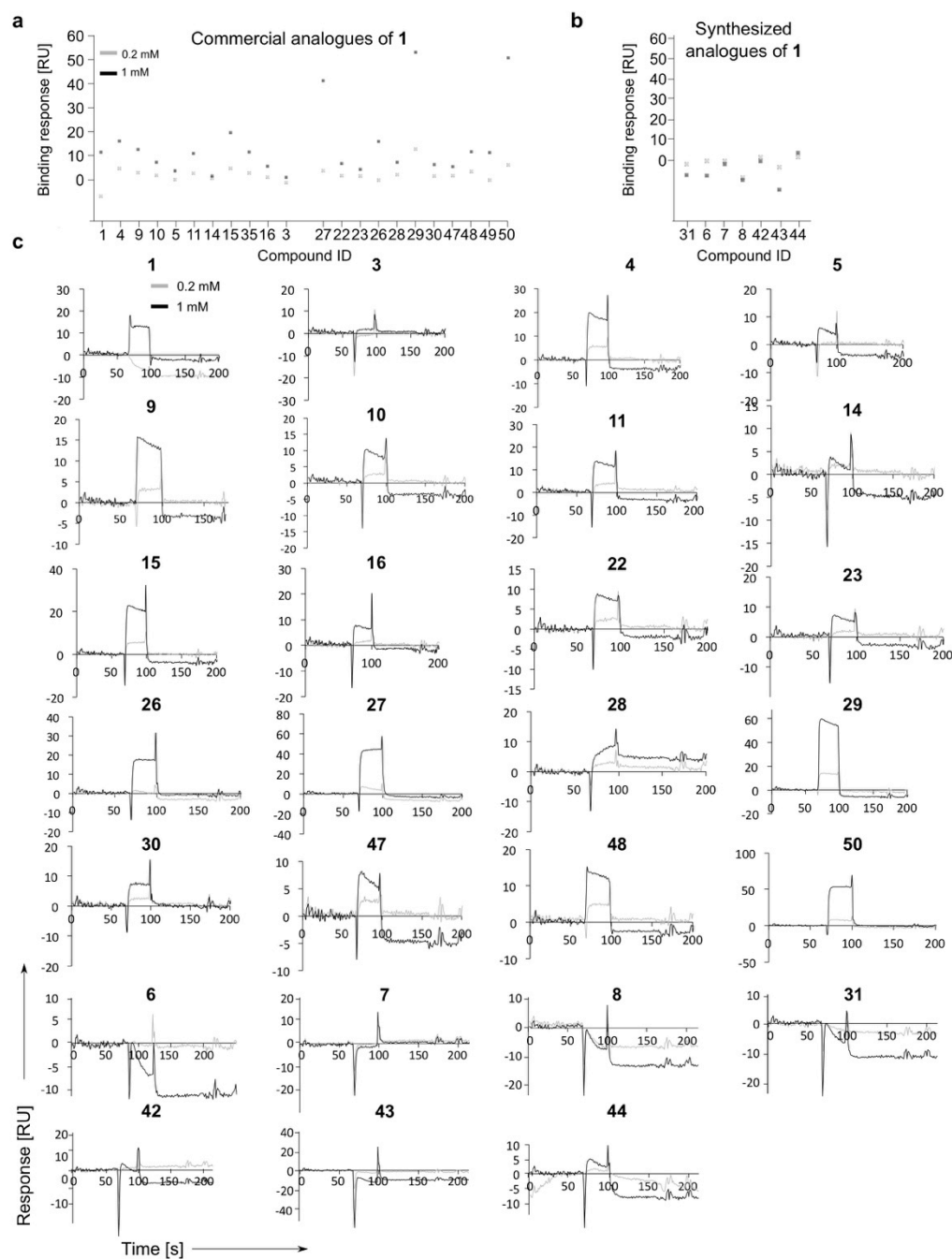
Figure 4.4-S6 SPR analysis of hydroxamate derivatives.

a and **b** The plots show a binding response of commercial and synthesized hydroxamate derivatives of **1** determined in SPR at 0.2 mM and 1 mM, respectively. The 'in-house' synthesized analogues of **1** did not show a dose-dependent binding in SPR. **c** SPR sensogramms show a dose-dependent binding of hydroxamate **1**, commercial (**3**-**5**, **9**-**11**, **14**-**16**, **22**, **23**, **26**-**30**), marketed (**47**, **48**, **50**) and 'in-house' synthesized (**6**-**8**, **31**, **42**-**44**) compounds at 0.2 (*gray*) and 1 mM (*black*).
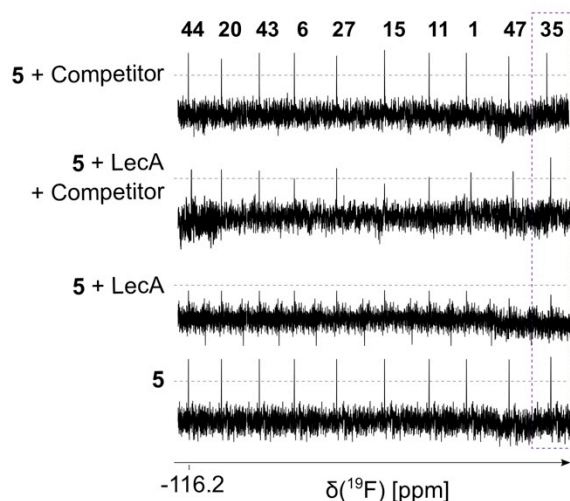
Figure 4.4-S7 Competitive ^{19}F T$_2$-filtered NMR study using hydroxamate derivative 5 as a reporter.

Shown is the ^{19}F T$_2$-filtered NMR spectrum of 0.1 mM **5** reporter alone and in presence of 0.02 mM LecA and 10 mM CaCl$_2$. The reduction in peak intensity indicates binding of **5** to LecA. The addition of 3 mM **35** displaced **5** compared to other compounds, which showed a weak competition. In particular, the initial hit **1** and the marketed hydroxamate derivative (SAHA, **47**) did not displace **5**.
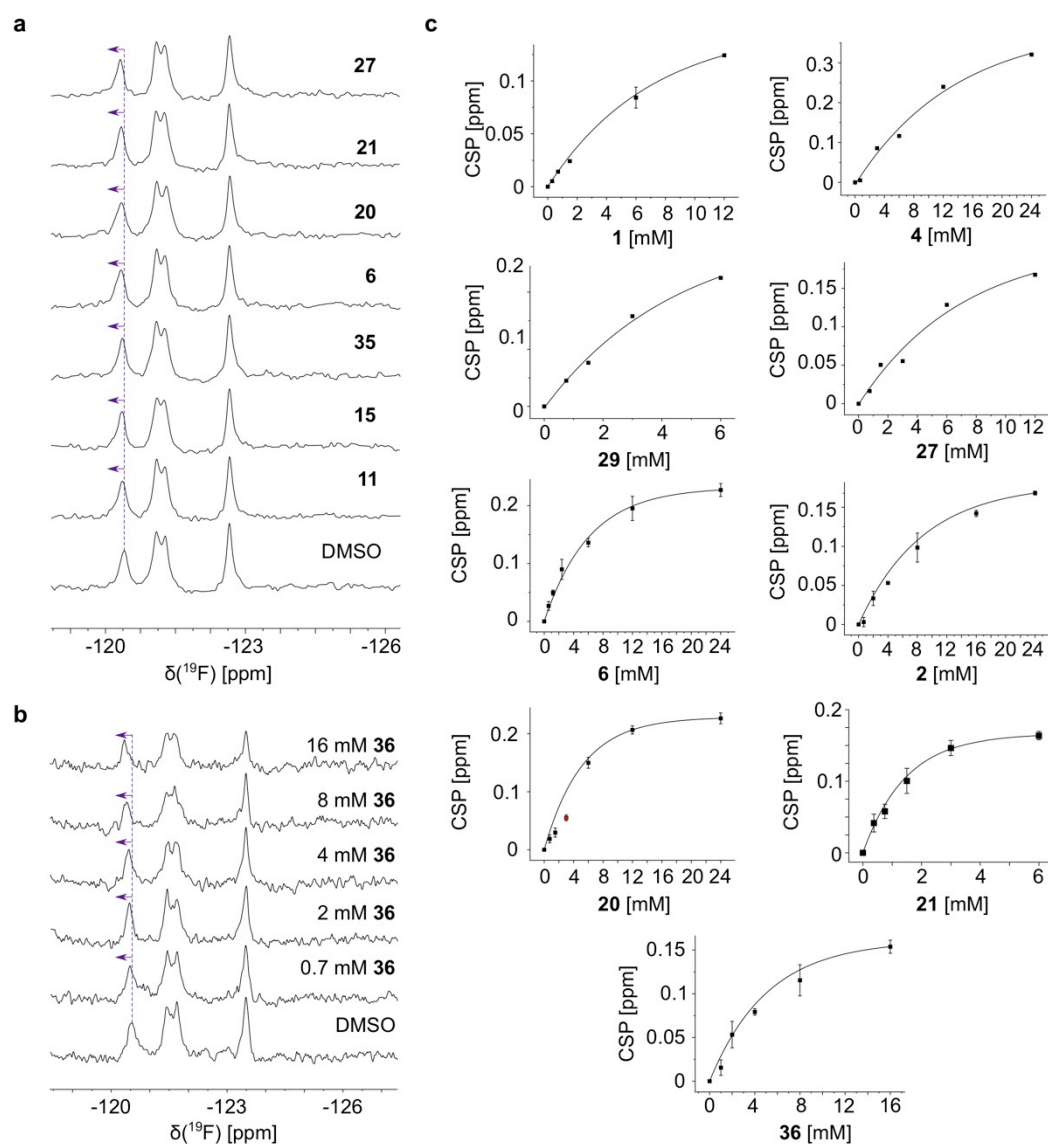
Figure 4.4-S8 PrOF NMR of 1 hydroxamate derivatives.

a Shown are PrOF NMR spectra of 0.15 mM 5FW LecA, where the addition of 2 mM hydroxamate derivatives promoted a chemical shift perturbation (CSP) of W42. **b** Changes in chemical shift of W42 were used to derive the affinities (K_d) of hydroxamates in PrOF NMR, as shown on example of **36**. **c** Titration data using CSPs of W42 was fitted to the one-site binding model to derive K_d values.
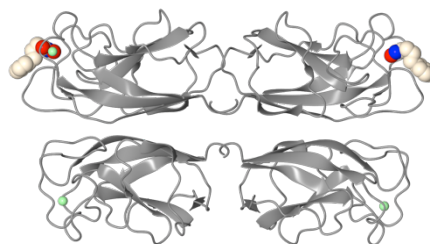
Figure 4.4-S9 Crystal structure of LecA in complex with 35.

Overall structure of LecA tetramer is shown in ribbon. Compound 35 is shown in sphere.
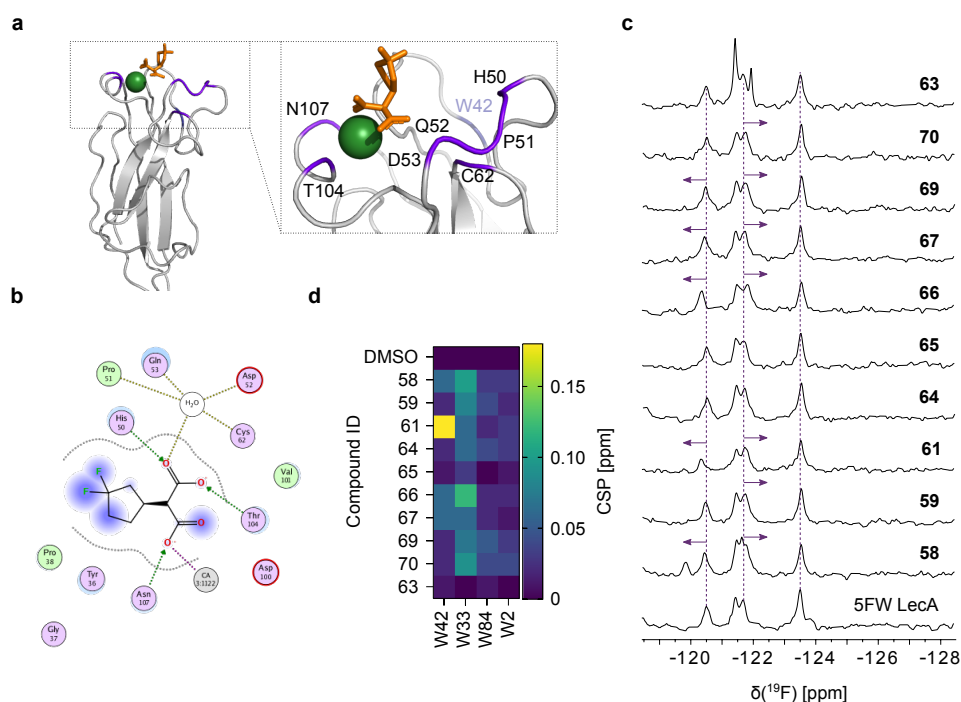
Figure 4.4-S10 Interactions of 58 malonate derivatives with LecA.

