

Developing Novel Glycomimetics Targeted to Bacterial Lectins LecA, LecB and BambL

Dissertation

Zur Erlangung des Grades des Doktors der Naturwissenschaften
der Naturwissenschaftlich-Technischen Fakultät
der Universität des Saarlandes

vorgelegt von

M. Sc. Eike Siebs

Saarbrücken

2022

Tag des Kolloquiums:	6. Dezember 2022
Dekan:	Prof. Dr. Ludger Santen
Berichterstatter:	Prof. Dr. Alexander Titz Prof. Dr. Uli Kazmaier
Vorsitz:	Prof. Dr. Anna K. H. Hirsch
Akad. Beisitzer:	PD Dr. Martin Frotscher

Die vorliegende Arbeit wurde von März 2018 bis März 2022 unter Anleitung von Herrn Univ.-Prof. Dr. Alexander Titz in der Fachrichtung Chemie der Naturwissenschaftlich-Technischen Fakultät der Universität des Saarlandes, sowie am Helmholtz-Institut für Pharmazeutische Forschung Saarland (HIPS) in der Arbeitsgruppe Chemische Biologie der Kohlenhydrate (CBCH) angefertigt.

Acknowledgement

Throughout the writing of this dissertation I have received a great deal of support and assistance.

My greatest thanks goes to Prof. Dr. Alexander Titz whose insightful feedback was invaluable and sharpened my scientific thinking and brought my work to a higher level.

I would like to thank the members of my PhD-Committee, Prof. Dr. Uli Kazmaier and Priv. Doz. Dr. Martin Frotscher, for the scientific discussions in our meetings and the evaluation of this thesis.

Great thanks to the GlycoMimeTeam, Prof. Dr. Anne Imberty, Prof. Dr. Christoph Rademacher, Prof. Dr. Didier Rognan, Dr. Sarkowan Kuhaudomlarp, Dr. Priscila da Silva Figueiredo Celestino Gomes and M.Sc. Elena Shanina for their suggestions, laughs, guidance and importantly the great collaboration.

In addition, I wish to thank Dr. Uwe Grether, Dr. Christian Lerner, Dr. Larissa Valeska Thiele, Dr. Bernd Kuhn and their Roche team for making one of my projects feasible.

Further, I would like to thank, Dr. Katharina Rox, Dr. Annabelle Varrot, M. Sc. Joscha Meiers and my students for their scientific inputs to my thesis.

I would like to thank the whole Chemical Biology of Carbohydrate Team for their support, discussions and adventures, especially the people from my office: Ghamdan Beshr, Eva Zahorska, Varvara Verkhova, Patrycja Mała, Marta Czekanska, and Aketza Romaniega Bilbao for all the time we spent together.

I gratefully acknowledge Dirk Hauk for maintaining the laboratory, the information technology team Michael Roth and Mark Caspari, the facility managers Frank Jacob and Christian Zeuner, and our secretaries Julia Mohr, Annette Herkströter, Lisa Marie Andre and Bahareh Kadkhodazadeh.

Special thanks goes to Sandra Johannsen for always having my back.

Zusammenfassung

Diese Dissertation beschreibt die Entwicklung neuer, (nicht)-kohlenhydratbasierter Lektininhibitoren als Grundlage für Glycomimetika. ESKAPE-Erreger *Pseudomonas aeruginosa* und *Burkholderia ambifaria* gehören beide zu den gramnegativen Bakterien, die in Mukoviszidose-Patienten gravierende Einflüsse haben, die sogar zu deren Tod führen können. Beide Bakterienarten exprimieren verschiedene kohlenhydratbindende Proteine, die so genannten Lektine. Diese sind zum Teil bei der Adhärenz und der Biofilmbildung involviert oder dienen als Virulenzfaktor. *P. aeruginosa* exprimiert zwei Calciumionen-abhängige Lektine, LecA und LecB, während *B. ambifaria* ein metallfreies Lektin, BambL, exprimiert. In dieser Arbeit wurden verschiedene Strategien verwendet, um diese Lektine mit neuen Liganden in ihrer Funktion zu inhibieren, dabei wurde entweder vom virtuellen Screening oder von einem Liganden, einer bekannten Co-Kristallstruktur des gewünschten Lektins, ausgegangen. Die neuen Liganden-Hits wurden anschließend in verschiedenen biophysikalischen Assays studiert und validiert. Dies führte zu neuen, zum Teil deutlich verbesserten (nicht)-kohlenhydratbasierten Lektininhibitoren.

Summary

This dissertation describes the study of new, (non)-carbohydrate based lectin inhibitors as basis for the development of glycomimetics. The ESKAPE pathogens *Pseudomonas aeruginosa* and *Burkholderia ambifaria* are both Gram-negative bacteria that are especially dangerous for cystic fibrosis patients and are prominently involved in the patients' deaths. These bacteria express lectins, which are carbohydrate binding proteins, that are partially involved in adherence, biofilm formation or serve as a virulence factor. *P. aeruginosa* expresses two calcium-ion dependent lectins, LecA and LecB, whereas BambL, a non-metalated lectin, belongs to *B. ambifaria*. In this work, various strategies were used for addressing these lectins with newly developed ligands for inhibiting their functions. The development started by either virtual screening or from a ligand of a known co-crystal structure of the desired lectin. The new ligand-hits were then further studied and validated in different biophysical assays leading to new, improved (non)-carbohydrate based lectin inhibitors.

Table of Contents

Acknowledgement	I
Zusammenfassung	II
Summary	II
Table of Contents	III
Publications Included in this Dissertation	V
Contribution Report	VII
Abbreviations	X
Introduction	1
A Brief History of Antibiotics	1
Bacterial Adherence and Biofilm Formation	1
Antibiotic Effects on Biofilms	2
<i>Pseudomonas aeruginosa</i> and <i>Burkholderia ambifaria</i> in Cystic Fibrosis	3
New Antimicrobial Strategies against <i>P. aeruginosa</i>	4
Lectins LecA, LecB and BambL	4
Lectin Antagonists in Infection, Immunity, and Inflammation	9
Results	26
Non-Carbohydrate Glycomimetics as Inhibitors of Calcium(II)-binding Lectins	26
Tolcapone-like Glycomimetics as Potent LecA Inhibitors	37
Targeting the Central Pocket of The <i>Pseudomonas aeruginosa</i> Lectin LecA	57
Discovery of <i>N</i> - β -L-Fucosyl Amides as High-Affinity Ligands for the <i>Pseudomonas aeruginosa</i> Lectin LecB	65
Ligand-Based Design of <i>Burkholderia ambifaria</i> Lectin BambL Inhibitors	92
Conclusion and Outlook	106
Supporting Information	109
Supporting Information to Non-Carbohydrate Glycomimetics as Inhibitors of Calcium(II)-Binding Lectin	109

Supporting Information to Tolcapone-like Glycomimetics as Potent LecA Inhibitors	142
Supporting Information to Targeting the Central Pocket of the <i>Pseudomonas aeruginosa</i> Lectin LecA	154
Supporting Information to Discovery of <i>N</i> - β -Fucosyl Amides as High-Affinity Ligands for the <i>Pseudomonas aeruginosa</i> Lectin LecB	221
Supporting Information to Ligand-Based Design for Targeting <i>Burkholderia ambifaria</i> Lectin BamBL	233
References	254
Curriculum Vitae	264

Publications Included in this Dissertation

One publication supports the introduction and three publications are in the result chapter of which two were published and one has recently been submitted.

Lectin Antagonists in Infection, Immunity, and Inflammation

Joscha Meiers#, Eike Siebs#, Eva Zahorska#, Alexander Titz

Ref.: *Curr. Opin. Chem. Biol.* **2019**, 53, 51–67.

Link: <https://doi.org/10.1016/j.cbpa.2019.07.005>.

#shared first authorship

Non-Carbohydrate Glycomimetics as Inhibitors of Calcium(II)-Binding Lectins

Sakonwan Kuhaudomlarp#, Eike Siebs#, Elena Shanina, Jérémie Topin, Ines Joachim, Priscila da Silva Figueiredo Celestino Gomes, Annabelle Varrot, Didier Rognan, Christoph Rademacher, Anne Imberty, Alexander Titz

Ref.: *Angew. Chem. Int. Ed.* **2021**, 60, 8104–8114.

Link: <https://doi.org/10.1002/anie.202013217>.

#shared first authorship

Targeting the Central Pocket of the *Pseudomonas aeruginosa* Lectin LecA

Eike Siebs, Elena Shanina, Sakonwan Kuhaudomlarp, Priscila da Silva Figueiredo Celestino Gomes, Cloé Fortin, Didier Rognan, Christoph Rademacher, Anne Imberty, and Alexander Titz

Ref.: *ChemBioChem.* **2022**, 23, e202100563.

Link: <https://doi.org/10.1002/cbic.202100563>.

Discovery of *N*- β -L-Fucosyl Amides as High-Affinity Ligands for the *Pseudomonas aeruginosa* Lectin LecB

Patrycja Mała#, Eike Siebs#, Joscha Meiers, Katharina Rox, Annabelle Varrot, Anne Imberty, Alexander Titz

Submitted to: *J. Med. Chem.*

Status: Returned with minor revisions.

#shared first authorship

Contribution Report

Lectin antagonists in infection, immunity, and inflammation

In this review, Eike Siebs is a shared first author together with Joscha Meiers and Eva Zahorska. He wrote the mammalian, Eva Zahorska the bacterial and Joscha Meiers the viral lectin part under the revision of all co-authors. Alexander Titz wrote the introduction and supervised/guided this publication.

Non-Carbohydrate Glycomimetics as Inhibitors of Calcium(II)-Binding Lectins

Eike Siebs is a shared first author together with Sakonwan Kuhaudomlarp in this “Angewandte Chemie” publication. He contributed by synthesizing the necessary fluorescent dye for the initial experimental screening in the competitive binding assay based on fluorescence polarization (FP). He expressed LecA, performed the FP and the thermal shift assay experiments and he wrote together with Alexander Titz several sections in this publication under the revision of all co-authors. Sakonwan Kuhaudomlarp was together with Annabelle Varrot responsible for the x-ray studies, she performed the SPR experiments and wrote the corresponding chapters. Elena Shanina performed the 19F PrOF and TROSY, and Ines Joachim the T1ρ-relaxations NMR experiments. Jérémie Topin performed the initial *in silico* experiments. Anne Imberty wrote together with Alexander Titz the introduction, and Didier Rognan, Christoph Rademacher and Priscila da Silva Figueiredo Celestino Gomes were improving the publication with their opinions and feedback. Christoph Rademacher also supervised the experiments from Elena Shanina.

Tolcapone-like Glycomimetics as Potent LecA Inhibitors

Eike Siebs performed the screening studies in the competitive binding assay based on fluorescence polarization of the dose-dependent (initial screening) and the single point measurements of the catechol-like library provided by Roche. He performed the SAR and wrote the whole chapter. Sakonwan Kuhaudomlarp performed the surface plasmon resonance experiments and together with Annabelle Varrot the x-ray crystallography studies including the data analysis and refinements of the crystal structure. Elena Shanina was responsible for the protein overserved 19F NMR experiments. Didier Rognan performed the *in silico* studies. Uwe Grether, Christian Lerner, Bernd Kuhn and Larissa Valeska together with their team (Isabella Kaufmann,

Alexandre Zimmermann, Philippe Hartz) prepared and provided the catechol-like library of 3564 compounds on 384-well plates in four replicates each. Alexander Titz supervised this project and gave support during the experiments and input to this chapter. Christoph Rademacher and Anne Imberty supervised Elena and Sakonwan.

Targeting the Central Pocket of the *Pseudomonas aeruginosa* Lectin LecA

Eike Siebs synthesized the designed compounds together with DAAD-intern Cloé Fortin who provided some building blocks. He expressed LecA and tested the compounds in different biophysical assays (FP and isothermal titration calorimetry). Additionally, he wrote this publication together with Alexander Titz and the input of all co-authors. Elena Shanina performed the NMR (¹⁹F PrOF and TROSY), Sakonwan Kuhaudomlarp the SPR and together with Annabelle Varrot the x-ray studies, and Priscila da Silva Figueiredo Celestino Gomes the *in silico* experiments. Anne Imberty, Christoph Rademacher, Didier Rognan and Alexander Titz were supervising and coordinating this project.

Discovery of *N*-β-L-Fucosyl Amides as High-Affinity Ligands for the *Pseudomonas aeruginosa* Lectin LecB

Patrycja Mala synthesized all fucopyranosides. Eike Siebs performed the *in silico* studies, the biophysical experiments of the fucopyranosides in the competitive binding assay and isothermal titration calorimetry, analyzed the data, visualized the electron density of the co-crystal structures with PyMol, analyzed the interactions and wrote this chapter together with Alexander Titz and the input of all co-authors. Annabelle Varrot, Yasmina Grimoire and Joscha Meiers co-crystallized promising ligands with LecB and they, mainly Joscha Meiers, performed the data analysis and refinements of the structures. Katharina Rox performed the ADMET studies. Alexander Titz was coordinating this project.

Ligand-Based Design for Targeting *Burkholderia ambifaria* Lectin BambL

Eike Siebs planed this project to develop new BambL inhibitors. He performed the biophysical experiments, the *in silico* studies, the final purification of the synthesized L-galactosides **16a** and **16b**, visualized the electron density of the co-crystal structures with PyMol, analyzed the interactions and wrote this chapter. Maylis Norlöf (ERASMUS student) was involved in the implementation and optimization of the synthesis of the L-galactosides **16a** and **16b**. Matas Martisauskas (DAAD intern) synthesized compound **17** together with Eike Siebs. BambL was expressed in Anne

Imberty's laboratory and provided by Elena Shanina and Sakonwan Kuhaudomlarp. The x-ray crystallography studies were conducted by Joscha Meiers together with Annabelle Varrot including the data analysis and refinements of the structures. The whole project was supervised by Alexander Titz who also gave supportive input in writing this chapter.

Abbreviations

%	Percent
°C	Degree celsius
2'-FL	2'-Fucosyl lactose
2ME	2-Mercaptoethanol
Å	Angstrom
ACN	Acetonitrile
<i>al.</i>	<i>Alii</i>
ANR	Agence nationale de la recherche
aq.	aqueous
<i>B. ambifaria</i>	<i>Burkholderia ambifaria</i>
BBB	Blood brain barrier
Bcc	Burkholderia cepacian complex
biol.	Biology
br	Broad
ca.	Circa
cal	Calorie
calcd.	Calculated
cat.	Catalogue
CBCH	Chemical Biology of Carbohydrates
CBS	Carbohydrate biding site
CD	Crohn's disease
CERMAV	Centre de Recherches sur les Macromolécules Végétales
CF	Cystic fibrosis
CNS	Central nervous system
COMT	Catechol-O-Methyltransferase
COPD	Chronic obstructive pulmonary disease
CRD	Carbohydrate recognition domain
CSP	Chemical shift perturbation
CTB	Cholera toxin B-subunit
CTL	C-type lectin
d	Doublet
DAAD	Deutscher Akademischer Austauschdienst

DAMP	Damage-associated molecular patterns
DC-SIGN	Dendritic Cell-Specific Intercellular adhesion molecule-3-Grabbing Non-integrin
dev.	Deviation
DFG	Deutsche Forschungsgemeinschaft
DIPEA	<i>N,N</i> -Diisopropylethylamine
DMF	Dimethylformamide
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
DZIF	Deutsches Zentrum für Infektionsforschung
e.g.	<i>Exemplia gratia</i>
EC ₅₀	Half maximal effective concentration
ECD	Extracellular domain
EDC	1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide
EDG	Electron donation group
EHEC	Enterohemorrhagic <i>E. coli</i>
em.	Emission
eq.	Equivalent
ESI	Electrospray ionization
ESKAPE	<i>Enterococcus faecium</i> , <i>Staphylococcus aureus</i> , <i>Klebsiella pneumoniae</i> , <i>Acinetobacter baumannii</i> , <i>Pseudomonas aeruginosa</i> , <i>Enterobacter</i> several species
EWG	Electron withdrawing group
ex.	Excitation
FP	Competitive binding assay based on fluorescence polarization
HBGA	Human blood group antigens
HIPS	Helmholtz-Institut für Pharmazeutische Forschung Saarland
HIV	Human immunodeficiency viruses
HOBT	Hydroxybenzotriazole
HPLC	High Performance Liquid Chromatography
HR-MS	High resolution mass spectrometry
HSQC	Heteronuclear single quantum coherence
Hz	Hertz
IC ₅₀	Half maximal inhibitory concentration

ID	Identification
IPTG	Isopropyl- β -D-thiogalactopyranosid
ITC	Isothermal Titration Calorimetry
J	Joule
K_d	Dissociation constant
LB	Lysogeny broth
LC-MS	Liquid Chromatography Mass Spectrometry
M	Molar concentration
MAG	Myelin-associated glycoprotein
MD	Molecular dynamics
MDCK	Madin-Darby Canine Kidney
N	Stoichiometry
n.s	Not significant
NCI	National Cancer Institute
NMR	Nuclear magnetic resonance
no.	Number
NOD-like receptors	Nucleotide-binding oligomerization domain-like receptors
org.	Organic
<i>P. aeruginosa</i>	<i>Pseudomonas aeruginosa</i>
PAINS	Pan-assay interference compounds
PAMP	Pathogen-associated molecular patterns
PBS	Phosphate-buffered saline
PDB	Protein data base
pH	Potential of hydrogen
PMSF	Phenylmethylsulfonylfluorid
pNPG(al)	<i>para</i> -Nitrophenyl β -D-galactoside
ppm	Parts-per-million
PrOF	Ligand- and protein-observed ^{19}F
PRR	Pattern recognition receptors
q	Quartet
r.t.	Room temperature
Ref	Reference
rf	Reflux

RMSE	Root-mean-square deviation
RNA	Ribonucleic acid
rpm	Revolutions per minute
s	Singlet
satd.	Saturated
sLeX	Lewis X
SPR	Surface Plasmon Resonance
STD	Saturation transfer difference
std.	Standard
t	Triplet
TBS	Tris-buffered saline
T_m	Melting temperature
tol	Toluene
TRIS	Tris(hydroxymethyl)aminomethane
TROSY	Transverse relaxation optimized spectroscopy
TSA	Thermal shift assay
UTI	Urinary tract infections
vs.	Versus
ΔH	Enthalpy
ΔS	Entropy

Introduction

A Brief History of Antibiotics

In 1929, the discovery of penicillin, an antibiotic, which is still used in current treatments, and of streptomycin in 1943, initiated the golden age of antibiotics. Suddenly, it was possible to perform routine surgeries on seriously injured people, organ transplantations, chemotherapeutic cancer treatments etc..^[1] Until 1960s numerous new antibiotic classes (e.g.: macrolides, glycopeptides, quinolones) were discovered,^[2] but their excessive use in human healthcare, agri- and aquaculture led to a steady rise in antimicrobial resistances.^[3] Today, there are six pathogens, the so called ESKAPE pathogens (*Enterococcus faecium*, *Staphylococcus aureus*, *Klebsiella pneumoniae*, *Acinetobacter baumannii*, *Pseudomonas aeruginosa*, and *Enterobacter* species) that play a leading role in nosocomial infections worldwide. Some isolated strains are multi-drug resistant (resistance against at least three antibiotics), which makes them one of the three, main global threats to human health.^[4] Every year, 700 000 people worldwide die of multi-drug resistant infections, and the number is predicted to rise to 10 million each year in 2050 by the World Health Organization.^[5]

Bacterial Adherence and Biofilm Formation

Bacteria have developed strategies to become resistant against antimicrobials due to target modifications, altered permeability, antibiotic inactivation and by-passing metabolic pathways.^[6] Pathogenic microorganisms such as *Pseudomonas aeruginosa* possess different microbial virulence factors for evading the host defense mechanism. These virulence factors are molecules produced by the pathogen such as toxins, enzymes, exopolysaccharides or cell surface structures e.g.: capsules, lipopolysaccharides, glyco- and lipoproteins. Some are able to enter the host cell where they interfere with its machinery and enable infections. To combat microbial infections, it is important to understand the functions of the virulence factors on the microbiological level and will be further described below.^[7]

The polysaccharides of a host-cell surface (glycocalyx) serve as a natural barrier between the cell wall and its surroundings, yet, bacteria have developed strategies to circumvent this obstacle and adhere to the cell surface initializing biofilm formation at their colonization sites.^[8] During the initial step, the planktonic bacterium cell has to overcome changes in physicochemical properties (pH, ionic strength, osmolarity, nutrient availability) to reach the surface of the host cell.^[9] Then, the cells use their cell

appendages, e.g.: flagella and pili if available, to locate a suitable spot for adhesion. At the target site, the flexible flagella form hydrophobic interactions with the eukaryotic cell surfaces and send signals to the cell to prevent its rotations after attachment.^[9] Pili are also involved in the attachment to the surface eg.: the *E. coli* type I pilus is equipped with a carbohydrate binding protein, FimH, for binding to mannoses from the host's glycocalyx. *P. aeruginosa* uses the cup pilus that besides mediating cell-surface interactions supports cell-cell aggregation from other stages of the biofilm formation.^[9,10] The attachment of the bacterial-cell body occurs mainly by Van-der Waals, electrostatic and acid-base interactions (**Figure 1A**).^[9] This initial, reversible attachment during the biofilm formation can lead to a permanent colonization by changing from a mobile to a sessile state of the bacteria. For that, the bacteria start producing an exopolymeric matrix consisting of polysaccharides (in *P. aeruginosa*: alginate, Psl and Pel), proteins (in *P. aeruginosa*: LecA, LecB, CdrA), nucleic acids and lipids^[11] to which they consolidate by binding to the O-antigen part of their lipopolysaccharides containing cell surface (**Figure 1B**). In case of *P. aeruginosa*, it is proposed that the initiation of the biofilm formation is triggered by envelope stress that activates the Wsp pathway, which represses flagella biosynthesis genes and activates the production of the biofilm matrix components.^[12,13] Stabilized in the matrix of the extracellular polymeric substances, the bacteria start forming microcolonies by filling the spaces between each other until the plane space is covered. Afterwards, they start growing upwards to form 3D biofilm structures known as mature biofilm, in case of *P. aeruginosa*, they are mushroom-like under static conditions^[14]. The matrix network hampers the transport of e.g. essential nutrients from the surface to the inner biofilm resulting in an environmental change causing mainly a starvation of the bacteria, which reduces their metabolic activity into a reversible persister state.^[15] In the end, the biofilm can release planktonic cells from the sessile structure to start a new colonization cycle.^[16]

Antibiotic Effects on Biofilms

In case of a treatment with microbicides (e.g.: antibiotics) only the outer cells in the biofilm are killed. The cells in the inner part of the biofilm cannot be addressed due to mechanical barriers of the matrix^[17]. Electrostatic charges of the matrix can interact with antimicrobials and prevent them from diffusion through the system. In addition, most antimicrobials that target metabolic processes such as replication, transcription, translation and cell wall biosynthesis are ineffective against persister cells, due to their

reduced metabolism.^[18] Furthermore, bacteria in biofilms are able to exchange genes and transfer resistance plasmids that lead to resistance and promote the biofilm stability.^[19] Studies revealed that those persister cells are up to 1000-fold more resistant against antibiotics than planktonic cultures.^[20]

***Pseudomonas aeruginosa* and *Burkholderia ambifaria* in Cystic Fibrosis**

Cystic Fibrosis (CF) is an autosomal recessive inherited disorder. The mutation occurs in the gene of the cystic fibrosis transmembrane regulator (CFTR) protein located in chromosome 7. This protein is involved in the transport of chloride and bicarbonate ions in epithelial cell surfaces.^[21] The dysfunction often leads to chronic sinopulmonary infections and pancreatic insufficiency, resulting in a high concentration of chloride in sweat and thick and sticky mucus due to hyper-secretion and lower mucus clearance.^[22,23] Due to the thick mucus, CF patients' immune system is not able to kill bacteria/germs. In case of an attack by a threat, the immune system sends white blood cells with neutrophils to the attacked sites, which in turn release oxidants and proteases (mainly elastase) for the phagocytosis of the invader.^[24–26] In CF patients, this mechanism destroys nearby tissue and leads to a high neutrophil influx, which can later, especially in presence of bacteria/fungi result in a bronchiectasis.^[26–28] The weakened lungs of CF patients (or wounds/inflammations of immunocompromised patients) are perfect targets for bacteria and can lead to chronic infections, where the bacteria are able to persist and the host defense is too weak to eradicate these microorganisms.^[29] This results in a low life expectancy of 46 years for CF people born between 2015–2019.^[30]

The two pathogens, *Pseudomonas aeruginosa* and *Burkholderia ambifaria*, which belongs to the *Burkholderia cepacia* complex (Bcc) (basonym *Pseudomonas cepacia*^[31]), play a crucial role in chronic CF patients. They are both rod shaped Gram-negative, opportunistic bacteria that are present in moist environments and are able to cause nosocomial infections.^[32] *P. aeruginosa* colonizes the lungs of CF patients and is involved in 81% of their deaths.^[33] *Burkholderia cepacia* is responsible for 20% of their deaths due to a sudden fatal pneumonia within weeks/month, the so called *cepacia* syndrome.^[32,34]

So far, there are several treatment strategies to eradicate *P. aeruginosa* in the early colonization states (12 weeks after detection) from the infested sites of CF patients.^[35] A tobramycin inhalation solution (TIS) led to a successful *P. aeruginosa* eradication in

more than 90% of CF patients with no re-grows occurring within one month after treatment. A combination of inhaled colistin and oral ciprofloxacin led to the same results as TIS, but the treatment took two months longer. In another trial, it was shown that intravenous treatment with antibiotics decreases also lung inflammation. In chronic *P. aeruginosa* infections, TIS is the main antibiotic that is used in a four week on- and off-cycle to reduce the amount of bacteria in CF patients. New approaches tackling the biological cause of CF by fixing flaws of CFTR with adjuvants such as ivacaftor a potentiator that keeps the chlorine channel open^[36], or lumacaftor a corrector that keeps the protein in the correct shape^[37].

New Antimicrobial Strategies against *P. aeruginosa*

To circumvent the self-defence mechanisms of *P. aeruginosa*, new strategies have to be developed. At the moment, there are different approaches under exploration such as antifouling or antimicrobial surfaces used on surgical tools to prevent adhesion or killing bacteria in the vicinity.^[38] New antimicrobials are being developed that address different targets in bacteria mainly the cell wall-, protein- and nucleic acid biosynthesis or interfere with cell metabolism.^[39] Another strategy is the development of inhibitors against biofilm formation by addressing the matrix proteins using chelating agents, lantibiotics or synthetic compounds derived from natural binding partners. In this work, I am addressing proteins (lectins) that are crucial for the biofilm formation and are present in the biofilm matrix of *P. aeruginosa*, LecA and LecB and the *B. ambifaria* lectin, BambL. An inhalation study of fucose/galactose, which are the natural ligands of these lectins, showed a reduction of *P. aeruginosa* in CF patients' sputum without side effects as a promising starting point.^[40]

Lectins LecA, LecB and BambL

Lectins (lat.: *lego*, eng.: choose) were first described in 1954 and were classified as proteins and glycoproteins that are able to agglutinate erythrocytes (binding to exposed carbohydrates of red blood cells) such as several snake venoms^[41] and ricin^[42].^[43] Later, the definition was specified to proteins that are able to bind carbohydrates and are not antibodies or enzymes.^[44]

P. aeruginosa expresses two soluble lectins, LecA (previously: PA-IL^[45], renamed to comply with standard genetic nomenclature^[46]) and LecB (previously: PA-IIL^[47]) that are involved in the attachment of the bacteria to a host cell^[48], hence, playing a crucial role in biofilm formation^[49,50] (**Figure 1B**). They are both homotetramers, calcium ion

dependent, heat, acid and proteolysis resistant. LecA (51 kDa) consists of four identical subunits with 121 amino acids and has a low isoelectric point ($pI = 4.94$). Every monomer has one carbohydrate binding site (CBS) with one calcium ion, which is involved in binding D-galactosides ($K_d \text{ D-Gal} = 88 \mu\text{M}^{[51]}$, **Figure 1C**). The natural ligand glycosphingolipid Gb3, a cell receptor, that is responsible for the engulfment of *P. aeruginosa* into host cells.^[52]

The tetrameric LecB (47 kDa, $pI = 3.88$) is slightly different to LecA. It is composed out of four subunits with 114 amino acids and one CBS per monomer containing two calcium ions, which are involved in binding to L-fucosides and D-mannosides (**Figure 1C**). This results in low micromolar affinities for the natural ligand fucose ($K_d \text{ Me-}\alpha\text{-L-Fuc} = 0.43 \mu\text{M}^{[53]}$). One natural ligand of LecB is Psl an exopolysaccharide stabilizing the biofilm matrix whose pentasaccharide subunit consists of mannose, rhamnose and glucose (3:1:1).^[54]

Different *P. aeruginosa* strains express LecA and LecB with different amino acid sequences. A study of 147 strains revealed that LecA diverged marginally (<3%), though the PA7 clade lacked LecA, and LecB significantly (<16%).^[55] In LecB, those changes also occur in the CBS and thus have an impact on the binding affinity of fucose.^[55,56] In current LecA and LecB studies there are two clinical *P. aeruginosa* strains of interest, PAO1 and PA14. PAO1 was the first isolated strain, which is moderately virulent, belongs to a rare clonal group and has been used as the most common strain in research^[57], whereas, PA14 is a highly virulent isolate and represents a common clonal group.^[58,59] Within these two strains, the LecB sequence differs by 13% and showed a modification in the CBS (Ser23Ala) whereas the carbohydrate binding site of LecA stayed conserved.^[55,60] In general, the LecB_{PA14} ligand affinities are 2–7-fold higher compared to LecB_{PAO1} ($K_d \text{ (PAO1) Lewis}^\alpha = 213 \text{ nM}^{[61]}$, $K_d \text{ (PA14) Lewis}^\alpha = 70 \text{ nM}^{[60]}$).^[60] This phenomenon has not been fully understood yet, however, LecB_{PA14} most probably compensates the lack of the *pslA–D* operon in the PA14 strain responsible for the expression of the exopolysaccharide Psl that is involved in the biofilm stabilization.^[60]

BambL is expressed by *B. ambifaria*, a species that belongs to the Bcc, which occurs besides inside humans in the rhizosphere of crop plants where it showed antifungal properties.^[62,63]

BambL is a homotrimeric six bladed, β -propeller lectin consisting of 87 amino acids per protomer (**Figure 1C**, bottom). BambL has six CBS, one per protomer and additional three at its contact points between monomers. BambL binds to fucosides ($K_{d(\alpha\text{MeFuc})} = 0.96 \pm 0.03 \mu\text{M}$) and is able to bind human oligosaccharides epitopes and has a preference for the blood groups H-type-2 ($7.5 \pm 0.1 \mu\text{M}$) and Lewis^y ($11.1 \pm 0.1 \mu\text{M}$), and might be involved in infections.^[63,64] Recently, it was shown that BambL and LecB are able to activate B cells (white blood cells) by binding to the B cell antigen receptor, its co-receptor CD19 and to the spleen tyrosine kinase leading to an intracellular calcium ion uptake, which releases cytokines and ends with necrosis.^[65]

For preventing the biofilm formation caused by LecA or LecB or the necrosis of B cells in case of BambL several groups have tried to address these lectins with different molecules and strategies (see review below for LecA and LecB). BambL was only targeted with a few synthesized compounds, penta/deca-glycocluster resulting in low nanomolar affinities ($K_d = 14\text{--}57 \text{ nM}$)^[66,67] and a synthesized monovalent α -L-fucopyranose 1-O,2-S-2,1-phenyl-2-thiol compound which showed two binding modes ($K_d = 0.24/0.23 \mu\text{M}$) (**Page: 93**).^[68]

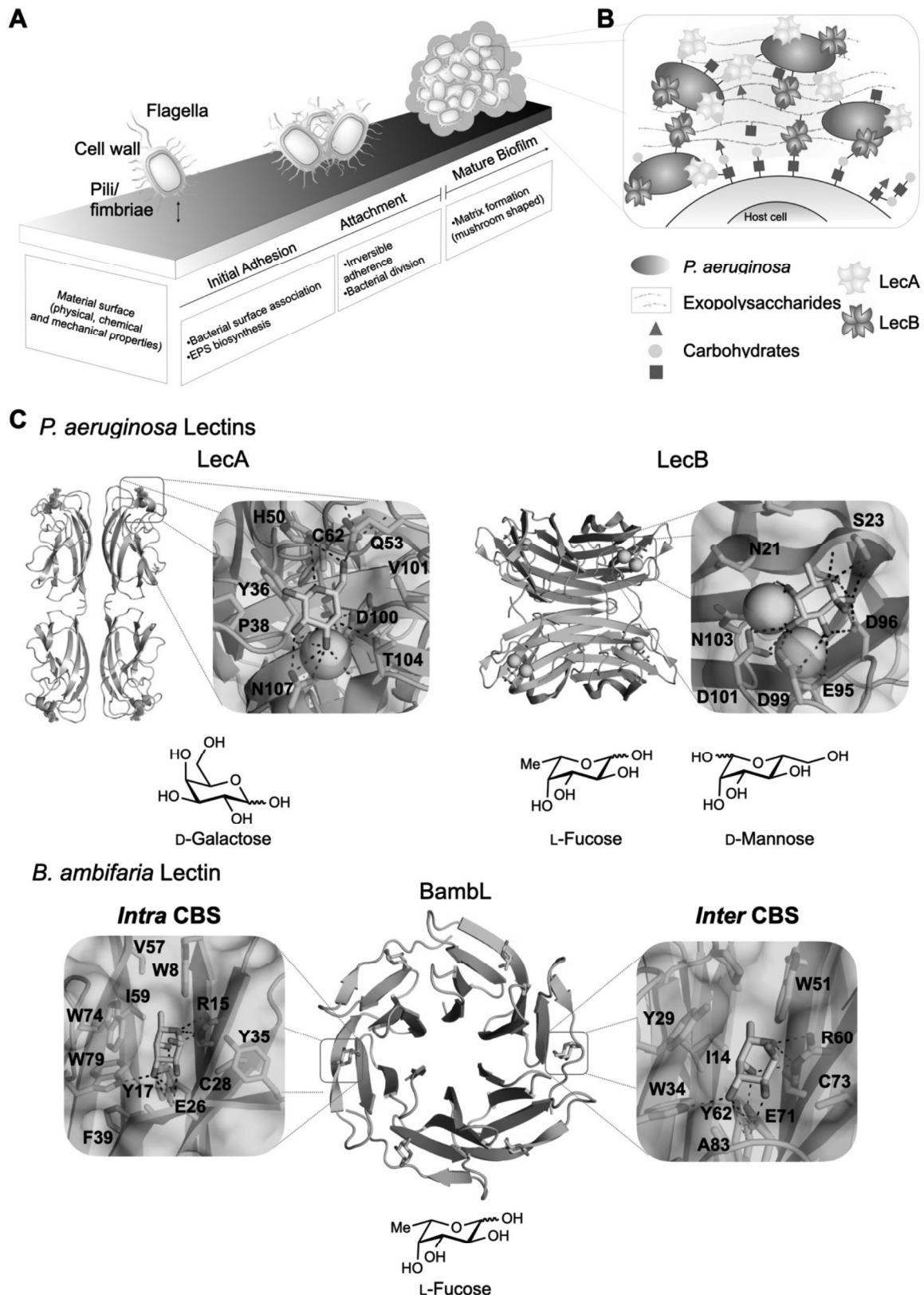


Figure 1: A) Biofilm formation, starting with the reversible attachment of a planktonic cell to a surface (initial adhesion), followed by the attachment (the regression of pili/fimbriae and flagella) and final maturation of the biofilm with developed matrix and cell-cell interactions. **B)** Model of the biofilm matrix – Binding interaction of the lectins (red/yellow) with the exopolysaccharides (lines), and the glycocalyx of the bacteria

(brown) and host (yellow) (by Stefanie Wagner). **C)** Top: *P. aeruginosa* lectins LecA (pdb: 1oko) and LecB (pdb: 1oxc) with their natural ligands (cyan) D-galactose and L-fucose/D-mannose with view of the carbohydrate binding sites and their interactions (black). Bottom: *B. ambifaria* lectin, BambL (pdb: 3zwe), in complex with L-fucose (cyan) and its interactions (black).



Lectin antagonists in infection, immunity, and inflammation

Joscha Meiers^{1,2,3,4}, Eike Siebs^{1,2,3,4}, Eva Zahorska^{1,2,3,4} and Alexander Titz^{1,2,3}

Lectins are proteins found in all domains of life with a plethora of biological functions, especially in the infection process, immune response, and inflammation. Targeting these carbohydrate-binding proteins is challenged by the fact that usually low affinity interactions between lectin and glycoconjugate are observed. Nature often circumvents this process through multivalent display of ligand and lectin. Consequently, the vast majority of synthetic antagonists are multivalently displayed native carbohydrates. At the cost of disadvantageous pharmacokinetic properties and possibly a reduced selectivity for the target lectin, the molecules usually possess very high affinities to the respective lectin through ligand epitope avidity. Recent developments include the advent of glycomimetic or allosteric small molecule inhibitors for this important protein class and their use in chemical biology and drug research. This evolution has culminated in the transition of the small molecule GMI-1070 into clinical phase III. In this opinion article, an overview of the most important developments of lectin antagonists in the last two decades with a focus on the last five years is given.

Addresses

¹ Chemical Biology of Carbohydrates, Helmholtz Institute for Pharmaceutical Research Saarland (HIPS), Helmholtz Centre for Infection Research, D-66123 Saarbrücken, Germany

² Deutsches Zentrum für Infektionsforschung (DZIF), Standort Hannover-Braunschweig, Germany

³ Department of Pharmacy, Saarland University, D-66123 Saarbrücken, Germany

Corresponding author: Titz, Alexander (alexander.titz@helmholtz-hzi.de)

⁴ These authors contributed equally.

Current Opinion in Chemical Biology 2019, 53:51–67

This review comes from a themed issue on **Mechanistic biology**

Edited by **Hermen S Overkleeft** and **David J Vocadlo**

<https://doi.org/10.1016/j.cbpa.2019.07.005>

1367-5931/© 2019 Elsevier Ltd. All rights reserved.

Introduction

Lectins are a highly diverse family of proteins found in all domains of life [1,2]. Various folds and classes have been identified and the common functional feature is their specificity for carbohydrate ligands. These glycan-binding

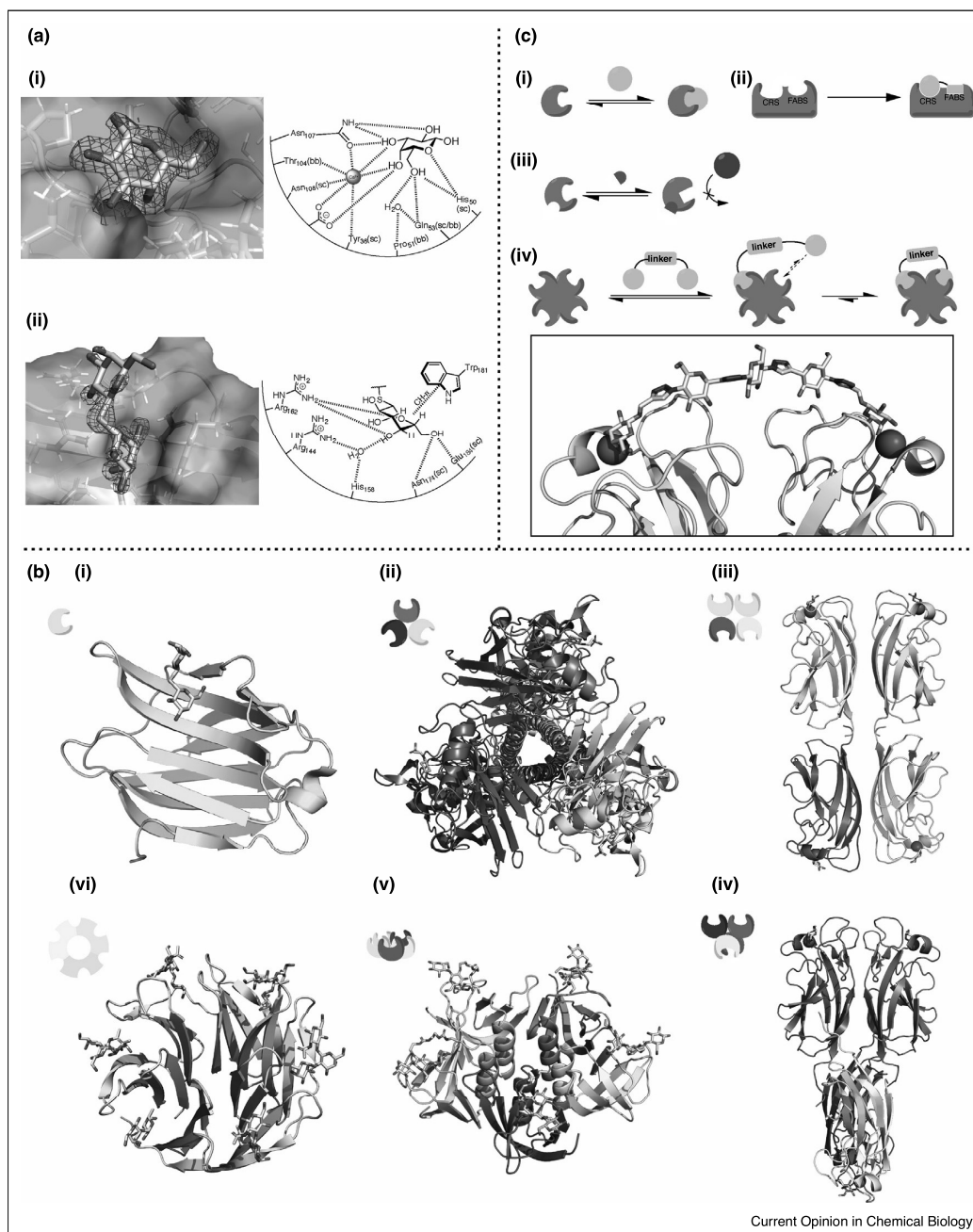
proteins have many important roles in infection, cell recognition, communication and various intracellular processes, such as protein folding and protein targeting.

Numerous viral, bacterial, fungal, and parasitic pathogens employ lectins for initiation and maintenance of an infection by adhering to surface-exposed glycoconjugates of their host organisms [3–5]. In contrast, the mammalian host has developed a plethora of lectin-containing pattern recognition receptors of the innate immune system recognizing glycan structures on intruders [6–8]. In addition to recognizing these non-self structures, other mammalian lectins bind to self-epitopes and thus mediate cell-recognition processes like inflammation and cancer metastasis [9–11].

The natural ligands of lectins are mostly bacterial or fungal polysaccharides, bacterial lipopolysaccharide and peptidoglycan, or eukaryotic glycoconjugates of lipids or proteins [1,12]. Except for bacteria which can have a high diversity among their monosaccharides, generally a relatively small set of different monosaccharide subunits are shared between animals, plants, fungi, parasites, bacteria, and other organisms. These building blocks are assembled into more diverse oligosaccharides where a very high complexity can be achieved due to many possible stereoisomers and regioisomers. In many cases, this leads to organism-specific oligosaccharides, which can then be recognized by innate immunity as non-self antigens and induce neutralization of the intruder [13], or elicit allergic reactions as observed for insect glycans, for example, in bee venom [14]. The opposite phenomenon that pathogen and host have identical glycoconjugates is also observed. The latter has been termed molecular mimicry or glycomimicry, a stealth process of the pathogen believed to be an evolutionary adaptation for evasion of immune surveillance of the host [15,16].

Despite the complexity of those oligosaccharide structures, lectins often recognize terminal monosaccharides or smaller oligosaccharides on a given glycoconjugate. Two common binding modes of carbohydrate ligands are shown in Figure 1a: (i) vicinal hydroxyl groups chelate a Ca^{2+} -ion present in the binding site, or (ii) carbon-bound hydrogen atoms of the carbohydrate ring interact via $\text{CH}-\pi$ stacking with aromatic amino acids in the binding site. Because of the recognition of rather small epitopes, common ligand specificity of different lectins with diverse functional roles often occurs. An example are

Figure 1



(a) Schematic representation of two important recognition modes of carbohydrates by lectins: (i) calcium-ion mediated binding of the ligands, example β -galactoside and LecA (PDB: 1OKO) (ii) tryptophan-mediated stacking on hydrophobic faces of carbohydrates, example galactoside with galectin-3 (PDB: 4JC1). **(b)** Various strategies for domain/binding site orientation: (i) monomeric in galectin-3 (4JC1), (ii) trimeric virus hemagglutinin (6CF5), (iii) tetrameric LecA (1OKO), (iv) tetrameric LecA ortholog PIIA with altered domain orientation (5ODU), (v) pentameric Shiga-like toxin B subunit (1QNU), (vi) trimeric BambL containing 6 carbohydrate binding sites in and between subunits (3ZW2). **(c)** Schematic representation of different lectin inhibition approaches: (i) direct inhibition of carbohydrate binding sites, (ii) growing toward non-carbohydrate binding sites, (iii) allosteric inhibition (iv) multivalent inhibition which refers to clustered binding sites, either multivalent proteins or monovalent lectins clustering on cell membranes.

the functionally different human DC-SIGN and the bacterial lectin LecB with shared specificity for Lewis blood group antigens [17–19]. A large data set for the glycan specificity of many lectins using microarrays is provided by the Consortium for Functional Glycomics (see <http://www.functionalglycomics.org>).

Specificity of the lectins can be further tuned by recognizing functional groups attached to the essential carbohydrate, and, for example, lipids are recognized by a secondary site of the lectin Mincle [20,21], *O*-methylation is required for recognition by the tectonins [22,23], sulfates on nearby amino acids enhance binding of P-selectin to the Lewis-blood groups on glycoproteins [24] and phosphates are required for intracellular trafficking of proteins by the mannose-6-phosphate receptor [25].

Lastly, the spatial presentation of ligands and/or lectin's carbohydrate binding sites (Figure 1b), as well as clustering of several lectin protomers into oligomeric bundles or membrane embedded protein complexes can contribute significantly to specificity by augmentation of apparent binding affinity through avidity [7,26].

Carbohydrate specificity, requirements of additional functional groups and spatial presentation of binding sites are important aspects for the design and success of lectin-targeting probes in chemical biology and drug research. Therefore, the design of lectin antagonists usually follows various approaches from (i) competitive inhibition of a carbohydrate recognition site, (ii) targeting adjacent binding sites, (iii) allosteric inhibition, to (iv) multivalent competitive inhibition of two or more binding sites (Figure 1c).

Consequently, lectins have developed into attractive targets for chemical biology and medicinal chemistry over the past two decades [27,28]. Very active areas of research are the targeting of (i) lectins of pathogenic origin to interfere with mechanisms of infection by viruses and bacteria, and to a smaller extent also fungi and parasites, (ii) the selectins as a family of three closely related proteins crucial for cell migration in inflammation and cancer, as well as (iii) immunotherapeutic or immunomodulatory approaches for the mammalian lectins langerin in vaccine delivery, DC-SIGN in HIV infection or the galectins in cancer and immune modulation. Lectins discussed in this opinion article are summarized in Table 1.

Bacterial lectin antagonists

Bacterial antibiotic resistance is increasing worldwide at an alarming rate. As one consequence, antivirulence drugs have gained considerable research interest as alternative treatment approach with the aim to avoid the rapid onset of resistance [50]. In this context, the inhibition of bacterial lectins to prevent infection and persistence is a newly exploited strategy [3,27]. Targeting lectins involved in

the formation of bacterial biofilms are of particular interest since bacteria embedded in their self-produced biofilm matrix exhibit increased antimicrobial resistance compared to free floating planktonic bacteria. Biofilm-associated bacterial infections are responsible for a broad range of chronic/recurring diseases [51].

The Gram-negative bacterium *Escherichia coli* is the prime pathogen in urinary tract infections (UTIs) and important for intestinal infections as a consequence of Crohn's disease (CD). *E. coli* can build various organelles called pili and fimbriae which are oligomeric cell appendices built up of several proteins. These organelles are often employed for bacterial adhesion. The pilus or fimbria lectins FimH and FmlH, localized on the top of the different organelles, play decisive roles in host colonization, invasion, and biofilm formation [52]. Thus, inhibition of these lectins to antagonize infections presents a viable therapeutic strategy [53,54].

FimH is located on the tip of fimbriae and usually binds to mannosylated glycoconjugates in the bladder endothelium. Pathogenicity of *E. coli* clinical isolates expressing different *fimH* alleles varies, but the mannose binding pocket is invariant [52,55,56]. Hultgren's group demonstrated the activity of a high affinity mannoside FimH inhibitor against different uropathogenic *E. coli* strains [57]. In recent years, several research groups have been developing FimH antagonists for treatment of urinary tract infections and gut inflammations associated with CD. X-ray crystallography guided drug design focused on optimization of interactions with the so-called tyrosine gate adjacent to the mannose binding site. Introduction of aryl and alkyl aglycons increased the binding affinity significantly compared to simple mannose [58–60]. Nanomolar binding affinities were achieved by introducing biaryl aglycons that are tightly coordinated by the tyrosine gate [61–63]. High affinity biaryl mannosides were further optimized to increase metabolic stability by replacing the labile *O*-glycosidic bond with carbon-based linkers to the aglycon [29**,64]. Ester and phosphorylated prodrugs were successfully explored to improve oral bioavailability of both *O*-mannosides and *C*-mannosides [29**,65,66*]. Rational design and optimization of FimH antagonists are summarized in a recent review by Mydock-McGrane *et al.* [67]. The promising preclinical candidate **1** ($EC_{90} = 31$ nM, Figure 2) is one example of a highly optimized FimH inhibitor with good metabolic stability and high efficacy in mouse models of acute and chronic UTI [29**]. Recent optimization attempts yielded thiomannosides (e.g. **2**, $EC_{90} = 0.31$ μ M, Figure 2) with improved metabolic stability compared to respective *O*-mannosides, ability to inhibit biofilm formation *in vitro* and with a prophylactic effect in a mouse UTI model [30]. The first FimH antagonist entering clinical trials was EB8018 from Enterome (Paris, France) designed for the treatment of CD, but its structure has not been

Table 1

Overview of bacterial, viral, and mammalian lectins discussed in this opinion article

	Origin	Binding specificity	Key roles	Status of development/indicator
Bacterial lectins				
FimH	<i>E. coli</i>	Man	Adhesion, biofilm formation	Lead optimization (1, 2) [29**,30], EB8018 in Phase I clinical trials (www.clinicaltrials.gov , NCT03709628)
FmlH	<i>E. coli</i>	Gal, GalNAc	Adhesion, biofilm formation	Hit optimization (3) [31]
LecA	<i>P. aeruginosa</i>	Gal	Adhesion, biofilm formation	Exploratory studies
LecB	<i>P. aeruginosa</i>	Man, Fuc	Adhesion, biofilm formation	First covalent lectin inhibitor (5) [32**] Lead optimization (6, 7) [33,34**]
Shiga toxins	<i>S. dysenteriae</i> , <i>E. coli</i>	Gal, Glc	Toxin	Lead optimization on hold, First peptide-based inhibitor [35]
Cholera toxin	<i>V. cholerae</i>	Gal, Fuc	Toxin	Hit optimization (8) [36]
Viral Lectins				
Hemagglutinin	Human influenza virus	Neu5Ac	Adhesion, cell entry	Hit optimization (12) [37–39] and exploratory studies (10, 11) [40*,41*,42**]
Hemagglutinin–neuraminidase	Human parainfluenza virus	Neu5Ac	Adhesion and detachment, cell entry	Hit optimization [43,44]
Capsid protein P domain	Norovirus	HBGAs	Adhesion, cell entry	Exploratory studies (14, citric acid) [45–47]
Mammalian Lectins				
Langerin	Langerhans cells	Man, Fuc, GlcNAc, sulfated Gal, Glc	Immune response	Exploratory studies First allosteric mammalian lectin inhibitor (15) [48**]
DC-SIGN	Dendritic cells	Man, Fuc, GlcNAc sLe ^x	Immune response	Exploratory studies
Selectins	L-selectin: leukocytes P-selectin: platelets and endothelial cells E-Selectin: endothelial cells	P/L-selectins: Man, Gal and Sulfation [49]	Cell adhesion	GMI-1070 (20) in Phase III clinical trials against vaso-occlusive anemia (www.clinicaltrials.gov , NCT02187003)
Mincle	Immune system	Glycolipids with terminal Glc or Man	Immune response	Exploratory studies
Galectin	Circulating proteins	Gal, for example, <i>N</i> -acetylglucosamine	Regulate cell death	TD139 (24) in Phase II clinical trials against idiopathic pulmonary fibrosis (www.clinicaltrials.gov , NCT03832946)
Siglecs	Immune-cells	Neu5Ac	Cell-cell signaling, immune response and adhesion	Exploratory studies

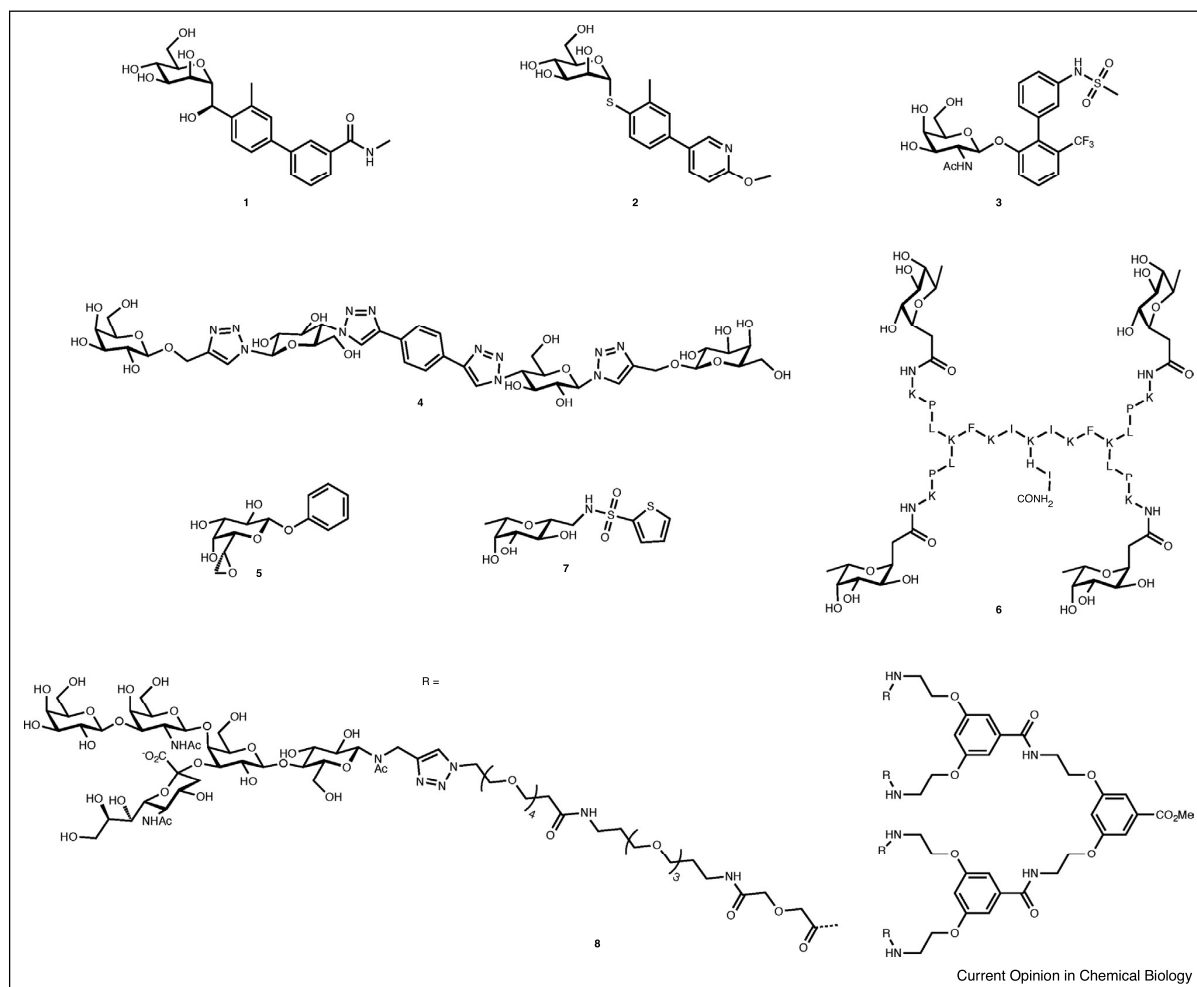
disclosed. In collaboration with Takeda, EB8018 has completed Phase Ia and the Phase Ib trial is ongoing in early 2019 (www.clinicaltrials.gov, NCT03709628). Furthermore, Fimbrion Therapeutics (St. Louis, MO) has announced the selection of a not further specified clinical candidate as antibiotic sparring molecule against UTIs in collaboration with GSK (www.fimbrion.com, press release Dec 06, 2018).

As a secondary target of uropathogenic *E. coli*, the FimH-like adhesin FmlH recognizes Gal(β1-3)GalNAc epitopes on bladder epithelium and enhances *E. coli* urinary tract colonization [54]. Recently, first structure-based inhibitor design approaches for FmlH have been reported [31,68**]. To date, the best FmlH inhibitor 3 (Figure 2) is based on *N*-acetyl galactosamine carrying a further substituted biphenyl aglycon and displays very high binding affinity (IC₅₀ = 34 nM), good aqueous solubility and high metabolic stability. Unfortunately, 3 showed

only low oral bioavailability in rats of less than 1% and further optimization is therefore mandatory [31,68**].

The opportunistic pathogen *Pseudomonas aeruginosa* has two soluble lectins, the extracellularly secreted proteins LecA (Figure 1) and LecB, both mediating bacterial virulence and being crucial components for biofilm formation [69–71]. Consequently, both proteins have been subject to intense research toward biofilm modulators and in drug discovery for antivirulence drugs [27,28,72–74]. LecA binds to various α-galactoside-terminating glycoconjugates with the glycosphingolipid Gb3 as proposed natural ligand [75]. This homotetrameric lectin was later shown to mediate bacterial uptake via Gb3 where it acts as a lipid zipper [76,77]. The affinity of LecA to galactose and simple glycosides thereof is rather weak in the 50–100 μM range. Consequently, development of LecA antagonists mainly focused on multivalent display of galactosides using many different linkers and maximizing

Figure 2

Inhibitors targeting lectins of pathogenic bacteria in *E. coli* (1-3), *P. aeruginosa* (4-7), and toxins of *V. cholerae* (8).

the number of presented epitopes [28,78]. Very potent tetraivalent galactoclusters with low nanomolar binding affinities toward LecA have been developed [79^{**},80,81^{*},82,83]. In contrast to the high target-binding affinity, they showed only moderate inhibition of biofilm growth in the micromolar range *in vitro*.

The Pieters group has undertaken a different approach and focused on divalent galactosides oriented in a perfect manner to bridge two adjacent binding sites in the LecA tetramer. Several highly potent divalent inhibitors with the rigid spacers consisting of glucose and triazole groups were obtained, including the most potent LecA inhibitor reported so far with a K_d of 12 nM (4, Figure 2) [84^{**},85]. Again, recent optimization of these highly potent molecules on the target revealed a need for additional

multimerization and rather high micromolar concentrations for biofilm blocking [82,86].

Monovalent galactose-derived ligands with binding affinities in low micromolar range could be obtained after introduction of a β -aryl aglycon which establishes a π -stacking interaction with an imidazole-CH of His50 adjacent to the carbohydrate binding site (Figure 1a) [87^{**},88,89]. However, the specificity for further variations appears relaxed and changing substituents at the phenyl aglycon did not lead to significant potency improvements. As an alternative approach to the generally employed glycosides of unmodified galactose residues in LecA ligands, we have embarked on the modification of the galactose residue itself. A cysteine residue in the carbohydrate binding site of LecA was targeted with

the aim to develop a covalent lectin inhibitor using a small electrophilic headgroup in a modified galactose [32**]. Despite the fact that covalent inhibitors are widespread for many other protein classes, epoxide **5** (Figure 2) was established as the first-in-class covalent lectin inhibitor. Because of its moderate affinity toward LecA ($IC_{50} = 64 \mu M$), the molecule was converted into a tool compound after synthetic derivatization and conjugation to fluorescein enabling the visualization of *P. aeruginosa* biofilm aggregates by confocal fluorescence microscopy [32**].

The second *P. aeruginosa* lectin LecB also forms a homotetrameric quarternary structure, binds broadly to fucosides and mannosides and the highest affinity was determined for Lewis blood group antigens [17,90]. In contrast to LecA, the protein sequence of LecB varies among clinical isolates and two important types occurring in the clinical isolates PAO1 and PA14 have been identified as representative for all studied isolates [18,91]. Despite the observed amino acid sequence differences in LecB between strains, its carbohydrate binding specificity is conserved, underpinning the suitability of LecB as a drug target with conserved specificity among all isolates. Also for LecB, multivalent inhibitors have been the first choice for inhibition [28,78]. However, because of a sterically more distant and less favorable orientation of binding sites in LecB compared to LecA, the obtained multivalent ligands could not achieve a comparable boost in affinity. Nevertheless, two types of multivalent ligands carrying fucosides stand out of the very broad field: tetravalent glycopeptide dendrimer **6** ($IC_{50} = 140 nM$, Figure 2) was able to efficiently prevent biofilm formation of *P. aeruginosa* at a concentration of $20 \mu M$ *in vitro*; [33] furthermore, a calixarene carrying four fucose residues was tested in an infection model in mice [79**]. This compound significantly reduced the number of bacteria colonizing lung and spleen, but was unable to inhibit bacterial biofilms *in vitro* at a concentration of $100 \mu M$ despite its high affinity at the target ($K_d = 48 nM$).

To overcome the intrinsic disadvantages associated with large molecules and multidirectional valency in biofilm formation, we have used the small molecule LecB ligand mannose as a starting point for the rational design of monovalent biofilm targeting glycomimetics [92]. These compounds exhibited rather good target-binding potency ($K_d = 3\text{--}20 \mu M$) and prevented bacterial adhesion to a glycosylated surface at $100 \mu M$. Further optimization [93] and removal of the anomeric center [94] finally yielded C-glycosidic inhibitors of LecB (e.g. **7**, Figure 2) with good target-binding potency ($K_d = 290 nM$) and very long receptor residence times ($t_{1/2} = 28 min$) [34**]. Glycomimetic **7** showed approx. 85% inhibition of biofilm growth *in vitro* at $100 \mu M$, which contrasts the lack of antibiofilm activity of the natural LecB binder methyl α -L-fucoside, despite its very high target binding affinity ($K_d = 430 nM$). Furthermore, glycomimetic **7** is orally

bioavailable which is not possible for large multivalent molecules.

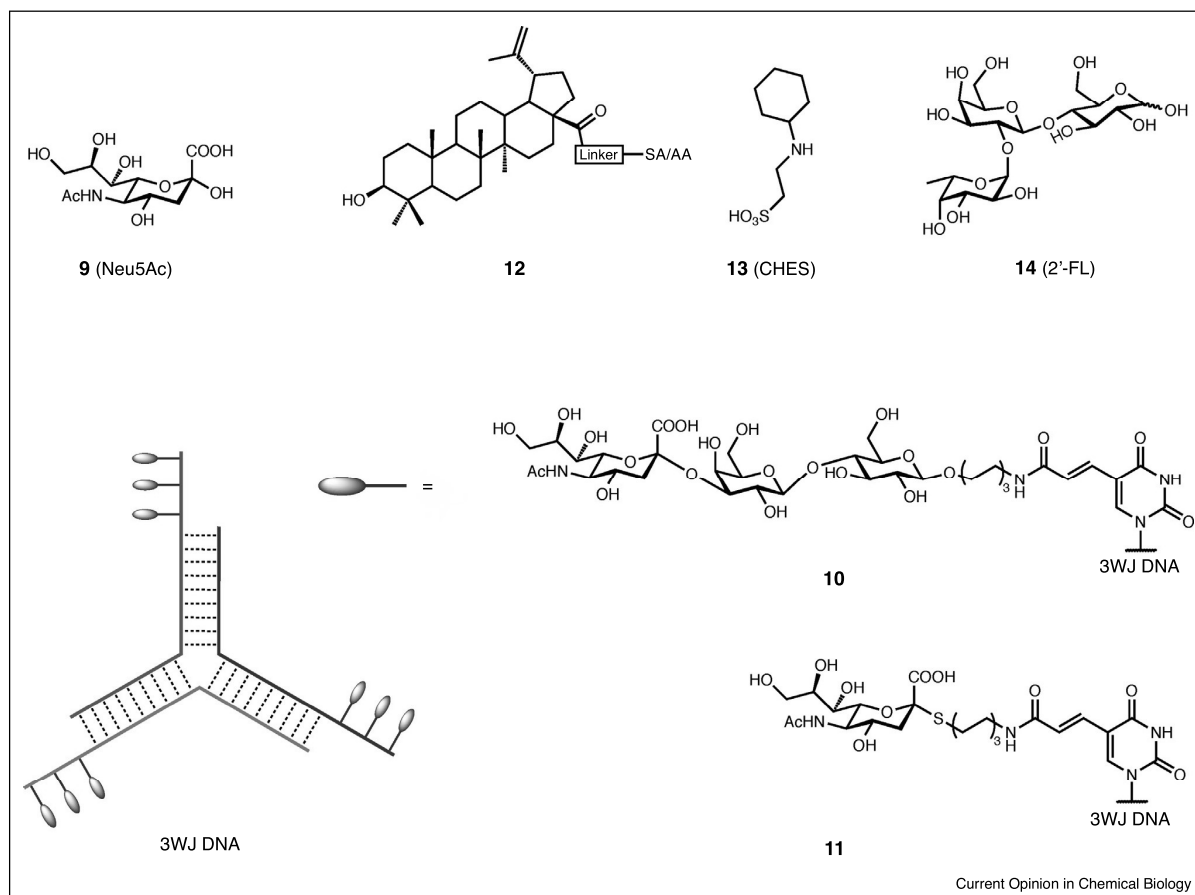
Shiga and cholera toxins are bacterial proteins responsible for severe symptoms in gastrointestinal infections. These so-called AB₅ toxins consist of one catalytic A-subunit and five lectin-like B-subunits (Figure 1b) which are responsible for the binding of the complex to the host cell surface in the gut. Inhibition of the B-subunits and thereby preventing adhesion is a potential treatment strategy [95].

Shiga toxins (Stxs) are produced by *Shigella dysenteriae* and some enteropathogenic *E. coli* strains, for example, enterohemorrhagic *E. coli* (EHEC). Kitov *et al.* designed the pentavalent ligand STARFISH to match the carbohydrate binding sites of the five B-subunits with subnanomolar inhibitory activity against Shiga-like toxins I and II (Stx1 and Stx2) [96]. A modified version of STARFISH, called DAISY, improved the *in vivo* activity and provided full protection against the toxins when administered simultaneously in a mouse model despite its lower target-binding potency [97]. However, further development of DAISY-based inhibitors appears halted (no further publications) since the compound proved ineffective in a treatment scenario, that is, drug administration after infecting mice with the Shiga toxin producing strain *E. coli* O91:H21. Nishikawa *et al.* designed a series of carbosilane dendrimers called SUPERTWIG. The most potent compound of the series was able to completely neutralize Stxs in the blood stream and protect mice against a fatal dose of the Shiga toxin producing strain *E. coli* O157:H7 even when administered after establishment of infection [98]. The rather complex synthesis of multivalent-trisaccharide inhibitors is hindering further clinical development.

From a peptide library, the branched proline and arginine rich high molecular weight peptide Ac-PPP-tet was identified to bind to Stx2 B-subunit and inhibit Stx2 cytotoxicity [35]. This peptide affects the intracellular transport of Stx2 and protected mice from a fatal dose of *E. coli* O157:H7 even when administered after an established infection; this molecule further protected rabbit intestines *ex vivo* against the toxic effect of Stx2 [35,99]. Recent efforts include the synthesis of sugar-amino acid hybrid polymers with highly clustered globotriaosyl residues that showed low micromolar affinities to both Stxs with the ability to neutralize the toxic effects on Vero cells [100].