a Docking pose of **58** with LecA. The residues of the binding site interacting with **58** are highlighted (*violet*). **b** Interaction map of LecA calcium binding site with malonic acid **58**. **c** One-point-titration experiments in PrOF NMR using 0.15 mM 5FW LecA and 2 mM malonic acid derivatives showing the malonates, which perturbed W42 and W33. DMSO and MeGal served as negative and positive controls, respectively. **d** Heat map shows the magnitude of CSP on 5FW resonances in presence of **58** analogues. Malonates perturbed W42, whereas compounds with an acetic group influenced only W33 suggesting a Ca^{2+}-dependent binding of malonates only.
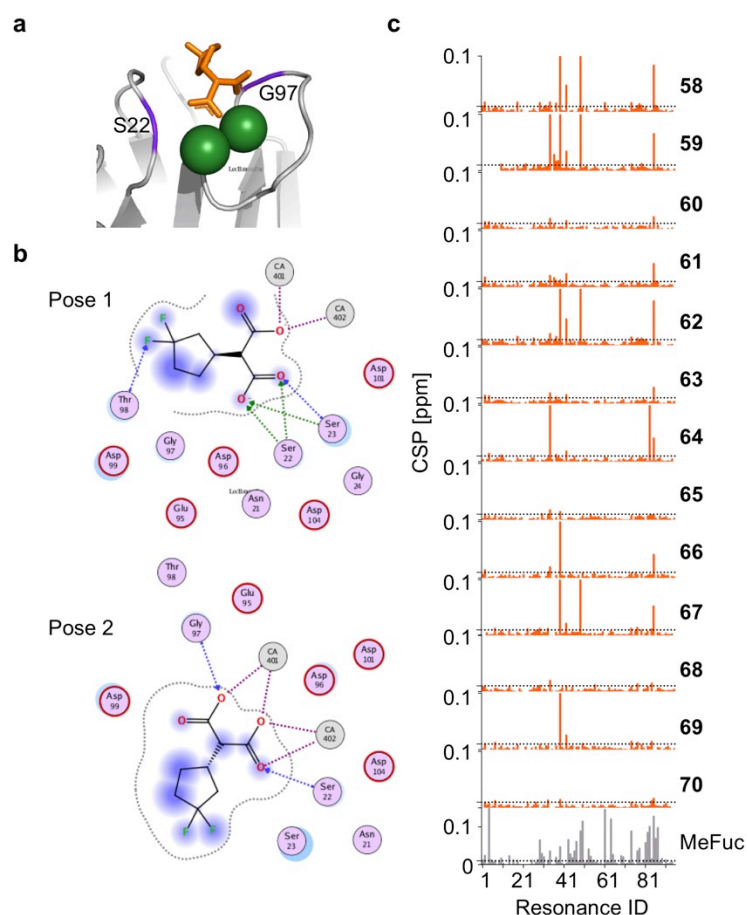
Figure 4.4-S11 Interactions 58 analogues with LecB.

a Docking pose 2 of **58** to LecB, where Ca^{2+} is shown in *green* and the interacting residues in *violet***. b** Interaction maps of poses 1 and 2 of **58** and LecB. Pose 1 demonstrates a potential interaction of CF$_2$-group with T98. **c** CSP plots of 0.07 mM ^{15}N LecB resonances in ^{1}H-^{15}N TROSY NMR in presence of 2 mM malonic acid derivatives of **58** and MeFuc. MeFuc is a positive control verifying the protein is active. Compounds harboring a malonate group **58**, **59**, **62**, **64**, **66**, **67** and **69** perturbed the most resonances in ^{15}N LecB. Notably, some of resonances were perturbed similar to MeFuc suggesting these compounds target the carbohydrate-binding site of ^{15}N LecB. Compared to **64**, **58** and **59** showed a larger magnitude of CSPs in ^{15}N LecB suggesting a role of an electronegative group in the interaction with ^{15}N LecB as it has been proposed in docking analysis (pose 1).
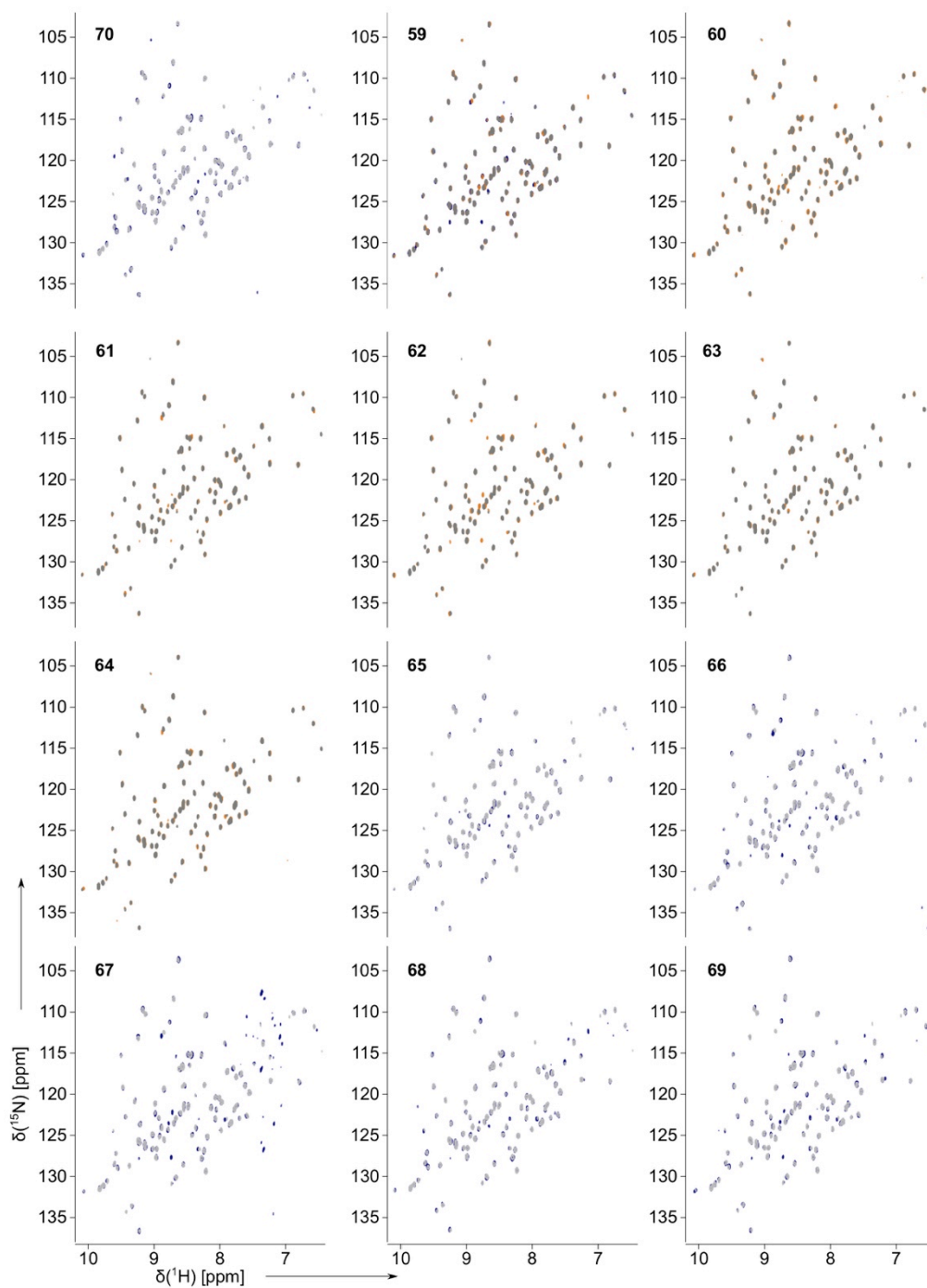
Figure 4.4-S12 Supplementary figure for Figure S11.

Shown are [15]N LecB fingerprints (*gray*) in presence of 2 mM malonic acid (**58)** derivatives (*blue and orange*) in ^{1}H-^{15}N TROSY NMR. Additionally, **59** (*blue*) was overlaid with **58** (*orange*) for a comparison demonstrating a similarity in the binding mode of both compounds as both compounds perturbed similar ^{15}N resonances.
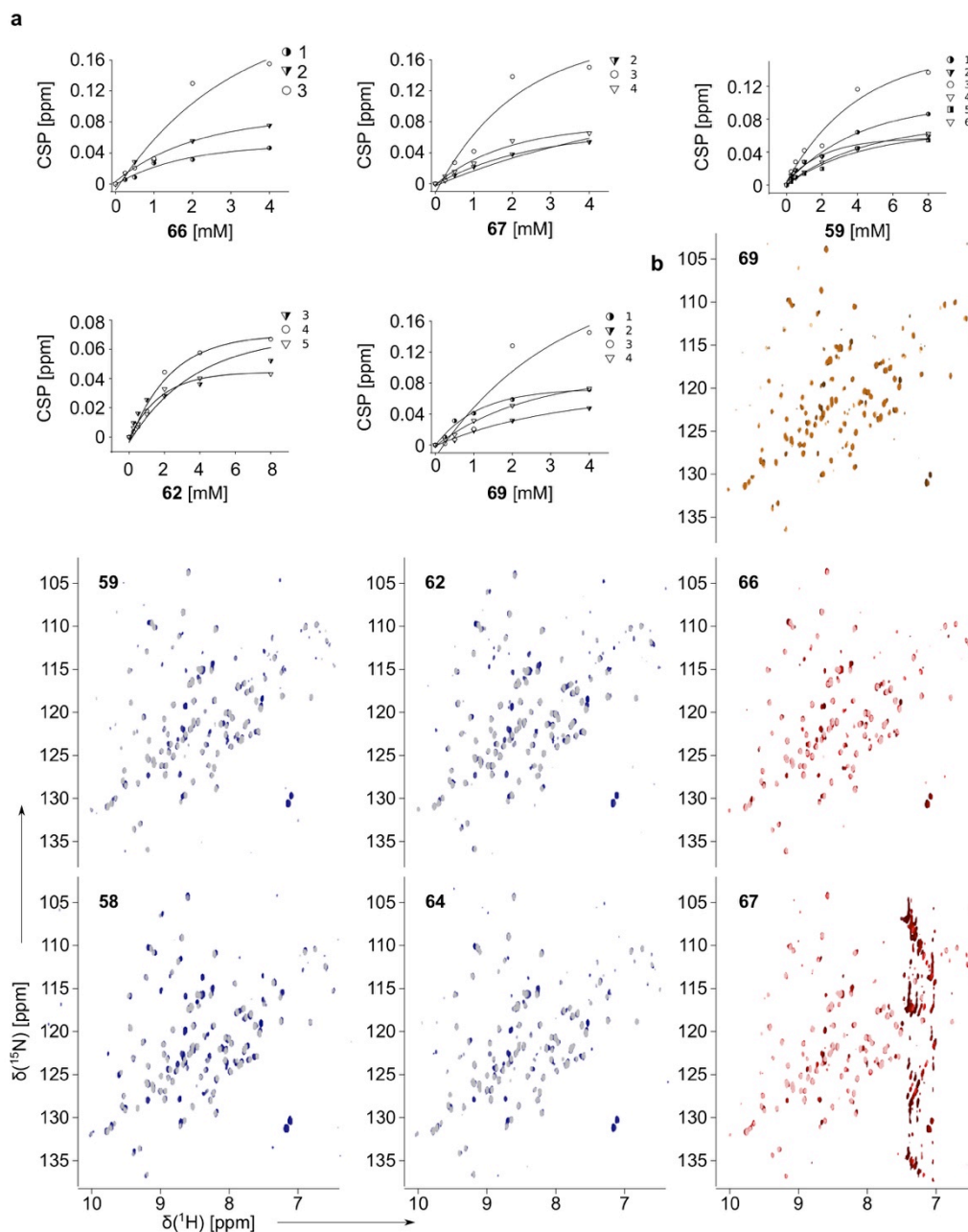
Figure 4.4-S13 Titration ^{1}H-^{15}N TROSY NMR studies of 58 derivatives and LecB.