Vibrio cholerae produces cholera toxin where each B-subunit (CTB) has two binding sites – one primary binding site recognized by the ganglioside GM₁ and a secondary low affinity site recognized by fucosylated glycans [101]. A number of derivatives mimicking the terminal galactose from GM₁ has been screened and m-nitrophenyl α -D-galactoside and 3,5-disubstituted phenylgalactosides were identified as monovalent CTB inhibitors [102,103]. Numerous multivalent inhibitors targeting the primary

Figure 3



Inhibitors of influenza hemagglutinin: Neu5Ac (**9**), macromolecular sialylated three way junctioned DNA **10** and **11** and small molecules **12–13**; or. Norovirus spike protein can be blocked using the trisaccharide 2'-fucosyl lactose **14**. SA: sialic acid, AA: ascorbic acid.

site with down to picomolar binding affinities (e.g. **8**, $IC_{50} = 34 \text{ pM}$, Figure 2) [36] have been developed and were summarized in a recent review by Kumar and Turnbull [104]. Targeting the fucose binding site as new strategy was published by Wands *et al.* who reported inhibition of CTB binding to cell surfaces with 2'-fucosyllactose and a fucosylated polymer [105**].

Viral lectin inhibitors

Viral infections are difficult to treat, control and prevent. Frequent antigen variation, for which the influenza virus is a perfect example, prevents efficient protection and virus clearance by the human immune system. In many viruses, lectin-carbohydrate interactions are crucial for an efficient infection of the host. Hemagglutinin is the sialic acid binding lectin on the surface of the influenza viral envelope and plays a key role in the host cell-virus interaction. Sialic acids are defined as a family of acidic sugars with a nine

carbon atom backbone and the most abundant member found in vertebrates is *N*-acetyl neuraminic acid (**9**, Neu5Ac, Figure 3) [106]. Because the binding interaction of one monomeric hemagglutinin to sialylated glycans is weak ($K_d > 1 \text{ mM}$) [107], trimerization of hemagglutinin on the viral envelope and a high sialic acid density on the host cell lead to an increased avidity. This binding event then triggers the internalization of the virus by endocytosis [108]. Therefore, inhibition of the hemagglutinin-sialic acid interaction could yield prophylactic as well as therapeutic treatments of an influenza virus infection.

For this purpose, Strauch *et al.* [42**] developed a trimeric influenza neutralizing protein, targeting the hemagglutinin receptor binding site. This protein was designed to mimic the key interactions of broadly neutralizing antibodies and its optimization led to a highly avid protein with a trimeric binding mode and nanomolar apparent K_d

values. *In vivo*, using an H3 HK68 influenza infection mouse model, prophylactic and therapeutic treatment significantly protected mice from establishing disease and weight loss. Unfortunately, this designed protein does not show broad spectrum activity since it does not bind to the pathogenic ‘bird flu’ subtype H5N1. Limitations in high scale production and price, together with challenging pharmacokinetic properties will impact on its commercial use as an anti-influenza drug.

A recent review by Li *et al.* describes a wide range of chemical scaffolds and strategies to inhibit the hemagglutinin – host cell interaction. Mostly, trimeric sialosides are presented as binders to the receptor binding site [109].

2,3-Sialyllactose (2,3-SL) conjugated to three way junction (3WJ) DNA, with each DNA strand presenting one, three or five 2,3-SL molecules complementary to the hemagglutinin trimer geometry was reported by Yamabe *et al.* [40*,41*]. Hemagglutinin inhibition revealed 3WJ DNA with three sialic acid residues per arm in compound **10** as best inhibitor with a $K_i = 0.25 \mu\text{M}$, which corresponds to an 80 000-fold increase compared to monomeric 2,3-SL and an eightfold increase compared to 3WJ DNA with only one sialic acid per strand. Surprisingly, 3WJ DNA presenting five sialic acid per strand led to a reduction in activity ($K_i^{\text{HAI}} > 4.0 \mu\text{M}$) which probably originates from an altered orientation of the carbohydrate epitopes induced by steric hindrance. In contrast to the neuraminidase labile *O*-linked **10**, the more stable thio-linked sialic acid derivative **11** was synthesized as a follow up. For **11**, an increased stability toward influenza neuraminidase present on the viral envelope was observed, while its activity was retained. However, in presence of the full virus both derivatives, that is, *O*-glycoside and *S*-glycoside, were stable under the conditions tested. Another approach using a macromolecular scaffold by Nagao *et al.* yielded a trimeric star-shaped glycopolymer presenting 6'-sialyllactose on each of the three arms, synthesized by reversible addition-fragmentation chain transfer polymerization [110]. The degree of polymerization dictated the length of each arm. Hemagglutinin inhibition clearly depended on the arm-length, resulting in a $K_i = 21 \mu\text{M}$ for their best glycopolymer.

Conjugation of sialic acid or ascorbic acid derivatives onto pentacyclic triterpenes by Zhou and co-workers [37,38] was inspired by the broad antiviral activity of *Dipsacus asperoides* triterpenes and the corresponding synthetic leads [39]. In both cases, conjugation to betulinic acid as in **12** led to a strong reduction of infection by influenza A/WSN/33 in MDCK cells. Cytotoxicity of the triterpenes was also reduced by conjugation to sialic acid or ascorbic acid and a hemagglutination assay and SPR experiments with immobilized hemagglutinin suggested hemagglutinin as the putative target ($K_d = 17 \mu\text{M}$ for the sialic acid conjugate, $K_d = 8.0 \mu\text{M}$ for the ascorbic acid conjugate). Interestingly, the synthetic 2,3-di-*O*-benzyl

ascorbic acid intermediate showed a higher affinity for hemagglutinin ($K_d = 3.78 \mu\text{M}$) and improved inhibition of viral plaque formation (IC_{50} 's of $8.7 \mu\text{M}$ versus $41.3 \mu\text{M}$).

Small molecules possess superior pharmacokinetic properties for drug development than the rather large structures described above. Kadam and Wilson [111] identified the common buffer molecule CHES (**13**) by X-ray crystallography in complex with hemagglutinin. The molecule's binding mode with hemagglutinin mimics the one of sialic acid and its sulfonic acid superimposes with the carboxylate of sialic acid in the complex. Furthermore, the cyclohexyl moiety of CHES forms a $\text{CH}-\pi$ interaction with W153 of hemagglutinin which is normally established by the *N*-acetyl group of sialic acid. As binding of CHES, although in slightly different binding modes, was confirmed for H3-hemagglutinin and H5-hemagglutinin, Kadam and Wilson proposed this non-carbohydrate molecule as a starting point for fragment growing to overcome its very low affinity ($K_d > 20 \text{ mM}$) in the discovery of new types of hemagglutinin inhibitors.

The human parainfluenza virus causes respiratory tract diseases in children and elderly patients. In contrast to other influenza viruses, its multifunctional hemagglutinin–neuraminidase protein possesses both receptor-binding (hemagglutinin-function) and receptor-processing (neuraminidase-function) functionalities in one binding site [112]. Usually, lectins are defined as carbohydrate binding proteins without catalytic activity. However, this multifunctionality makes this parainfluenza virus protein an interesting topic for this review. Von Itzstein and co-workers synthesized a set of enzymatic intermediate-like *N*-acylated Neu-2-en and substrate-like *N*-acylated 2,3-difluoro-Neu derivatives to block both functionalities with a single molecule [43,44]. Especially the *N*-isobutyramido Neu-2-en derivatives showed potent hemagglutinin inhibition ($\text{IC}_{50} = 1.15 \mu\text{M}$) as well as inhibition of neuraminidase activity and virus growth.

Norovirus, a worldwide cause of mild to severe acute gastroenteritis, can lead to life-threatening infections for pediatric and geriatric patients and outbreaks, especially in day care centers or nursing homes, which are particularly problematic. To date, therapy of norovirus infections is only supportive and limited to reversal of dehydration and loss of electrolytes [113]. Thus, to control and prevent outbreaks, new drugs are needed. The human norovirus capsid protein P domain interacts with human blood group antigens (HBGA) and plays an important role in infection [114]. This virus–host interaction can be blocked by human milk oligosaccharides such as 2'-fucosyl lactose (**14**, 2'-FL) as shown by Hansman and co-workers [45,46]. The very high concentrations of 2'-FL needed to inhibit the interaction of virus like particles with HBGA *in vitro* ($\text{IC}_{50} = 13\text{--}50 \text{ mM}$), could be achieved because of the low toxicity of 2'-FL, its

metabolic stability and low gastrointestinal absorption [115]. Indeed, 2'-FL is a major constituent of human milk with a concentration in the mM range and has been postulated to prevent infections in breast-fed newborns [116]. Another commonly used and safe food supplement, citrate, was shown to bind norovirus in a HBGA-like manner [47].

Mammalian lectin antagonists

There are numerous mammalian lectins and the three important classes, siglecs, galectins and the C-type lectins, are currently addressed in chemical biology and medicinal chemistry. Sialic acid-binding immunoglobulin-like lectins, siglecs, are cell-surface receptors, mainly expressed by cells of the immune system. They are involved in various processes ranging from self-/non-self discrimination to regulating inflammation caused by damage-associated or pathogen-associated molecular patterns (DAMP/PAMP) [117,118]. Galectins, a family of soluble secreted lectins with 14 members, generally bind to β -galactosides [119]. Their functions are diverse and comprise mediation of cell–cell interactions, cell–matrix adhesion and transmembrane signaling [120–122]. C-type lectins are the largest and most diverse lectin family which share a conserved protein fold. The name giving Ca^{2+} -ion present in all carbohydrate recognizing family members directly mediates the binding to the glycan ligand [7]. Only a few examples exist for which Ca^{2+} is dispensable for carbohydrate recognition with dectin-1 being the most prominent example. The C-type lectin receptor family in mammals contains 17 members and many are part of innate immunity [123,124].

Langerin, DC-SIGN

All cells of the innate immune system express a variety of pattern recognition receptors (PRR) such as toll-like receptors, NOD-like receptors and C-type lectin receptors, which allow the orchestration of an appropriate biological response to an incoming microbial threat. These PRRs are specialized to recognize PAMPs such as bacterial cell wall structures, fungal polysaccharides, the viral envelope and foreign RNA/DNA [127,128]. The signaling cascades initiated by these recognition events as well as the antigen uptake and processing pathways eventually lead to activation of cells of the adaptive immune system and hence are central elements bridging these two arms of immunity. For example, PAMPs recognized and processed by dendritic cells can lead to differentiation of CD4^+ -cells into T-helper cells [123,126]. Important C-type lectin receptors are langerin, DC-SIGN and dectin-1 [123].

The homotrimeric protein langerin is expressed on Langerhans cells in epithelial and mucosal tissues and binds to D-mannose, L-fucose, and D-GlcNAc as well as sulfated D-galactose. Langerin mediates the uptake of *Yersinia pestis* and influenza A virus amongst others in host infection [7,8]. Capitalizing on these carbohydrate-mediated

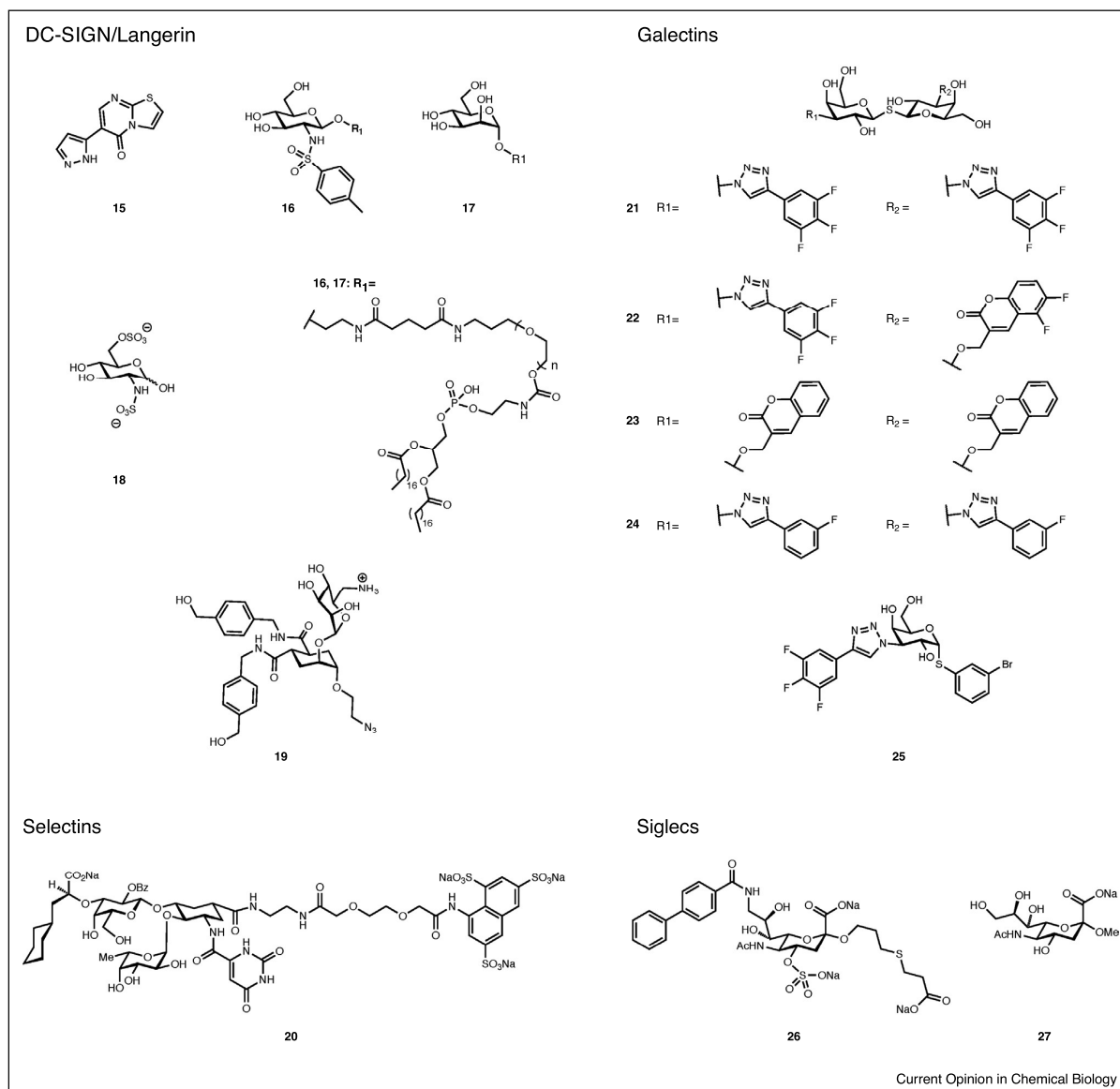
antigen uptake and processing pathways, langerin has also been described as an attractive target for targeted drug-delivery approaches to Langerhans cells [129,130]. This raised the interest in specific langerin ligands and, for example, Aretz *et al.* reported the discovery of thiazolopyrimidines as murine langerin antagonists, revealing the first allosteric inhibition of a mammalian lectin [48**]. Optimization of the initial hit **15** (Figure 4) was found beneficial at position 6 and led to up to 10-fold lower K_d and IC_{50} -values (K_d (**15**) = 0.7 mM; IC_{50} = 0.6 mM). Overall, a large series of langerin inhibitors was presented with IC_{50} values ranging in the two digit micromolar range.

Furthermore, it is well known that langerin has high affinity for sulfated polysaccharides or large oligosaccharides, for example, heparin (K_d = ~ 2.4 nM). As the binding affinity is electrostatically driven, no binding was detected with pH values below 4 or at high salt concentrations above 0.5 M [131]. A screening for langerin binding molecules revealed a sulfonamide of glucosamine as weakly binding langerin ligand [132**,133**]. Considering this screening hit, the modified phospholipids **16** and **17** were synthesized with the aim to produce glycomimetic modified liposomes for langerin targeting. These were tested against Langerin⁺, DC-SIGN⁺ or Dectin-1⁺ Raji cells. Liposomes consisting of mannosylated phospholipid **17** bound specifically to DC-SIGN⁺ cells and those consisting of sulfonamide **16** specifically to Langerin⁺ cells. Intracellular trafficking of the langerin targeting liposomes consisting of **16** was then observed in Langerin⁺ COS-7 cells by confocal microscopy.

Tetrameric DC-SIGN is expressed by myeloid dendritic cells and macrophages. Since DC-SIGN shares the same EPN amino acid motif with langerin, both proteins recognize similar monosaccharide ligands. While langerin was reported to be protective against HIV infections [134], DC-SIGN promotes viral dissemination via a process called *trans*-infection. Targeting DC-SIGN is therefore of interest to stop the transmission of HIV [135].

One common approach to increase affinity for DC-SIGN is the multivalent presentation of monosaccharide ligands. Following such an avidity-driven strategy, a dodecavalent fuco-dendrimer with a 420-fold potency increase compared to fucose was reported [136]. However, unspecific binding to langerin due to its similar binding specificity imposes a selectivity issue. GlcNAc is recognized by both C-type lectins but sulfation of position six and replacement of the *N*-acetyl group by a *N*-sulfate led to a favored recognition of the negatively charged compound **18** by langerin [125*]. The development of positively charged amino species in the pseudo-1,2-mannobioside **19** favored the selectivity toward DC-SIGN (IC_{50} = 254 μM ; (langerin IC_{50} > 4400 μM) [125*]. Pseudo-1,2-mannobiosides were shown to bind to the carbohydrate recognition domain in DC-SIGN using X-ray crystallography [137]. As an alternative approach to

Figure 4

Allosteric (**15**) and carbohydrate-binding site directed (**16–27**) mammalian lectin antagonists.

generate specificity, a recent report highlighted the presence of five secondary binding sites on DC-SIGN. These sites recognize drug-like compounds unrelated to carbohydrates, and hence constitute a potential starting point for future development [138*].

Dectin-1, a mammalian lectin of the innate immune system, recognizes β -glucans found on fungal cell walls and is able to function as a PRR in fungal-infection [124]. Liposomes carrying the currently used antifungal drug

amphotericin B intercalated into the lipid membrane reduce the antifungal's toxicity compared to detergent-solubilized drugs. Coating of these liposomes with dectin-1 for the specific targeting toward fungal cells showed a 200-fold higher affinity to those cells than untargeted liposomes [139]. These dectin-modified delivery vehicles also reduced growth and viability of the mold *Aspergillus fumigatus* with higher efficiency and thus provide a new opportunity to fight those resistant and difficult to treat infections.

Selectins

Selectins are a subfamily of the C-type lectins consisting of three single-chain transmembrane glycoproteins, which are found on endothelial cells (E-selectin or CD62E), leukocytes (L-selectin or CD62L) and platelets (P-selectin or CD62P). They are involved in constitutive lymphocyte homing, chronic and acute inflammation processes and their minimal common binding epitope is the blood group antigen sialyl Lewis X (sLe^x). [140]

In accordance with the bioactive conformation of the tetrasaccharide sLe^x for E-selectin, this carbohydrate lead was successively optimized in a series of papers from Ernst and co-workers [141–145]. NMR screening of fragments allowed the identification of a second site binder and upon merging with the first site sLe^x mimic, 30 nM lectin antagonists were obtained from a 1 mM lead [146]. Subsequent addressing of the additional sulfate-binding domain in P-selectins/L-selectins led to the successful pan-selectin antagonist Rivipansel (GMI-1070, **20**) out of the development program by Ernst and Magnani that started in the mid-1990s, despite the common fashion to drop selectin research in pharmaceutical industry in the early 2000s [147]. Since June 2015, Rivipansel is in clinical phase III studies against vaso-occlusive anemia in hospitalized subjects with sickle cell disease (trial end date: June 2019, clinicaltrials.gov Identifier: NCT02187003).

Mincle

Mincle has been identified as a C-type lectin receptor of the innate immune system with glycolipid binding specificity that plays an important role in infection by mycobacteria. Mincle binds the mycobacterial glycolipid trehalose dimycolate [20,21] and has recently been addressed by a number of groups describing synthetic molecules based on the bacterial glycolipid [148^{**},149,150,151].

Galectins

Galectin-3, the best described member of the galectin family, is involved in many biological processes, *inter alia*, cell growth, cell adhesion and apoptosis. Consequently, it plays an important role in many diseases, among them are cancer, inflammation, fibrosis, heart disease and stroke [152–154]. For that reason, galectin-3 became an important drug target, recently reviewed by Cagnoni *et al.* [11].

Symmetric C3-aryltriazolyl-substituted thiodigalactosides have shown high affinities for galectin-3 down to $K_d = 1\text{--}2\text{ nM}$. However, most of the compounds also bound to galectin-1 raising concerns about the specificity (e.g.: **21**, K_d (galectin-1) = 69 nM; K_d (galectin-3) = 2.3 nM). After combining C3 aryltriazolyl groups with O3-coumaryl groups into asymmetrical thiodigalactosides the selectivity toward galectin-3 increased: specificity of compound **22** toward galectin-3 was achieved with a high affinity (K_d (galectin-1) = 340 nM; K_d (galectin-3) = 7.5 nM) [155^{**}]. Dicoumaryl digalactoside **23** (K_d (galectin-1) = 16 μM ; K_d (galectin-

3) = 91 nM) was then analyzed *in vivo* in mice against bleomycin-induced lung fibrosis. At a dose of 3.5 mg/kg of digalactoside **23** the fibrosis score could be reduced but no effect on the inflammatory score was observed [156]. TD139 (**24**) is a derivative of **21** with a single fluorine atom in *meta*-position of the phenyl rings which is in clinical trials phase II as a galectin-3 inhibitor in idiopathic pulmonary fibrosis since February 2019 using the pulmonary route of administration (www.clinicaltrials.gov, NCT03832946) [157,158]. Oral administration of these disaccharides is impeded by their poor membrane permeability. Currently, various research groups are optimizing this property and a new galectin-inhibitor class with only one sugar residue and low nanomolar affinity was discovered, for example, **25**, $K_d = 37\text{ nM}$ [159].

Siglecs

A number of siglecs have attracted the attention in the past decades and several antibodies targeting siglecs are approved drugs or in clinical trials [160,161]. Many publications report the development of antagonists for siglec-4, also called myelin-associated glycoprotein (MAG) [162–164]. This protein is important for glial scar formation after central nervous system lesions and inhibition of MAG is considered one therapeutic approach to prevent scar formation and enable axonal regeneration [165,166].

Siglec-2 (CD22) is a target receptor in anti-cancer therapy of lymphoma, leukemia as well as in the treatment of autoimmune diseases such as lupus and rheumatoid arthritis [167]. Biphenylcarboxamidated sialic acid derivative **26** ($\text{IC}_{50} = 2\text{ nM}$) was developed with an over 500 000-fold stronger binding affinity compared to the minimal siglec ligand $\alpha\text{Me-Neu5Ac}$ (**27**, $\text{IC}_{50} = 1.5\text{ mM}$) against siglec-2 [168^{**}]. Despite the fact that this protein is a monomeric protein, divalent or trivalent *N*-glycans show a very high affinity in the low nM/high pM range. The group by Paulson suggest that this high affinity in their assays originates from simultaneous binding to several CD22 lectins clustering on the cell surface within 30–50 Å to each other [169^{**}].

Conclusions

Lectins are a large family of proteins that are present in each domain of life. These carbohydrate-binding proteins possess numerous functions, both intracellularly and outside the cell. Research toward lectin antagonists has developed rapidly over the past two decades focusing on lectins from selected fields, mainly related to immunity and infection involving mammalian lectins and those from pathogenic bacteria and viruses. The largest block of literature focusses on the assembly of native carbohydrates onto a plethora of different multivalent scaffolds. With some important exceptions discussed here, these publications usually center around the chemical synthesis and compounds are only evaluated in a target-binding assay and not employed further for questions of chemical biology and drug research.

However, in the last decade, a number of strategies toward glycomimetic lectin antagonists has been published that led to drug-like structures which proved equally useful in chemical biology research and early preclinical drug discovery. Antibacterial glycomimetic drugs applied alone or in combination with conventional antibiotics will provide new effective therapies for multi-resistant bacterial infections. And because of an increasing resistance toward established drugs and the absence of effective drugs against several, so far untreated viruses, viral lectins have become attractive targets in recent years and further research will likely yield new tools for chemical biology and drug therapy. Despite the intrinsic difficulty of developing probes/therapeutics for these low affinity carbohydrate–protein interactions, the field is developing rapidly and the first lectin antagonist currently in phase III clinical trials is GMI-1070 (20, Figure 4).

Many new lectins are being uncovered every year providing a large playground for new lectin antagonists for chemical biology and potentially as therapeutic targets. Lectins from other organisms, such as fungi or bacteria that are not pathogenic to humans are active areas of research. It will be interesting to probe, for example, fungal lectins [22,23,170,171] with a distinct specificity for methylated glycans or those of bacteria [172–174] that live in symbiosis with nematodes and kill invaded insects. Furthermore, a large number of bacterial adhesins in pathogenic bacteria are being uncovered, for example, the *Burkholderia* lectins [175–178] or carbohydrate binding adhesins from *Salmonella enterica* [179], and thus, there is a bright future for the chemical biology of lectin antagonists ahead.

Conflict of interest statement

Nothing declared.

Acknowledgements

The authors thank Dr. Christoph Rademacher for constructive comments on the manuscript. We further acknowledge funding by the Helmholtz Association (VH-NG-934).

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

- of special interest
- of outstanding interest

1. Cummings RD, Schnaar RL, Esko JD, Drickamer K, Taylor ME: **Principles of glycan recognition**. In *Essentials of Glycobiology*. Edited by Varki A, Cummings RD, Esko JD, Stanley P, Hart GW, Aebi M, Darvill AG, Kinoshita T, Packer NH, Prestegard JH, Schnaar RL, Seeberger PH. Cold Spring Harbor Laboratory Press; 2015.
2. Lis H, Sharon N: **Lectins: carbohydrate-specific proteins that mediate cellular recognition**. *Chem Rev* 1998, **98**:637–674.
3. Sharon N: **Carbohydrates as future anti-adhesion drugs for infectious diseases**. *Biochim Biophys Acta* 2006, **1760**:527–537.
4. Rodrigues JA, Acosta-Serrano A, Aebi M, Ferguson MAJ, Routier FH, Schiller I, Soares S, Spencer D, Titz A, Wilson IBH, Izquierdo L: **Parasite glycobiology: a bittersweet symphony**. *PLoS Pathog* 2015, **11**:e1005169.
5. Thompson AJ, de Vries RP, Paulson JC: **Virus recognition of glycan receptors**. *Curr Opin Virol* 2019, **34**:117–129.
6. van Kooyk Y, Rabinovich GA: **Protein-glycan interactions in the control of innate and adaptive immune responses**. *Nat Immunol* 2008, **9**:593–601.
7. Drickamer K, Taylor ME: **Recent insights into structures and functions of C-type lectins in the immune system**. *Curr Opin Struct Biol* 2015, **34**:26–34.
8. Dam TK, Brewer CF: **Lectins as pattern recognition molecules: the effects of epitope density in innate immunity**. *Glycobiology* 2010, **20**:270–279.
9. McEver RP: **Selectins: initiators of leucocyte adhesion and signalling at the vascular wall**. *Cardiovasc Res* 2015, **107**:331–339.
10. Borsig L: **Selectins in cancer immunity**. *Glycobiology* 2018, **28**:648–655.
11. Cagnoni AJ, Pérez Sáez JM, Rabinovich GA, Mariño KV: **Turning-off signaling by siglecs, selectins, and galectins: chemical inhibition of glycan-dependent interactions in cancer**. *Front Oncol* 2016, **6**:109.
12. Varki A, Gagneux P: **Biological functions of glycans**. In *Essentials of Glycobiology*. Edited by Varki A, Cummings RD, Esko JD, Stanley P, Hart GW, Aebi M, Darvill AG, Kinoshita T, Packer NH, Prestegard JH, Schnaar RL, Seeberger PH. Cold Spring Harbor Laboratory Press; 2015.
13. Fujita T: **Evolution of the lectin–complement pathway and its role in innate immunity**. *Nat Rev Immunol* 2002, **2**:346–353.
14. Hoffmann-Sommergruber K, Paschinger K, Wilson IBH: **Glycomarkers in parasitic infections and allergy**. *Biochem Soc Trans* 2011, **39**:360–364.
15. Moran AP: **Molecular mimicry of host glycosylated structures by bacteria**. In *Microbial Glycobiology*. Edited by Holst O, Brennan PJ, von Itzstein M. Academic Press; 2010:847–870.
16. Comstock LE, Kasper DL: **Bacterial glycans: key mediators of diverse host immune responses**. *Cell* 2006, **126**:847–850.
17. Perret S, Sabin C, Dumon C, Pokorná M, Gautier C, Galanina O, Ilia S, Bovin N, Nicaise M, Desmadril M et al.: **Structural basis for the interaction between human milk oligosaccharides and the bacterial lectin PA-IIL of *Pseudomonas aeruginosa***. *Biochem J* 2005, **389**:325–332.
18. Sommer R, Paulson JC, Titz A, Varrot A, Wagner S, Khaledi A, Häussler S, Nycholai CM, Imbert A: **The virulence factor LecB varies in clinical isolates: consequences for ligand binding and drug discovery**. *Chem Sci* 2016, **7**:4990–5001.
19. Guo Y, Feinberg H, Conroy E, Mitchell DA, Alvarez R, Blixt O, Taylor ME, Weis WI, Drickamer K: **Structural basis for distinct ligand-binding and targeting properties of the receptors DC-SIGN and DC-SIGNR**. *Nat Struct Mol Biol* 2004, **11**:591–598.
20. Williams SJ: **Sensing lipids with mincle: structure and function**. *Front Immunol* 2017, **8**:1662.
21. Furukawa A, Kamishikiryo J, Mori D, Toyonaga K, Okabe Y, Toji A, Kanda R, Miyake Y, Ose T, Yamasaki S, Maenaka K: **Structural analysis for glycolipid recognition by the C-type lectins Mincle and MCL**. *Proc Natl Acad Sci U S A* 2013, **110**:17438–17443.
22. Wohlschläger T, Buttschi A, Grassi P, Sutov G, Gauss R, Hauck D, Schmieder SS, Knobel M, Titz A, Dell A, Haslam SM, Hengartner MO, Aebi M, Künzler M: **Methylated glycans as conserved targets of animal and fungal innate defense**. *Proc Natl Acad Sci U S A* 2014, **111**:E2787–E2796.
23. Sommer R, Makshakova ON, Wohlschläger T, Hutin S, Marsh M, Titz A, Künzler M, Varrot A: **Crystal structures of fungal tectonin in complex with O-methylated glycans suggest key role in innate immune defense**. *Structure* 2018, **26**:391–402.e4.
24. Wilkins PP, Moore KL, McEver RP, Cummings RD: **Tyrosine sulfation of P-selectin glycoprotein ligand-1 is required for**

- high affinity binding to P-selectin. *J Biol Chem* 1995, **270**:22677-22680.
25. Dahms NM, Olson LJ, Kim J-JP: **Strategies for carbohydrate recognition by the mannose 6-phosphate receptors.** *Glycobiology* 2008, **18**:664-678.
 26. Weis WI, Drickamer K: **Structural basis of lectin-carbohydrate recognition.** *Annu Rev Biochem* 1996, **65**:441-473.
 27. Ernst B, Magnani JL: **From carbohydrate leads to glycomimetic drugs.** *Nat Rev Drug Discov* 2009, **8**:661-677.
 28. Cecioni S, Imberty A, Vidal S: **Glycomimetics versus multivalent glycoconjugates for the design of high affinity lectin ligands.** *Chem Rev* 2015, **115**:525-561.
 29. Mydock-McGrane L, Cusumano Z, Han Z, Binkley J, Kostakioti M, Hannan T, Pinkner JS, Klein R, Kalas V, Crowley J et al.: **Antivirulence C-mannosides as antibiotic-sparing, oral therapeutics for urinary tract infections.** *J Med Chem* 2016, **59**:9390-9408.
- Design, synthesis, *in vitro* and *in vivo* evaluation of the new class of C-mannosides as FimH inhibitors. The lead compounds showed improved PK and metabolic stability compared to O-mannosides and high efficacy in mouse UTI model.
30. Sattigeri JA, Garg M, Bhatteja P, Soni A, Rauf ARA, Gupta M, Deshmukh MS, Jain T, Alekar N, Barman TK et al.: **Synthesis and evaluation of thiomannosides, potent and orally active FimH inhibitors.** *Bioorg Med Chem Lett* 2018, **28**:2993-2997.
 31. Kalas V, Hibbing ME, Maddirala AR, Chugani R, Pinkner JS, Mydock-McGrane LK, Conover MS, Janetka JW, Hultgren SJ: **Structure-based discovery of glycomimetic FimH ligands as inhibitors of bacterial adhesion during urinary tract infection.** *Proc Natl Acad Sci U S A* 2018, **115**:E2819-E2828.
 32. 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.** *Angew Chem—Int Ed* 2017, **56**:16559-16564.
- Design and synthesis of the first covalent lectin inhibitor. The covalent inhibitor targeting a cysteine residue of LecA was conjugated to fluorescein and used for LecA-specific staining of *P. aeruginosa* biofilm aggregates.
33. Johansson EMV, Cruz SA, Kolomiets E, Buts L, Kadam RU, Cacciarini M, Bartels KM, Diggle SP, Cámara M, Williams P et al.: **Inhibition and dispersion of *Pseudomonas aeruginosa* biofilms by glycopeptide dendrimers targeting the fucose-specific lectin LecB.** *Chem Biol* 2008, **15**:1249-1257.
 34. Sommer R, Wagner S, Rox K, Varrot A, Hauck D, Wamhoff EC, Schreiber J, Ryckmans T, Brunner T, Rademacher C et al.: **Glycomimetic, orally bioavailable lecB inhibitors block biofilm formation of *Pseudomonas aeruginosa*.** *J Am Chem Soc* 2018, **140**:2537-2545.
- Development of small molecule LecB inhibitors with high potency, excellent receptor binding kinetics, thermodynamics, selectivity, and pharmacokinetic properties. These glycomimetic inhibitors showed inhibition of *P. aeruginosa* biofilm formation *in vitro* and are promising leads for drug development.
35. Nishikawa K, Watanabe M, Kita E, Igai K, Omata K, Yaffe MB, Natori Y: **A multivalent peptide library approach identifies a novel Shiga toxin inhibitor that induces aberrant cellular transport of the toxin.** *FASEB J* 2006, **20**:2597-2599.
 36. Fu O, Pukin AV, Vanufford HCQ, Branson TR, Thies-Weesie DME, Turnbull WB, Visser GM, Pieters RJ: **Tetra-versus pentavalent inhibitors of cholera toxin.** *Chem Open* 2015, **4**:471-477.
 37. Han X, Shi Y, Si L, Fan Z, Wang H, Xu R, Jiao P, Meng K, Tian Z, Zhou X et al.: **Design, synthesis and biological activity evaluation of novel conjugated sialic acid and pentacyclic triterpene derivatives as anti-influenza entry inhibitors.** *MedChemComm* 2016, **7**:1932-1945.
 38. Wang H, Xu R, Shi Y, Si L, Jiao P, Fan Z, Han X, Wu X, Zhou X, Yu F et al.: **Design, synthesis and biological evaluation of novel l-ascorbic acid-conjugated pentacyclic triterpene derivatives as potential influenza virus entry inhibitors.** *Eur J Med Chem* 2016, **110**:376-388.
 39. Yu M, Si L, Wang Y, Wu Y, Yu F, Jiao P, Shi Y, Wang H, Xiao S, Fu G et al.: **Discovery of pentacyclic triterpenoids as potential entry inhibitors of influenza viruses.** *J Med Chem* 2014, **57**:10058-10071.
 40. Yamabe M, Kaihatsu K, Ebara Y: **Sialyllactose-modified three-way junction DNA as binding inhibitor of influenza virus hemagglutinin.** *Bioconjug Chem* 2018, **29**:1490-1494.
- Sialic acid presented on a three-way junction DNA matches the hemagglutinin receptor binding site. The authors studied the structure activity relationship and showed highly active hemagglutinin inhibitors.
41. Yamabe M, Fujita A, Kaihatsu K, Ebara Y: **Synthesis of neuraminidase-resistant sialoside-modified three-way junction DNA and its binding ability to various influenza viruses.** *Carbohydr Res* 2019, **474**:43-50.
- As a follow-up study of Ref. 40, the introduction of a S-glycosidic bond instead of an O-glycosidic bond increased the stability against neuraminidase.
42. Strauch E-M, Bernard SM, La D, Bohn AJ, Lee PS, Anderson CE, Niesuma T, Holstein CA, Garcia NK, Hooper KA et al.: **Computational design of trimeric influenza-neutralizing proteins targeting the hemagglutinin receptor binding site.** *Nat Biotechnol* 2017, **35**:667-671.
- A highly avid trimeric protein, specifically *in silico* designed to match the binding site architecture of hemagglutinin. The resulting hemagglutinin inhibitor shows prophylactic and therapeutic activity against H3N2 in a mouse model.
43. Guillon P, Dirr L, El-Deeb IM, Winger M, Bailly B, Haselhorst T, Dyason JC, Von Itzstein M: **Structure-guided discovery of potent and dual-acting human parainfluenza virus haemagglutinin-neuraminidase inhibitors.** *Nat Commun* 2014, **5**:5268.
 44. Dirr L, El-Deeb IM, Chavas LMG, Guillon P, Von Itzstein M: **The impact of the butterfly effect on human parainfluenza virus haemagglutinin-neuraminidase inhibitor design.** *Sci Rep* 2017, **7**.
 45. Koromyslova A, Tripathi S, Morozov V, Schrotten H, Hansman GS: **Human norovirus inhibition by a human milk oligosaccharide.** *Virology* 2017, **508**:81-89.
 46. Weichert S, Koromyslova A, Singh BK, Hansman S, Jennewein S, Schrotten H, Hansman GS: **Structural basis for norovirus inhibition by human milk oligosaccharides.** *J Virol* 2016, **90**:4843-4848.
 47. Koromyslova AD, White PA, Hansman GS: **Treatment of norovirus particles with citrate.** *Virology* 2015, **485**:199-204.
 48. Aretz J, Anumala UR, Fuchsberger FF, Molavi N, Ziebart N, Zhang H, Nazaré M, Rademacher C: **Allosteric inhibition of a mammalian lectin.** *J Am Chem Soc* 2018, **140**:14915-14925.
- The first allosteric inhibition of a mammalian lectin (langerin) using thiazolopyrimidines led to binding affinities in a double-digit micromolar range.
49. Moeveer RP: **Selectins: initiators of leucocyte adhesion and signalling at the vascular wall.** *Cardiovasc Res* 2015, **107**:331-339.
 50. Clatworthy AE, Pierson E, Hung DT: **Targeting virulence: a new paradigm for antimicrobial therapy.** *Nat Chem Biol* 2007, **3**:541-548.
 51. Davies D: **Understanding biofilm resistance to antibacterial agents.** *Nat Rev Drug Discov* 2003, **2**:114-122.
 52. Hung C, Bouckaert J, Hung D, Pinkner J, Widberg C, Defusco A, Augustine CG, Strouse R, Langermann S, Waksman G, Hultgren SJ: **Structural basis of tropism of *Escherichia coli* to the bladder during urinary tract infection.** *Mol Microbiol* 2002, **44**:903-915.
 53. Hartmann M, Lindhorst TK: **The bacterial lectin FimH, a target for drug discovery – carbohydrate inhibitors of Type 1 fimbriae-mediated bacterial adhesion.** *Eur J Org Chem* 2011, **2011**:3609.
 54. Conover MS, Ruer S, Taganna J, Kalas V, De Greve H, Pinkner JS, Dodson KW, Remaut H, Hultgren SJ: **Inflammation-induced adhesin-receptor interaction provides a fitness advantage to uropathogenic *E. coli* during chronic infection.** *Cell Host Microbe* 2016, **20**:482-492.

55. Chen SL, Hung CS, Pinkner JS, Walker JN, Cusumano CK, Li Z, Bouckaert J, Gordon JI, Hultgren SJ: **Positive selection identifies an in vivo role for FimH during urinary tract infection in addition to mannose binding.** *Proc Natl Acad Sci U S A* 2009, **106**:22439-22444.
56. Schwartz DJ, Kalas V, Pinkner JS, Chen SL, Spaulding CN, Dodson KW, Hultgren SJ: **Positively selected FimH residues enhance virulence during urinary tract infection by altering FimH conformation.** *Proc Natl Acad Sci U S A* 2013, **110**:15530-15537.
57. Spaulding CN, Klein RD, Ruer S, Kau AL, Schreiber IVHL, Cusumano ZT, Dodson KW, Pinkner JS, Fremont DH, Janetka JW *et al.*: **Selective depletion of uropathogenic *E. coli* from the gut by a FimH antagonist.** *Nature* 2017, **546**:528-532.
58. Firon N, Ashkenazi S, Mirelman D, Ofek I, Sharon N: **Aromatic alpha-glycosides of mannose are powerful inhibitors of the adherence of type 1 fimbriated *Escherichia coli* to yeast and intestinal epithelial cells.** *Infect Immun* 1987, **55**:472-476.
59. Bouckaert J, Berglund J, Schembri M, De Genst E, Cools L, Wuhler M, Hung CS, Pinkner J, Slättegård R, Zavalov A *et al.*: **Receptor binding studies disclose a novel class of high-affinity inhibitors of the *Escherichia coli* FimH adhesin.** *Mol Microbiol* 2005, **55**:441-455.
60. Sivignon A, Yan X, Dorta DA, Bonnet R, Bouckaert J, Fleury E, Bernard J, Gouin SG, Darfeuille-Michaud A, Barnich N: **Development of heptylmannoside-based glycoconjugate antiadhesive compounds against adherent-invasive *Escherichia coli* bacteria associated with crohn's disease.** *mBio* 2015, **6**:1-9.
61. Chalopin T, Alvarez Dorta D, Sivignon A, Caudan M, Dumych TI, Billy RO, Deniaud D, Barnich N, Bouckaert J, Gouin SG: **Second generation of thiazolymannosides, FimH antagonists for *E. coli*-induced Crohn's disease.** *Org Biomol Chem* 2016, **14**:3913-3925.
62. Jarvis C, Han DZ, Kalas V, Klein R, Pinkner JS, Ford B, Binkley J, Cusumano CK, Cusumano Z, Mydock-McGrane L *et al.*: **Antivirulence isouquinolone mannosides: optimization of the biaryl aglycone for FimH lectin binding affinity and efficacy in the treatment of chronic UTI.** *ChemMedChem* 2016, **11**:367-373.
63. Schönmemann W, Cramer J, Mühlethaler T, Fiege B, Silbermann M, Rabbani S, Dätwyler P, Zihlmann P, Jakob RP, Sager CP *et al.*: **Improvement of aglycone π -stacking yields nanomolar to sub-nanomolar FimH antagonists.** *ChemMedChem* 2019, **14**:749-757.
64. Alvarez Dorta D, Sivignon A, Chalopin T, Dumych TI, Roos G, Billy RO, Deniaud D, Krammer EM, De Ruyck J, Lensink MF *et al.*: **The antiadhesive strategy in crohn's disease: orally active mannosides to decolonize pathogenic *Escherichia coli* from the gut.** *ChemBioChem* 2016, **17**:936-952.
65. Schönmemann W, Kleeb S, Dätwyler P, Schwardt O, Ernst B: **Prodruggability of carbohydrates – oral FimH antagonists.** *Can J Chem* 2016, **94**:909-919.
66. Kleeb S, Jiang X, Frei P, Sigl A, Bezençon J, Bamberger K, Schwardt O, Ernst B: **FimH antagonists: phosphate prodrugs improve oral bioavailability.** *J Med Chem* 2016, **59**:3163-3182.
Phosphate-prodrugs have been synthesized and increased drug availability at the site of infection.
67. Mydock-McGrane LK, Hannan TJ, Janetka JW: **Rational design strategies for FimH antagonists: new drugs on the horizon for urinary tract infection and Crohn's disease.** *Expert Opin Drug Discov* 2017, **12**:711-731.
68. Maddirala AR, Klein R, Pinkner JS, Kalas V, Hultgren SJ, Janetka JW: **Biphenyl Gal and GalNAc FimH lectin antagonists of uropathogenic *E. coli* (UPEC): optimization through iterative rational drug design.** *J Med Chem* 2019, **62**:467-479.
Structure-guided optimization was used to develop very potent FimH inhibitors with excellent metabolic stability and good PK, but low oral bioavailability. This represents a good starting point for drug development for the recently identified target FimH.
69. Tielker D, Hacker S, Loris R, Strathmann M, Wingender J, Wilhelm S, Rosenau F, Jaeger KE: ***Pseudomonas aeruginosa* lectin LecB is located in the outer membrane and is involved in biofilm formation.** *Microbiology* 2005, **151**:1313-1323.
70. Diggle SP, Stacey RE, Dodd C, Cámara M, Williams P, Winzer K: **The galactophilic lectin, LecA, contributes to biofilm development in *Pseudomonas aeruginosa*.** *Environ Microbiol* 2006, **8**:1095-1104.
71. Gilboa-Garber N: ***Pseudomonas aeruginosa* lectins.** *Methods Enzymol* 1982, **83**:378-385.
72. Wagner S, Sommer R, Hinsberger S, Lu C, Hartmann RW, Empting M, Titz A: **Novel strategies for the treatment of *Pseudomonas aeruginosa* infections.** *J Med Chem* 2016, **59**:5929-5969.
73. Calvert MB, Jumde VR, Titz A: **Pathoblockers or antivirulence drugs as a new option for the treatment of bacterial infections.** *Beilstein J Org Chem* 2018, **14**:2607-2617.
74. Titz A: **Carbohydrate-based anti-virulence compounds against chronic *Pseudomonas aeruginosa* infections with a focus on small molecules.** *Top Med Chem* 2014, **12**:169-186.
75. Blanchard B, Nurisso A, Hollville E, Tétaud C, Wiels J, Pokorná M, Wimmerová M, Varrot A, Imberty A: **Structural basis of the preferential binding of globo-series glycosphingolipids displayed by *Pseudomonas aeruginosa* Lectin I.** *J Mol Biol* 2008, **383**:837-853.
76. Eierhoff T, Bastian B, Thuenauer R, Madl J, Audfray A, Aigal S, Juillot S, Rydell GE, Muller S, de Bentzmann S *et al.*: **A lipid zipper triggers bacterial invasion.** *Proc Natl Acad Sci U S A* 2014, **111**:12895-12900.
77. Imberty A, Wimmerová M, Mitchell EP, Gilboa-Garber N: **Structures of the lectins from *Pseudomonas aeruginosa*: insights into the molecular basis for host glycan recognition.** *Microbes Infect* 2004, **6**:221-228.
78. Bernardi A, Jiménez-Barbero J, Casnati A, De Castro C, Darbre T, Fieschi F, Finne J, Funken H, Jaeger K-E, Lahmann M *et al.*: **Multivalent glycoconjugates as anti-pathogenic agents.** *Chem Soc Rev* 2013, **42**:4709-4727.
79. Boukerb AM, Rousset A, Galanos N, Méar JB, Thépaut M, Grandjean T, Gillon E, Cecioni S, Abderrahmen C, Faure K *et al.*: **Antiadhesive properties of glycoclusters against *Pseudomonas aeruginosa* lung infection.** *J Med Chem* 2014, **57**:10275-10289.
- Lectin-targeting clusters show beneficial effects in a *Pseudomonas aeruginosa* co-instillation mouse model of acute lung infection.
80. Kadam RU, Bergmann M, Hurley M, Garg D, Cacciarini M, Swiderska MA, Nativi C, Sattler M, Smyth AR, Williams P *et al.*: **A Glycopeptide dendrimer inhibitor of the galactose-specific lectin LecA and of *Pseudomonas aeruginosa* biofilms.** *Angew Chem Int Ed* 2011, **50**:10631-10635.
81. Michaud G, Visini R, Bergmann M, Salerno G, Bosco R, Gillon E, Richichi B, Nativi C, Imberty A, Stocker A *et al.*: **Overcoming antibiotic resistance in *Pseudomonas aeruginosa* biofilms using glycopeptide dendrimers.** *Chem Sci* 2016, **7**:166-182.
Lectin-antagonistic peptide dendrimers targeting LecB restored efficacy of the antibiotic tobramycin in biofilms of *Pseudomonas aeruginosa*.
82. Visini R, Jin X, Bergmann M, Michaud G, Pertici F, Fu O, Pukin A, Branson TR, Thies-Weesie DME, Kemmink J *et al.*: **Structural insight into multivalent galactoside binding to *Pseudomonas aeruginosa* Lectin LecA.** *ACS Chem Biol* 2015, **10**:2455-2462.
83. Ligeour C, Vidal O, Dupin L, Casoni F, Gillon E, Meyer A, Vidal S, Vergoten G, Lacroix J, Souteyrand E *et al.*: **Mannose-centered aromatic galactoclusters inhibit the biofilm formation of *Pseudomonas aeruginosa*.** *Org Biomol Chem* 2015, **13**:8433-8444.
84. Pertici F, Pieters RJ: **Potent divalent inhibitors with rigid glucose click spacers for *Pseudomonas aeruginosa* lectin LecA.** *Chem Commun* 2012, **48**:4008-4010.
Synthesis of divalent LecA inhibitors using azide-alkyne click chemistry. The most potent divalent inhibitor showed 545-fold increased potency compared to the monovalent alkyne ligand.
85. Yu G, Vicini AC, Pieters RJ: **Assembling of divalent ligands and their effect on divalent binding to *Pseudomonas aeruginosa* lectin LecA.** *J Org Chem* 2019, **84**:2470-2488.

86. Yu G, Thies-Weesie DME, Pieters RJ: **Tetravalent *Pseudomonas aeruginosa* adhesion lectin leca inhibitor for enhanced biofilm inhibition.** *Helv Chim Acta* 2019, **102**:e1900014.
87. Kadam RU, Garg D, Schwartz J, Visini R, Sattler M, Stocker A, ●● Darbre T, Reymond JL: **CH- π "t-shape" interaction with histidine explains binding of aromatic galactosides to *Pseudomonas aeruginosa* lectin LecA.** *ACS Chem Biol* 2013, **8**:1925-1930.
- CH- π interaction between galactoside aryl aglycon and His50 increases ligand binding potency. Identification and exploiting such interactions may help with design of small molecule inhibitors.
88. Rodrigue J, Ganne G, Blanchard B, Saucier C, Giguère D, Shiao TC, Varrot A, Imberty A, Roy R: **Aromatic thioglycoside inhibitors against the virulence factor LecA from *Pseudomonas aeruginosa*.** *Org Biomol Chem* 2013, **11**:6906-6918.
89. 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*.** *Org Biomol Chem* 2016, **14**:7933-7948.
90. Mitchell E, Houles C, Sudakevitz D, Wimmerova M, Gautier C, Pérez S, Wu AM, Gilboa-Garber N, Imberty A: **Structural basis for oligosaccharide-mediated adhesion of *Pseudomonas aeruginosa* in the lungs of cystic fibrosis patients.** *Nat Struct Biol* 2002, **9**:918-921.
91. Boukerb AM, Decor A, Ribun S, Tabaroni R, Rousset A, Commin L, Buff S, Doléans-Jordheim A, Vidal S, Varrot A et al.: **Genomic rearrangements and functional diversification of lecA and lecB lectin-coding regions impacting the efficacy of glycomimetics directed against *Pseudomonas aeruginosa*.** *Front Microbiol* 2016, **7**:1-16.
92. Hauck D, Joachim I, Frommeyer B, Varrot A, Philipp B, Möller HM, Imberty A, Exner TE, Titz A: **Discovery of two classes of potent glycomimetic inhibitors of *Pseudomonas aeruginosa* LecB with distinct binding modes.** *ACS Chem Biol* 2013, **8**:1775-1784.
93. Sommer R, Hauck D, Varrot A, Wagner S, Audfray A, Prestel A, Möller HM, Imberty A, Titz A: **Cinnamide derivatives of d-mannose as inhibitors of the bacterial virulence factor LecB from *Pseudomonas aeruginosa*.** *ChemistryOpen* 2015, **4**:756-767.
94. Sommer R, Exner TE, Titz A: **A biophysical study with carbohydrate derivatives explains the molecular basis of monosaccharide selectivity of the *Pseudomonas aeruginosa* lectin lecB.** *PLoS One* 2014, **9**:1-22.
95. Fan E, Merritt EA, Verlinde CLMJ, Hol WGJ: **AB5 toxins: structures and inhibitor design.** *Curr Opin Struct Biol* 2000, **10**:680-686.
96. Kitov PI, Sadowska JM, Mulvey G, Armstrong GD, Ling H, Pannu NS, Read RJ, Bundle DR: **Shiga-like toxins are neutralized by tailored multivalent carbohydrate ligands.** *Nature* 2000, **403**:669-672.
97. Mulvey GL, Marcato P, Kitov PI, Sadowska J, Bundle DR, Armstrong GD: **Assessment in mice of the therapeutic potential of tailored, multivalent shiga toxin carbohydrate ligands.** *J Infect Dis* 2003, **187**:640-649.
98. Nishikawa K, Matsuoka K, Kita E, Okabe N, Mizuguchi M, Hino K, Miyazawa S, Yamasaki C, Aoki J, Takashima S et al.: **A therapeutic agent with oriented carbohydrates for treatment of infections by Shiga toxin-producing *Escherichia coli* O157:H7.** *Proc Natl Acad Sci U S A* 2002, **99**:7669-7674.
99. Watanabe-Takahashi M, Sato T, Dohi T, Noguchi N, Kano F, Murata M, Hamabata T, Natori Y, Nishikawa K: **An orally applicable Shiga toxin neutralizer functions in the intestine to inhibit the intracellular transport of the toxin.** *Infect Immun* 2010, **78**:177-183.
100. Matsuoka K, Nishikawa K, Goshu Y, Koyama T, Hatano K, Matsushita T, Watanabe-Takahashi M, Natori Y, Terunuma D: **Synthetic construction of sugar-amino acid hybrid polymers involving globotriaose or lactose and evaluation of their biological activities against Shiga toxins produced by *Escherichia coli* O157:H7.** *Bioorg Med Chem* 2018, **26**:5792-5803.
101. Cervin J, Wands AM, Casselbrant A, Wu H, Krishnamurthy S, Cvjetkovic A, Estelius J, Dedec B, Sethi A, Wallom KL et al.: **GM1 ganglioside-independent intoxication by Cholera toxin.** *PLoS Pathog* 2018, **14**:e1006862.
102. Merritt EA, Sarfaty S, Feil IK, Hol WGJ: **Structural foundation for the design of receptor antagonists targeting *Escherichia coli* heat-labile enterotoxin.** *Structure* 1997, **5**:1485-1499.
103. Mitchell DD, Pickens JC, Korotkov K, Fan E, Hol WGJ: **3,5-Substituted phenyl galactosides as leads in designing effective cholera toxin antagonists: synthesis and crystallographic studies.** *Bioorg Med Chem* 2004, **12**:907-920.
104. Kumar V, Turnbull WB: **Carbohydrate inhibitors of cholera toxin.** *Beilstein J Org Chem* 2018, **14**:484-498.
105. Wands AM, Cervin J, Huang H, Zhang Y, Youn G, Brautigam CA, ●● Matson Dzebo M, Björklund P, Wallenius V, Bright DK et al.: **Fucosylated molecules competitively interfere with cholera toxin binding to host cells.** *ACS Infect Dis* 2018, **4**:758-770.
- Inhibition of cholera toxin by targeting the neglected fucose binding site compared to the well studied GM1 primary binding site. For the first time, the fucosylated polymers were used to inhibit cholera toxin binding to human cells *in vitro*.
106. Varki A, Schnaar RL, Schauer R: *Sialic Acids and Other Nonulosonic Acids*. Cold Spring Harbor Laboratory Press; 2015.
107. Sauter NK, Bednarski MD, Wurzburg BA, Hanson JE, Whitesides GM, Skehel JJ, Wiley DC: **Hemagglutinins from two influenza virus variants bind to sialic acid derivatives with millimolar dissociation constants: a 500-MHz proton nuclear magnetic resonance study.** *Biochemistry* 1989, **28**:8388-8396.
108. Nizet V, Varki A, Aebi M: **Microbial lectins: hemagglutinins, adhesins, and toxins.** In *Essentials of Glycobiology*. Edited by Varki A, Cummings RD, Esko JD, Stanley P, Hart GW, Aebi M, Darvill AG, Kinoshita T, Packer NH, Prestegard JH, Schnaar RL, Seeberger PH. Cold Spring Harbor Laboratory Press; 2015.
109. Li F, Ma C, Wang J: **Inhibitors targeting the influenza virus hemagglutinin.** *Curr Med Chem* 2015, **22**.
110. Nagao M, Matsubara T, Hoshino Y, Sato T, Miura Y: **Topological design of star glycopolymers for controlling the interaction with the influenza virus.** *Bioconj Chem* 2019, **30**:1192-1198 <http://dx.doi.org/10.1021/acs.bioconjchem.9b00134>.
111. Kadam RU, Wilson IA: **A small-molecule fragment that emulates binding of receptor and broadly neutralizing antibodies to influenza A hemagglutinin.** *Proc Natl Acad Sci U S A* 2018, **115**:4240-4245.
112. Moscona A: **Entry of parainfluenza virus into cells as a target for interrupting childhood respiratory disease.** *J Clin Invest* 2005, **115**:1688-1698.
113. Robilotti E, Deresinski S, Pinsky BA: **Norovirus.** *Clin Microbiol Rev* 2015, **28**:134-164.
114. Taube S, Mallagaray A, Peters T: **Norovirus, glycans and attachment.** *Curr Opin Virol* 2018, **31**:33-42.
115. Coulet M, Phothirath P, Allais L, Schiller B: **Pre-clinical safety evaluation of the synthetic human milk, nature-identical, oligosaccharide 2'-O-Fucosyllactose (2'FL).** *Regul Toxicol Pharmacol* 2014, **68**:59-69.
116. Morrow AL, Ruiz-Palacios GM, Altaye M, Jiang X, Lourdes Guerrero M, Meinen-Derr JK, Farkas T, Chaturvedi P, Pickering LK, Newburg DS: **Human milk oligosaccharides are associated with protection against diarrhea in breast-fed infants.** *J Pediatr* 2004, **145**:297-303.
117. Pillai S, Netravali IA, Cariappa A, Mattoo H: **Siglecs and immune regulation.** *Annu Rev Immunol* 2012, **30**:357-392.
118. Macauley MS, Crocker PR, Paulson JC: **Siglec-mediated regulation of immune cell function in disease.** *Nat Rev Immunol* 2014, **14**:653-666.
119. Barondes SH, Cooper DN, Gitt MA, Leffler H: **Galectins. Structure and function of a large family of animal lectins.** *J Biol Chem* 1994, **269**:20807-20810.

120. Thiemann S, Baum LG: **Galectins and immune responses-just how do they do those things they do?** *Annu Rev Immunol* 2016, **34**:243-264.
121. Compagno D, Jaworski FM, Gentilini L, Contrufo G, González Pérez I, Elola MT, Pregi N, Rabinovich GA, Laderach DJ: **Galectins: major signaling modulators inside and outside the cell.** *Curr Mol Med* 2014, **14**:630-651.
122. Rabinovich GA, Toscano MA: **Turning "sweet" on immunity: galectin-glycan interactions in immune tolerance and inflammation.** *Nat Rev Immunol* 2009, **9**:338-352.
123. van den Berg LM, Gringhuis SI, Geijtenbeek TBH: **An evolutionary perspective on C-type lectins in infection and immunity.** *Ann N Y Acad Sci* 2012, **1253**:149-158.
124. Brown GD, Willment JA, Whitehead L: **C-type lectins in immunity and homeostasis.** *Nat Rev Immunol* 2018, **18**:374-389.
125. Porkolab V, Chabrol E, Varga N, Ordanini S, Sutkevičiūtė I, Thépaut M, García-Jiménez MJ, Girard E, Nieto PM, Bernardi A, Fieschi F: **Rational-differential design of highly specific glycomimetic ligands: targeting DC-SIGN and excluding langerin recognition.** *ACS Chem Biol* 2018, **13**:600-608.
- GlcNAc is recognized by both lectins—langerin and DC-SIGN. Selectivity toward langerin was achieved by sulfation on position six.
126. Wilson NJ, Boniface K, Chan JR, McKenzie BS, Blumenschein WM, Mattson JD, Basham B, Smith K, Chen T, Morel F *et al.*: **Development, cytokine profile and function of human interleukin 17-producing helper T cells.** *Nat Immunol* 2007, **8**:950-957.
127. Yang K, Park CG, Cheong C, Bulgheresi S, Zhang S, Zhang P, He Y, Jiang L, Huang H, Ding H *et al.*: **Host Langerin (CD207) is a receptor for *Yersinia pestis* phagocytosis and promotes dissemination.** *Immunol Cell Biol* 2015, **93**:815-824.
128. Ng WC, Londrigan SL, Nasr N, Cunningham AL, Turville S, Brooks AG, Reading PC: **The C-type lectin langerin functions as a receptor for attachment and infectious entry of Influenza A virus.** *J Virol* 2016, **90**:206-221.
129. Idoyaga J, Suda N, Suda K, Park CG, Steinman RM: **Antibody to Langerin/CD207 localizes large numbers of CD8alpha+ dendritic cells to the marginal zone of mouse spleen.** *Proc Natl Acad Sci U S A* 2009, **106**:1524-1529.
130. Flacher V, Tripp CH, Stoitzner P, Haid B, Ebner S, Del Frari B, Koch F, Park CG, Steinman RM, Idoyaga J, Romani N: **Epidermal Langerhans cells rapidly capture and present antigens from C-type lectin-targeting antibodies deposited in the dermis.** *J Invest Dermatol* 2010, **130**:755-762.
131. Zhao J, Liu X, Kao C, Zhang E, Li Q, Zhang F, Linhardt RJ: **Kinetic and structural studies of interactions between glycosaminoglycans and langerin.** *Biochemistry* 2016, **55**:4552-4559.
132. Wamhoff E-C, Schulze J, Bellmann L, Bachem G, Fuchsberger FF, Rademacher J, Hermann M: **A specific, glycomimetic Langerin ligand for human Langerhans cell targeting.** *bioRxiv* 2018 <http://dx.doi.org/10.1101/286021>.
- Intracellular trafficking of synthesized langerin targeting liposomes was observed in Langerin+ COS-7 cells by confocal microscopy. It paves the way for trans-cutaneous vaccinations using these liposomes in therapeutic applications.
133. Wamhoff E-C, Schulze J, Bellmann L, Rentzsch M, Bachem G, Fuchsberger FF, Rademacher J, Hermann M, Del Frari B, van Dalen R *et al.*: **A specific, glycomimetic langerin ligand for human langerhans cell targeting.** *ACS Cent Sci* 2019, **5**:808-820 <http://dx.doi.org/10.1021/acscentsci.9b00093>.
- Intracellular trafficking of synthesized langerin targeting liposomes was observed in Langerin+ COS-7 cells by confocal microscopy. It paves the way for trans-cutaneous vaccinations using these liposomes in therapeutic applications.
134. de Witte L, Nabatov A, Pion M, Fluitsma D, de Jong MAWP, de Gruijl T, Piguet V, van Kooyk Y, Geijtenbeek TBH: **Langerin is a natural barrier to HIV-1 transmission by Langerhans cells.** *Nat Med* 2007, **13**:367-371.
135. Geijtenbeek TB, Kwon DS, Torensma R, van Vliet SJ, van Duijnhoven GC, Middel J, Cornelissen IL, Nottet HS, KewalRamani VN, Littman DR *et al.*: **DC-SIGN, a dendritic cell-specific HIV-1-binding protein that enhances trans-infection of T cells.** *Cell* 2000, **100**:587-597.
136. Bertolotti B, Sutkevičiūtė I, Ambrosini M, Ribeiro-Viana R, Rojo J, Fieschi F, Dvořáková H, Kašáková M, Parkan K, Hlaváčková M *et al.*: **Polyvalent C-glycomimetics based on l-fucose or d-mannose as potent DC-SIGN antagonists.** *Org Biomol Chem* 2017, **15**:3995-4004.
137. Thépaut M, Guzzi C, Sutkevičiūtė I, Sattin S, Ribeiro-Viana R, Varga N, Chabrol E, Rojo J, Bernardi A, Angulo J *et al.*: **Structure of a glycomimetic ligand in the carbohydrate recognition domain of C-type Lectin DC-SIGN. Structural requirements for selectivity and ligand design.** *J Am Chem Soc* 2013, **135**:2518-2529.
138. Aretz J, Baukman H, Shanina E, Hanske J, Wawrzinek R, Zapolskii VA, Seeberger PH, Kaufmann DE, Rademacher C: **Identification of multiple druggable secondary sites by fragment screening against DC-SIGN.** *Angew Chem Int Ed* 2017, **56**:7292-7296.
- Increased number of druggable pockets on DC-SIGN allows the development of new multivalent compounds with higher binding affinities. The inhibition of the cell-surface receptor DC-SIGN is important due to pathogenic threats.
139. Ambati S, Ferraro AR, Kang SE, Lin J, Lin X, Momany M, Lewis ZA, Meagher RB: **Dectin-1-targeted antifungal liposomes exhibit enhanced efficacy.** *mSphere* 2019, **4**:e00025-19.
140. Ley K: **The role of selectins in inflammation and disease.** *Trends Mol Med* 2003, **9**:263-268.
141. Peters T, Scheffler K, Ernst B, Katopodis A, Magnani JL, Wang WT, Weisemann R: **Determination of the bioactive conformation of the carbohydrate ligand in the E-selectin/Sialyl LewisX complex.** *Angew Chem Int Ed Engl* 1995, **34**:1841-1844.
142. Thoma G, Magnani JL, Patton JT, Ernst B, Jahnke W: **Preorganization of the bioactive conformation of Sialyl LewisX analogues correlates with their affinity to E-selectin.** *Angew Chem Int Ed* 2001, **40**:1941-1945.
143. Norman KE, Anderson GP, Kolb HC, Ley K, Ernst B: **Sialyl Lewis (x) (sLe(x)) and an sLe(x) mimetic, CGP69669A, disrupt E-selectin-dependent leukocyte rolling in vivo.** *Blood* 1998, **91**:475-483.
144. Schwizer D, Patton JT, Cutting B, Smieško M, Wagner B, Kato A, Weckerle C, Binder FPC, Rabbani S, Schwaradt O, Magnani JL, Ernst B: **Pre-organization of the core structure of E-selectin antagonists.** *Chem – A Eur J* 2012, **18**:1342-1351.
145. Kolb HC, Ernst B: **Development of tools for the design of selectin antagonists.** *Chem – A Eur J* 1997, **3**:1571-1578.
146. Egger J, Weckerle C, Cutting B, Schwaradt O, Rabbani S, Lemme K, Ernst B: **Nanomolar E-selectin antagonists with prolonged half-lives by a fragment-based approach.** *J Am Chem Soc* 2013, **135**:9820-9828.
147. Chang J, Patton JT, Sarkar A, Ernst B, Magnani JL, Frenette PS: **GMI-1070, a novel pan-selectin antagonist, reverses acute vascular occlusions in sickle cell mice.** *Blood* 2010, **116**:1779-1786.
148. Decout A, Silva-Gomes S, Drocourt D, Barbe S, André I, Cueto FJ, Lioux T, Sancho D, Pérouzel E, Vercellone A *et al.*: **Rational design of adjuvants targeting the C-type lectin Mincle.** *Proc Natl Acad Sci U S A* 2017, **114**:2675-2680.
- Structure-based design of mincle inhibitors as promising vaccine adjuvants.
149. Feinberg H, Rambaruth NDS, Jégouzo SAF, Jacobsen KM, Djurhuus R, Poulsen TB, Weis WI, Taylor ME, Drickamer K: **Binding sites for acylated trehalose analogs of glycolipid ligands on an extended carbohydrate recognition domain of the macrophage receptor mincle.** *J Biol Chem* 2016, **291**:21222-21233.
150. Matsumaru T, Ikeno R, Shuchi Y, Iwamatsu T, Tadokoro T, Yamasaki S, Fujimoto Y, Furukawa A, Maenaka K: **Synthesis of glycerolipids containing simple linear acyl chains or aromatic rings and evaluation of their Mincle signaling activity.** *Chem Commun (Camb)* 2019, **55**:711-714.