a Titration data using chemical shift perturbations (CSPs) of resonances perturbed in presence of malonic acid derivatives were fitted to the one-site binding model to derive K_d values. B Shown are ^{15}N LecB fingerprints (*gray*) in presence of various concentrations of malonic acid derivatives of **58** in ^{1}H-^{15}N TROSY NMR.
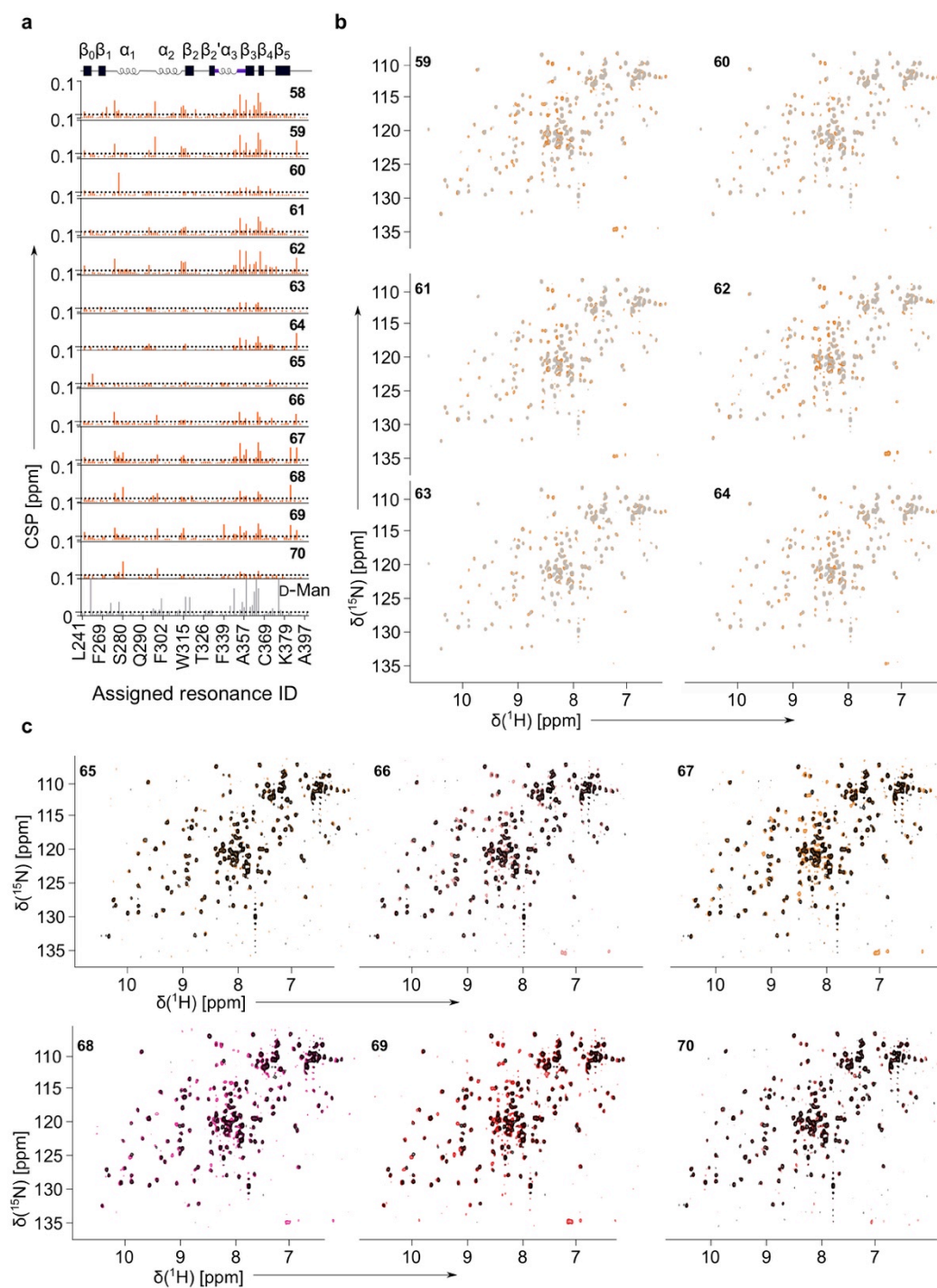
Figure 4.4-S14 ^{1}H-^{15}N HSQC NMR of 58 derivatives with ^{15}N DC-SIGN CRD.

a Shown are chemical shift perturbation (CSP) plots of 0.1 mM ^{15}N DC-SIGN CRD resonances in ^{1}H-^{15}N HSQC NMR in presence of 2 mM malonic acid derivatives of **58** or D-mannose as a positive control. Compounds harboring a malonate group **58**, **59**, **62** and **67** perturbed the most resonances in ^{15}N DC-SIGN. Notably, malonates promoted the CSPs in a similar manner to D-mannose suggesting these compounds target the carbohydrate-binding site of ^{15}N DC-SIGN CRD. **b** Shown are ^{15}N DC-SIGN CRD fingerprints (*gray*) in ^{1}H-^{15}N HSQC NMR in presence of malonates (*orange*). **c** To derive the affinities of **58** derivatives to DC-SIGN-CRD, titration experiments were performed in ^{1}H-^{15}N HSQC NMR. Shown are ^{15}N

DC-SIGN CRD fingerprints (*gray*) in presence of various concentrations of malonic acid derivatives of **58**.
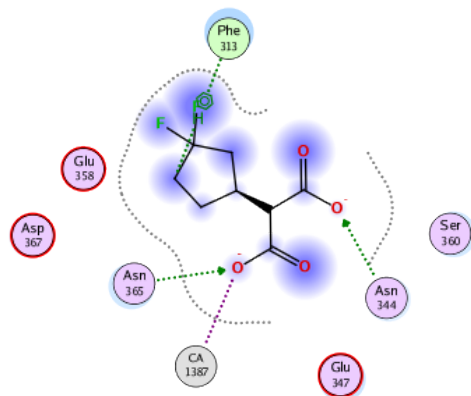
Figure 4.4-S15 Interaction map of 58 with DC-SIGN CRD.
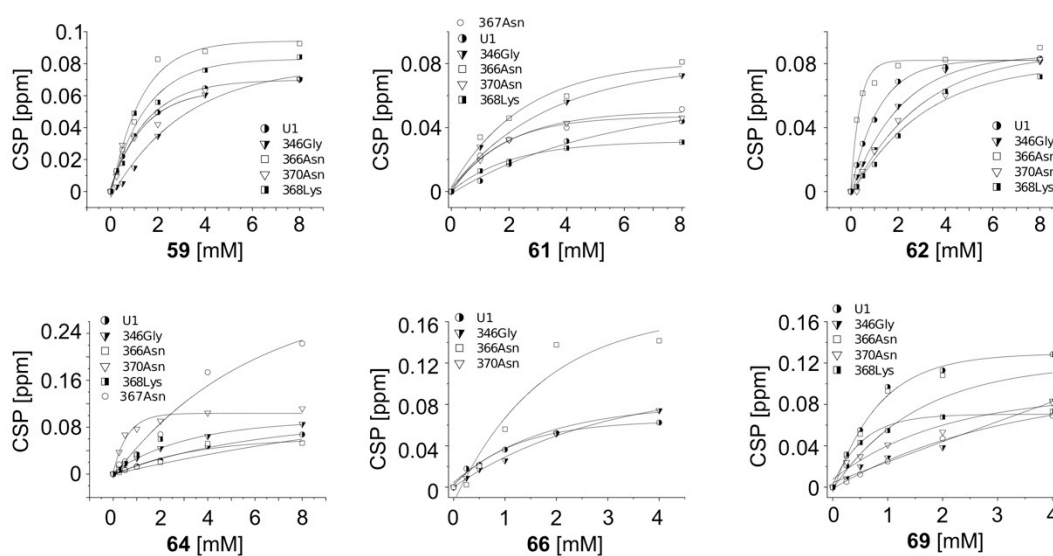
Figure 4.4-S16 Titration study by ^{1}H-^{15}N HSQC NMR to derive affinities of malonates for DC-SIGN CRD.

Titration data using chemical shift perturbations (CSPs) of resonances perturbed in presence of malonic acid derivatives in **Figure S14c** were fitted to the one-site binding model to derive K_d values.

a
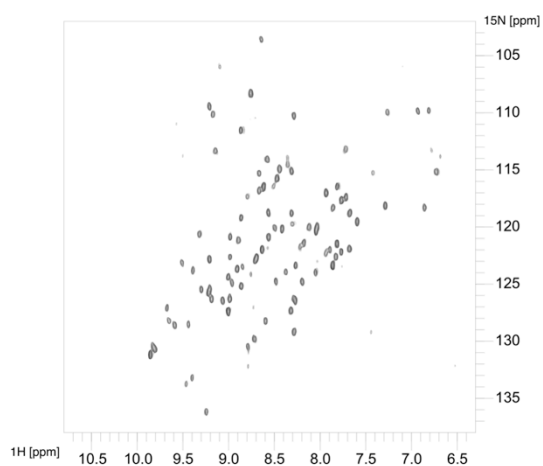
b
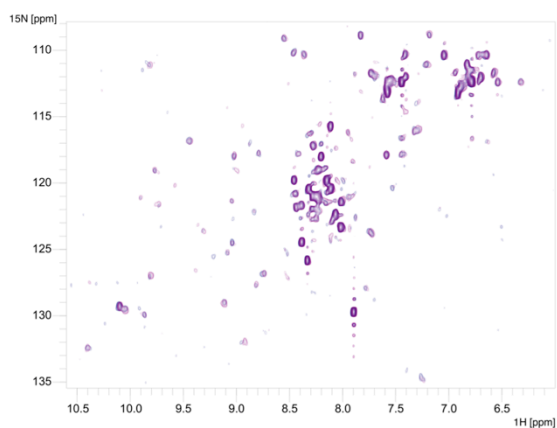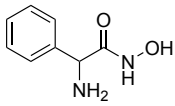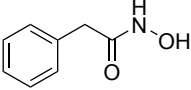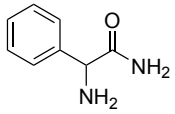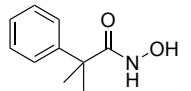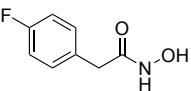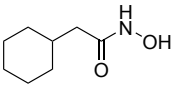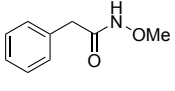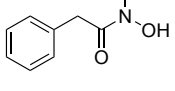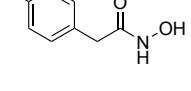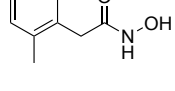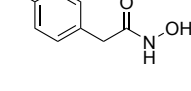
Figure 4.4-S17 Ca^{2+}-dependent binding of 58 to LecB and DC-SIGN CRD in ^{1}H-^{15}N HSQC/TROSY NMR.

a ^{1}H-^{15}N TROSY NMR of 0.07 mM ^{15}N LecB in presence of 10 mM EDTA and DMSO (*gray*) or 2 mM **58** (*black*). Malonate **58** did not bind ^{15}N LecB in absence of Ca^{2+}. **b** The ^{1}H-^{15}N HSQC NMR spectra of 0.1 mM ^{15}N DC-SIGN CRD in presence of 10 mM EDTA in combination with 2 mM **58** (*violet*) or DMSO (*blue*) as negative control. Malonate **58** did not bind ^{15}N DC-SIGN in absence of Ca^{2+}.
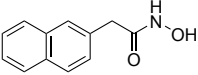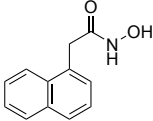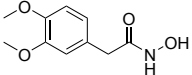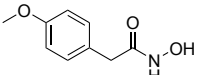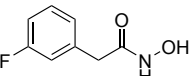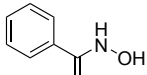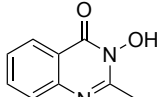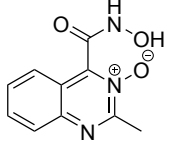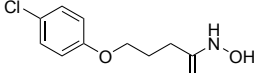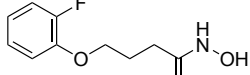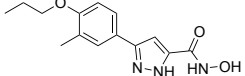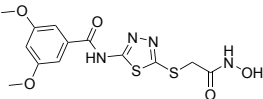
7.4.4. Supplementary Tables

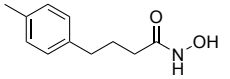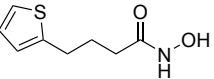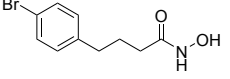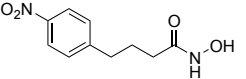**Table 4.4-S1 Commercial and synthesized hydroxamates.**
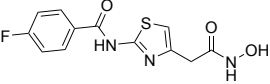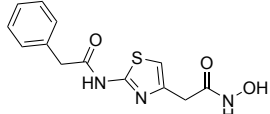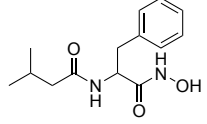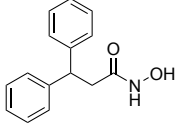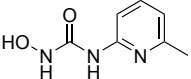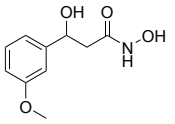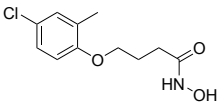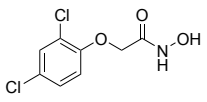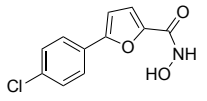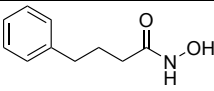
Shown are the structures of hydroxamates and its affinities derived in PrOF NMR, [%] inhibition in competitive fluorescence polarization (FP) assay and [%] efficiency in SPR towards LecA. Groups: 1) commercial and synthesized derivatives of the hydroxamate **1**, 2) diversity oriented, 3) derivatives of the hydroxamate **35** and 4) marketed drugs for metalloenzymes.

ID	Compound	PA-IL (LecA)			
		K_d in PrOF NMR [mM]	LE [kcal mol^{-1} HA^{-1})	Inhibition in FP [%]	Efficiency in SPR [%]
	Group 1				
1		4.4 ± 0.7	0.24	6 *	11.6
2		6.1 ± 0.9	0.29	18.3 ± 1.2	*n.d.*
3		*n.d.*	*n.d.*	*No inhibition*	*No dose response*
4		9.4 ± 2.5	0.22	10%	14.8
5		*n.d.*	*n.d.*	*n.d.*	4.1
6		4.4 ± 0.6	0.30	21.4 ± 0.6	*n.d.*
7		*n.d.*	*n.d.*	*No inhibiton*	*n.d.*
8		*n.d.*	*n.d.*	19.9 ± 2.7	*n.d.*
9		*n.d.*	*n.d.*	16.4 ± 2.4	12.8
10		*n.d.*	*n.d.*	13.6 ± 4.1	7.7
11		*n.d.*	*n.d.*	14.3 ± 1.1	14
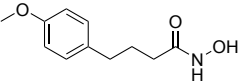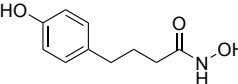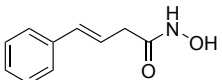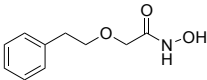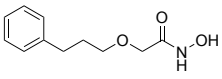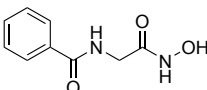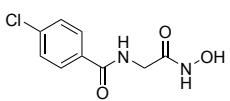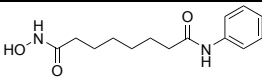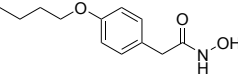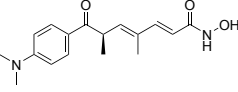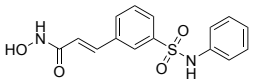
12		n.d.	n.d.	n.d.	n.d.
13		n.d.	n.d.	n.d.	n.d.
14		n.d.	n.d.	n.d.	3.5
15		n.d.	n.d.	16 ± 4	10.2
16		n.d.	n.d.	n.d.	6
Group 2					
17		n.d.	n.d.	n.d.	14.2
18		n.d.	n.d.	n.d.	n.d.
19		n.d	n.d	37.2 ± 1.8	n.d.
20		4.5 ± 0.2	0.26	32.7 ± 0.1	n.d.
21		2.4 ± 0.4	0.23	n.d.	n.d.
22		n.d.	n.d.	n.d.	5.2
23		n.d.	n.d.	n.d.	3.4
24		n.d.	n.d.	n.d.	n.d.
25		n.d.	n.d.	n.d.	3.1

26		*Precipitated*	*n.d.*	*n.d.*	9.2
27		4.8 ± 1.3	0.17	15.5	23.1
28		*n.d.*	*n.d.*	*n.d.*	4.9
29		3.1 ± 0.9	0.20	*n.d.*	36.1
30		*n.d.*	*n.d.*	*n.d.*	6.8
31		*n.d.*	*n.d.*	8.6 ± 2.3	*n.d.*
32		*n.d.*	*n.d.*	5.4 ± 3.2	*n.d.*
33		*n.d.*	*n.d.*	*Precipitated*	*n.d.*
34		*n.d.*	*n.d.*	*Precipitated*	*n.d.*
	Group 3				
35		4.6 ± 0.9	0.25	26 ± 0.7	10.9
36		4.4 ± 0.7	0.24	35.2 ± 3.3	*n.d.*
37		*n.d.*	*n.d.*	24.7 ± 3.3	*n.d.*
38		*n.d.*	*n.d.*	25.2 ± 1.3	*n.d.*
39		*n.d.*	*n.d.*	4.1 ± 0.2	*n.d.*

#	Structure				
40		*n.d.*	*n.d.*	26.6 ± 1	*n.d.*
41		*n.d.*	*n.d.*	39 ± 1.2	*n.d.*
42		*n.d.*	*n.d.*	12.7 ± 2.7	*n.d.*
43		*n.d.*	*n.d.*	7.6 ± 0.3	*n.d.*
44		*n.d.*	*n.d.*	16.7 ± 2.4	*n.d.*
45		*n.d.*	*n.d.*	*No inhibiton*	*n.d.*
46		*n.d.*	*n.d.*	*n.d.*	*n.d.*
Group 4					
47		*n.d.*	*n.d.*	*n.d.*	3.8
48		*Precipitated*	*n.d.*	*n.d.*	9.0
49		*n.d.*	*n.d.*	*n.d.*	27.7
50		*n.d.*	*n.d.*	*n.d.*	7.5

Inhibition [%] compared to MeGal at 10 mM (16 h).

* Measured at 4 mM (not enough material)

n.d. = not determined

Table 4.4-S2 Statistics for data collection and refinement of LecA-35 complex.

Data collection	LecA-35
Beamline	PROXIMA1 (SOLEIL)
Wavelength (Å)	0.9788
Detector	Pilatus 6M
Resolution (Å)	46.85-1.79 (1.83-1.79)
Space group	$P2_12_12_1$
a, b, c (Å)	48.98, 51.68, 160.53
α, β, γ (º)	90.0, 90.0, 90.0
Total observations	332865
Unique reflections	39017
Multiplicity	8.5 (7.5)
Mean $I/\sigma(I)$[a]	14.2 (2.2)
Completeness (%)[a]	99.4 (89.9)
R_{merge}[a,b]	0.073 (0.717)
$CC_{1/2}$[a,c]	0.999 (0.949)
Refinement	
Reflections: working/free[d]	38945/2019
R_{work}/R_{free}[e]	0.180/0.221
Ramachandran plot:	
allowed/favoured/outliers (%)[f]	3/97/1
r.m.s bond deviations (Å)	0.0113
r.m.s angle deviations (º)	1.556
Mean *B*-factors: protein/ligand[f]/water (Å^2)	35/39/38

[a] Values for the outer resolution shell are given in parentheses.

[b] Rmerge = $\sum_{hkl} \sum_i |I_i(hkl) - <I(hkl)>|/ \sum_{hkl} \sum_i I_i(hkl)$.

[c] $CC_{1/2}$ is the correlation coefficient between symmetry-related intensities taken from random halves of the dataset.
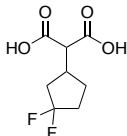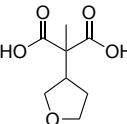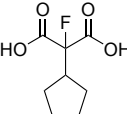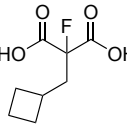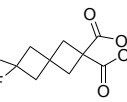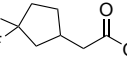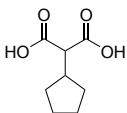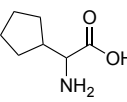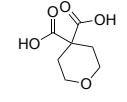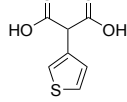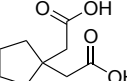
[d] The data set was split into "working" and "free" sets consisting of 95 and 5% of the data, respectively. The free set was not used for refinement.

[e] The R-factors R_{work} and R_{free} are calculated as follows: R = å$\sum(| F_{obs} - F_{calc} |)/\sum| F_{obs} |$, where F_{obs} and F_{calc} are the observed and calculated structure factor amplitudes, respectively
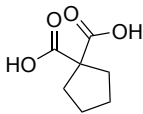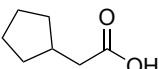
[f] refers to ligands bound in the active site and potential surface binding sites

Table 4.4-S3 Commercial derivatives of malonic acid 58.

Shown are the structures of malonates, its affinities and LE values for DC-SIGN CRD and LecB derived in ^{1}H-^{15}N HSQC and TROSY NMR, respectively.

ID	Compound	DC-SIGN CRD (CD209)		PA-IIL (LecB)	
		K_d [mM] in HSQC NMR	LE [kcal mol^{-1} HA^{-1}]	K_d [mM] in TROSY NMR	LE [kcal mol^{-1} HA^{-1}]
58		1.2 ± 0.5	0.28	1.2 ± 04	0.29
59		1.2 ± 0.4	0.31	2.7 ± 0.6	0.28
60		n.d.	n.d.	n.d.	n.d.
61		1.9 ± 0.8	0.28	n.d.	n.d.
62		1.5 ± 0.7	0.26	1.5 ± 0.2	0.27
63		n.d.	n.d.	n.d.	n.d.
64		2.6 ± 0.9	0.29	2.6 ± 0.6	0.31
65		n.d.	n.d.	n.d.	n.d.
66		1.3 ± 0.4	0.33	1.6 ± 0.4	0.33
67		1.2 ± 0.5	0.33	1.6 ± 0.3	0.31
68		n.d.	n.d.	n.d.	n.d.

69		1.0 ± 0.6	0.37	1.7 ± 0.6	0.36
70		*n.d.*	*n.d.*	*n.d.*	*n.d.*

n.d. = not determined

Table 4.4-S4 Quantitative analysis of 5FW resonances in PrOF NMR in presence of 2 mM fragments.

ID	W42 [ppm]	W33 [ppm]	W2 [ppm]	W84 [ppm]
DMSO	120.50	121.67	121.43	123.50
58	120.44	121.77	121.47	123.53
59	120.48	121.75	121.46	123.54
60	*n.d*	*n.d*	*n.d*	*n.d*
61	120.32	121.73	121.46	123.52
62	*n.d*	*n.d*	*n.d*	*n.d*
63	120.49	121.67	121.44	123.51
64	120.52	121.73	121.47	123.54
65	120.51	121.70	121.44	123.50
66	120.43	121.80	121.46	123.52
67	120.44	121.74	121.44	123.52
68	120.48	121.74	121.47	123.55
69	120.51	121.76	121.48	123.54
70	120.50	121.67	121.43	123.50

n.d = not measured

Table 4.4-S5 List of LecA PDB structures and co-crystalized ligands analyzed for virtual screening.

Ligand	PDB ID	Resolution (Å)
Apo	1l71	1.5
Apo	1uoj	2.4
β-D-galactose / α-D-galactose	1oko	1.6
αGal1-3βGal1-4Glc	2vxj	1.9
αGal1-2βGal-O-Met	2wyf	2.4
βGal+P-hydroxybenzoic acid	3zyb	2.29
1-O-[P-Nitrophenyl]-β-D-galactopyranose	3zyf	1.94
3-(β-D-galactopyranosylthio)propanoic acid	3zyh	1.5
Naphthalen-2-YL-Thio-β-D-galactopyranoside	4a6s	2.15
Melibiose (αGal1-6Glc)	4al9	1.75
β-D-galactose + (4S)-N-ethyl-4-{[N-methyl-3-(1-{2-[(4-sulfanylbenzoyl)aminoethyl}-1H-1,2,3-triazol-4-yl)-L-alanyl]amino}-Lprolinamide	4cp9	1.65
β-D-galactose + (4S)-N-ethyl-4-{[N-methyl-3-(1-{2-[(4-sulfanylbenzoyl)aminoethyl}-1H-1,2,3-triazol-4-yl)-L-alanyl]amino}-Lprolinamide	4cpb	1.57
β-D-galactose + 1-methyl-1H-indol-3-ol	4ljh	1.45
β-D-galactose + Chlorophenol Red (2-[(E)-(3-chloro-4hydroxyphenyl)(3-chloro-4 oxocyclohexa-2,5-dien-1ylidene)methyl]benzenesulfonic acid	4lk6	2.86

237

β-D-galactose + resorufin (7-hydroxy-3H-phenoxazin-3-one)	4lk7	1.76
β-D-galactose +P-hydroxybenzoic acid	4lkd	2.31
β-D-galactose +P-hydroxybenzoic acid	4lke	1.65
β-D-galactose +P-hydroxybenzoic acid	4lkf	1.64
N-[(2S)-6-amino-1-oxo-1-(pyrrolidin-1-yl)hexan-2-yl]-4-(beta-Dgalactopyranosyloxy)benzamide	4yw6	1.4
(2R,3R,4S,5R,6R,2'R,3'R,4'S,5'R,6'R)-2,2'-([(2R,3R,4S,5S,6S)-3,4dihydroxy-6-(hydroxymethyl)tetrahydro-2H-pyran-2,5-diyl]bis{1H-1,2,3-triazole-1,4-diyl[(2S,3R,4S,5S,6S)-3,4-dihydroxy-6-(hydroxymethyl)tetrahydro-2H-pyran-2,5-diyl]-1H-1,2,3-triazole-1,4diylpropane-3,1-diyloxy})bis[6-(hydroxymethyl)tetrahydro-2H-pyran3,4,5-triol]	4yw7	1.82
(2R,3R,4S,5R,6R,2'R,3'R,4'S,5'R,6'R)-2,2'-([(2R,3R,4S,5S,6S)-3,4dihydroxy-6-(hydroxymethyl)tetrahydro-2H-pyran-2,5-diyl]bis{1H-1,2,3-triazole-1,4-diyl[(2S,3R,4S,5S,6S)-3,4-dihydroxy-6-(hydroxymethyl)tetrahydro-2H-pyran-2,5-diyl]-1H-1,2,3-triazole-1,4diylmethanediyloxy})bis[6-(hydroxymethyl)tetrahydro-2H-pyran3,4,5-triol]	4ywa	1.19
Phenyl β-D-galactopyranoside	5d21	1.9
Phenyl 6,7-dideoxy-6,7-epoxy-beta-D-galacto-heptopyranoside(6D)	5mih	1.8

Table 4.4-S6 List of LecB PDB structures and co-crystalized ligands analyzed for virtual screening.