151. Bird JH, Khan AA, Nishimura N, Yamasaki S, Timmer MSM, Stocker BL: **Synthesis of branched trehalose glycolipids and their mincle agonist activity.** *J Org Chem* 2018, **83**:7593-7605.
152. Dumic J, Dabelic S, Flögel M: **Galectin-3: an open-ended story.** *Biochim Biophys Acta – Gen Subj* 2006, **1760**:616-635.
153. Sharma UC, Pokharel S, van Brakel TJ, van Berlo JH, Cleutjens JPM, Schroen B, Andre' S, Crijns HJGM, Gabius H-J, Maessen J, Pinto YM: **Galectin-3 marks activated macrophages in failure-prone hypertrophied hearts and contributes to cardiac dysfunction.** *Circulation* 2004, **110**:3121-3123128.
154. Raimond J, Zimonjic DB, Mignon C, Mattei M-G, Popescu NC, Monsigny M, Legrand A: **Mapping of the galectin-3 gene (LGALS3) to human chromosome 14 at region 14q21-22.** *Mamm Genome* 1997, **8**:706-707.
155. Peterson K, Kumar R, Stenström O, Verma P, Verma PR, Håkansson M, Kahl-Knutsson B, Zetterberg F, Leffler H, Akke M et al.: **Systematic tuning of fluoro-galectin-3 interactions provides thiodigalactoside derivatives with single-digit nM affinity and high selectivity.** *J Med Chem* 2018, **61**:1164-1175.
Selectivity for galectin-3 inhibitors over galectin-1 is important for the targeting of galectin-3, for example, in cancer, inflammation and fibrosis. Asymmetric thiodigalactosides were designed and synthesized for the selective inhibition of galectin-3.
156. Rajput VK, MacKinnon A, Mandal S, Collins P, Blanchard H, Leffler H, Sethi T, Schambye H, Mukhopadhyay B, Nilsson UJ: **A selective galactose-coumarin-derived Galectin-3 inhibitor demonstrates involvement of Galectin-3-glycan interactions in a pulmonary fibrosis model.** *J Med Chem* 2016, **59**:8141-8147.
157. Delaine T, Collins P, MacKinnon A, Sharma G, Stegmayr J, Rajput VK, Mandal S, Cumpstey I, Larumbe A, Salameh BA et al.: **Galectin-3-binding glycomimetics that strongly reduce bleomycin-induced lung fibrosis and modulate intracellular glycan recognition.** *ChemBioChem* 2016, **17**:1759-1770.
158. Chen W-S, Cao Z, Leffler H, Nilsson UJ, Panjwani N: **Galectin-3 inhibition by a small-molecule inhibitor reduces both pathological corneal neovascularization and fibrosis.** *Invest Ophthalmol Vis Sci* 2017, **58**:9.
159. Zetterberg FR, Peterson K, Johnsson RE, Brimert T, Håkansson M, Logan DT, Leffler H, Nilsson UJ: **Monosaccharide derivatives with low-nanomolar lectin affinity and high selectivity based on combined fluorine-amide, phenyl-arginine, sulfur- π , and halogen bond interactions.** *ChemMedChem* 2018, **13**:133-137.
160. Angata T, Nycholat CM, Macauley MS: **Therapeutic targeting of siglecs using antibody- and glycan-based approaches.** *Trends Pharmacol Sci* 2015, **36**:645-660.
161. O'Reilly MK, Paulson JC: **Siglecs as targets for therapy in immune cell mediated disease.** *Trends Pharmacol Sci* 2009, **30**:240.
162. Schwardt O, Kelm S, Ernst B: **SIGLEC-4 (MAG) antagonists: from the natural carbohydrate epitope to glycomimetics.** *Top Curr Chem* 2013:151-200.
163. Zaccai NR, Maenaka K, Maenaka T, Crocker PR, Brossmer R, Kelm S, Jones EY: **Structure-guided design of sialic acid-based Siglec inhibitors and crystallographic analysis in complex with sialoadhesin.** *Structure* 2003, **11**:557-567.
164. Zeng Y, Rademacher C, Nycholat CM, Futakawa S, Lemme K, Ernst B, Paulson JC: **High affinity sialoside ligands of myelin associated glycoprotein.** *Bioorg Med Chem Lett* 2011, **21**:5045-5049.
165. Lopez PHH: **Role of myelin-associated glycoprotein (Siglec-4a) in the nervous system.** *Adv Neurobiol* 2014:245-262.
166. Schnaar RL, Collins BE, Wright LP, Kiso M, Tropak MB, Roder JC, Crocker PR: **Myelin-associated glycoprotein binding to gangliosides. Structural specificity and functional implications.** *Ann N Y Acad Sci* 1998, **845**:92-105.
167. Macauley MS, Crocker PR, Paulson JC: **Siglec-mediated regulation of immune cell function in disease.** *Nat Rev Immunol* 2014, **14**:653-666.
168. Prescher H, Schweizer A, Kuhfeldt E, Nitschke L, Brossmer R: **Discovery of multifold modified sialosides as Human CD22/ Siglec-2 ligands with nanomolar activity on B-cells.** *ACS Chem Biol* 2014, **9**:1444-1450.
Modified sialoside inhibitors against CD22 were synthesized and showed increased binding affinities. These compounds are useful for further investigation of the function of CD22.
169. Peng W, Paulson JC: **CD22 ligands on a natural N-glycan scaffold efficiently deliver toxins to B-lymphoma cells.** *J Am Chem Soc* 2017, **139**:12450-12458.
A chemically defined natural N-linked glycan scaffold showed 1500-fold increase in potency compared to the monovalent ligand. Conjugates of auristatin and saporin toxins with this scaffold resulted in efficient killing of the B-cell lymphoma cells. This represents an alternative strategy to the antibody and nanoparticle mediated approaches for drug delivery.
170. Cabanettes A, Perkams L, Spies C, Unverzagt C, Varrot A: **Recognition of complex core-fucosylated N-glycans by a mini lectin.** *Angew Chem Int Ed Engl* 2018, **57**:10178-10181.
171. Varrot A, Basheer SM, Imberty A: **Fungal lectins: structure, function and potential applications.** *Curr Opin Struct Biol* 2013, **23**:678-685.
172. Kumar A, Sýkorová P, Demo G, Dobeš P, Hyřil P, Wimmerová M: **A novel fucose-binding lectin from *Photobacterium luminescens* (PLL) with an unusual heptabladed β -propeller tetrameric structure.** *J Biol Chem* 2016, **291**:25032-25049.
173. Jančaříková G, Houser J, Dobeš P, Demo G, Hyřil P, Wimmerová M: **Characterization of novel bangle lectin from *Photobacterium symbiotica* with dual sugar-binding specificity and its effect on host immunity.** *PLoS Pathog* 2017, **13**: e1006564.
174. Beshr G, Sikandar A, Jemiller E-M, Klymiuk N, Hauck D, Wagner S, Wolf E, Koehnke J, Titz A: ***Photobacterium luminescens* lectin A (PLA): a new probe for detecting α -galactoside-terminating glycoconjugates.** *J Biol Chem* 2017, **292**:19935-19951.
175. Lameignere E, Malinová L, Sláviková M, Duchaud E, Mitchell EP, Varrot A, Sedo O, Imberty A, Wimmerová M: **Structural basis for mannose recognition by a lectin from opportunistic bacteria *Burkholderia cenocepacia*.** *Biochem J* 2008, **411**:307-318.
176. Beshr G, Sommer R, Hauck D, Siebert DCB, Hofmann A, Imberty A, Titz A: **Development of a competitive binding assay for the *Burkholderia cenocepacia* lectin BC2L-A and structure activity relationship of natural and synthetic inhibitors.** *MedChemComm* 2016, **7**:519-530.
177. Šulák O, Cioci G, Delia M, Lahmann M, Varrot A, Imberty A, Wimmerová M: **A TNF-like trimeric lectin domain from *Burkholderia cenocepacia* with specificity for fucosylated human histo-blood group antigens.** *Structure* 2010, **18**:59-72.
178. Šulák O, Cioci G, Lameignere E, Balloy V, Round A, Gutsche I, Malinová L, Chignard M, Kosma P, Aubert DF et al.: ***Burkholderia cenocepacia* BC2L-C is a super lectin with dual specificity and proinflammatory activity.** *PLoS Pathog* 2011, **7**: e1002238.
179. Wagner C, Barlag B, Gerlach RG, Deiwick J, Hensel M: **The *Salmonella enterica* giant adhesin SiiE binds to polarized epithelial cells in a lectin-like manner.** *Cell Microbiol* 2014, **16**:962-975.

Glycosides

How to cite: *Angew. Chem. Int. Ed.* **2021**, *60*, 8104–8114

International Edition: doi.org/10.1002/anie.202013217

German Edition: doi.org/10.1002/ange.202013217

Non-Carbohydrate Glycomimetics as Inhibitors of Calcium(II)-Binding Lectins

Sakonwan Kuhaudomlarp⁺, Eike Siebs⁺, Elena Shanina, Jérémie Topin, Ines Joachim, Priscila da Silva Figueiredo Celestino Gomes, Annabelle Varrot, Didier Rognan, Christoph Rademacher, Anne Imberty,* and Alexander Titz*

Abstract: Because of the antimicrobial resistance crisis, lectins are considered novel drug targets. *Pseudomonas aeruginosa* utilizes LecA and LecB in the infection process. Inhibition of both lectins with carbohydrate-derived molecules can reduce biofilm formation to restore antimicrobial susceptibility. Here, we focused on non-carbohydrate inhibitors for LecA to explore new avenues for lectin inhibition. From a screening cascade we obtained one experimentally confirmed hit, a catechol, belonging to the well-known PAINS compounds. Rigorous analyses validated electron-deficient catechols as millimolar LecA inhibitors. The first co-crystal structure of a non-carbohydrate inhibitor in complex with a bacterial lectin clearly demonstrates the catechol mimicking the binding of natural glycosides with LecA. Importantly, catechol **3** is the first non-carbohydrate lectin ligand that binds bacterial and mammalian calcium(II)-binding lectins, giving rise to this fundamentally new class of glycomimetics.

Introduction

Glycoconjugates on the cell surface of host tissues serve as recognition patterns for microbial carbohydrate-binding proteins, lectins, in the initial steps of an infection with a pathogen.^[1] The co-evolution between host glycoconjugates and pathogen receptors resulted in the ability of lectins to decipher the three-dimensional structure of complex

branched oligosaccharides, referred to as “glycocode”.^[2] Among pathogens, *Pseudomonas aeruginosa* is an opportunistic Gram-negative bacterium inter alia responsible for pneumonia in immunocompromised patients, those suffering from cystic fibrosis (CF), or patients under ventilation.^[3] *P. aeruginosa* starts forming a biofilm on its colonization site to shield the bacterium from antimicrobial treatment and the immune response and thus hampers combating the bacteria. Host recognition, tissue adhesion and biofilm formation of *P. aeruginosa* are mediated by two lectins: D-galactose-specific and L-fucose/b-mannose-specific LecA and LecB, respectively.^[4] Both lectins are virulence factors and have been considered as targets for the development of carbohydrate-based anti-infective compounds and for the delivery of antibiotics via carbohydrate targeting moieties.^[5] LecA is of special interest because of its ability to damage cell membranes and to facilitate bacterial cell internalization.^[6]

LecA is a homotetramer containing one calcium ion in the carbohydrate-binding site coordinating amino acids and galactosides (Figure 1 A,B).^[7] The target of this lectin is likely the globotriaosylceramide glycolipid that presents the Gal α 1-4Gal epitope.^[8] The relatively low affinity for galactose or galactobiose ($K_d = 50 \mu\text{M}$) is overcome by high avidity with divalent or multivalent ligands that can cross-link the neighbouring binding sites ($K_d = 1\text{--}20 \text{ nM}$).^[5a,9] However, because of the presence of a hydrophobic surface nearby

[*] S. Kuhaudomlarp,^[1] J. Topin, A. Varrot, A. Imberty
Université Grenoble Alpes, CNRS, CERMAV
38000 Grenoble (France)
E-mail: anne.imberty@cermav.cnrs.fr
E. Siebs,^[1] I. Joachim, A. Titz
Chemical Biology of Carbohydrates (CBCH)
Helmholtz-Institute for Pharmaceutical Research Saarland (HIPS),
Helmholtz Centre for Infection Research
66123 Saarbrücken (Germany)
and
Department of Chemistry, Saarland University
66123 Saarbrücken (Germany)
and
Deutsches Zentrum für Infektionsforschung (DZIF)
Hannover-Braunschweig (Germany)
E-mail: alexander.titz@helmholtz-hzi.de
E. Shanina, C. Rademacher
Department of Biomolecular Systems, Max Planck Institute of
Colloids and Interfaces
14424 Potsdam (Germany)
and
Institute of Chemistry and Biochemistry

Department of Biology, Chemistry and Pharmacy
Freie Universität Berlin
14195 Berlin (Germany)

J. Topin
Institute of Chemistry-Nice, UMR 7272 CNRS
Université Côte d'Azur
06108 Nice (France)

P. da Silva Figueiredo Celestino Gomes, D. Rognan
Laboratoire d'Innovation Thérapeutique
UMR 7200 CNRS-Université de Strasbourg
67400 Illkirch (France)

[†] 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.202013217.



© 2020 The Authors. Angewandte Chemie International Edition published by Wiley-VCH GmbH. This is an open access article under the terms of the Creative Commons Attribution Non-Commercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

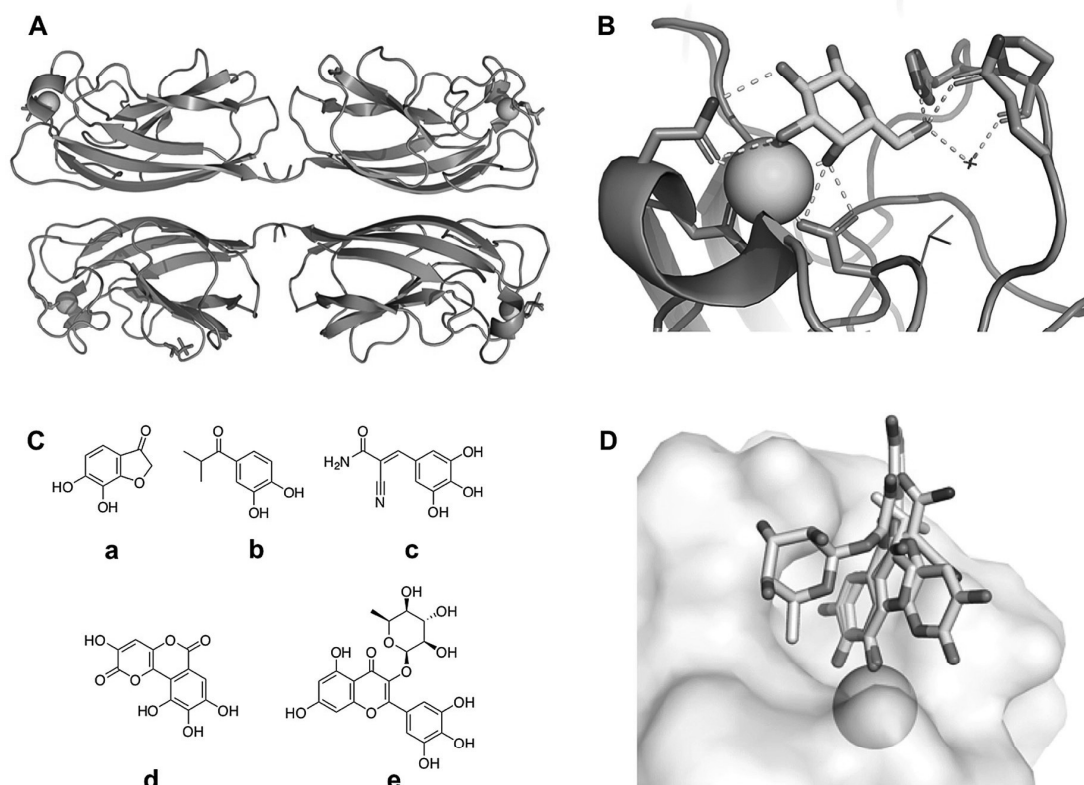


Figure 1. A) Crystal structure of the homotetrameric LecA in complex with galactose (PDB code: 1OKO). B) View of calcium-ion dependent galactose coordination in the carbohydrate-binding site of LecA. C/D) Shown are docking poses of five catechol-containing compounds from the 60 best virtual screening hits. Compound b is identical to **21** in Table 2.

the galactose-binding site, monovalent galactosides with aromatic aglycons demonstrated high affinity for the lectin in the low micromolar range.^[10] Recently, we reported an epoxide-carrying galactoside that binds covalently to a cysteine residue present in the carbohydrate-binding site of LecA.^[11]

Targeting LecA is a promising way to counteract *P. aeruginosa* infections. The chemical properties of the binding pocket of LecA were used as an inspiration for the synthesis of many compounds, all maintaining the presence of a galactose residue for the primary calcium-dependent binding site.^[5a,c,12] The development of non-carbohydrate mimics would be an alternative strategy, opening a large chemical space for novel inhibitors. Indeed, since carbohydrate-protein interactions are mostly governed by a complex arrangement of hydrogen bonds and hydrophobic contacts, the success of non-carbohydrate analogues is challenging. Non-carbohydrate glycomimetics were obtained for the model plant lectin Concanavalin A^[13] and virtual screening resulted in ligands with higher affinities than the natural ligand mannose.^[14] Apart from this work on a model lectin, only C-type lectins, that is, calcium-dependent animal lectins with roles in inflammation and immunity, have been targeted. Several compounds for the selectins have been obtained.^[12,15] Glyco-

mimetics based on a shikimic acid scaffold have shown good inhibitory activity against the C-type lectin DC-SIGN.^[16] The presence of a calcium ion is therefore a promising factor for a glycomimetic strategy against LecA.

Virtual screening is a powerful technique to identify potential ligands from large chemical databases.^[17] Therefore, this approach was applied to screen the National Cancer Institute (NCI) Diversity IV database for identifying non-carbohydrate ligands directed to the galactose binding site of LecA. The best virtual screening hits were experimentally tested in a competitive binding assay based on fluorescence polarization,^[18] identifying catechol-containing compounds as binder. Further, catechols were confirmed in orthogonal biophysical methods: a thermal shift assay (TSA), ligand- and protein-observed ¹⁹F (ProF) and ¹H-¹⁵N TROSY NMR spectroscopy, surface plasmon resonance (SPR) and X-ray crystallography was employed to elucidate their binding modes to LecA. This work demonstrates the first example of a novel group of non-carbohydrate glycomimetic compounds targeting a bacterial lectin, which can be subsequently developed into a novel class of LecA inhibitors.

Results and Discussion

Virtual Screening of Ligands for LecA from NCI and Validation of Catechol-Containing Hits in a Competitive Binding Assay

A docking protocol was designed using Glide^[19] and the 1.5 Å resolution crystal structure of LecA in complex with a galactoside (PDB code 3ZYH).^[10a] After internal validation with ligands of known structures and affinities, the E-model score function was selected for ranking of orientations together with MM-GBSA^[20] for ranking of energies (Figure S1). In total 1597 molecules from the National Cancer Institute (NCI) diversity set IV were docked into the binding site of LecA. Only those poses presenting a minimum of four contacts in the site, that is, hydrogen bonds and coordination to the calcium ion, were retained and ranked based on MM-GBSA rescoring function which resulted in the top 60 poses (Table S1). Those poses with at least two contacts to the calcium ion were manually clustered according to their similarity in binding mode into 15 groups and ranked as a function of their lower energy pose. We identified the 10 top clusters shown in Figure S2 that illustrate the encountered variety of chemical space. Interestingly, clusters 1, 5 and 6 corresponded to analogs of nucleotides, binding to LecA through 5-membered furanose rings. Other molecules bound to the calcium ion through glycerol side chains (cluster 3), sulfate groups (cluster 4) or a carboxylate (cluster 9). Of special interest was cluster 2 with five molecules (Figure 1 C) that have a catechol motif to coordinate to the calcium ion with the two vicinal oxygen atoms in a very conserved geometry (Figure 1 D). Finally, the lowest energy hit, 4-isobutyryl catechol (compound b in Figure 1 C, **21** in Table 2), allows coordination to the calcium ion and an appropriate

balance of hydrogen bonds and hydrophobic contacts that mimics the binding of galactose (see Figure S3).

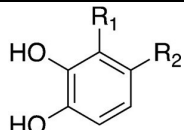
Initially, 40 compounds derived from virtual screening were selected based on structural diversity (Table S1), purchased and screened in the established competitive binding assay based on fluorescence polarization at a single concentration of 3.3 mM. Among the 40 compounds tested, 4-isobutyryl catechol (b in Figure 1 C, compound **21** in Table 2) was identified as weak inhibitor (approx. 10% inhibition at 3.3 mM). It was the only molecule able to displace the fluorescent galactoside ligand in this assay. Therefore, both virtual and experimental screening suggest the catechol as a new scaffold for LecA.

Experimental Screening of a Catechol Library

Despite the fact that some catechols are approved drugs, for example, the β_1 -adrenergic agonist dobutamine or the antibiotic cefiderocol, this class of compounds is well known as frequent hitter and pan-assay interference compounds (PAINS).^[21] Catechols are prone to oxidation into highly reactive quinones, which are reactive towards nucleophiles and may lead to an unspecific interaction with proteins.^[22] However, the concept of a general exclusion of PAINS has been recently questioned based on extensive analysis of publicly available assay data^[23] and a case-to-case consideration using orthogonal assay activity has been stated necessary.

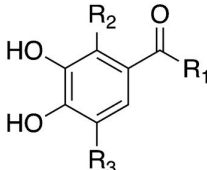
The structure–activity relationship of catechols was followed up by purchasing 29 further commercial catechols with different substitution patterns and biophysical analysis using various orthogonal methods (Table 1, Table 2). Unsubstituted

Table 1: Catechols and LecA binding by fluorescence polarization (FP), thermal shift assay (TSA), and surface plasmon resonance (SPR).

CAS no.			FP	FP with 2ME	TSA	SPR	ProF NMR	
	R ₁	R ₂	Inhibition [%] at 4 mM	Inhibition [%] at 4 mM	ΔT [K]	Efficiency [%] at 1 mM	CSP of W42 [ppm]	
1	120-80-9	H	H	inconcl.	none	−4.5	5.2	<i>n.s.</i>
2	4018-65-9	Cl	H	inconcl.	none	–	2.4	<i>n.s.</i>
3	67984-81-0	CN	H	23 ± 2	23 ± 2	+1.8	20.6	0.045
4	363-52-0	F	H	inconcl.	none	–	1.3	<i>n.s.</i>
5	488-17-5	Me	H	inconcl.	none	none	8	0.055
6	14235-77-9	CH ₂ OH	H	inconcl.	11 ± 4	none	13.8	0.035
7	2144-08-3	OH	CHO	inconcl.	inconcl.	−3.3	17.2	<i>n.s.</i>
8	2138-22-9	H	Cl	inconcl.	none	–	2.5	<i>n.s.</i>
9	17345-61-8	H	CN	4 ± 1	7 ± 2	+0.5	6.1	0.030
10	367-32-8	H	F	inconcl.	none	–	2.6	<i>n.s.</i>
11	452-86-8	H	Me	inconcl.	none	−11.0	6.9	<i>n.s.</i>
12	3316-09-4	H	NO ₂	37 ± 1	33 ± 2	autofluor.	56.1	0.065
13	98-29-3	H	<i>tert</i> -butyl	inconcl.	none	−15.4	7.9	0.130
14	331-39-5	H	<i>E</i> -CHCHCO ₂ H	inconcl.	none	–	7.3	<i>n.s.</i>
15	3843-74-1	H	<i>E</i> -CHCHCO ₂ Me	inconcl.	none	–	7.8	0.045
16	133550-30-8	H	<i>E</i> -CHC(CN)CONHBn	inconcl.*	inconcl.*	–	15.8 at 0.5 mM	<i>n.s.</i>

Protein-observed ¹⁹F (ProF) NMR spectroscopy was performed with 5FW-LecA. FP averages and std. dev. from 3 experiments, TSA averages from two experiments, * = tested at 25% DMSO, 2ME = 2-mercaptoethanol, CSP = chemical shift perturbation, n.s. = not significant, std. dev. < 0.03 ppm.

Table 2: Acylated catechols and LecA binding by fluorescence polarization (FP), thermal shift assay (TSA), and surface plasmon resonance (SPR).

CAS no.	R ₁	R ₂	R ₃						CSP of W42 [ppm]
				FP	FP with 2ME	TSA	SPR	ProF NMR	
				Inhibition [%] at 4 mM	Inhibition [%] at 4 mM	ΔT [K]	Efficiency [%] at 1 mM		
17	1197-09-7	Me	H	11 ± 3	10 ± 3	+ 0.5	14.4		0.035
18	99-50-3	OH	H	none	none	–	5.0		n.s.
19	99-40-1	CH ₂ Cl	H	inconcl.	none	–	22.8		0.095
20	3943-89-3	OEt	H	none	none	–	9.0		n.s.
21	5466-89-7	ⁱ Pr	H	10 ± 2	11 ± 2	+ 0.35	9.0		0.08
22	62-13-5	(CH ₂) ₂ NHMe*HCl	H	inconcl.	insoluble	–	9.6		0.065
23	10425-11-3	Ph	H	39 ± 5	39 ± 3	+ 0.36	50.9		0.110
24	13 4612-84-3	pCl-C ₆ H ₄ -	H	29 ± 3*	22 ± 2*	–	insoluble		0.185
25	61 445-50-9	2,4-(HO) ₂ -C ₆ H ₃ -	H	14 ± 5	13 ± 2	inconcl.	61.8		0.105
26	52 479-85-3	3,4,5-(HO) ₃ -C ₆ H ₂ -	OH	inconcl.*	inconcl.*	–	142.6		0.125
27	1143-72-2	Ph	OH	inconcl.	inconcl.	–	13.2 at 0.5 mM		0.045
28	31127-54-5	pHO-C ₆ H ₄ -	OH	inconcl.	inconcl.*	–	38.5		0.035
29	5995-86-3	OH	H	inconcl.	none	–	24.0		n.s.
30	3856-05-1	OCH ₂ ⁱ Pr	H	8 ± 2	8 ± 1	–	16.4		0.055

Protein-observed ¹⁹F (ProF) NMR spectroscopy was performed with 5FW-LecA. FP averages and std. dev. from 3 experiments, TSA averages from two experiments, * = tested at 25 % DMSO, 2ME = 2-mercaptoethanol, CSP = chemical shift perturbation, n.s. = not significant, std. dev. < 0.03 ppm.

catechol **1** and various derivatives carrying electron-donating or -withdrawing substituents in different positions were obtained (**2–30**) among which are fourteen acylated catechols (**17–30**, including the hit 4-isobutyryl catechol **21**).

All compounds were tested in a modified competitive binding assay with a Cy5 dye conjugated to a galactoside **S7** (synthesis of **S7** see Scheme S1, binding of **S7** to LecA see Figure S4). This change in fluorescent dye was necessary to circumvent spectral overlap of some catechols with the previously used fluorescein conjugate.^[18] After 1 h incubation, various catechols bearing electron-withdrawing substituents showed a weak inhibition of LecA. Surprisingly, the polarization measured for some of the tested compounds has changed dramatically after extended incubation (16 h). While some catechols showed a constant inhibition over time, in particular catechols bearing electron-donating substituents such as *tert*-butyl or methyl groups showed the strongest increase in inhibition.

To avoid the false-positive detection of an apparent inhibition due to an unspecific reactivity of potentially formed quinones, 2-mercaptoethanol was added to the competitive binding assay as an agent to trap those reactive species at equimolar concentration to the highest catechol concentration. Interestingly, the previously active electron-rich catechols were inactive under these conditions (Figure S5). In contrast, LecA inhibition of catechols carrying electron withdrawing substituents, such as nitrile **3**, nitro **12**, and ketone **21**, has not changed in presence of 2-mercaptoethanol and over time (Table 1, Table 2).

At a concentration of 4 mM, ten out of the 30 catechols tested showed a quantifiable and 2-mercaptoethanol-independent inhibition of LecA in this competitive binding assay

(Table 1, Table 2, Figure 2A). Benzoylcatechol **23** was the most active compound with an inhibition of 39%. Docking of **23** into the LecA binding site suggests that the benzoyl ring can interact with Gln53 (Figure S3A). In the competitive binding assay, this compound's activity is followed by nitro-catechol **12** and cyanocatechol **3** with inhibition of 33% and 23%, respectively. All three compounds were two- to fourfold more active than the initial in silico hit, isobutyryl catechol **21** (11% inhibition). Even though some of the catechols, such as 4-parachlorobenzoyl catechol (**24**), were only soluble at the test concentration of 4 mM in presence of 25% DMSO as a cosolvent, LecA bound successfully to the Cy5 galactoside **S7** and **24** showed 22% inhibition of LecA. These results support the notion that electron-deficient catechols can either allosterically inhibit the LecA-galactoside interaction or competitively bind to its carbohydrate-binding site.

Secondary Assays Verifying Binding of Catechols to LecA.

First, a subset of electron-rich and electron-deficient catechols was tested in a thermal shift assay to verify their interaction with LecA. In case of a ligand binding event, the thermal stability of a protein increases due to stabilizing interactions in the ligated state.^[24] Therefore, the melting point of LecA was measured in absence of ligand (*T_m* 89.4 °C), in presence of a galactoside (*T_m* 91.1 °C) or thirteen catechols (Figure S6).

A general trend could be observed where the electron-deficient catechols, that is, nitriles **3** and **9**, acetyl **17**, isobutyryl **21** and benzoyl **23**, led to an increase of the melting point of LecA. In particular, 4-benzoyl catechol **23** and 3-

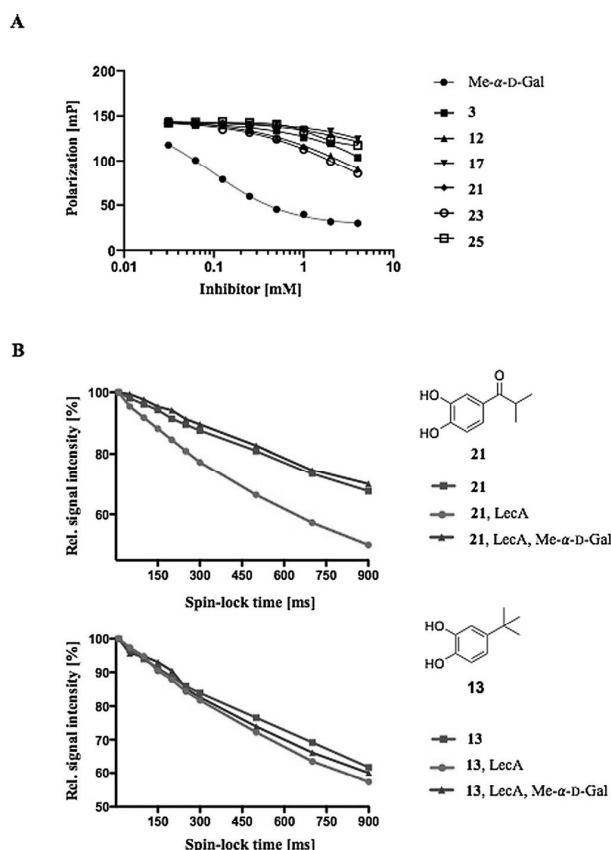


Figure 2. Analysis of catechol binding to LecA using: A) a competitive binding assay based on fluorescence polarization with catechols **3**, **12**, **17**, **21**, **23** and **25** and methyl α -D-galactoside as positive control; B) $T_{1\rho}$ relaxation NMR spectroscopy of isobutyryl catechol **21** and *tert*-butyl catechol **13**. Signal intensity decay of methyl protons plotted for **21** and **13** vs. spin-lock times.

cyano catechol **3** had the largest effect on the melting point of LecA with an increase of $+0.4^{\circ}\text{C}$ and $+1.8^{\circ}\text{C}$, respectively (Tables 1, 2, Figure S6). In contrast, the electron-rich catechols (methyl **11** and *tert*-butyl **13**) and the unsubstituted catechol **1** destabilized LecA and thus, lowered its melting temperature. A strong destabilization was observed in case of 4-*tert*-butyl catechol **13** (-15.4°C) and 4-methyl catechol **11** (-11.0°C).

As an orthogonal binding assay ligand-observed NMR spectroscopy experiments monitoring the $T_{1\rho}$ relaxation of the electron-deficient catechol, virtual screening hit isobutyryl **21**, and the electron-rich catechol 4-*tert*-butyl catechol **13** were measured. The $T_{1\rho}$ relaxation rates of a small molecule depend on its correlation time and hence can be used to detect binding to higher molecular weight receptors.^[25] Therefore, $T_{1\rho}$ relaxation of isobutyryl catechol **21** or *tert*-butyl catechol **13** were measured in presence and absence of LecA, and both compounds showed an increase in relaxation with LecA (Figure 2B, Figure S7). However, the relaxation rate of isobutyryl catechol **21** was strongly increased in presence of LecA (red) and could be fully reversed to the free

state in presence of an excess of Me- α -D-Gal (blue). The effect of *tert*-butyl catechol **13** with LecA was much weaker and addition of Me- α -D-Gal had no significant effect. Since the binding of isobutyryl catechol **21** to LecA can be fully competed with a galactoside, this data suggests that electron-deficient catechols bind LecA in the carbohydrate-binding site. Moreover, this result is in agreement with our competitive binding assay with a fluorescent tracer, where isobutyryl catechol **21** inhibited LecA and the inhibition of *tert*-butyl catechol **13** was unspecific and absent in presence of a scavenging nucleophile.

Both assays, $T_{1\rho}$ relaxation NMR spectroscopy and the thermal shift assay, have further validated electron-deficient catechols as ligands of LecA and NMR gave strong hints for a direct galactose-competition as mode of action. Next, we quantified the catechols' binding affinity using surface plasmon resonance (SPR) as an orthogonal method and localized their binding site by protein-observed NMR spectroscopy and X-ray crystallography.

Quantification of direct catechol binding to LecA was performed using SPR analysis. Binding analyses were performed by injecting various catechols at two concentrations (0.2 and 1 mM, except for **16** and **27** which were injected at 0.2 and 0.5 mM). We identified 14 catechol hits by SPR, six of which were in agreement with other biophysical assays (nitrile **3**, hydroxymethyl **6**, nitro **12**, acetyl **17**, benzoyl **23** and dihydroxybenzoyl **25**, Figure 3A, Figure S8). Subsequently, multicycle kinetic analyses were performed on these six prioritized hits to obtain the corresponding dissociation constants (K_d) (Figure 3B, Figure S9). These data showed that hydroxymethyl **6**, acetyl **17** and dihydroxybenzoyl **25** have very weak affinity towards LecA (extrapolated K_d of >100 M) (Figure S9), whereas compounds nitrile **3**, nitro **12** and benzoyl **23** catechols have promising K_d values in the low millimolar range (Figure 3). The lowest K_d value was obtained for nitro **12** ($K_d = 0.56 \pm 0.34$ mM) followed by nitrile **3** ($K_d = 1.11 \pm 0.07$ mM) and the most active compound in the competitive binding assay, benzoyl **23**, was somewhat less active ($K_d = 3.46 \pm 0.41$ mM). Finally, the initial virtual screening hit isobutyryl **21** had a K_d of approx. 145 M corresponding to the results from the competitive binding assay. For weaker binding hits, the efficiency was used for comparison of the affinity, among which **26** showed the highest efficiency (142.6%), which is suspiciously high for a small molecule. We hypothesized that this unexpected high binding response observed on SPR is due to a potential oxidation and unspecific interaction with the protein.

Protein-Observed NMR Reveals Catechols Targeting the Carbohydrate-Binding Site of LecA

Protein-observed ^{19}F (PrOF) NMR spectroscopy is a sensitive technique to spot weak binders in a low millimolar affinity range. Here, we used previously established PrOF NMR with 5-fluorotryptophan-labeled (5FW)-LecA^[26] to determine the binding region of catechols. First, we measured 5FW-LecA in presence of *para*-nitrophenyl β -D-galactoside (pNPGal) as well-known ligand for the carbohydrate-binding

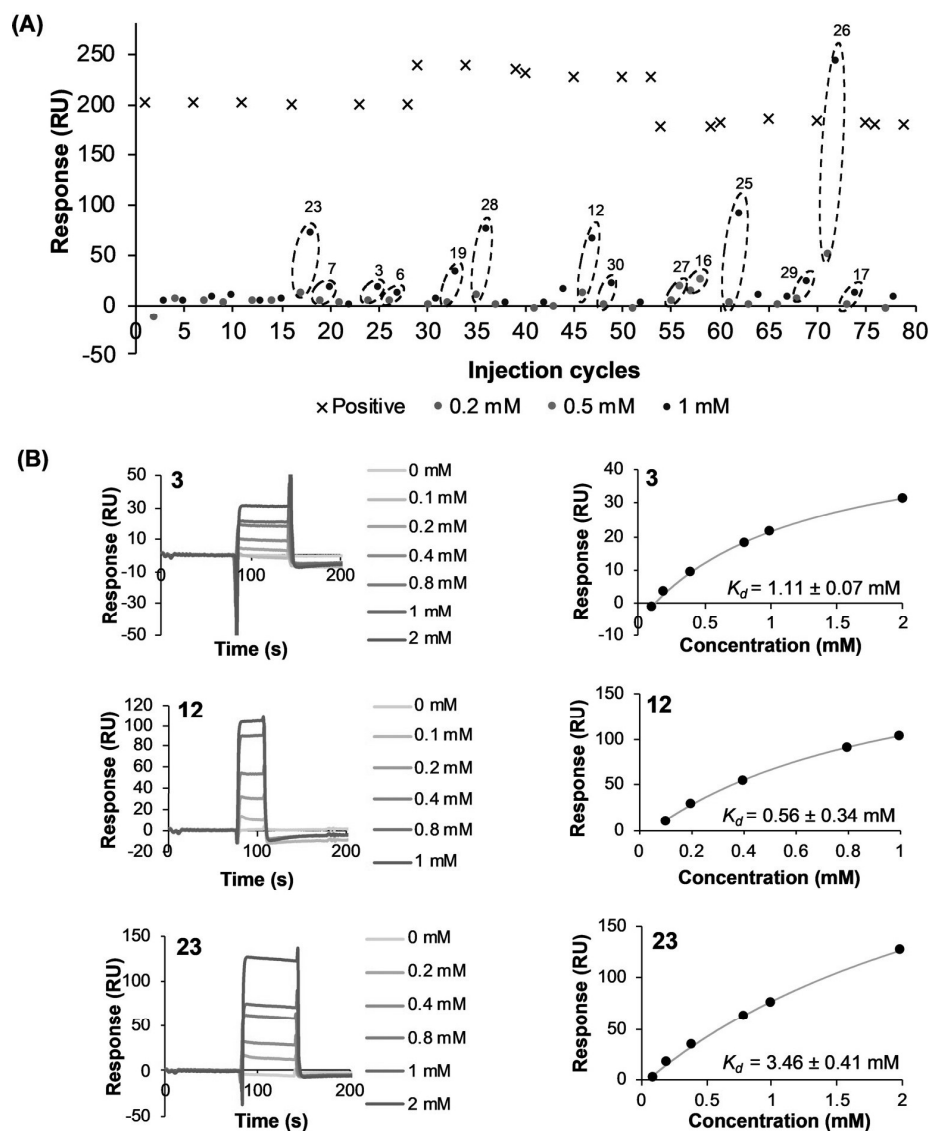


Figure 3. SPR analysis of catechol direct binding to LecA. A) The binding responses at steady state on the sensorgrams, when 0.2, 0.5 or 1 mM of catechols were injected (blue, orange and black dots, respectively), were plotted against injection cycles of compounds. Dose-response increases more than twofold are indicated by dotted circles. The catechol entries corresponding to the dose-response binding are indicated. Positive control (0.1 mM pNPGal represented by crosses) was injected to monitor the activity of immobilized LecA and enable data normalization. B) Multi-cycle kinetic analyses of the prioritized hits nitrile **3**, nitro **12**, benzoyl **23** by SPR. Left: Sensorgrams, Right: affinity analyses based on the data from the sensorgrams.

site of LecA. As expected, the resonance of the 5-fluorotryptophan located in the carbohydrate-binding site, W42, showed a clear chemical shift perturbation (CSP) towards a ligand-bound form in presence of pNPGal indicating that ProOF NMR can be used to spot the binding site of ligands (Figure 4).

Next, we validated 30 catechols in ProOF NMR (Table S2) and observed the perturbation of W42 in presence of 18 catechols that target the carbohydrate-binding site (Figure S10). Notably, the chemical shift perturbation of W42 in presence of the electron-deficient benzoyl catechol **23** was comparable to pNPGal (Figure S11) and nitrile **3** also induced

a chemical shift perturbation of the W42 resonance (Figure 4).

We further analyzed uniformly ^{15}N -labelled LecA in ^1H , ^{15}N TROSY NMR experiments in presence and absence of two galactosides or four catechols (Figure 4). The fingerprint of the resulting chemical shift perturbations was highly similar for D-Gal and pNPGal compared to the three tested catechols **3**, **23** and **25** (Figure 4).

Taken together, compounds nitrile **3** and benzoyl **23** show a comparable binding of LecA in protein-observed NMR spectroscopy as the positive control pNPGal allowing us to

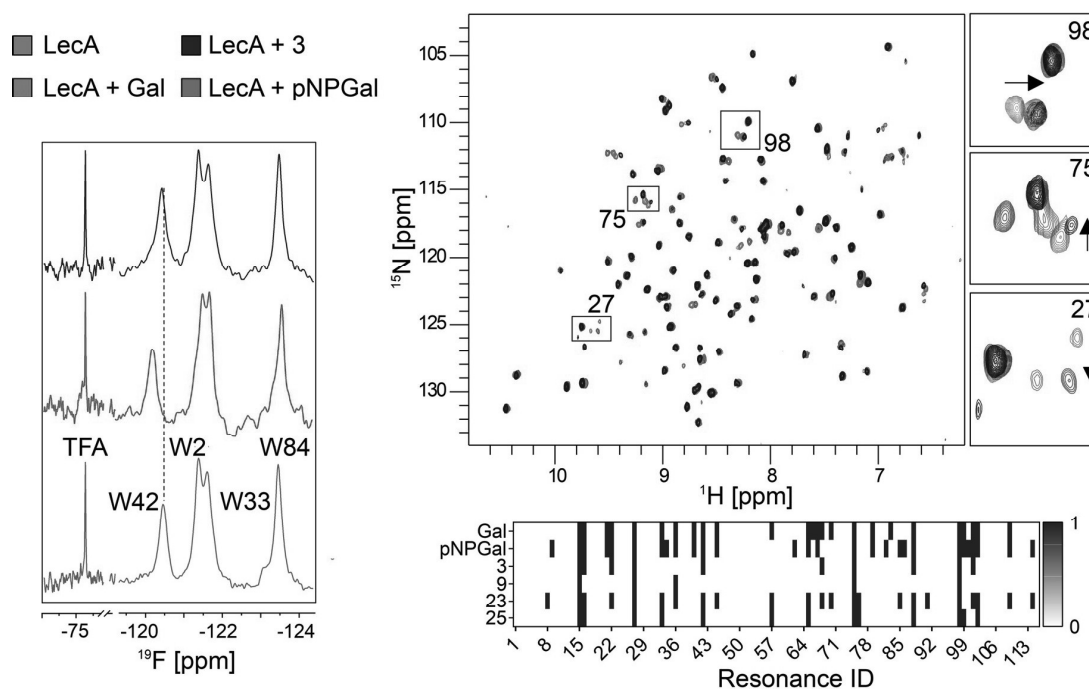


Figure 4. (left) Protein-observed ^{19}F (ProF) NMR spectroscopy of nitrile **3** and pNPGal with 5FW-LecA; (right) ^1H , ^{15}N TROSY-HSQC NMR spectra (upper panel) of uniformly ^{15}N -labelled LecA in absence and presence of ligands D-Gal, pNPGal or nitrile **3** and a plot with a 1:0 barcode (bottom panel) shows CSPs and changes in peak intensity of arbitrarily numbered resonances in LecA observed for ligands D-Gal, pNPGal, nitriles **3**, **9**, benzoyl **23** and dihydroxybenzoyl **25**. This data indicated catechols **3**, **9**, **23**, **25** bind LecA in a region similar to D-Gal and pNPGal suggesting binding to the carbohydrate-binding site of LecA.

hypothesize that both compounds target the carbohydrate-binding site of LecA.

X-Ray Crystallography Reveals Glycomimetic Binding Mode of One Catechol Hit

To further investigate the interaction between the catechol hits and LecA at atomic resolution, we co-crystallized LecA with the six hits that were identified among all biophysical assays, nitrile **3**, hydroxymethyl **6**, nitro **12**, acetyl **17**, benzoyl **23** and dihydroxybenzoyl **25**. We adapted a dry co-crystallization approach previously reported by Gelin et al., which involves deposition of the compound onto the crystallization slide and gentle evaporation of DMSO prior to dispensing the protein onto the dried compound.^[27] This approach enables the deposition of high amounts of the dried compounds onto the slide, whilst avoiding complicated crystal manipulation and use of high DMSO concentrations. Protein crystals were obtained for co-crystallization of LecA with nitrile **3**, benzoyl **23** and dihydroxybenzoyl **25**. X-ray diffraction data were collected from crystals of LecA in complex with nitrile **3** (LecA-**3**) and in complex with dihydroxybenzoyl **25** (LecA-**25**) whereas the crystals of LecA with benzoyl **23** diffracted poorly (resolution $> 4 \text{ \AA}$) and were not further investigated. Both complex structures were successfully solved at 1.84 \AA in space group I 2 (for LecA-**3**) and $P2_122_1$ (for LecA-**25**) (Table S3).

The overall structure of LecA-**25** is very similar to LecA-galactose (global RMSD = 0.29 \AA). Intriguingly, the electron density corresponding to compound **25** was found remotely from the galactose binding site (Figure S12A), which was unexpected considering its ability to outcompete the galactose-derived fluorescent probe. We hypothesized that the interaction of **25** with the remote site on LecA maybe a crystallization artefact since the molecule locates at the interface of the protein crystal contacts and forms only minimal interactions with the protein (Figure S12B).

The LecA-**3** structure is also very similar to the LecA-galactose complex (PDB: 1OKO) with a global RMSD of 0.8 \AA (superposition of equivalent Ca atoms). Here, however, catechol **3** occupies the previously reported galactose binding site in monomers A and C of the LecA tetramer (Figure 5A and B). The two vicinal hydroxy groups of **3** coordinate the Ca^{2+} ion and form hydrogen-bonded interactions with the side chains of Asp100 and Asn107 and the backbone oxygen of Tyr36 (Figure 5D), the interaction of which imitates the roles of OH3 and OH4 of galactose in the LecA-galactose interaction (Figure 5C). A conserved water molecule (WAT1) forms hydrogen bond bridges between the nitrogen atom of the nitrile in catechol **3** and the oxygen atom of the carbonyl group of Pro51 (Figure 5D). The nitrogen atom thus assumes the role of OH6 in the LecA-galactose interaction, where hydrogen bond bridges through WAT1 were identified between the OH6 group and Pro51 (Figure 5C). Shifting the nitrile group to the neighboring carbon atom, that is, to

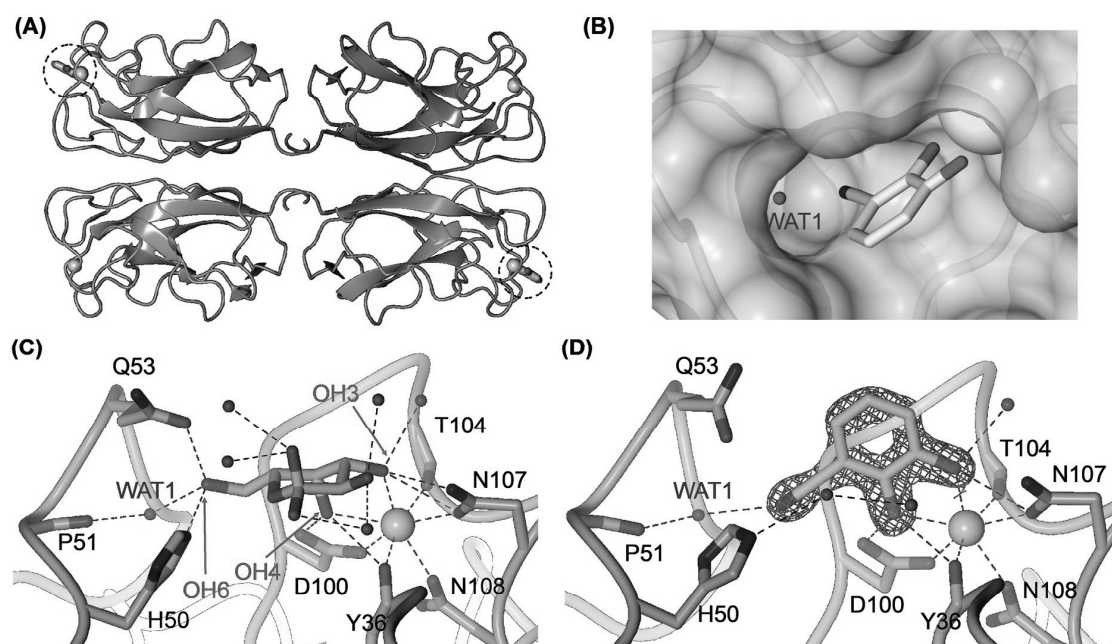


Figure 5. A) Overall structure of LecA-3 complex (PDB code: 6YO3). 2 molecules of compound **3** are shown as sticks and the Ca²⁺ ions as green spheres. Dotted circles indicate the binding sites of **3**. Each LecA monomer is depicted as cartoon in different colors. B) Surface display of compound **3** at the binding site. C) Interaction of galactose with LecA (PDB code: 1OKO). D) Interaction of **3** with LecA, electron density is displayed at 1σ.

position R2 in compound **9** instead of position R1 in compound **3**, resulted in approximately a threefold-reduction of the SPR efficiency and inhibition in the FP assay (Table 1), which further substantiates the essence of the positioning and the role of the nitrile group in the LecA-3 interaction.

To investigate if catechols interact with other Ca²⁺-binding lectins and could be used as a general motif for glycomimetics, we performed ¹H-¹⁵N HSQC NMR with ¹⁵N-labeled Langerin CRD (Figure 6A). We observed that **3** promoted chemical shift perturbations (CSPs) of amino acids located in the carbohydrate-binding site of Langerin (Figure 6B), but also affected amino acids that are located at a certain distance. The magnitude of CSPs observed for **3** (tested at 1 mM due to solubility restrictions) was generally weaker than mannose which was tested at 30 mM. Given the allosteric nature of some calcium binding lectins such as Langerin,^[16a,28] changes in the resonances located far away from the carbohydrate-binding site can be correlated to changes in the binding site. Potentially, **3** can be located in the binding site, or in a remote site that triggers the allosteric network.

To demonstrate that **3** targets the carbohydrate-binding site of Langerin, we performed ¹H STD competition NMR experiments (Figure 6C). For this, we recorded STD spectra of **3** without or with Langerin ECD and upon addition of mannose as the competitor. As result, we observed a 65% reduction in peak intensities of resonances from compound **3** in STD NMR spectra in presence of mannose indicating its partial competition from the carbohydrate-binding. This is in agreement with ¹H-¹⁵N HSQC data, where **3** has been observed to alter resonances from both the carbohydrate-

and secondary binding sites in Langerin. Thus, these data reveal catechol **3** as a general glycomimetic motif for a bacterial and a mammalian calcium-binding lectin.

Conclusion

In this work, a non-carbohydrate glycomimetic has been identified for a relevant anti-infective drug target. We have established catechols as mimicry of the carbohydrate ligand, binding to the protein via the crucial calcium-ion present in the carbohydrate-binding site of LecA, a key lectin involved in infection mechanisms of the ESKAPE pathogen *P. aeruginosa*. Therefore, the catechols have good potential beyond LecA and may also serve as glycomimetics for the large family of animal C-type lectins and other bacterial lectins relying on calcium-mediated glycan recognition. Target specificity required for a selective drug can be achieved by further optimizing a central catechol motif addressing the carbohydrate-binding site.

Despite the widespread belief, catechols are not necessarily PAINS when care is taken in analysing their binding properties in orthogonal assays under appropriate experimental conditions. We showed that electron-deficient catechols are stable under the conditions tested and their interaction with LecA has been confirmed in several orthogonal assays. In contrast, electron-rich catechols proved to be unspecific binders and their observed effect on LecA could be reversed in presence of a competing nucleophile, 2-mercaptoethanol.

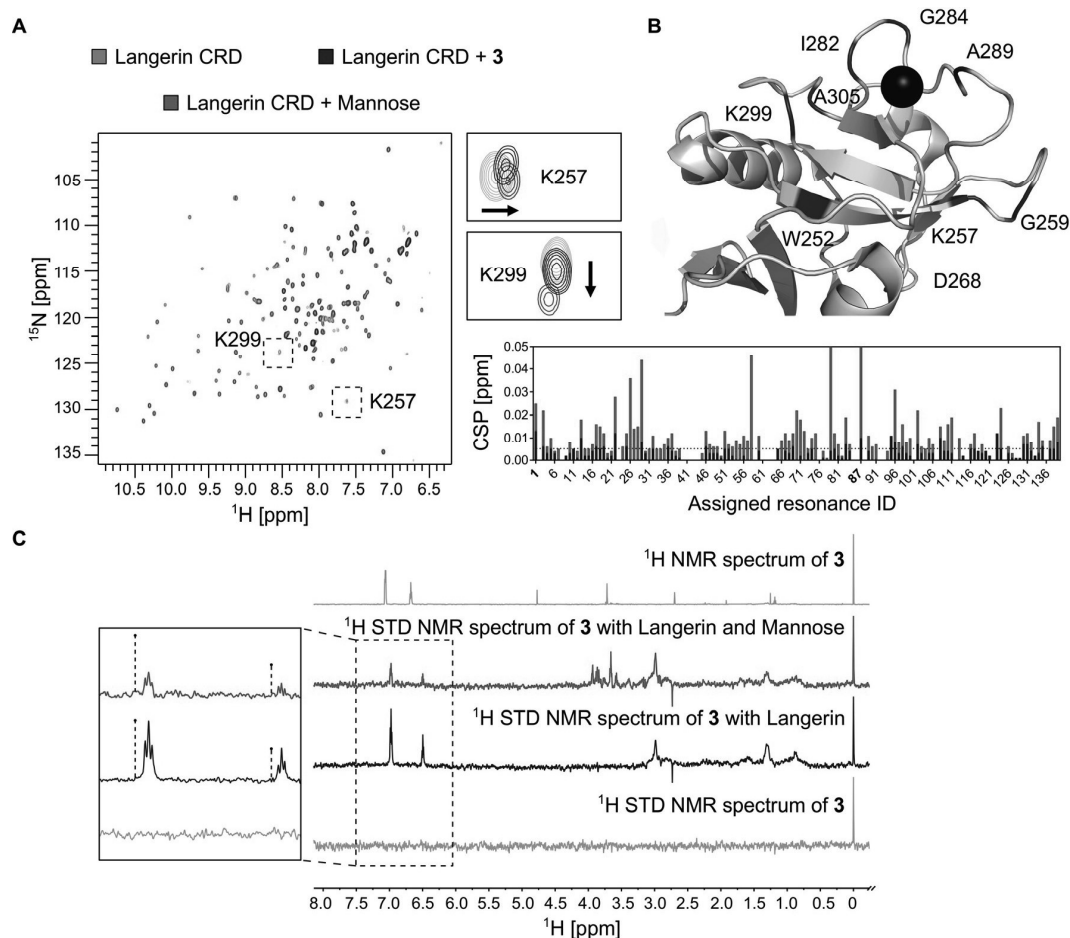


Figure 6. A) ^1H - ^{15}N HSQC NMR spectra of uniformly ^{15}N -labelled human Langerin CRD in absence of ligands (grey) and in presence of mannose (red) or catechol **3** (blue). The two resonances K257 and K299 are magnified and shown as an example for the interaction of **3** (blue) with Langerin, which is in the similar manner to mannose (red). B) The perturbed resonances in the CSP plot (CSP > 0.005 p.p.m., for correlation of resonance ID and amino acid see Table S4) were used for mapping the binding site of **3** on a structure of Langerin CRD (PDB: 3P5D). K257 and K299 are highlighted in the CSP plot as 1 and 87, respectively. C) The competition ^1H STD NMR experiment with mannose was performed with **3** in presence and absence of Langerin ECD. Binding of **3** to the carbohydrate-binding site of Langerin has been confirmed as indicated by decreasing peak intensities in the STD NMR spectrum upon addition of mannose.

From an initial virtual screening approach for the bacterial lectin LecA, we have identified electron-deficient catechols as non-carbohydrate inhibitors for the lectin LecA. We validated our hits in various orthogonal biophysical assays. The binding of the initial catechol hit isobutyryl **21** derived from virtual screening was confirmed in a competitive binding assay based on fluorescence polarization. This assay and a thermal shift assay were used to demonstrate that only catechols with electron-withdrawing groups are robust inhibitors of LecA, while catechols with electron-donating substituents showed protein destabilization and unspecific inhibition which was absent in presence of a nucleophilic scavenger. This specificity of catechols was further confirmed in ligand-observed $T_{1\rho}$ -relaxation NMR experiments for electron-deficient isobutyryl catechol **21** and electron-rich 4-*tert*-butylcatechol **13**. The binding affinity of catechols was determined by SPR and led to K_d -values in the millimolar range, for example, nitrile **3** with a K_d of 1.11 ± 0.07 mM and

nitro **12** with a K_d of 0.56 ± 0.34 mM. These affinities of the fragment hits (MW 135–155 g mol^{-1}) result in ligand efficiencies of 0.40 for both compounds, providing a good basis for future fragment growing. ^{19}F -protein observed NMR spectroscopy revealed that the catechols with electron-withdrawing substituents were binding to the carbohydrate-binding site of LecA. Finally, X-ray analysis of LecA in complex with nitrile **3** unambiguously confirmed catechols as galactose-mimicking ligands located in the carbohydrate-binding site. The structure revealed that catechols were coordinating to the calcium ion via their vicinal hydroxy groups, and the nitrile group underwent a hydrogen-bonding interaction with a crystal bound water molecule, highly homologous to the primary hydroxy group of galactose in complex with LecA.

This work sets the basis for future optimization of the catechol scaffold to increase binding potency and drug-like properties of a possible antimicrobial agent. Therefore, this work sets a new starting point for the development of future

LecA inhibitors against biofilm-associated *P. aeruginosa* infections by developing new non-carbohydrate LecA glycomimetics and provides a structure-based rationale for the development of lectin-inhibitors in general.

Moreover, catechol **3** proved to target the carbohydrate-binding site of both Ca²⁺-binding bacterial (LecA) and mammalian (Langerin) lectins. Given that Ca²⁺-binding lectins have been considered “undruggable” for a long time,^[29] we believe catechols challenge this paradigm as a novel class of non-carbohydrate glycomimetics. Further growing of this low affinity fragment hit, catechol **3**, into lead structures will provide high target affinity. Structure-based design of additional functional groups will provide the specificity towards the lectin of interest for the desired indication.

Acknowledgements

The work was supported by the ANR/DFG French-German project (ANR-AAPG-2017). The authors acknowledge financial support from Agence Nationale de la Recherche (grant no. ANR-17-CE11-0048) and from Deutsche Forschungsgemeinschaft (grant no. Ti756/5-1 and RA1944/7-1). J.T. was supported by Labex Arcane/ CBH-EUR-GS (ANR-17-EURE-0003) and we also thank Glyco@Alps (ANR-15-IDEX02) for support. We acknowledge synchrotron SOLEIL (Saint Aubin, France) for access and technical support at beamline PROXIMA1. We thank Emilie Gillon (CERMAV) for providing recombinant LecA for SPR and X-ray crystallographic experiments and Dr. Alberto Plaza (HIPS) for implementing the T1p pulse sequence provided by Dr. Marcel J.J. Blommers (Saverna Pharmaceuticals, Switzerland). Open access funding enabled and organized by Projekt DEAL.

Conflict of interest

The authors declare no conflict of interest.

Keywords: carbohydrates · catechol · glycomimetic · lectin · PAINS

- [1] a) A. Imberty, M. Wimmerová, E. P. Mitchell, N. Gilboa-Garber, *Microbes Infect.* **2004**, *6*, 221–228; b) H. Lis, N. Sharon, *Chem. Rev.* **1998**, *98*, 637–674.
- [2] D. Solís, N. V. Bovin, A. P. Davis, J. Jiménez-Barbero, A. Romero, R. Roy, K. Smetana, Jr., H. J. Gubius, *Biochim. Biophys. Acta Gen. Subj.* **2015**, *1850*, 186–235.
- [3] S. de Bentzmann, P. Plesiat, *Environ. Microbiol.* **2011**, *13*, 1655–1665.
- [4] a) S. P. Diggle, R. E. Stacey, C. Dodd, M. Camara, P. Williams, K. Winzer, *Environ. Microbiol.* **2006**, *8*, 1095–1104; b) D. Tielker, S. Hacker, R. Loris, M. Strathmann, J. Wingender, S. Wilhelm, F. Rosenau, K.-E. Jaeger, *Microbiology* **2005**, *151*, 1313–1323.
- [5] a) S. Cecioni, A. Imberty, S. Vidal, *Chem. Rev.* **2015**, *115*, 525–561; b) J. Meiers, E. Zahorska, T. Röhrig, D. Hauck, S. Wagner, A. Titz, *J. Med. Chem.* **2020**, *63*, 11707–11724; c) R. Sommer, K. Rox, S. Wagner, D. Hauck, S. Henrikus, S. Newsad, T. Arnold, T. Ryckmans, M. Bronstrup, A. Imberty, A. Varrot, R. W. Hartmann, A. Titz, *J. Med. Chem.* **2019**, *62*, 9201–9216; d) R. Sommer, S. Wagner, K. Rox, A. Varrot, D. Hauck, E.-C. Wamhoff, J. Schreiber, T. Ryckmans, T. Brunner, C. Rademacher, R. W. Hartmann, M. Brönstrup, A. Imberty, A. Titz, *J. Am. Chem. Soc.* **2018**, *140*, 2537–2545; e) S. Wagner, R. Sommer, S. Hinsberger, C. Lu, R. W. Hartmann, M. Empting, A. Titz, *J. Med. Chem.* **2016**, *59*, 5929–5969.
- [6] T. Eierhoff, B. Bastian, R. Thuenauer, J. Madl, A. Audfray, S. Aigal, S. Juillot, G. E. Rydell, S. Müller, S. de Bentzmann, A. Imberty, C. Fleck, W. Römer, *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 12895–12900.
- [7] G. Cioci, E. P. Mitchell, C. Gautier, M. Wimmerová, D. Sudakevitz, S. Pérez, N. Gilboa-Garber, A. Imberty, *FEBS Lett.* **2003**, *555*, 297–301.
- [8] B. Blanchard, A. Nurisso, E. Hollville, C. Tétaud, J. Wiels, M. Pokorná, M. Wimmerová, A. Varrot, A. Imberty, *J. Mol. Biol.* **2008**, *383*, 837–853.
- [9] a) M. Bergmann, G. Michaud, R. Visini, X. Jin, E. Gillon, A. Stocker, A. Imberty, T. Darbre, J. L. Reymond, *Org. Biomol. Chem.* **2016**, *14*, 138–148; b) A. Novoa, T. Eierhoff, J. Topin, A. Varrot, S. Barluenga, A. Imberty, W. Römer, N. Winsinger, *Angew. Chem. Int. Ed.* **2014**, *53*, 8885–8889; *Angew. Chem.* **2014**, *126*, 9031–9035; c) E. Zahorska, S. Kuhaudomlarp, S. Minervini, S. Yousaf, M. Lepsik, T. Kinsinger, A. K. H. Hirsch, A. Imberty, A. Titz, *Chem. Commun.* **2020**, *56*, 8822–8825.
- [10] a) R. U. Kadam, M. Bergmann, M. Hurley, D. Garg, M. Cacciarini, M. A. Swiderska, C. Nativi, M. Sattler, A. R. Smyth, P. Williams, M. Camara, A. Stocker, T. Darbre, J. L. Reymond, *Angew. Chem. Int. Ed.* **2011**, *50*, 10631–10635; *Angew. Chem.* **2011**, *123*, 10819–10823; b) R. U. Kadam, D. Garg, J. Schwartz, R. Visini, M. Sattler, A. Stocker, T. Darbre, J. L. Reymond, *ACS Chem. Biol.* **2013**, *8*, 1925–1930; c) J. Rodrigue, G. Ganne, S. Blanchard, C. Saucier, D. Giguère, T. C. Shiao, A. Varrot, A. Imberty, R. Roy, *Org. Biomol. Chem.* **2013**, *11*, 6906–6918.
- [11] S. Wagner, D. Hauck, M. Hofmann, I. Joachim, R. Sommer, R. Müller, A. Imberty, A. Varrot, A. Titz, *Angew. Chem. Int. Ed.* **2017**, *56*, 16559–16564; *Angew. Chem.* **2017**, *129*, 16786–16791.
- [12] J. Meiers, E. Siebs, E. Zahorska, A. Titz, *Curr. Opin. Chem. Biol.* **2019**, *53*, 51–67.
- [13] M. Goel, D. Jain, K. J. Kaur, R. Kenoth, B. G. Maiya, M. J. Swamy, D. M. Salunke, *J. Biol. Chem.* **2001**, *276*, 39277–39281.
- [14] K. T. Welch, T. A. Turner, C. E. Preast, *Bioorg. Med. Chem. Lett.* **2008**, *18*, 6573–6575.
- [15] a) N. Kaila, B. E. Thomas IV, *Med. Res. Rev.* **2002**, *22*, 566–601; b) R. Kranich, A. S. Busemann, D. Bock, S. Schroeter-Maas, D. Beyer, B. Heinemann, M. Meyer, K. Schierhorn, R. Zahlten, G. Wolff, E. M. Aydt, *J. Med. Chem.* **2007**, *50*, 1101–1115.
- [16] a) J. Aretz, H. Baukman, 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; b) M. J. Borrok, L. L. Kiessling, *J. Am. Chem. Soc.* **2007**, *129*, 12780–12785; c) K. C. Garber, K. Wangkanont, E. E. Carlson, L. L. Kiessling, *Chem. Commun.* **2010**, *46*, 6747–6749.
- [17] O. Slater, M. Kontoyianni, *Expert Opin. Drug Discovery* **2019**, *14*, 619–637.
- [18] I. Joachim, S. Rikker, D. Hauck, D. Ponader, S. Boden, R. Sommer, L. Hartmann, A. Titz, *Org. Biomol. Chem.* **2016**, *14*, 7933–7948.
- [19] T. A. Halgren, R. B. Murphy, R. A. Friesner, H. S. Beard, L. L. Frye, W. T. Pollard, J. L. Banks, *J. Med. Chem.* **2004**, *47*, 1750–1759.
- [20] Z. Yu, M. P. Jacobson, R. A. Friesner, *J. Comput. Chem.* **2006**, *27*, 72–89.
- [21] J. B. Baell, G. A. Holloway, *J. Med. Chem.* **2010**, *53*, 2719–2740.

- [22] L. Sleno, R. F. Staack, E. Varesio, G. Hopfgartner, *Rapid Commun. Mass Spectrom.* **2007**, *21*, 2301–2311.
- [23] S. J. Capuzzi, E. N. Muratov, A. Tropsha, *J. Chem. Inf. Model.* **2017**, *57*, 417–427.
- [24] M. C. Lo, A. Aulabaugh, G. Jin, R. Cowling, J. Bard, M. Malamas, G. Ellestad, *Anal. Biochem.* **2004**, *332*, 153–159.
- [25] M. Pellicchia, I. Bertini, D. Cowburn, C. Dalvit, E. Giralt, W. Jahnke, T. L. James, S. W. Homans, H. Kessler, C. Luchinat, B. Meyer, H. Oschkinat, J. Peng, H. Schwalbe, G. Siegal, *Nat. Rev. Drug Discovery* **2008**, *7*, 738–745.
- [26] E. Shanina, E. Siebs, H. Zhang, D. V. Silva, I. Joachim, A. Titz, C. Rademacher, *Glycobiology* **2020**, *31*, 159–165.
- [27] M. Gelin, V. Delfosse, F. Allemand, F. Hoh, Y. Sallaz-Damaz, M. Pirocchi, W. Bourguet, J.-L. Ferrer, G. Labesse, J.-F. Guichou, *Acta Crystallogr. Sect. D* **2015**, *71*, 1777–1787.
- [28] J. Hanske, S. Aleksic, M. Ballaschk, M. Jurk, E. Shanina, M. Beerbaum, P. Schmieder, B. G. Keller, C. Rademacher, *J. Am. Chem. Soc.* **2016**, *138*, 12176–12186.
- [29] a) A. L. Hopkins, C. R. Groom, *Nat. Rev. Drug Discovery* **2002**, *1*, 727–730; b) A. P. Russ, S. Lampel, *Drug Discovery Today* **2005**, *10*, 1607–1610.