Ligand	PDB-ID	Resolution (Å)
apo	1ous	1.2
apo	1oux	2.0
apo	5a6q	1.7
α-L-fucose	1gzt	1.3
α-D-mannose	1our	1.42

238

β-D-fructopyranose	1ovp	1.4
α-D-mannose	1ovs	1.75
Fucose (both)	1oxc	1.2
α-L-fucose	1uzv	1.0
α-L-fucose +β-D-glucose+β-D-galactose+N-acetyl-D-glucosamine	1w8f	1.05
α-L-fucose + (N-acetyl-D-glucosamine or 2acetylamino-2-deoxy-A-D-glucopyranose)+β-D-galactose /Lewis[A] trisaccharide	1w8h	1.75
Methyl β-D-arabinopyranoside	2boj	1.8
α-L-galactopyranose OR GAL	2bp6	2.5
α-L-fucose +N-acetyl-D-glucosamine + 2H-1,2,3-Triazol-4-YLmethanol	2jdh	1.1
α-L-fucose +N-acetyl-D-glucosamine + methyl 2H-1,2,3-triazole-4-carboxylate	2jdk	1.1
α-L-methyl-fucose	2jdm	1.7
O1-methyl-mannose	2jdn	1.3
α-L-methyl-fucose	2jdp	1.3
α-L-methyl-fucose	2jdu	1.5
O1-methyl-mannose	2jdy	1.7
[(3E)-3-(1-hydroxyethylidene)-2,3-dihydroisoxazol-5yl]methyl 6-deoxy-alpha-L-galactopyranoside	2vuc	1.3
[1-(2-oxoethyl)-1H-1,2,3-triazol-5-yl]methyl 6-deoxyalpha-L-galactopyranoside	2vud	1.7
(2S)-1-[(2S)-6-amino-2-({[(2S,3S,4R,5S,6S)-3,4,5trihydroxy-6-methyltetrahydro-2H-pyran-2yl]acetyl}amino)hexanoyl]-N-[(1S)-1-carbamoyl-3methylbutyl]pyrrolidine-2-carboxamide	3dcq	1.8
O1-methyl-mannose + 2,4,6 - trimethyl-benezenesulfonamide	3zdv	1.41
α-L-fucose	4ce8	0.9
O1-methyl-mannose + cinnamide	5a3o	1.6
α-L-methyl-fucose	5a6x	1.55
α-D-mannose + α-D-mannose	5a6y	1.4
α-L-fucose +N-acetyl-D-glucosamine + β-D-galactose +α-D-galactose	5a6z	1.5
α-L-fucose +N-acetyl-D-glucosamine + β-D-galactose +α-D-galactose	5a70	1.6
3,7-anhydro-2,8-dideoxy-L-glycero-D-gluco-octonic acid	5d2a	2.13

(6S)-2,6-anhydro-1-deoxy-6-(2-{[(S)-hydroxy(oxido)lambda~5~-phosphanyl]oxy}ethyl)-D-galactitol + Bound DNA-duplex	5hch	2.9
3,7-anhydro-2,8-dideoxy-L-glycero-D-gluco-octonic acid	5i8m	2.13
3,7-anhydro-2,8-dideoxy-L-glycero-D-gluco-octonic acid	5i8x	1.89
3,7-anhydro-2,8-dideoxy-L-glycero-D-gluco-octonic acid	5nes	1.61
3,7-anhydro-2,8-dideoxy-L-glycero-D-gluco-octonic acid	5ney	1.55
3,7-anhydro-2,8-dideoxy-L-glycero-D-gluco-octonic acid	5nf0	1.27
3,7-anhydro-2,8-dideoxy-L-glycero-D-gluco-octonic acid	5ngq	1.17
β-L-fucose + N-methyl-2-thiophenesulfonamide	5may	1.65
β-L-fucose + N,2,5-trimethyl-3-thiophenesulfonamide	5maz	1.45
β-L-fucose + ~{N},2,4,6 - tetramethyl-benezenesulfonamide	5mb1	1.65

Table 4.4-S7 List of assigned resonances in ^{15}N DC-SIGN CRD.

Assigned resonance ID	^{1}H [ppm]	^{15}N [ppm]
241Leu	7.86	123.52
242Val	7.10	121.48
260Trp	8.43	120.81
261Thr	9.79	119.27
262Phe	9.13	129.22
263Phe	8.82	127.47
264Gln	8.76	127.15
266Asn	7.79	118.38
267Cys	9.40	116.79
268Tyr	9.91	121.05
269Phe	8.54	124.96
270Met	7.79	128.09
271Ser	7.77	120.49
273Ser	7.35	113.65
274Gln	8.32	116.32

275Arg	9.72	123.15
276Asn	9.46	119.69
277Trp	8.22	122.07
278His	7.44	114.82
279Asp	9.03	118.97
280Ser	8.11	124.48
281Ile	7.07	122.05
282Thr	7.34	116.19
283Ala	7.79	123.60
284Cys	7.77	113.18
285Lys	8.27	124.36
286Glu	8.22	118.87
287Val	7.18	108.85
288Gly	7.83	108.91
289Ala	8.32	122.35
290Gln	8.05	117.44
291Leu	8.93	132.33
292Val	7.86	122.73
293Val	7.02	126.95
295Lys	9.05	122.43
296Ser	7.95	116.33
298Glu	9.00	117.46
299Glu	8.01	122.45
300Gln	7.82	121.63
301Asn	8.20	117.05
302Phe	7.47	119.76
304Gln	8.91	122.02
305Leu	7.38	117.77
306Gln	7.20	115.79
307Ser	7.44	113.37
308Ser	8.59	117.80
310Ser	7.12	110.97
311Asn	8.00	118.59
312Arg	7.52	116.49
313Phe	8.54	127.65
315Trp	10.02	127.32

316Met	8.46	113.23
317Gly	9.74	111.28
318Leu	8.27	128.26
319Ser	8.45	118.78
320Asp	7.86	123.58
321Leu	6.65	119.47
322Asn	7.93	115.42
323Gln	7.70	120.63
324Glu	8.06	128.40
326Thr	8.55	121.22
327Trp	8.83	129.05
328Gln	9.08	123.37
329Trp	9.47	128.87
330Val	8.92	115.00
331Asp	7.58	117.81
332Gly	8.57	109.13
333Ser	8.29	119.38
335Leu	7.83	120.43
336Leu	8.76	132.12
339Phe	8.33	120.52
340Lys	7.48	118.84
341Gln	7.30	113.96
342Tyr	6.69	118.67
343Trp	6.47	117.55
344Asn	10.06	123.48
346Gly	8.01	121.43
347Glu	8.16	118.88
352Gly	8.33	112.55
356Cys	8.06	116.52
357Ala	7.99	125.80
358Glu	9.08	116.63
360Ser	9.03	116.35
361Gly	8.97	116.57
363Gly	7.58	113.28
364Trp	8.51	122.38
365Asn	9.13	117.80

366Asp	9.08	121.51
367Asp	9.91	125.47
368Lys	9.61	120.43
369Cys	8.46	119.89
370Asn	7.87	111.75
371Leu	8.07	124.17
372Ala	8.26	122.78
373Lys	8.55	123.68
374Phe	7.27	117.80
375Trp	8.91	117.05
376Ile	6.57	116.99
377Cys	9.03	121.38
378Lys	9.27	123.66
379Lys	9.07	125.34
380Ser	8.77	117.67
381Ala	7.66	123.70
382Ala	9.03	124.56
383Ser	7.74	112.53
386Arg	8.43	121.81
388Glu	8.28	120.03
390Gln	8.15	119.85
392Leu	8.02	123.43
393Ser	8.20	118.01
397Ala	8.38	124.45
398Thr	8.11	115.71

Table 4.4-S8 List of resonance IDs in ^{15}N LecB.

Resonance ID	^{1}H [ppm]	^{15}N [ppm]
1	9.24	136.17
2	9.45	133.97
3	9.37	133.16
4	10.10	131.56
5	9.86	131.19
6	9.79	130.74
7	9.73	130.21
8	8.70	129.77

9	8.25	129.06
10	9.58	128.71
11	9.40	128.35
12	8.60	128.21
13	9.62	128.16
14	8.32	127.51
15	9.00	127.39
16	9.61	126.92
17	8.28	126.58
18	8.36	126.35
19	9.05	126.32
20	8.98	126.33
21	9.16	126.10
22	9.17	125.74
23	9.26	125.40
24	8.88	125.23
25	8.22	124.83
26	8.95	124.92
27	8.46	124.69
28	9.05	124.57
29	9.64	124.16
30	8.02	124.04
31	8.81	123.93
32	8.91	123.78
33	8.40	123.86
34	8.67	123.96
35	9.24	123.39
36	7.88	123.46
37	8.74	123.30
38	7.87	123.28
39	8.36	122.80
40	8.11	122.75
41	8.70	122.64
42	7.93	122.64
43	7.74	122.67
44	9.46	122.48

45	8.97	122.65
46	7.63	122.26
47	9.02	122.05
48	8.63	122.00
49	8.59	121.67
50	8.10	121.66
51	7.81	121.50
52	8.18	121.40
53	8.51	121.09
54	7.79	121.22
55	8.58	121.10
56	9.01	120.79
57	8.01	120.72
58	9.32	120.58
59	8.05	120.19
60	8.11	120.14
61	8.41	119.94
62	7.60	119.64
63	8.87	119.09
64	9.50	118.90
65	8.57	118.68
66	8.28	118.72
67	7.66	118.72
68	6.85	118.31
69	7.28	118.18
70	7.78	117.66
71	7.72	117.28
72	7.93	117.01
73	8.65	116.74
74	7.82	116.60
75	8.61	116.45
76	8.48	116.31
77	7.28	115.18
78	8.53	114.96
79	8.47	115.13
80	8.32	115.13

81	9.54	115.09
82	7.40	113.64
83	7.96	113.62
84	9.27	112.98
85	8.88	112.25
86	6.62	111.58
87	8.80	111.13
88	9.16	110.09
89	8.26	110.24
90	6.93	109.96
91	6.78	109.72
92	9.20	109.56
93	8.73	108.30
94	8.66	103.62
95	9.25	125.49

Table 4.4-S9 List of resonances in ^{15}N LecA.

Resonance ID	^{1}H [ppm]	^{15}N [ppm]
1	8.66	132.29
2	10.44	131.27
3	8.77	131.12
4	8.53	130.16
5	8.69	129.87
6	8.66	129.72
7	9.88	129.63
8	9.13	129.49
9	9.74	129.41
10	8.30	129.43
11	10.36	128.83
12	7.33	128.87
13	7.10	128.52
14	8.97	128.42
15	8.80	128.18
16	8.62	127.56
17	7.68	127.33
18	8.47	127.06

19	8.65	126.79
20	8.88	126.74
21	9.73	126.75
22	7.35	126.63
23	9.31	125.71
24	9.18	125.52
25	9.75	125.26
26	8.92	125.15
27	9.59	124.75
28	8.18	124.62
29	8.36	124.16
30	6.77	123.75
31	8.95	123.71
32	8.31	123.66
33	8.50	123.16
34	8.16	123.21
35	9.01	123.04
36	8.34	122.98
37	8.97	122.99
38	7.59	122.98
39	8.64	122.80
40	7.42	122.79
41	9.14	122.45
42	6.54	122.36
43	8.67	122.16
44	9.41	122.10
45	7.08	121.93
46	8.13	121.72
47	9.33	121.44
48	8.59	121.36
49	7.15	121.38
50	9.94	121.11
51	8.89	121.03
52	8.21	120.48
53	8.14	120.42
54	9.52	120.32

55	7.40	120.06
56	9.28	119.84
57	7.82	119.77
58	7.77	119.58
59	7.24	119.30
60	9.05	119.14
61	8.48	118.90
62	8.03	118.54
63	8.75	118.52
64	7.39	117.89
65	8.24	117.74
66	8.05	117.72
67	7.89	117.71
68	9.18	117.54
69	8.84	117.49
70	7.48	117.42
71	8.34	117.31
72	6.97	116.86
73	7.73	116.60
74	8.90	116.54
75	9.12	116.20
76	8.84	115.56
77	9.18	115.39
78	7.15	115.03
79	8.42	114.46
80	8.07	114.36
81	9.28	113.86
82	9.04	113.55
83	6.96	112.87
84	8.09	112.85
85	8.43	112.77
86	6.87	112.58
87	7.30	113.51
88	9.42	112.48
89	6.73	112.54
90	6.88	112.39

91	7.30	112.24
92	6.78	112.25
93	7.47	112.00
94	7.19	111.28
95	6.72	111.38
96	7.28	111.04
97	6.62	110.91
98	8.30	110.95
99	8.54	110.59
100	7.55	110.45
101	6.96	110.48
102	8.76	110.05
103	8.21	109.95
104	8.97	109.13
105	8.94	108.67
106	9.01	108.20
107	8.45	107.41
108	7.79	106.90
109	8.50	106.77
110	8.16	104.95
111	6.91	104.43
112	8.38	120.11
113	7.71	122.25
114	6.74	105.68
115	6.78	113.52