Manuscript received: September 30, 2020
Accepted manuscript online: December 13, 2020
Version of record online: March 3, 2021

Tolcapone-like Glycomimetics as Potent LecA Inhibitors

Eike Siebs^{1,2,3}, Sakonwan Kuhaudomlarp^{4,5}, Elena Shanina^{6,7}, Annabelle Varrot⁴, Priscila da Silva Figueiredo Celestino Gomes⁸, Christian Lerner⁹, Uwe Grether⁹, Christoph Rademacher^{6,7,10,11}, Didier Rognan⁸, Anne Imberty⁴, Alexander Titz^{1,2,3}

¹Chemical Biology of Carbohydrates (CBCH) Helmholtz-Institute for Pharmaceutical Research Saarland (HIPS) Helmholtz Centre for Infection Research, 66123 Saarbrücken (Germany);

²Department of Chemistry, Saarland University 66123 Saarbrücken (Germany);

³Deutsches Zentrum für Infektionsforschung (DZIF) Standort Hannover-Braunschweig (Germany);

⁴Université Grenoble Alpes, CNRS, CERMAV 38000 Grenoble (France);

⁵Department of Biochemistry and Centre for Excellence in Protein and Enzyme Technology, Faculty of Science Mahidol University, Bangkok (Thailand);

⁶Department of Biomolecular Systems Max Planck Institute of Colloids and Interfaces, 14424 Potsdam (Germany);

⁷Institute of Chemistry and Biochemistry Department of Biology, Chemistry and Pharmacy Freie Universität Berlin, 14195 Berlin (Germany);

⁸Laboratoire d'Innovation Thérapeutique UMR 7200 CNRS-Université de Strasbourg, Strasbourg, 67400 Illkirch (France);

⁹Roche Innovation Center Basel, F. Hoffman-La Roche Ltd., Basel 4070, Switzerland;

¹⁰Department of Pharmaceutical Sciences University of Vienna, 1090 Vienna (Austria);

¹¹Department of Microbiology, Immunology and Genetics University of Vienna, Max F. Perutz Labs, 1030 Vienna (Austria).

Corresponding e-mail: alexander.titz@helmholtz-hzi.de

Keywords: glycomimetics, tolcapone, LecA

Abstract

The opportunistic ESKAPE pathogen *Pseudomonas aeruginosa* is associated with biofilm formation in chronic infections. The biofilm matrix protects the bacteria from external threats such as the host's immune response and antibiotics. Crucial for the biofilm formation are the two lectins LecA and LecB. Both carbohydrate-binding proteins are additionally involved in different infection processes and hence promiscuous pathoblocker targets. Carbohydrate-based LecA and LecB inhibitors have been proven to reduce the biofilm formation and thus restore the antimicrobial susceptibility. Here, we report about the optimization of non-carbohydrate based LecA ligands as potential glycomimetics. We screened catechol-like compounds derived from a structure-based design by fluorescence polarization (FP) that resulted in tolcapone, the first micromolar non-carbohydrate LecA ligand equally potent as natural LecA inhibitors. We validated its affinity by surface plasmon resonance and revealed its interactions with LecA at the atomic level by x-ray diffraction analysis. Afterwards, we screened a catechol library of 3564 compounds by FP which led to 48 tolcapone-like hits with affinities in the low micromolar range.

Introduction

Multi-drug resistant pathogens are a general threat to our society and predicted to cause ten million deaths worldwide every year by 2050 by the World Health Organization (WHO).^[5] *Pseudomonas aeruginosa* is listed as one of the priority one pathogens for which new antibiotics are urgently needed.^[69] This opportunistic, Gram-negative bacterium is a threat especially for immunocompromised patients suffering from cystic fibrosis or those under ventilation and causes e.g: pneumonia, urinary tract infections, septicemia, meningitis or malignant otitis externa.^[70] Furthermore, several of the *P. aeruginosa* strains became multi-drug resistant^[71], whereupon, novel strategies for the treatment of *P. aeruginosa* are under investigation.^[39]

One strategy is the development of pathoblockers that instead of killing, disarm the bacteria by blocking their virulence factors.^[72] *P. aeruginosa* forms a biofilm at its colonization site at the host's glycocalyx.^[73] This biofilm protects the bacteria by orders of magnitudes compared to planktonic cells from environmental threats, the host's immunity, but also from antimicrobial treatments such as antibiotics.^[8,74] Host recognition, tissue adhesion and biofilm formation are mediated by two lectins, LecA

and LecB.^[49,50] Both virulence factors are soluble homotetramers and contain calcium ions in each of their four carbohydrate binding sites (CBS) (LecA: 1/CBS, **Figure 2A**; LecB: 2/CBS).^[47] In addition, the galactose binding LecA is involved in the internalization of *P. aeruginosa* into the host cell by binding to the glycosphingolipid Gb3.^[52] During the engulfment process, LecA interacting with Gb3 triggers a phosphorylation of Tyr221 of the adaptor protein CrkII which most probably initiates the internalization.^[75] The fucose/mannose binding LecB is mainly involved in adhesion processes to its environment such as epithelia cells and stabilizes the biofilm where it binds to the exopolysaccharide Psl from the matrix.^[54,76]

Both lectins are considered as targets for pathoblockers. In cystic fibrosis patients, it was demonstrated that inhalation of a fucose/galactose solution reduces the amount of *P. aeruginosa* in sputum.^[40] In addition, an increased reduction of the bacterial load with antibiotics in presence of carbohydrates was shown in *P. aeruginosa* infected mouse trachea sections.^[77]

The binding affinity of LecB to fucosides is favored over mannosides (K_d of Me- α -L-Fuc = 0.43 μ M, K_d of Me- α -D-Man = 71 μ M).^[53] However, different inhibitors were synthesized starting from both carbohydrates or even fucose/mannose-hybrids.^[78–81] So far, the best LecB inhibitor class are the β -fucosyl amides with affinities in the nano-molar range (IC_{50} = 85 \pm 16 nM (**Page**: 74, unpublished) and are even more potent than the best identified natural LecB ligand Lewis^a (K_d = 213 nM).^[61] Additional to monovalent LecB inhibitors, divalent and oligovalent compounds have been developed reaching affinities in the two digit nano-molar range.^[81]

In case of LecA, the affinity of monovalent inhibitors has been slightly increased compared to naturally derived galactosides (α MeGal = 50 μ M^[82]) to the single-digit micromolar range ($K_d \geq 3 \mu$ M^[83]).^[81,84] A more promising approach is the use of divalent ligands addressing, two in the same direction facing, CBSs which led to affinities in the low nano-molar range ($K_d \geq 11$ nM).^[85–87] Also, oligovalent inhibitors showed affinities in the nano-molar range but their affinities were weaker than those of the best divalent ligands.^[81] Moreover, multivacancy could lead to aggregation, cross-linking or interfering with the host's pathogen recognition receptors such as human lectins.^[39]

For the exploration of new binding interactions and identification of new inhibitor classes, the development of monovalent LecA ligands is still of interest due to easier accessibilities (synthetically and commercially). Afterwards, the monovalent ligands

can be linked to form divalent compounds leading to affinity boosts of approximate three orders of magnitude.^[88] Currently, different approaches are under investigation such as the development of an epoxide that binds LecA covalently^[89], the discovery of a new binding pocket, the central pocket, between two adjacent CBSs^[90] or the development of non-carbohydrate glycomimetics with higher derivatization potential^[91,92].

In a recent study, we showed that catachols bind to different calcium-ion dependent lectins such as LecA and langerin.^[91] Catechols are considered to be pan-assay interference compounds (PAINS), however, we demonstrated in a careful validation cascade and later by x-ray diffraction analysis that electron-poor catechols are LecA ligands. These compounds mimic OH3 and OH4 of galactose with their diol function when coordinating to the calcium ion. 3-Cyanocatechol showed an additional interaction with its nitrile function to Pro51 via a water bridge by mimicking OH6 resulting in an affinity in the low millimolar range ($K_d = 1.1 \pm 0.1$ mM).^[91]

In general, different catechols are of medicinal interest e.g.: enterobactin, cefiderocol or tolcapone (**Figure 2B**). Enterobactin is a siderophore with the highest ferric ion (Fe^{3+}) affinity known.^[93] Inherently, it is used by Gram-negative bacteria such as *Escherichia coli* or *Salmonella typhimurium* as an iron transporter^[94], and is considered to be used as an anti-cancer drug by iron deprivation that could attenuate cancer progression.^[95] Cefiderocol (Fetroja) is a beta lactam antibiotic with a cephalosporin nucleus and an attached siderophore that chelates iron for an active uptake by bacteria cells.^[96] It was introduced to the market in 2019 and is used against complicated urinary tract infections and pneumonia.^[97,98] Tolcapone (Tasmar) is a medication for Parkinson's disease and was first marketed 1997.^[99] Only one year later, it was taken off the European market due to hepatotoxicity but readmitted 2004 requiring monitoring.^[99,100] Tolcapone is a catechol-O-methyltransferase (COMT) inhibitor that prevents the biodegradation of levodopa, a prodrug of dopamine, to 3-O-methyldopa. Levodopa is able to cross the blood brain barrier (BBB) and is converted by the aromatic L-amino acid decarboxylase to dopamine in the central nervous system (CNS). In addition, tolcapone is the only inhibitor that crosses the BBB and blocks COMT in the CNS and not only in the periphery.^[101]

In this work, we demonstrate the optimization of catechols to become potent LecA inhibitors by validating compounds in different biophysical assays, x-ray

crystallography and a follow up screening of 3564 catechol-like compounds *in vitro* by fluorescence polarization.

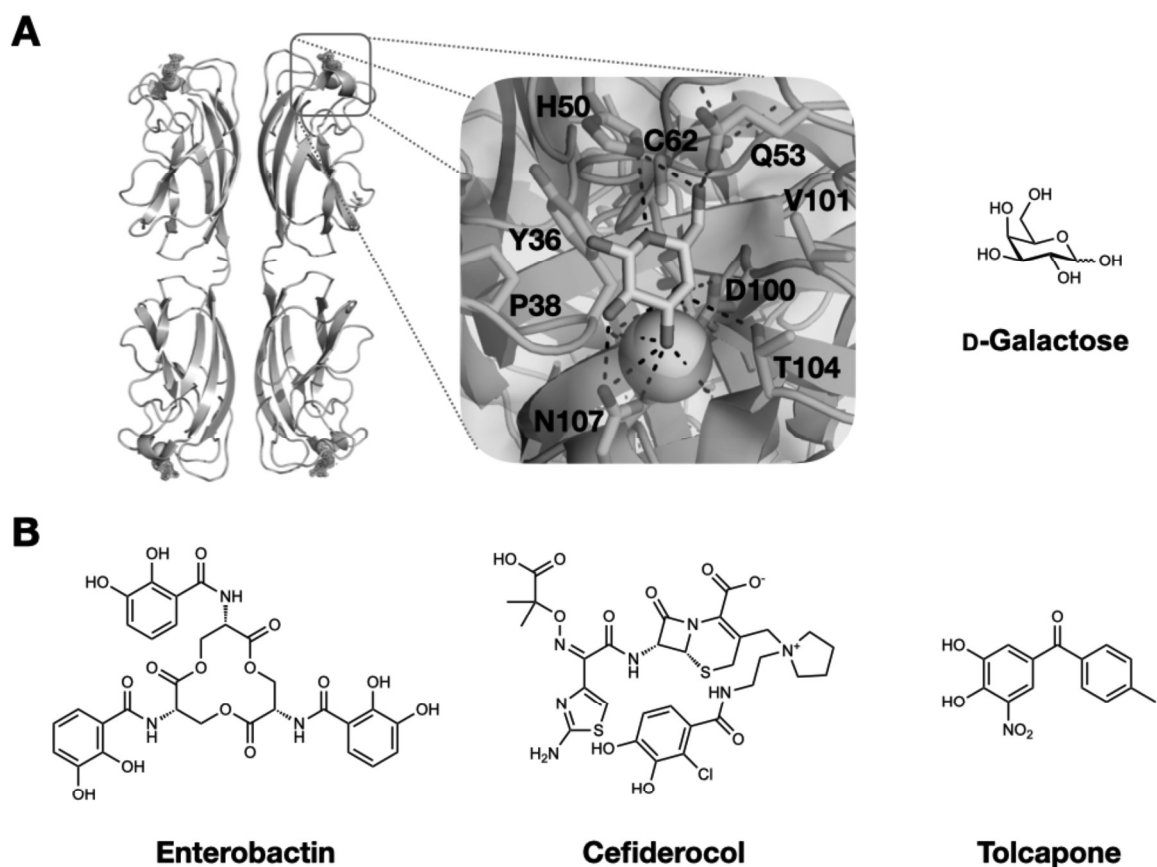


Figure 2: **A)** Top: co-crystal structure of galactose in complex with LecA (pdb: 1oko) with view of the carbohydrate binding site and hydrogen bond interactions (black dashed line) of galactose (cyan) with LecA (light blue, calcium ion = green sphere) **B)** Catechols of medicinal interest: enterobactin (siderophore), cefiderocol (β -lactam antibiotic with siderophore unit) and tolcapone (medication for Parkinson's disease).

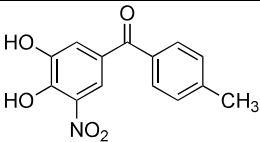
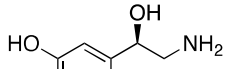
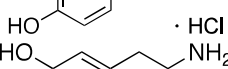
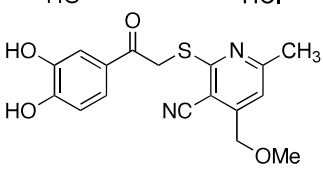
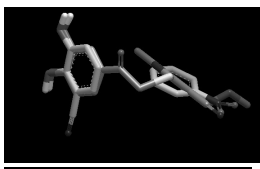
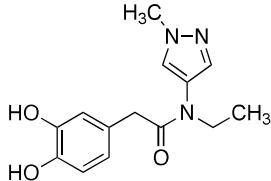
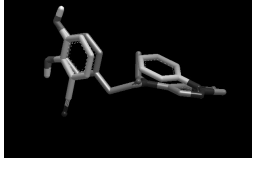
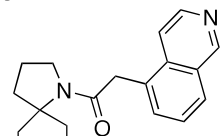
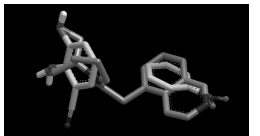
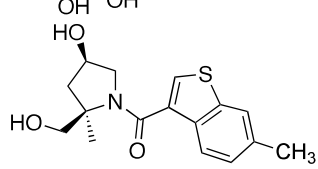
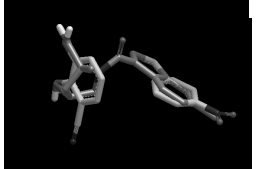
Results and Discussions

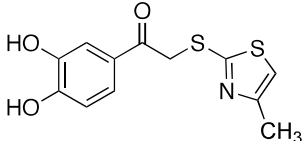
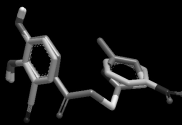
Selection of New Catechol-like Compounds

After we identified electron-deficient catechols as calcium-ion depended lectin inhibitors^[91], we focused on their affinity optimization towards LecA. It is known that the potency of carbohydrate based LecA inhibitors increases if an aromatic ring is introduced in proximity to the galactose moiety which forms a $\text{CH}\cdots\pi$ interaction with His50 as for *para*-nitrophenyl galactopyranoside (pNPG) ($K_d = 14.1 \mu\text{M}$).^[102] Thereupon, we looked into commercial compounds of different suppliers that carried a catechol-like moiety and an aromatic ring in close proximity. We then superimposed the structures with 3-cyanocatechol and the aromatic ring of pNPG from their

corresponding co-crystal structures (pdb: 6yo3, pdb: 3zyf) *in silico* by overlapping their pharmacophores accordingly. We identified 25 compounds for which the fit of the catechol function and aromatic ring was good (**Table S1**). From those, we chose a structurally diverse set of nine compounds of electron-deficient catechols, benzene-1,2-diamine, and compounds with a dihydroxy pyrrolidine unit (highlighted in green **Table S1, Table 1**).

Table 1: Catechol-like compounds tested against LecA *in silico* by overlapping their pharmacophores with the co-crystal structures of 3-cyanocatechol and pNPG (pdb: 6yo3, pdb: 3zyf) in the fluorescence polarization (FP) assay and protein-observed ^{19}F (PrOF) NMR.

No.	Structure	Docking	FP Inhibition [%] at 4 mM	^{19}F PrOF NMR CSP of W42 [ppm]
1		Not docked	Insoluble	Insoluble
2		Not docked	Inconclusive	Not significant
3		Not docked	Inconclusive	Not significant
4			Insoluble	Not significant
5			Inconclusive	0.10
6			No binding	W33 increment
7			No binding	Not significant

8			$49 \pm 7\%^{[*]}$	0.27 W33 increment
---	---	---	--------------------	-----------------------

Protein-observed ^{19}F (PrOF) NMR was performed with 5FW-LecA. $^{[*]}$ FP averages and standard deviation from two experiments, %inhibition compared to αMeGal at 4 mM. CSP = chemical shift perturbation.

Biophysical Validation of Catechol-like Compound Set

Five of those compounds were commercial available and afterwards tested together with three additional compounds: tolcapone (**1**), noradrenaline (**2**) and dopamine (**3**) in an established competitive binding assay based on fluorescence polarization (FP)^[91,103] (**Table 1**). Two compounds, tolcapone and catechol **4** precipitated under those assay conditions with 4% DMSO at 4 mM. The two catecholamines dopamine and noradrenalin (electron-rich) and catechol **5** with a dihydroxyphenyl acetamide moiety were unstable over time and were not further studied. The two dihydroxy pyrrolidines **6** and **7** did not inhibit LecA. Only catechol **8** inhibited LecA with $49 \pm 7\%$ at 4 mM compared to αMeGal and was slightly more active than 3,4-dihydroxybenzophenone our previously best non-carbohydrate LecA inhibitor in this assay (%inh.: $39 \pm 5^{[91]}$). Afterwards, we remeasured catechol **4** and tolcapone (**1**) with an increased DMSO concentration (25%) to avoid precipitation. Compound **4** showed no inhibition to LecA below 1 mM and interfered with Cy5-Gal^[91] reporter ligand at higher concentrations. Tolcapone on the other hand showed a full IC_{50} curve and was stable over time and thus considered as hit. To avoid a false positive, we followed this compound under assay conditions by HPLC-UV and did not see a decomposition over time (**Figure S1**). In addition, to avoid an unspecific chelation to calcium ions, we repeated this assay with a tenfold higher calcium ion concentration (10 mM) without seeing a negative inhibition effect (data not shown). Afterwards, we tested tolcapone together with two positive controls αMeGal ($\text{IC}_{50} = 39 \pm 16 \mu\text{M}^{[84]}$), and GalNAc ($\text{IC}_{50} = 1230 \pm 200 \mu\text{M}^{[104]}$) to qualify its IC_{50} value (**Figure 3A**). In presence of 25% DMSO, tolcapone showed an IC_{50} of $238 \pm 49 \mu\text{M}$ and was therefore threefold more potent than GalNAc ($\text{IC}_{50} = 674 \pm 182 \mu\text{M}$) but fourfold less active than αMeGal ($\text{IC}_{50} = 70 \pm 20 \mu\text{M}$).

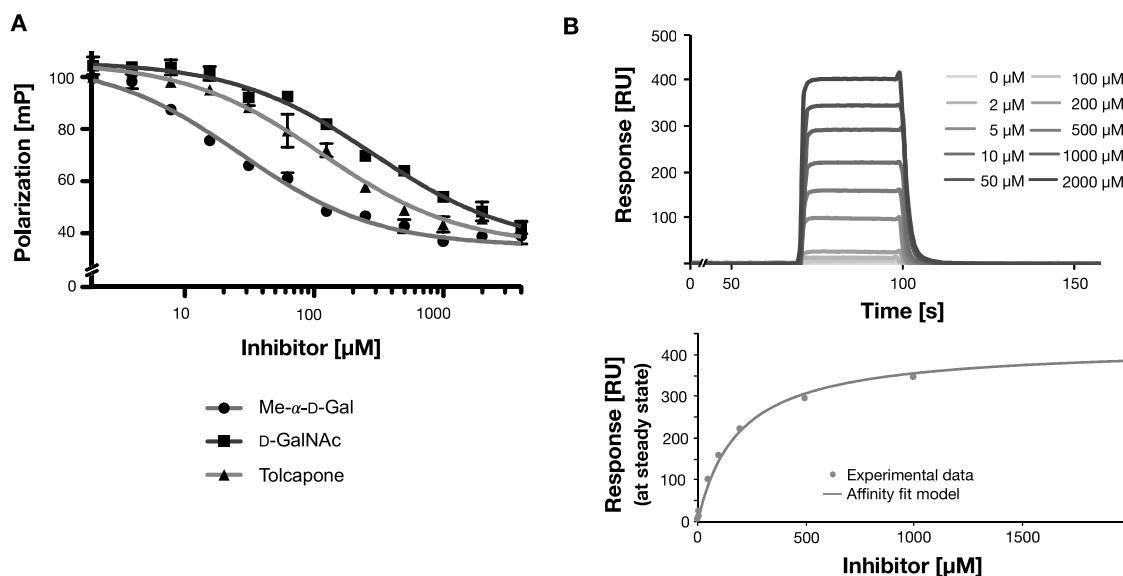


Figure 3: Analysis of tolcapone interacting with LecA using **A)** a competitive binding assay based on fluorescence polarization in presence of two positive controls α MeGal and GalNAc. Standard deviations of one technical replicate. **B)** Surface plasmon resonance using multi-cycle studies obtained by titration of tolcapone to LecA (top panel) with integration of peaks and fit (bottom panel). Both graphs show one representative titration of tolcapone.

Further, we tested the eight commercial compounds by protein-observed ^{19}F (PrOF) NMR with 5-fluorotryptophan-labeled LecA (5FW). A technique to identify weak ligands in the low millimolar binding range.^[104] Compounds binding to the carbohydrate binding site (CBS) show chemical shift perturbations (CSP) on Trp42 and an additional peak enhancement on Trp33 if they are moderately potent as shown for pNPG and GalNAc.^[91] As a result, both catechols **5** (0.10 ppm) or **8** (0.27 ppm) showed a CSP above the threshold of 0.01 ppm^[104] on Trp42 and in case of **8** an additional effect on Trp33 (**Table 1**, **Figure S2**). In presence of tolcapone (**1**), the fluorine signals were absent most probably due to crushed protein initiated by the precipitation of tolcapone. The two catecholamines **2** and **3**, catechol **4** and the dihydroxy pyrrolidine **7** did not show a CSP on Trp42 which fits to the binding results from the FP assay. The second dihydroxy pyrrolidine compound **6** showed only a peak enhancement on Trp33 which might come from hydrophobic interactions of its quinoline residue with LecA without binding to the calcium ion from the CBS. In summary, ^{19}F PrOF NMR confirmed our binding studies from the FP assay that catechol **8** bound to the CBS of LecA in low DMSO concentrations.

Then, tolcapone was studied by surface plasmon resonance (**Figure 3B**). Multi-cycle analysis was used by injecting tolcapone in concentrations between 0–2000 μM over an immobilized LecA. The K_d value was calculated with a steady state affinity model leading to a value of $181 \pm 6 \mu\text{M}$ that corresponded to the measured IC_{50} of $238 \pm 49 \mu\text{M}$ determined by FP. This makes tolcapone threefold more active than the previously best catechol ligand, 4-nitrobenzene-1,2-diol by SPR ($K_d = 560 \pm 340 \mu\text{M}^{[91]}$) and 19-fold more active than our previously best FP hit, 3,4-dihydroxybenzophenone ($K_d = 3460 \pm 410 \mu\text{M}^{[91]}$). The association and dissociation were too fast to determine the rate constants.

Table 2: Binding affinities of tolcapone determined in the competitive binding assay based on fluorescence polarization and by surface plasmon resonance from three independent experiments respectively.

Compound	FP (IC_{50} [μM])	SPR (K_d [μM])
Tolcapone	238 ± 49	181 ± 6

Tolcapone Binds to LecA Determined by X-ray Crystallography

We co-crystallized tolcapone with LecA to further understand the binding interactions at the atomic level. Crystals were obtained using a hanging drop method.^[105] The orthorhombic co-crystal was solved at a resolution of 1.32 Å with space group C 2 2 2₁ (pdb: 8GUV).

In protomer A, tolcapone was fully resolved although the electron density for the *p*-tolyl ring was not completely visible (**Figure 4**, protomer A). The ligand bound with its two hydroxy groups to the calcium ion mimicking OH3 and OH4 of galactose and with its nitro group to Pro51 *via* a water bridge mimicking OH6 and thus showed the same orientation and interactions as the previously co-crystallized 3-cyanocatechol with LecA (pdb: 6yo3, **Figure S4A**, superimposed). The carbonyl group together with the *p*-tolyl ring pointed into space where the aromatic ring formed weak hydrophobic contacts with the surface but without forming a $\text{CH}\cdots\pi$ interaction with His50 as shown for pNPG with LecA (**Figure S4B**, superimposed).

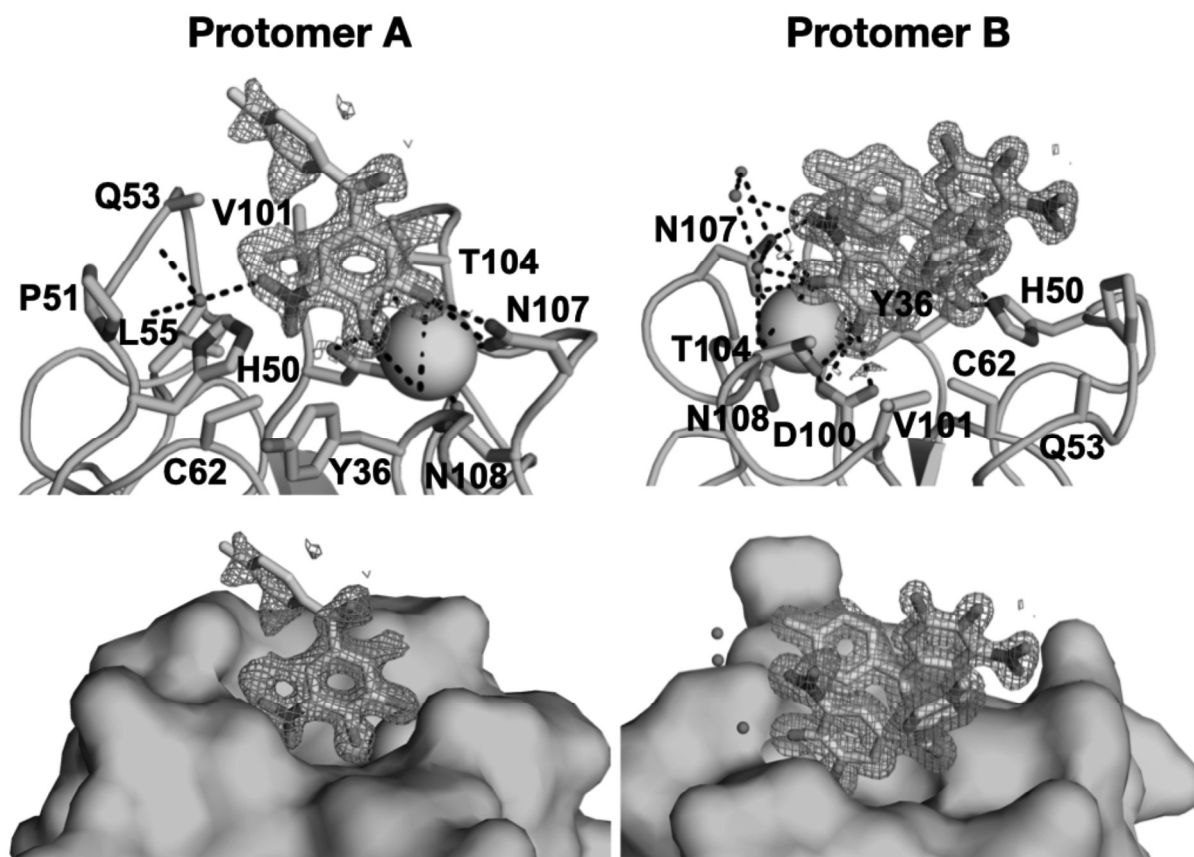


Figure 4: Tolcapone in complex with LecA (pdb: 8GUV); the ligand is presented as sticks (cyan), calcium ions as spheres (green), hydrogen bonds of the ligand with LecA as dotted lines (black), the electron density at 1σ . Top: interaction of tolcapone with LecA in protomer A and B, bottom: surface display of tolcapone at the carbohydrate binding sites.

In protomer B, the electron density represented two tolcapone structures (**Figure 4**, protomer B). One ligand interacted with its carbonyl function with His50, and the dihydroxy function pointed into space towards a neighboring crystal's calcium ion where its catechol ring formed an offset π/π -stacking with the catechol ring of the tolcapone ligand from protomer A. The *p*-tolyl ring formed a $\text{CH}\cdots\pi$ interaction with Pro38. However, we hypothesize that this hydrophobic ligand which sticks between two neighboring ligands is a crystal artifact. The other tolcapone ligand bound with its hydroxy groups to the calcium ion of LecA. The nitro group pointed into space (**Figure S4A**, superimposed with 3-cyanocatechol) and is exposed to the solvate layer, however, it still formed a hydrogen bond with Asn107. The oxygen of the carbonyl function formed a hydrogen bond with His50. The *p*-tolyl residue pointed away from the surface and did not contribute in binding but might form hydrophobic contacts with the second ligand in this protomer B. The superimposed complex of protomer B with

pNPG (pdb: 3zyf, **Figure S4B**) demonstrated that the hydroxy groups of tolcapone mimic OH3 and OH4 of pNPG by binding to the calcium ion, however, the *p*-tolyl part was orientated completely different without forming the CH $\cdots\pi$ interaction with His50 as shown for the *p*-nitrophenyl residue of pNPG.

Tolcapone-like Structure Motif Most Promising for Potent Non-Carbohydrate LecA Inhibitors

Our promising discovery of tolcapone being a micromolar LecA ligand led to a collaboration with Roche – the company who developed tolcapone as a Parkinson's disease medication. They provided two sets of compounds with molecular weights smaller 500 Da in three different concentrations (100 (125), 25 (32), 3 (4) μ M) and four replicates each, one with tolcapone-like compounds (342), and one bearing catechol-like motifs (3222) which resulted in several identical molecules in both sets. These compounds were then measured in the fluorescence polarization assay supplemented with 10% DMSO and 2-mercaptoethanol as a nucleophile to trap possibly formed quinones and prevent assay interference. pNPG served as a positive control and was added twice to this assay once equimolar to the compound concentration and at 500 μ M for a complete inhibition. Then, the protein solution together with a red-shifted reporter ligand (Cy5-Gal^[91]) was added and the polarization was measured after an incubation of 1 h and 24 h. Afterwards, the relative inhibition was calculated compared to pNPG at 500 μ M, the measured intensities were normalized to the negative control and compared. Compounds that showed polarization >105% due to spectral overlaps, precipitations or quench of the reporter ligand or a normalized intensity deviation >0.5 were excluded. Then, compounds with inhibitions >50% at the highest concentration after 1 h were selected which resulted in 48 hits that are listed in **Table S3**.

The first compound set was used for the optimization of this assay resulting in only three out of four replicates at the highest (125 μ M) and two for the second highest concentration (25 μ M). The readout at the lowest concentration (3 or (4) μ M) in both sets was difficult because the compound concentrations were below the protein concentration (5 μ M), hence, a fully outcompeted reporter ligand was not ensured and even the positive control showed very weak inhibition at that concentration. Performing this assay with lower protein concentrations narrows the polarization range and makes the readout not possible.

The first set resulted in 27/342 hits and the second led to 44/3222 compounds that fulfilled our criteria. 25 hits of the 44 compounds were identical to the hits from the first set. Therefore the following SAR will be mainly described on the second screening results at the highest measured concentration. The involved compounds are summarized in **Table 3**.

In general, all 48 screening hits showed a tolcapone-like structure motif with an electron withdrawing group next to the hydroxy groups. This was mainly a nitro function except in two cases where the nitro group was replaced by a triflyl **9** or an oxadiazole **10** moiety but keeping the general tolcapone motif. Surprisingly, tolcapone itself that was only included in the first screening set did not show an inhibition under these screening conditions and was not found by LC-MS after the experiment. The other compounds had a carbonyl group in *meta* position to the nitro group except two compounds **11** and **12** where the carbonyl group was located in *ortho* position. The additional phenyl ring of the screening hits was predominantly substituted with electron withdrawing groups (EWG) in different positions but there were also a few compounds where the phenyl ring was replaced by either substituted indoles **13** and **14**, quinolines **15** and **16** or rather a tertiary amide **12**.

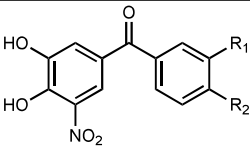
At 100 μM , the best inhibition was measured for compound **17** that contained an allyl ester in *para* position on the phenyl ring instead of the methyl group as in tolcapone, however, this compound showed only a moderate inhibition at 25 μM (inh. at 100 μM : $103 \pm 3\%$, inh. at 25 μM : $17 \pm 5\%$). Whereas compound **18** with a benzylester in *para* position showed the overall best inhibition (inh. at 100 μM : $94 \pm 8\%$, inh. at 25 μM : $88 \pm 7\%$) closely followed by **13** where the phenyl ring was replaced by a 5-bromo-3-methyl indole (inh. at 100 μM : $97 \pm 4\%$, inh. at 25 μM : $43 \pm 14\%$). In a direct comparison of *meta/para* substituted phenyl rings, it seemed that the *para* position was slightly favored. Comparing the pentyl esters **19** and **20**, the *meta* compound showed a slightly higher potency (inh. at 100 μM : *meta*-**19** = $92 \pm 10\%$, *para*-**20** = $89 \pm 1\%$), however, in case 4,4,4-trifluoro-3-hydroxybutyl ester **21** and **22**, the allyl esters **23** and **17** and the fluorinated compounds **24** and **25**, the *para* compounds showed higher potencies (inh. at 100 μM : *meta*-**21** = $60 \pm 3\%$, *para*-**22** = $71 \pm 4\%$, *meta*-**23** = $63 \pm 4\%$, *para*-**17** = $103 \pm 3\%$, *meta*-**24** = $52 \pm 5\%$, *para*-**25** = $55 \pm 5\%$). To further quantify the favored position to improve the potencies, the compounds have to be compared in dose-dependent experiments. The length of the alkyl ester in *ortho/meta* position had no significant influence on the affinity, however, linear alkyl chains were favored over branched

alkanes; esters were slightly favored over tertiary amides which were favored over secondary amides. Additionally, substituents in *ortho* position of the phenyl ring were tolerated, however, EWGs were favored over EDGs e.g.: trifluoromethyl **26** or methyl **27** (inh. at 100 μ M: **26** = $86 \pm 0\%$, **27** = $79 \pm 3\%$), though, the nitrile substituted compound **28** showed the weakest inhibition (inh. at 100 μ M (**28**): $72 \pm 2\%$).