4.4.5. Supplementary Schemes

A

	R		R			R	Yield	R		Yield
a = Ph(CH₂)₃-		f = 4-MeOPh(CH₂)₃-			a	35	(60%)	f	40	(38%)

a = Ph(CH₂)₃- f = 4-MeOPh(CH₂)₃- a **35** (60%) f **40** (38%)
b = Ph(CH₂)₂OCH₂- g = 4-HOPh(CH₂)₃- b **43** (36%) g **41** (22%)
c = Ph(CH₂)₃OCH₂- h = 4-NO₂Ph(CH₂)₃- c **44** (35%) h **39** (77%)
d = CyCH₂- i = 2-Thionyl(CH₂)₃- d **6** (29%) i **37** (29%)
e = 4-MePh(CH₂)₃- j = PhCHCHCH₂- e **36** (82%) j **42** (40%)

B

7 (61%)   **S2**   **8** (32%)
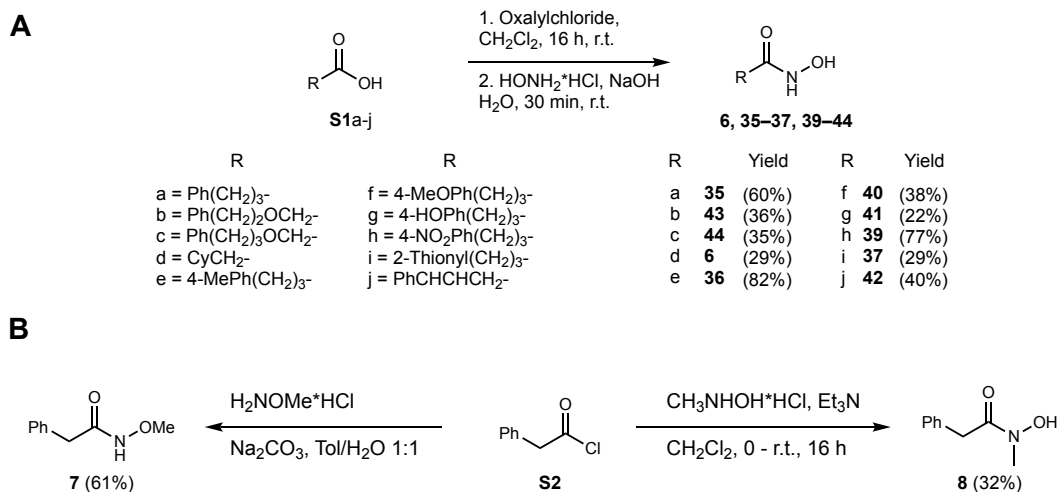
Scheme 4.4-1: Synthetic modifications of acids/ acyl chlorides for generating a library of hydroxamic acids: (A) Synthesis of hydroxamic acid derivatives. (B) Modifications at the hydroxamic acid functional group

A   **B**
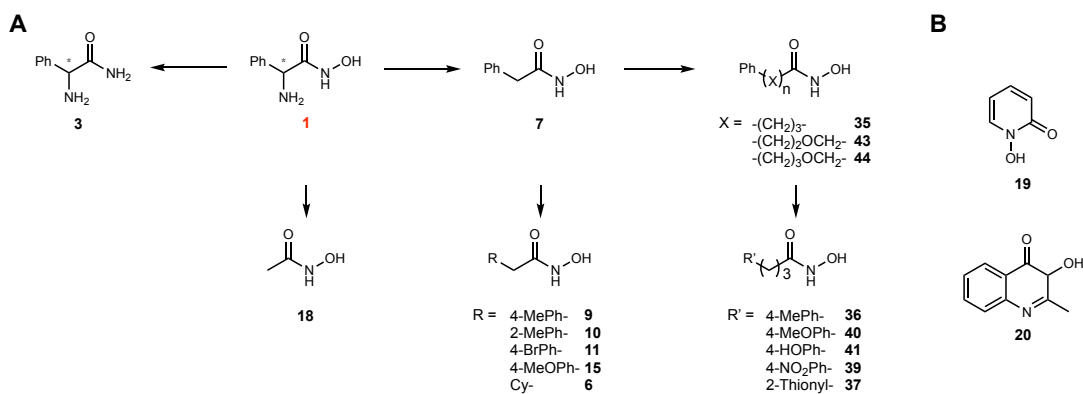
Scheme 4.4-2: Structures of the hydroxamic acid library (A) Optimization of the initial hit 1 leading to the *N*-hydroxy-4-phenylbutanamide series. (B) Structures with cyclic hydroxamic acid functional groups.

4.4.6. Supplementary References

1.   Harner, M. J.; Frank, A. O.; Fesik, S. W., Fragment-based drug discovery using NMR spectroscopy. *Journal of biomolecular NMR* **2013,** *56* (2), 65-75.

2.   Chen, A. Y.; Adamek, R. N.; Dick, B. L.; Credille, C. V.; Morrison, C. N.; Cohen, S. M., Targeting Metalloenzymes for Therapeutic Intervention. *Chemical reviews* **2019,** *119* (2), 1323-1455.

3.   Davis, B. J.; Erlanson, D. A., Learning from our mistakes: the 'unknown knowns' in fragment screening. *Bioorganic & medicinal chemistry letters* **2013,** *23* (10), 2844-52.

4.   Hermann, J. C.; Chen, Y.; Wartchow, C.; Menke, J.; Gao, L.; Gleason, S. K.; Haynes, N.-E.; Scott, N.; Petersen, A.; Gabriel, S.; Vu, B.; George, K. M.; Narayanan, A.; Li, S. H.; Qian, H.; Beatini, N.; Niu, L.; Gan, Q.-F. Metal impurities cause false positives in high-throughput screening campaigns *ACS Med Chem Lett* [Online], 2013, p. 197-200. PubMed.

5.   Aretz, J.; Baukmann, H.; Shanina, E.; Hanske, J.; Wawrzinek, R.; Zapol'skii, V. A.; Seeberger, P. H.; Kaufmann, D. E.; Rademacher, C., Identification of Multiple Druggable Secondary Sites by Fragment Screening against DC-SIGN. *Angewandte Chemie International Edition* **2017,** *56* (25), 7292-7296.

6.   Aretz, J.; Anumala, U. R.; Fuchsberger, F. F.; Molavi, N.; Ziebart, N.; Zhang, H.; Nazaré, M.; Rademacher, C., Allosteric Inhibition of a Mammalian Lectin. *Journal of the American Chemical Society* **2018,** *140* (44), 14915-14925.

7.   Shanina E., K. S., Lal K., Seeberger P.H., Imberty A. and Rademacher C., Identification of druggable allosteric pockets in β-propeller lectins. *in submission* **2021**.

8.   Denavit, V.; Laine, D.; Bouzriba, C.; Shanina, E.; Gillon, E.; Fortin, S.; Rademacher, C.; Imberty, A.; Giguere, D., Stereoselective Synthesis of Fluorinated Galactopyranosides as Potential Molecular Probes for Galactophilic Proteins: Assessment of Monofluorogalactoside-LecA Interactions. *Chemistry (Weinheim an der Bergstrasse, Germany)* **2019,** *25* (17), 4478-4490.

9.   Sommer, R.; Hauck, D.; Varrot, A.; Wagner, S.; Audfray, A.; Prestel, A.; Möller, H. M.; Imberty, A.; Titz, A., Cinnamide Derivatives of d-Mannose as Inhibitors of the Bacterial Virulence Factor LecB from Pseudomonas aeruginosa. *ChemistryOpen* **2015,** *4* (6), 756-767.

10.   Audfray, A.; Claudinon, J.; Abounit, S.; Ruvoën-Clouet, N.; Larson, G.; Smith, D. F.; Wimmerová, M.; Le Pendu, J.; Römer, W.; Varrot, A.; Imberty, A., Fucose-binding lectin from opportunistic pathogen Burkholderia ambifaria binds to both plant and human oligosaccharidic epitopes. *The Journal of biological chemistry* **2012,** *287* (6), 4335-4347.

11.   Hanske, J.; Aleksić, S.; Ballaschk, M.; Jurk, M.; Shanina, E.; Beerbaum, M.; Schmieder, P.; Keller, B. G.; Rademacher, C., Intradomain Allosteric Network Modulates Calcium Affinity of the C-Type Lectin Receptor Langerin. *Journal of the American Chemical Society* **2016,** *138* (37), 12176-12186.

12.   Bietz, S.; Urbaczek, S.; Schulz, B.; Rarey, M., Protoss: a holistic approach to predict tautomers and protonation states in protein-ligand complexes. *Journal of Cheminformatics* **2014,** *6* (1), 12.

13.   Jain, A. N., Surflex: fully automatic flexible molecular docking using a molecular similarity-based search engine. *J Med Chem* **2003,** *46* (4), 499-511.

14.   Desaphy, J.; Raimbaud, E.; Ducrot, P.; Rognan, D., Encoding protein-ligand interaction patterns in fingerprints and graphs. *Journal of chemical information and modeling* **2013,** *53* (3), 623-37.

15.   Hassan, M.; Brown, R. D.; Varma-O'brien, S.; Rogers, D., Cheminformatics analysis and learning in a data pipelining environment. *Molecular diversity* **2006,** *10* (3), 283-99.

16.   Aretz, J.; Wamhoff, E.-C.; Hanske, J.; Heymann, D.; Rademacher, C., Computational and Experimental Prediction of Human C-Type Lectin Receptor Druggability. *Frontiers in Immunology* **2014,** *5* (323).

17.   Troelsen, N. S.; Shanina, E.; Gonzalez-Romero, D.; Danková, D.; Jensen, I. S. A.; Śniady, K. J.; Nami, F.; Zhang, H.; Rademacher, C.; Cuenda, A.; Gotfredsen, C. H.; Clausen, M. H., The 3F Library: Fluorinated Fsp3-Rich Fragments for Expeditious 19F NMR Based Screening. *Angewandte Chemie International Edition* **2020,** *59* (6), 2204-2210.

18.   Delaglio, F.; Grzesiek, S.; Vuister, G. W.; Zhu, G.; Pfeifer, J.; Bax, A., NMRPipe: A multidimensional spectral processing system based on UNIX pipes. *Journal of biomolecular NMR* **1995,** *6* (3), 277-293.

19.   Vranken, W. F.; Boucher, W.; Stevens, T. J.; Fogh, R. H.; Pajon, A.; Llinas, M.; Ulrich, E. L.; Markley, J. L.; Ionides, J.; Laue, E. D., The CCPN data model for NMR spectroscopy: Development of a software pipeline. *Proteins: Structure, Function, and Bioinformatics* **2005,** *59* (4), 687-696.

20.   Williamson, M. P., Using chemical shift perturbation to characterise ligand binding. *Progress in nuclear magnetic resonance spectroscopy* **2013,** *73*, 1-16.

21.   Joachim, I.; Rikker, S.; Hauck, D.; Ponader, D.; Boden, S.; Sommer, R.; Hartmann, L.; Titz, A., Development and optimization of a competitive binding assay for the galactophilic low affinity lectin LecA from Pseudomonas aeruginosa. *Organic & biomolecular chemistry* **2016,** *14* (33), 7933-48.

22.   Gasteiger, E.; Gattiker, A.; Hoogland, C.; Ivanyi, I.; Appel, R. D.; Bairoch, A., ExPASy: The proteomics server for in-depth protein knowledge and analysis. *Nucleic acids research* **2003,** *31* (13), 3784-8.

23.   Kuhaudomlarp, S.; Siebs, E.; Shanina, E.; Topin, J.; Joachim, I.; da Silva Figueiredo Celestino Gomes, P.; Varrot, A.; Rognan, D.; Rademacher, C.; Imberty, A.; Titz, A., Non-Carbohydrate Glycomimetics as Inhibitors of Calcium(II)-binding Lectins. *Angewandte Chemie (International ed. in English)* **2020**.

24.   Shanina, E.; Siebs, E.; Zhang, H.; Varón Silva, D.; Joachim, I.; Titz, A.; Rademacher, C., Protein-observed 19F NMR of LecA from Pseudomonas aeruginosa. *Glycobiology* **2021**.

25.   Kabsch, W., XDS. *Acta crystallographica. Section D, Biological crystallography* **2010,** *66* (Pt 2), 125-32.

26.   Evans, P. R., An introduction to data reduction: space-group determination, scaling and intensity statistics. *Acta crystallographica. Section D, Biological crystallography* **2011,** *67* (Pt 4), 282-92.

27.   McCoy, A. J., Solving structures of protein complexes by molecular replacement with Phaser. *Acta crystallographica. Section D, Biological crystallography* **2007,** *63* (Pt 1), 32-41.

28.   Emsley, P.; Lohkamp, B.; Scott, W. G.; Cowtan, K., Features and development of Coot. *Acta crystallographica. Section D, Biological crystallography* **2010,** *66* (Pt 4), 486-501.

29.   Murshudov, G. N.; Skubak, P.; Lebedev, A. A.; Pannu, N. S.; Steiner, R. A.; Nicholls, R. A.; Winn, M. D.; Long, F.; Vagin, A. A., REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallographica Section D* **2011,** *67* (4), 355-367.

30.   Long, F.; Nicholls, R. A.; Emsley, P.; Graæulis, S.; Merkys, A.; Vaitkus, A.; Murshudov, G. N., AceDRG: a stereochemical description generator for ligands. *Acta Crystallogr D Struct Biol* **2017,** *73* (Pt 2), 112-122.

31.   Potterton, L.; Agirre, J.; Ballard, C.; Cowtan, K.; Dodson, E.; Evans, P. R.; Jenkins, H. T.; Keegan, R.; Krissinel, E.; Stevenson, K.; Lebedev, A.; McNicholas, S. J.; Nicholls, R. A.; Noble, M.; Pannu, N. S.; Roth, C.; Sheldrick, G.; Skubak, P.; Turkenburg, J.; Uski, V.; von Delft, F.; Waterman, D.; Wilson, K.; Winn, M.; Wojdyr, M., CCP4i2: the new graphical user interface to the CCP4 program suite. *Acta Crystallogr D Struct Biol* **2018,** *74* (Pt 2), 68-84.

32.   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.; Richardson, D. C., MolProbity: all-

atom structure validation for macromolecular crystallography. *Acta crystallographica. Section D, Biological crystallography* **2010,** *66* (Pt 1), 12-21.

33.   McNicholas, S.; Potterton, E.; Wilson, K. S.; Noble, M. E., Presenting your structures: the CCP4mg molecular-graphics software. *Acta crystallographica. Section D, Biological crystallography* **2011,** *67* (Pt 4), 386-94.

34.   Schulze, J.; Baukmann, H.; Wawrzinek, R.; Fuchsberger, F. F.; Specker, E.; Aretz, J.; Nazaré, M.; Rademacher, C., CellFy: A Cell-Based Fragment Screen against C-Type Lectins. *ACS chemical biology* **2018,** *13* (12), 3229-3235.

35.   Fulmer, G. R.; Miller, A. J. M.; Sherden, N. H.; Gottlieb, H. E.; Nudelman, A.; Stoltz, B. M.; Bercaw, J. E.; Goldberg, K. I., NMR Chemical Shifts of Trace Impurities: Common Laboratory Solvents, Organics, and Gases in Deuterated Solvents Relevant to the Organometallic Chemist. *Organometallics* **2010**, *29* (9), 2176-2179.

36.   Xiao, Z.-P.; Peng, Z.-Y.; Dong, J.-J.; Deng, R.-C.; Wang, X.-D.; Ouyang, H.; Yang, P.; He, J.; Wang, Y.-F.; Zhu, M.; Peng, X.-C.; Peng, W.-X.; Zhu, H.-L., Synthesis, molecular docking and kinetic properties of β-hydroxy-β-phenylpropionyl-hydroxamic acids as Helicobacter pylori urease inhibitors. *Eur J Med Chem* **2013**, *68*, 212-221.

37.   Ohtsuka, N.; Okuno, M.; Hoshino, Y.; Honda, K., A base-mediated self-propagative Lossen rearrangement of hydroxamic acids for the efficient and facile synthesis of aromatic and aliphatic primary amines. *Organic & Biomolecular Chemistry* **2016**, *14* (38), 9046-9054.
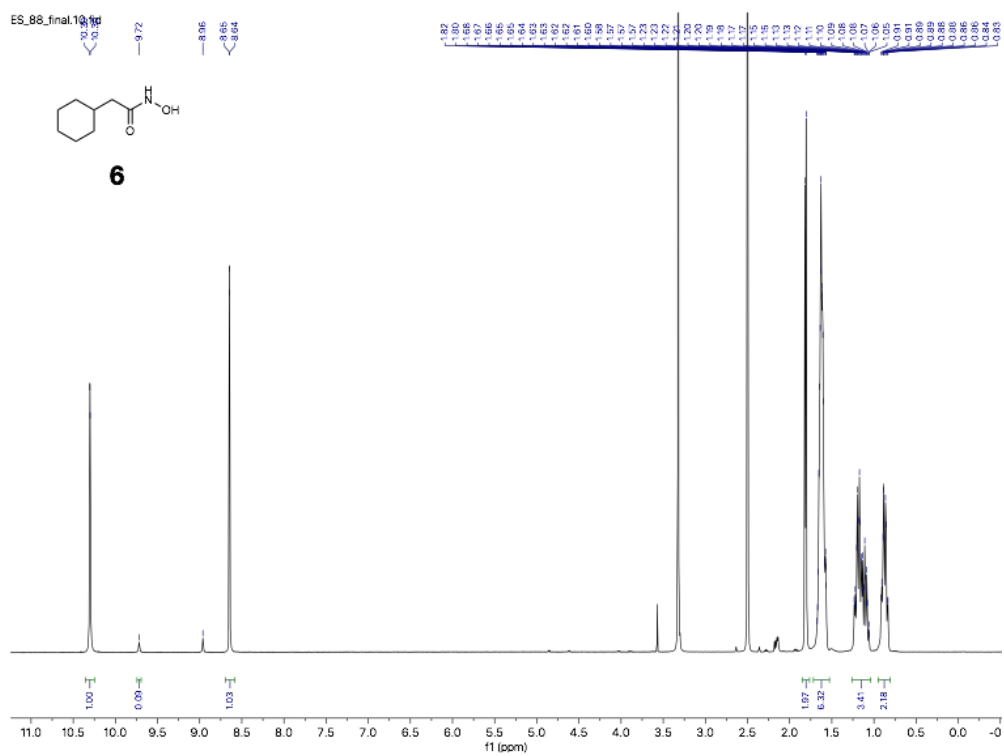
38.   Hermant, P.; Bosc, D.; Piveteau, C.; Gealageas, R.; Lam, B.; Ronco, C.; Roignant, M.; Tolojanahary, H.; Jean, L.; Renard, P. Y.; Lemdani, M.; Bourotte, M.; Herledan, A.; Bedart, C.; Biela, A.; Leroux, F.; Deprez, B.; Deprez-Poulain, R., Controlling Plasma Stability of Hydroxamic Acids: A MedChem Toolbox. *J Med Chem* **2017**, *60* (21), 9067-9089.

39.   Kawase, M.; Kitamura, T.; Kikugawa, Y., Electrophilic aromatic substitution with N-methoxy-N-acylnitrenium ions generated from N-chloro-N-methoxy amides: syntheses of nitrogen heterocyclic compounds bearing a N-methoxy amide group. *The Journal of Organic Chemistry* **1989**, *54* (14), 3394-3403.
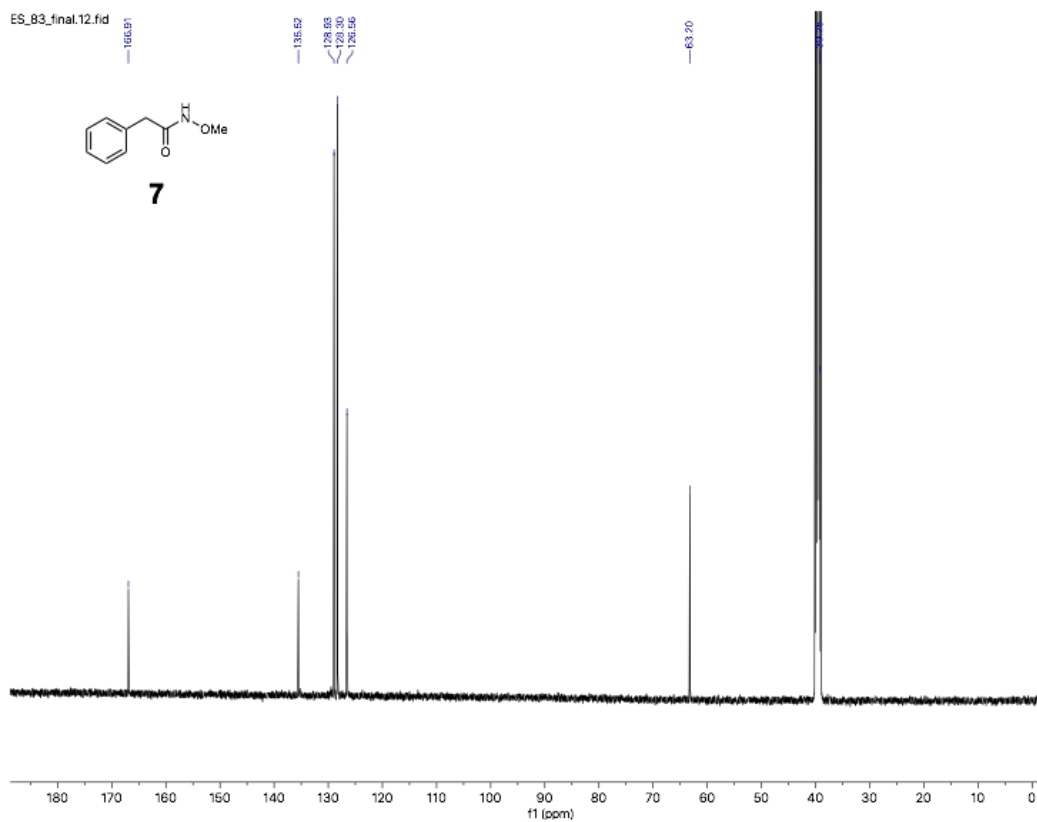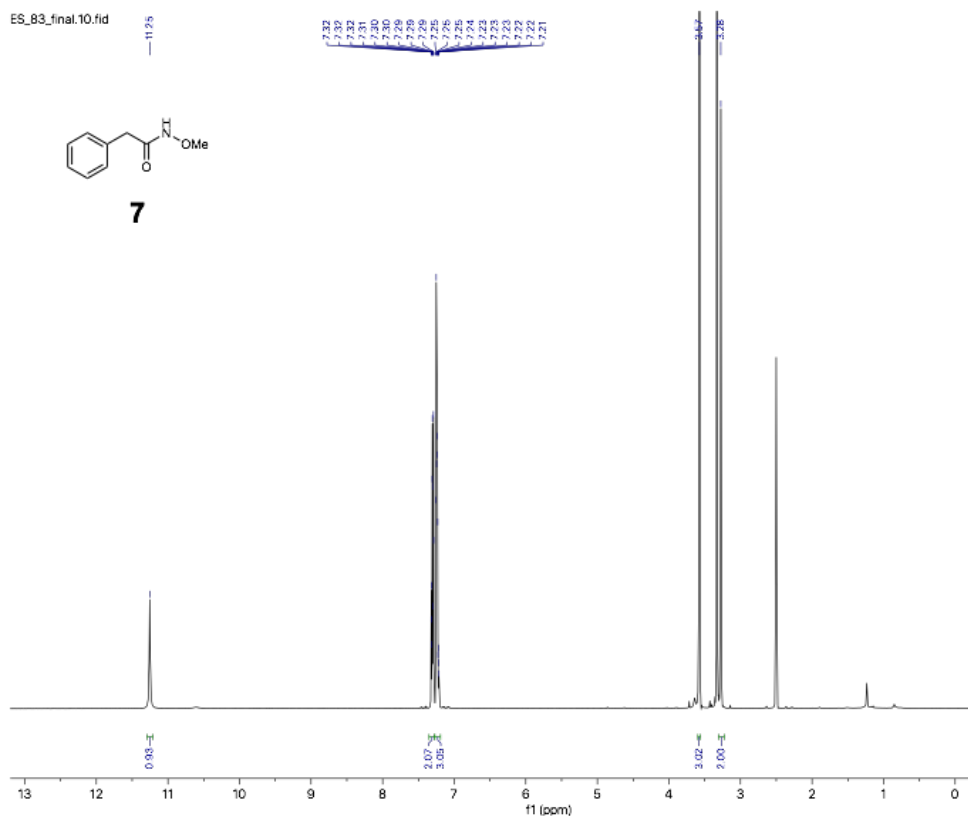
40.   Kukosha, T.; Trufilkina, N.; Belyakov, S.; Katkevics, M., Copper-Catalyzed Cross-Coupling of O-Alkyl Hydroxamates with Aryl Iodides. *Synthesis* **2012**, *44* (15), 2413-2423.

41.   Clark, A. J.; Al-Faiyz, Y. S. S.; Broadhurst, M. J.; Patel, D.; Peacock, J. L., Base catalysed rearrangement of N-alkyl-O-acyl hydroxamic acids: synthesis of 2-acyloxyamides. *Journal of the Chemical Society, Perkin Transactions 1* **2000**, (7), 1117-1127.
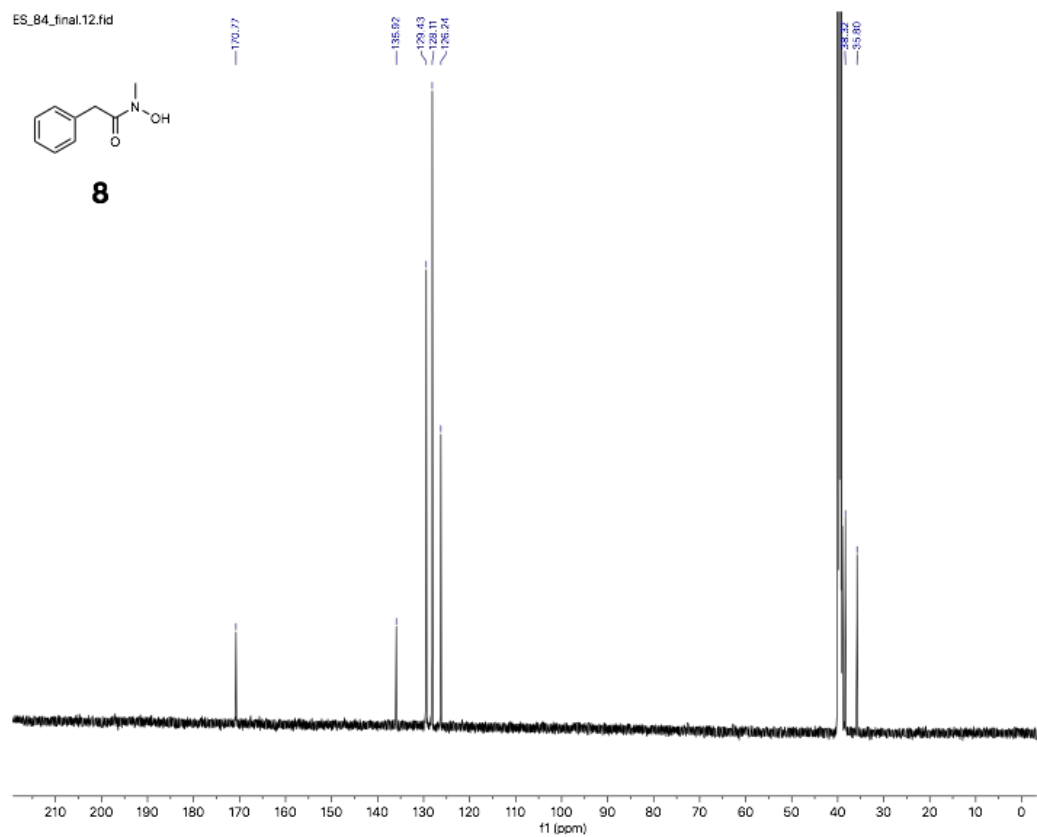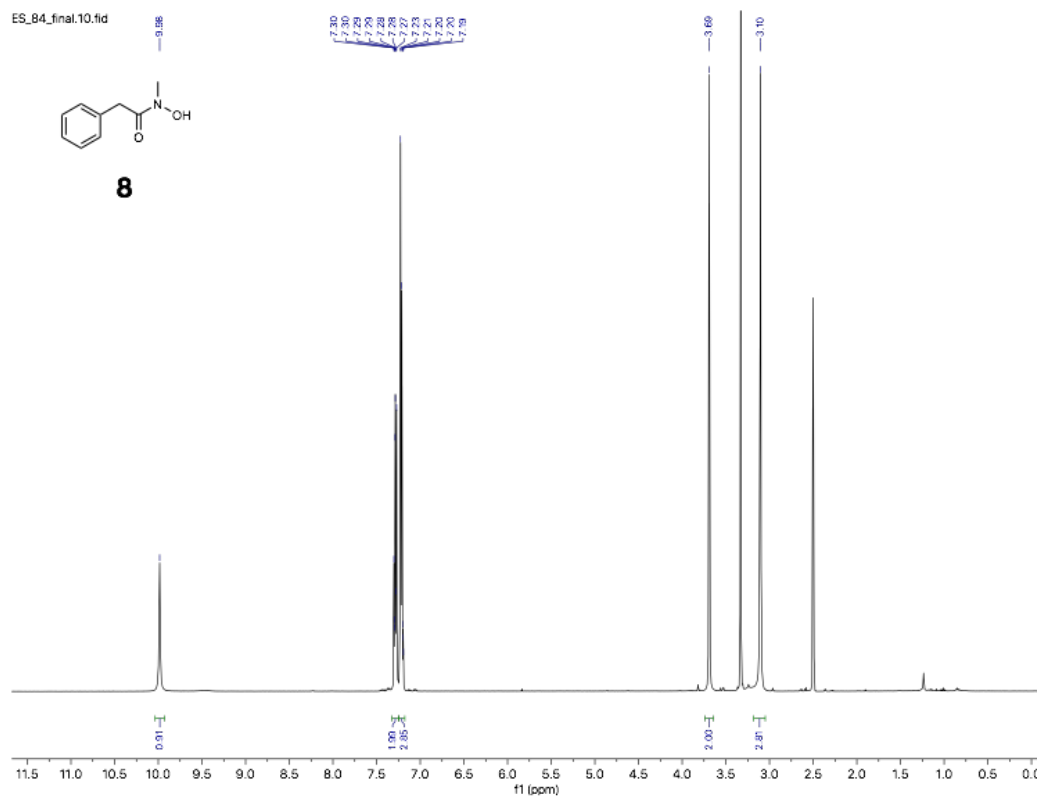
42.   Trabulsi, H.; Guillot, R.; Rousseau, G., Preparation of Imino Lactones by Electrophilic Cyclization of β,γ-Unsaturated Hydroxamates: Formation of 3-Cyanoprop-2-en-1-ones through Fragmentation Reactions. *European Journal of Organic Chemistry* **2010**, *2010* (30), 5884-5896.

43.   A. Publication, *Org. Synth.* **1969**, *49*, 50.

4.4.7. ^{1}H NMR
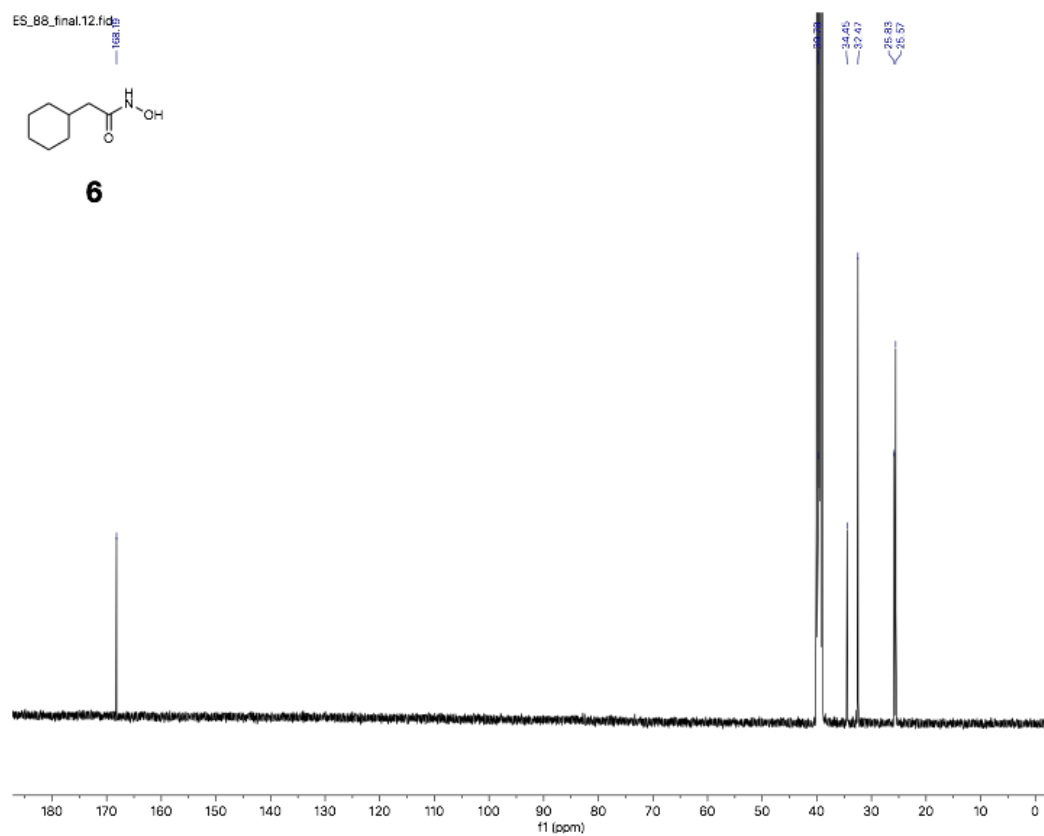
ES_83_final.10.fid
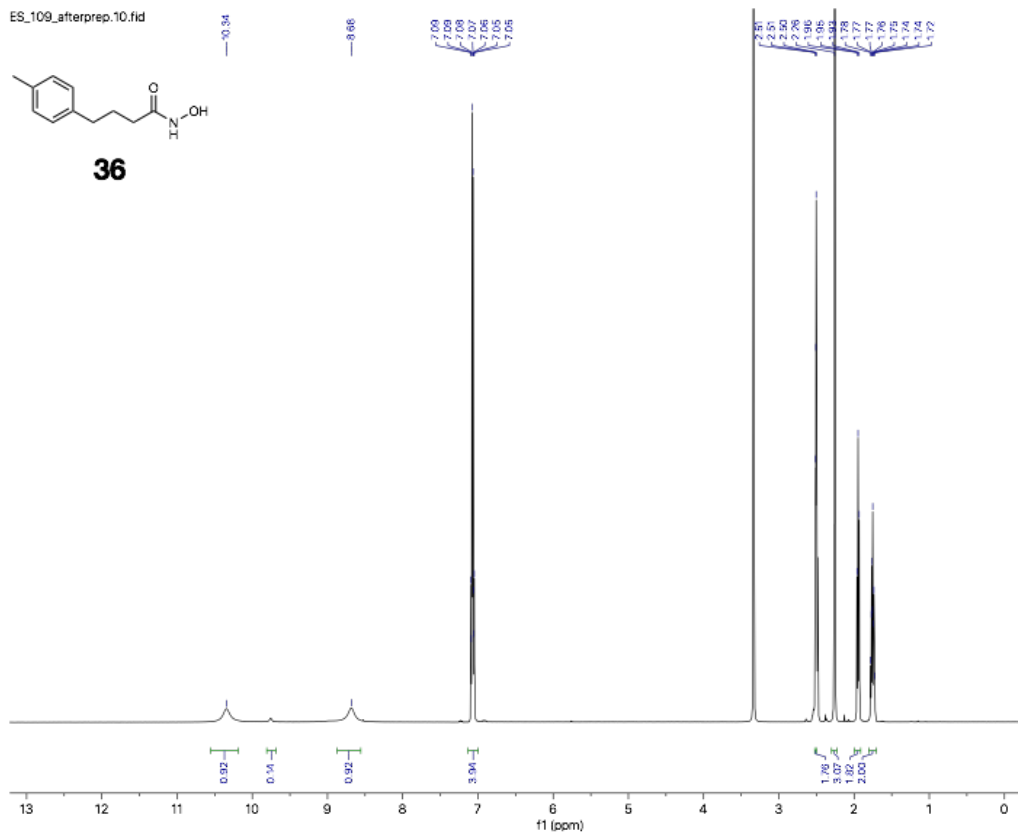
7

ES_83_final.12.fid

7

ES_84_final.10.fid

9.98

7.30 7.30 7.29 7.29 7.28 7.27 7.23 7.21 7.20 7.20 7.19

3.69

3.10

11.5 11.0 10.5 10.0 9.5 9.0 8.5 8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5 0.0
f1 (ppm)

ES_84_final.12.fid

170.77

135.92

129.43 128.11 126.24

38.52 35.80

210 200 190 180 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 0
f1 (ppm)

ES_118.10.fid

35

ES_118.11.fid

35

ES_109_afterprep.10.fid

36

ES_109_afterprep.12.fid

36

ES_110.10.fid

10.37

8.69

7.32
7.31
6.94
6.93
6.85
6.84

2.77

2.01
1.99
1.92

37

1.79   0.21   1.34   1.10 1.09 1.13   1.99   2.00 2.79

11.5 11.0 10.5 10.0 9.5 9.0 8.5 8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5 0.0
f1 (ppm)

ES_110.12.fid

169.13

144.65

127.40
125.01
124.03

32.00
29.10
27.63

37

210 200 190 180 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 0
f1 (ppm)

ES_113_afterprep.10.fid

38

ES_113_afterprep.12.fid

38

40

40

ES_129.10.fid

10.33

9.12

8.66

6.96
6.94
6.67
6.65

2.44
2.42
2.41
1.94
1.91
1.74
1.73
1.72
1.71
1.70
1.68

41

| 1.76 | 0.16 | 1.06 0.13 | 1.00 | 1.98 | 2.00 | 1.86 | 1.93 | 2.00 |

11.5 11.0 10.5 10.0 9.5 9.0 8.5 8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5 0.0
f1 (ppm)

ES_129.12.fid

168.97

155.35

131.67
129.12

115.04

33.78
31.77

27.31

41

180 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 0
f1 (ppm)

ES_90_final.10.fid

42

ES_90_final.12.fid

42

ES_95_final.10.fid

ES_95_final.12.fid

ES_96_final.10.fid

44

ES_96_final.20.fid

44