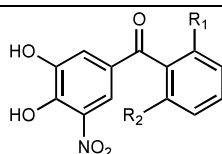
The carbonyl function between the catechol moiety and the phenyl ring was crucial but could be extended to a diketone as can be seen in **29** and *p*-tolyl compound **30**. The diketones were slightly more potent than the carbonyl compounds e.g.: 3,4-dihydroxy 5-nitrobenzophenone (**31**) (inh. at 100 μ M: **29** = $69 \pm 2\%$, **30** = $66 \pm 4\%$, **31** = $42 \pm 7\%$). Additionally, the carbonyl function was tolerated in *ortho* and *meta* position to the nitro group, however, the *meta* position was favored. (inh. at 100 μ M: **11** = $58 \pm 6\%$, **12** = $28 \pm 5\%$).

Replacement of the nitro group was accepted as in triflyl **9** or oxadiazole **10** but the compounds were slightly less active compared to the nitro compounds (inh. at 100 μ M: **9** = $60 \pm 7\%$, **10** = $20 \pm 4\%$). Based on the co-crystal structure, we would suggest one small substituent such as nitro or nitrile that fits into the small cleft from the CBS of LecA to form the additional hydrogen bond with Pro51 *via* a water bridge. Interestingly, the nitrile compounds showed affinities below our selection criteria. The best inhibition was obtained for 5-(2-fluorobenzoyl)-2,3-dihydroxybenzonitrile (**31**) followed by 2,3-dihydroxy-5-(4-methylbenzoyl)benzonitrile (**32**) (inh. at 125 μ M: **31** = $41 \pm 12\%$, **32** = $20 \pm 5\%$).

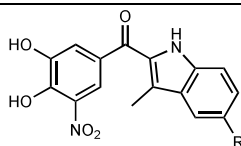
Table 3: Selection of the top 48 screening hits from the Roche library (3564 compounds) that showed at least 50% inhibition after 1 h compared to pNPG (500 μ M) in the fluorescence polarization assay against LecA. The experiment was conducted in four independent experiments for the concentrations 100 and 25 μ M, in three at 125 μ M and in two at 32 μ M.

				
No.	R1	R2	%inhibition at 100 μ M	%inhibition at 25 μ M
19	-CO ₂ (CH ₂) ₄ CH ₃	-H	92 \pm 10	49 \pm 5
23	-CO ₂ Allyl	-H	63 \pm 4	26 \pm 5
21	-CO ₂ (CH ₂) ₂ CHOH _(S) CCF ₃	-H	60 \pm 3	24 \pm 5
24	-F	-H	52 \pm 5	23 \pm 4
17	-H	-CO ₂ Allyl	103 \pm 3	17 \pm 5
18	-H	-CO ₂ CH ₂ Ph	94 \pm 8	88 \pm 7
20	-H	-CO ₂ (CH ₂) ₄ CH ₃	89 \pm 1	28 \pm 6

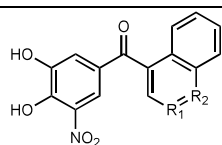
22	-H	-CO ₂ (CH ₂) ₂ CHOH _(S) CF ₃	71 ± 4	16 ± 4
25	-H	-F	55 ± 5	21 ± 5



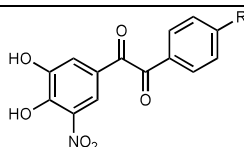
No.	R₁	R₂	at 100 μM	at 25 μM
26	-CF ₃	-H	86 ± 0	52 ± 5
27	-CH ₃	-H	79 ± 3	44 ± 7
28	-CN	-H	72 ± 2	48 ± 3



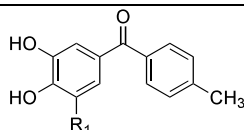
No.	R₁	at 100 μM	at 25 μM
13	-Br	97 ± 4	43 ± 14
14	-H	49 ± 5	18 ± 5



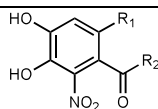
No.	R₁	R₂	at 100 μM	at 25 μM
15	-H	-N	78 ± 5	42 ± 2
16	-N	-H	62 ± 3	26 ± 4



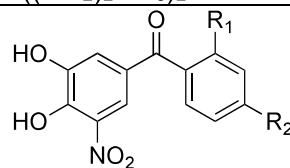
No.	R₁	at 100 μM (at 125 μM)	at 25 μM (at 32 μM)
30	-CH ₃	69 ± 2	36 ± 3
29	-H	(66 ± 4)	(31)



No.	R₁	at 100 μM	at 25 μM
9	-S(O) ₂ CF ₃	60 ± 7	20 ± 4
10	5-(3-methyl-1,2,4-oxadiazole)	58 ± 4	19 ± 4



No.	R₁	R₂	at 100 μM	at 25 μM
11	-H	-C ₄ H ₆ - <i>p</i> CH ₃	58 ± 6	35 ± 11
12	-Br	-N((CH ₂) ₂ CH ₃) ₂	58 ± 5	24 ± 5



No.	R₁	R₂	at 100 μM (at 125 μM)	at 25 μM (at 32 μM)
------------	----------------------	----------------------	----------------------------------	--------------------------------

31	-F	-H	(35 ± 6)	(4)
32	-H	-CH ₃	20 ± 5	4 ± 6

Based on the *in vitro* screening data of the 3564 compounds, the most active LecA inhibitor would be a tolcapone-like molecule with a strong EWG such as nitro in *ortho* position to the dihydroxy function, carrying one to two carbonyl groups between the additional phenyl ring that should also be substituted with small EWGs in *ortho* position and linear alkyl ester chains in the *meta* and *para* positions (**Figure 5**).

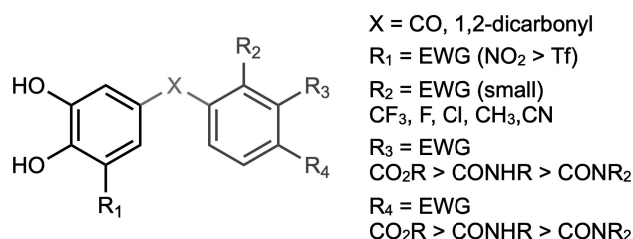


Figure 5: Structure activity relationship of 3564 compounds (<500 Da) against LecA in the competitive binding assay based on fluorescence polarization summarized in one model structure.

Summary and Outlook

In our research for novel non-carbohydrate LecA glycomimetics, we screened commercial catechol-like compounds based on a rational ligand based design *in vitro* in a fluorescence polarization assay leading to non-carbohydrate micromolar LecA ligand tolcapone that was further validated in different biophysical assays. We revealed its interactions at the atomic level by x-ray diffraction analysis.

Commercial compounds that consisted of a dihydroxy function and an aromatic residue at a concise distant were identified and superimposed by their functionalities *in silico* with 3-cyanocatechol (pdb: 6yo3) and the phenyl ring of pNPG (pdb: 3zyf). Together with three additional compounds, we tested a distinct compound set in a fluorescence polarization assay that resulted in two hits: catechol **8** and tolcapone that were further validated by PrOF NMR or SPR. Tolcapone is the first non-carbohydrate LecA ligand with comparable affinities to natural LecA ligands in the micromolar range (K_d (**1**) = $181 \pm 6 \mu\text{M}$, K_d (D-Gal) = $88 \mu\text{M}$ ^[51]). The interactions with LecA were revealed by crystallography where it bound to the calcium ion with its two hydroxy groups by mimicking OH3 and OH4 of galactose and its nitro group bound to Pro51 *via* a water bridge mimicking OH6 (protomer A). The *p*-tolyl ring was resolved fully in protomer A

although the electron density was not completely visible and its orientation might differ. In protomer B, tolcapone's *p*-tolyl ring was hindered to form the potential CH $\cdots\pi$ interaction with His50 as in pNPG by an additional tolcapone ligand stuck between it and a ligand from a neighboring LecA unit due to favorable hydrophobic contacts which was considered to be a crystal artifact.

Further, we screened a library of 3564 catechol-like compounds against LecA in the fluorescence polarization assay resulting in 48 tolcapone-like hits and whose structure activity relationship (SAR) was analyzed. Judging from the single point measurements, *para* benzylester **18** was more potent than the positive control pNPG (inh. at 100 μ M: **18** = 94 \pm 8%, pNPG = 90 \pm 5%; inh. at 25 μ M: **18** = 88 \pm 7%, pNPG = 59 \pm 4%). Combining the best features from the SAR, the most potent non-carbohydrate LecA ligand should have a strong EWG (eg.: nitro) in *ortho* position to their dihydroxy function and a carbonyl group in either *ortho* or *meta* position with an attached phenyl ring substituted with small EWGs in *ortho* position and alkyl ester chains in *meta* and *para* positions.

The next step will be the validation of the 48 screening hits in dose-dependent experiments by fluorescence polarization. The best hits will be further studied by surface plasma resonance, isothermal titration calorimetry and x-ray crystallography. A re-synthesis of the most promising hits and re-validation in biophysical assays will then hopefully result in a potent non-carbohydrate LecA lead structure.

Thus, the tolcapone-like compounds represent a promising LecA inhibitor class for future applications as pathoblockers against *P. aeruginosa* infections.

Material and Methods

Docking

All *in silico* experiments were performed by Didier Rognan.

Recombinant Expression and Purification of LecA

LecA was expressed and purified using *Escherichia coli* BL21 (DE3) and the plasmid pET25pa1I^[106] as described before.^[84]

Biophysical Evaluation

Competitive Binding Assay (dose-dependent)

The competitive binding assay based on fluorescence polarization was performed in analogy to a previous procedure.^[84] To a solution (10 μ L) of LecA (40 μ M) and fluorescent reporter ligand (Sulfo-Cy5-Gal^[91], 20 nM) in TBS/Ca²⁺-buffer containing 8% DMSO (20 mM Tris, 137 mM NaCl, 2.6 mM KCl, 1 mM CaCl₂, pH = 7.4), a 10 μ L serial dilution of the inhibitor in the same buffer (4000–31.25 μ M, dilution factor 2) was added in triplicates in a black 384-well plate (Greiner Bio-One, Germany, cat. no. 781900). The plate was sealed (EASYseal, Greiner Bio-One, cat. no. 676001), centrifuged (1500 x g, r.t., 1 min) and incubated in a dark chamber under shaking conditions for 24 h. Afterwards, the cover was removed and fluorescence was measured on a PHERAstar FS (BMG Labtech, Germany, filter ex.: 590 nm em.: 675 nm). After blank value (TBS/Ca²⁺-buffer with LecA) subtraction from the measured values, polarization was calculated and the data were analyzed using the four-parameter variable slope model in MARS software (BMG Labtech). Then, top and bottom plateaus were redefined using the values of the positive control α MeGal, and TBS/Ca²⁺-buffer containing LecA and reporter ligand as a negative control. Three independent experiments were performed, the data were averaged and visualized using GraphPad PRISM (version 5). The experiments containing tolcapone were conducted at a final DMSO concentration of 25%. The experiment with an enhanced calcium ion concentration was performed with the same buffer but supplemented with 10 mM CaCl₂.

Competitive Binding Assay (single point measurements)

Screening compounds were provided in black 384-well plates (Greiner Bio-One, Germany, cat. no. 781900) by Roche (1 μ L per well in DMSO, first set: 2500, 638.3, 86.8 μ M, second set: 2000, 500, 60 μ M). Then, the compounds were diluted with 9 μ L of TBS/Ca²⁺-buffer containing 11.1% DMSO (20 mM Tris, 137 mM NaCl, 2.6 mM KCl, 1 mM CaCl₂, pH = 7.4). Afterwards, a solution of LecA (10 μ M, 10 μ L) and fluorescent reporter ligand (Sulfo-Cy5-Gal^[91], 20 nM) was added. The plates were sealed (EASYseal, Greiner Bio-One, cat. no. 676001), gently shaken (IKA MS3 digital, 500 rpm, 30 s), centrifuged (1500 x g, r.t., 1 min) and incubated in a dark chamber under shaking conditions for 1 h and 24 h. The cover was removed and fluorescence was measured on a PHERAstar FS (BMG Labtech, Germany, filter ex.: 590 nm em.: 675 nm). The blank value (TBS/Ca²⁺-buffer containing 10% DMSO with LecA (5 μ M)) was subtracted from the measured values, polarization and intensities were calculated

with the MARS software (BMG Labtech) and the %inhibition (4) was calculated from the measured polarizations (P) of the analyte (X) by defining the top and bottom plateaus using the values of the positive control pNPG (at 500 μ M, PC), and TBS/Ca²⁺-buffer containing LecA and reporter ligand as a negative control (NC).

$$\%_{\text{Inhibition}} = \frac{P_X - P_{\text{PC}}}{P_{\text{PC}} - P_{\text{NC}}} * 100 \quad (1)$$

The intensities (I) of the individual compounds (X) were normalized to the intensity of NC with the following equation (2):

$$N_{\text{Intensity}} = \frac{I_X}{I_{\text{NC}}} * 100 \quad (2)$$

For the first screening set an additional linear regression was performed due to shrinking polarization and intensity values in the black well plate from top left to bottom right. A trendline (equation (3)) was created.

$$y = mx + b \quad (3)$$

Polarizations values were corrected by using the gradient (m) of the curve that was then multiplied with the position on the well (x) and subtracted from the measured polarization.

Protein-Observed ¹⁹F NMR

5FW-LecA was expressed recombinantly as reported previously.^[104] The NMR experiments were conducted on a Bruker Ascend™700 (AvanceIII HD) spectrometer equipped with a 5 mm TCI700 CryoProbe™ in 3 mm tubes (Norell S-3-800-7). PrOF NMR was recorded using 5FW-LecA (200 μ M) in Tris-HCl-buffer (20 mM, 150 mM NaCl, 10% D₂O, pH 7.8) and trifluoroacetic acid (100 μ M) at 310 K. Spectra changes upon addition of catechol-like compounds (1.0 mM) were processed, referenced and normalized to TFA as internal reference (-75.6 ppm). We considered only changes in chemical shift perturbation (CSP) >0.01 ppm of fluorine resonance.

Affinity Analysis by Surface Plasmon Resonance

CM7 chip was immobilised with LecA following standard amine coupling as previously described.^[91] Tolcapone was dissolved in 100% DMSO to the final concentration of 200 mM and diluted to 10 mM in PBS supplemented with 0.05% Tween® 20 and 100 μ M CaCl₂. The 10 mM stock was then diluted to required concentrations for affinity analyses in a running buffer (PBS supplemented with 0.05% Tween® 20, 100 μ M

CaCl₂, and 5% DMSO). Tolcapone was subjected to multi-cycle steady-state affinity studies (30 s association, 60 s dissociation, flow rate 30 μ L min⁻¹) consisting of injections of compounds at increasing concentrations onto the immobilized LecA. Binding was measured after reference subtraction of channel 1 (no immobilized LecA). The experiment was performed on BIACORE X100 at 25 °C and data evaluation was performed using BIACORE X100 evaluation software (version 2.0).

Protein Crystallization and X-ray Crystallography

LecA was dissolved in 20 mM HEPES pH = 7.5, 100 mM NaCl, 100 μ M CaCl₂ to the final concentration of 8 mg ml⁻¹. Tolcapone (200 mM in 100% DMSO) was mixed with LecA solution (1:20 ratio) and incubated for 1 h at room temperature. LecA-tolcapone mixture was then mixed with the crystallization solution (24% PEG 2kMME, 0.1 M KSCN, 0.1 M sodium acetate pH = 4.5) in 1:1 ratio and deposited on a siliconized glass circle cover slide (22 mm, Hampton research). Crystallization was performed by the hanging drop vapor diffusion method on a 24-well plate with sealant (Hampton Research) at 19 °C. The protein crystals were harvested without cryo protection and flashed cool in liquid nitrogen. X-ray diffraction data were collected at SOLEIL-PROXIMA1 (Saint Aubin, France) using a Pilatus 6M hybrid photon counting detector (Dectris). The recorded data were indexed, integrated and scaled using XDS^[107] and merged using AIMLESS^[108]. The structures were solved by molecular replacement using 1OKO as a searching template in PHASER^[109], followed by further iterations of manual rebuilding in COOT^[110] and restrained refinement in REFMAC5^[111]. Tolcapone structure was manually built in ACEDRG^[112]. The final models were validated with MOLPROBITY^[113], PDB-redo (<https://PDB-redo.eu/>) and wwPDB validation service (<http://validate-rcsb-1.wwPDB.org/>) prior to submission to the Protein Data Bank. Data processing and refinement statistics are summarised in **Table S2**.

Orcid

Eike Siebs: 0000-0002-7349-9872

Sakonwan Kuhaudomlarp: 0000-0003-4415-3781

Elena Shanina: 0000-0003-4235-9452

Annabelle Varrot: 0000-0001-6667-8162

Priscila da Silva Figueiredo Celestino Gomes: 0000-0001-7370-9596

Uwe Grether: 0000-0002-3164-9270

Christoph Rademacher: 0000-0001-7082-7239

Didier Rognan: 0000-0002-0577-641X

Anne Imberty: 0000-0001-6825-9527

Alexander Titz: 0000-0001-7408-5084

Notes

The authors declare no competing financial interest.

Acknowledgements

A.T., A.I., D.R. and C.R. acknowledge financial support of the French-German ANR/DFG project (ANR-AAPG-2017) funded by the Agence Nationale de la Recherche (grant no. ANR-17-CE11-0048) and Deutsche Forschungsgemeinschaft (grant no. Ti756/5-1 and RA1944/7-1). We acknowledge synchrotron SOLEIL (Saint Aubin, France) for access and technical support at beamline PROXIMA 1. We also acknowledge Dirk Hauck (HIPS) and Emilie Gillon (CERMAV) for excellent technical support. In addition, we like to thank the whole Roche team: Larissa Valeska Thiele, Bernd Kuhn, Isabelle Kaufmann, Alexandre Zimmermann and Philippe Hartz for their support in providing the 3564 screening compounds.

Targeting the Central Pocket of the *Pseudomonas aeruginosa* Lectin LecA

Eike Siebs,^[a, b, c] Elena Shanina,^[d, e] Sakonwan Kuhaudomlarp,^[f, g]
Priscila da Silva Figueiredo Celestino Gomes,^[h] Cloé Fortin,^[a] Peter H. Seeberger,^[d, e]
Didier Rognan,^[h] Christoph Rademacher,^[d, e, i, j] Anne Imberty,^[f] and Alexander Titz*^[a, b, c]

Pseudomonas aeruginosa is an opportunistic ESKAPE pathogen that produces two lectins, LecA and LecB, as part of its large arsenal of virulence factors. Both carbohydrate-binding proteins are central to the initial and later persistent infection processes, i.e. bacterial adhesion and biofilm formation. The biofilm matrix is a major resistance determinant and protects the bacteria against external threats such as the host immune system or antibiotic treatment. Therefore, the development of drugs against the *P. aeruginosa* biofilm is of particular interest to restore efficacy of antimicrobials. Carbohydrate-based inhibitors for LecA and LecB were previously shown to efficiently reduce

biofilm formations. Here, we report a new approach for inhibiting LecA with synthetic molecules bridging the established carbohydrate-binding site and a central cavity located between two LecA protomers of the lectin tetramer. Inspired by *in silico* design, we synthesized various galactosidic LecA inhibitors with aromatic moieties targeting this central pocket. These compounds reached low micromolar affinities, validated in different biophysical assays. Finally, X-ray diffraction analysis revealed the interactions of this compound class with LecA. This new mode of action paves the way to a novel route towards inhibition of *P. aeruginosa* biofilms.

Introduction

Pseudomonas aeruginosa belongs to the ESKAPE pathogens and is listed by the World Health Organization as the most critical bacterial pathogen. This rod-shaped, Gram-negative ubiquitous bacterium lives under moist conditions and thrives in hospitals, colonizing patients under ventilation or bearing catheters and cystic fibrosis patients.^[1] For colonizing the host, many bacteria exploit cell surface carbohydrates, the glycocalyx,^[2] through their carbohydrate-binding proteins (lectins), adhesins or capsid proteins.^[3] In case of *P. aeruginosa* two soluble virulence factors are involved: the D-galactose-binding LecA and D-mannose-/L-fucose-binding LecB.^[4] These two extracellular, calcium(II)-dependent and homotetrameric lectins are essential for establishing the biofilm matrix, a protective environment against environmental stress and antibiotic treatment and, thus, a major hurdle for therapy.^[5–7]

To overcome this resistance mechanism, alternative strategies are currently being developed to dismantle the pathogen and restore antibiotic efficacy via inhibition of the lectins LecA and LecB.^[8–10] Complementary to the numerous approaches for multivalent inhibition using native carbohydrates,^[11,12] we have developed glycomimetics targeting LecA and LecB. In case of LecB, we have developed small molecule glycomimetics starting from a mannoside^[13–15] via simple C-glycosides^[16] into orally available anti-biofilm lead compounds^[17,18] which are currently under further investigation. Potent monovalent glycomimetics have been proven difficult to obtain for galactophilic LecA,^[19–21] thus requiring divalent ligands that display two galactose residues to simultaneously bind to two adjacent binding sites in the LecA tetramer and yield low nanomolar LecA inhibitors.^[22,23] In addition, we have recently reported conceptionally new approaches for targeting LecA, i.e. covalent lectin inhibitors^[24]

- [a] E. Siebs, C. Fortin, A. Titz
Chemical Biology of Carbohydrates (CBCH)
Helmholtz-Institute for Pharmaceutical Research Saarland (HIPS)
Helmholtz Centre for Infection Research, 66123 Saarbrücken (Germany)
E-mail: alexander.titz@helmholtz-hzi.de
- [b] E. Siebs, A. Titz
Department of Chemistry, Saarland University
66123 Saarbrücken (Germany)
- [c] E. Siebs, A. Titz
Deutsches Zentrum für Infektionsforschung (DZIF)
Standort Hannover-Braunschweig (Germany)
- [d] E. Shanina, P. H. Seeberger, C. Rademacher
Department of Biomolecular Systems
Max Planck Institute of Colloids and Interfaces, 14424 Potsdam (Germany)
- [e] E. Shanina, P. H. Seeberger, C. Rademacher
Institute of Chemistry and Biochemistry
Department of Biology, Chemistry and Pharmacy
Freie Universität Berlin, 14195 Berlin (Germany)
- [f] S. Kuhaudomlarp, A. Imberty
Université Grenoble Alpes, CNRS, CERMAV
38000 Grenoble (France)
- [g] S. Kuhaudomlarp
Department of Biochemistry and Centre for Excellence in Protein and Enzyme Technology, Faculty of Science
Mahidol University, Bangkok (Thailand)
- [h] P. da Silva Figueiredo Celestino Gomes, D. Rognan
Laboratoire d'Innovation Thérapeutique
UMR 7200 CNRS-Université de Strasbourg, Strasbourg, 67400 Illkirch (France)
- [i] C. Rademacher
Department of Pharmaceutical Sciences
University of Vienna, Althanstrasse 14, 1090 Vienna (Austria)
- [j] C. Rademacher
Department of Microbiology, Immunology and Genetics
University of Vienna, Max F. Perutz Labs, Biocenter 5, 1030 Vienna (Austria)
- Supporting information for this article is available on the WWW under <https://doi.org/10.1002/cbic.202100563>

© 2021 The Authors. ChemBioChem published by Wiley-VCH GmbH. This is an open access article under the terms of the Creative Commons Attribution Non-Commercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

and non-carbohydrate glycomimetics mimicking the binding pattern of carbohydrates.^[25]

For LecA, inhibitor design generally started from the monosaccharide galactose (Figure 1A)^[10] which is the binding epitope from its natural ligand in human, glycosphingolipid Gb3, a host cell surface receptor mediating the engulfment of *P. aeruginosa* into host cells.^[26] Galactose forms an extensive hydrogen bonding network with LecA using all hydroxy groups. OH3 and OH4 are further coordinating to the calcium ion of LecA.^[27] OH6 is pointing into a small cavity formed by a loop from His50 to Gln53 and Val101 where it is coordinated through an extended hydrogen-bonding network. Introduction of β -linked aromatic aglycons led to a fivefold affinity increase compared to aliphatic analogues, which is explained by CH- π interaction of the aryl aglycon with the side chain of His50.^[28]

In the present work, we localized a cavity between two monomers of LecA (Figure 1B), designated as central pocket, and developed synthetic galactosides carrying aromatic moieties directed towards this cavity.

Results and Discussion

Identification of the central pocket in LecA

23 crystallographic structures of LecA from *P. aeruginosa* (Uniprot reference Q05097) were retrieved from the Protein Data Bank^[30] (PDB) and analyzed for the presence of the central pocket (Figure 1, Table S1). A central pocket of LecA can be identified between two adjacent monomers in the homotetrameric LecA structure. The entrance of the pocket is clearly visible (Figure 1B) and the upper part of the cavity including its entrance, formed by residues Gln40, Lys41, Asp47, Arg48 and Glu49, is highly polar. In contrast, the interior is more hydrophobic due to the presence of two tryptophan residues, Trp33 and Trp42. The central pocket's cavity was consistently detected in 25 out of the 31 dimers in the available crystallographic structures (Table S2).

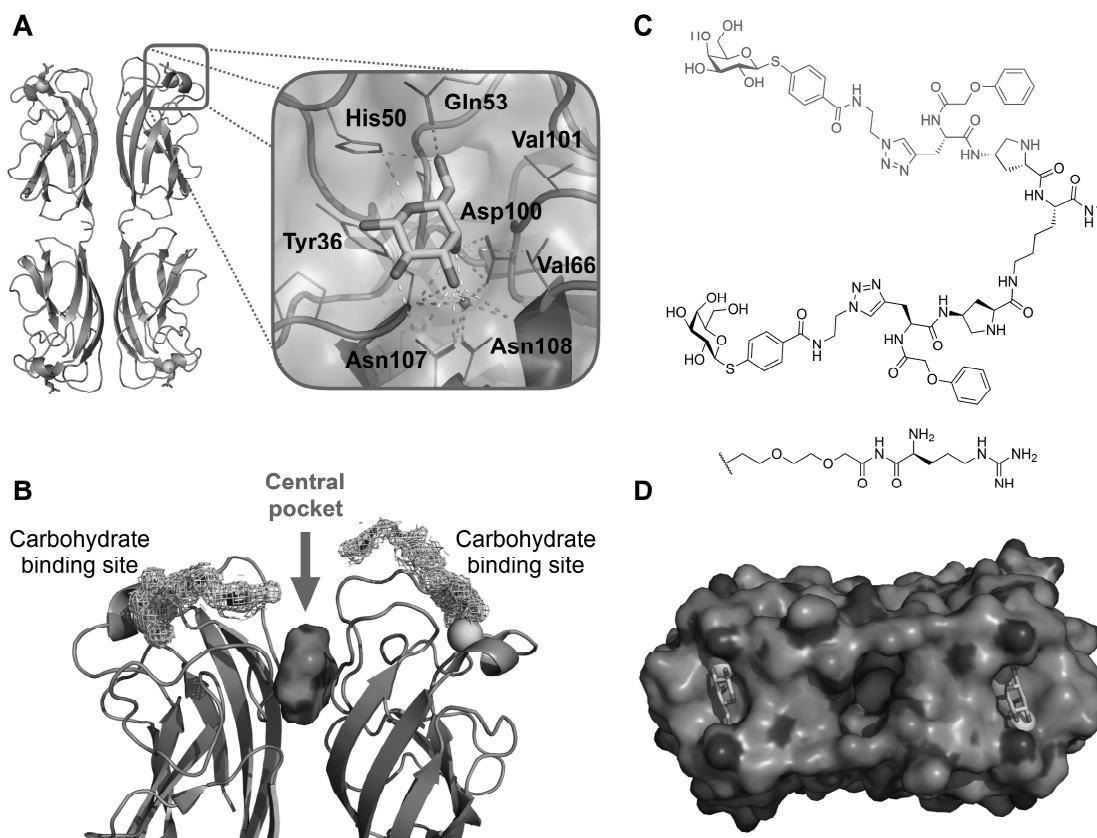


Figure 1. Rationale for targeting the central pocket in LecA: (A) Co-crystal structure showing the tetramer of LecA in complex with galactose (PDB: 1OKO). (B) Side view of one of the two adjacent dimers of LecA with the identified cavity (solid surface) between the two monomers defined as central pocket. The divalent ligand (PDB: 4CP9) reported by Winssinger *et al.*^[29] is represented as sticks in electron density and is pointing towards the central pocket. (C) Structure of the divalent LecA ligand.^[29] and the structural motif studied in this work highlighted in red. (D) Top view of the surface of one LecA dimer showing the cavity between the two monomers (PDB: 4LKE). The entrance of this cavity is polar due to the presence of Gln40, Lys41, Asp47, Arg48 and Glu49 and the interior is hydrophobic due to residues Trp33 and Trp42. Calcium ions in the carbohydrate binding sites are shown as green spheres.

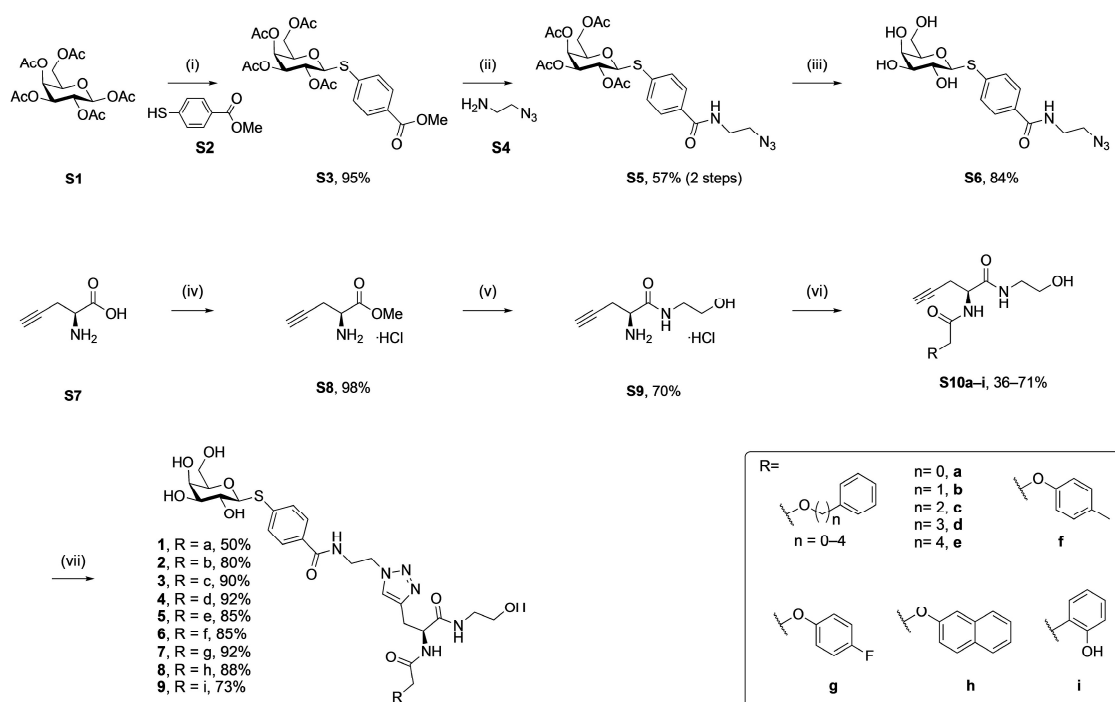
Design and *in silico* analysis by molecular dynamics simulation

As a starting point for ligand design, we were inspired by a bivalent peptide-based LecA inhibitor ($K_D = 82$ nM, Figure 1C) previously reported by Winssinger *et al.*^[29] As often observed for multivalent ligands, the crystal structure of the complex showed electron density only for part of the molecule, i.e. the galactose residue, the aromatic aglycon and the attached triazole ring which is in proximity to the entrance of the central pocket (Figure 1B). The authors reported the crystallographically invisible phenoxy acetyl substituent adjacent to the triazole to be important for LecA binding. We hypothesized that this phenoxyacetyl moiety may bind close to or into the central pocket. To test this hypothesis, we truncated the original divalent molecule and retained the galactose residue, its aglycon and the attached phenoxyacetyl residue (shown in red in Figure 1C) and designed derivatives carrying various substituents to target the central pocket *in silico*.

Compound **1** (Scheme 1) was then docked to LecA, and in the resulting poses, the phenoxy acetate part reaches towards the central pocket. Subsequently, molecular dynamics simulation was performed on one selected pose obtained from docking. Visual inspection of the 240 ns MD trajectory revealed significant fluctuation between the phenoxyacetate side chain and LecA with repeated contacts of the pharmacophore and

the central pocket (Figure 2). In general, the galactose moiety remained firmly bound inside the carbohydrate binding site with strong hydrogen bonds to residues His50, Gln53, Asp100 and Asn103, and hydrophobic contacts to Tyr36, Cys62 and Val101. The terminal phenyl ring was entering and leaving the central pocket repeatedly, possibly due to a suboptimal spacer length to position the phenyl group firmly within the pocket and a lack of interactions with the polar entrance of the cavity. Less frequent hydrogen bonds were also observed between the linker and Gln40 and Asp47. Last, hydrophobic contacts could be detected between the linker and Pro38, the phenoxy moiety and Trp42 (chains A and D), and the triazole ring and Pro51.

Based on these *in silico* considerations, we designed and synthesized derivatives **1–5** with stepwise increased linker length between the phenyl group and its ether oxygen to increase the possibility for hydrophobic contacts inside the central pocket. Furthermore, we modified the phenyl group in compounds **6–8** to assess the effect of increased lipophilicity and synthesized compound **9** carrying a hydroxy group to probe for hydrogen bonding with the cavity's polar entrance (see Scheme 1 for structures).



Scheme 1. Synthesis of LecA inhibitors targeting the central pocket. The different side chains were introduced in the penultimate step *via* amide coupling with alkyne **S9**. Final assembly was achieved by coupling azide **S6** with alkynes **S10a–i** in a copper(I)-catalyzed cycloaddition. Reagents and conditions: (i) $\text{BF}_3 \cdot \text{OEt}_2$, **S2**, CH_2Cl_2 , 0–25 °C, 18 h; (ii) 1. Lil, pyridine, 25 °C, 3 d; 2. **S4**, HOBT, EDC, DMF, 25 °C, 24 h; (iii) NaOMe, MeOH, 25 °C, 1.5 h; (iv) SOCl_2 , MeOH, 0–25 °C, 18 h; (v) ethanolamine, 25 °C, 18 h; (vi) various acetic acids, EDC, HOBT, DIPEA, DMF, 25 °C, 18 h; (vii) **S6**, CuSO_4 in H_2O , sodium ascorbate in H_2O , DMF, 25 °C, 2 h.

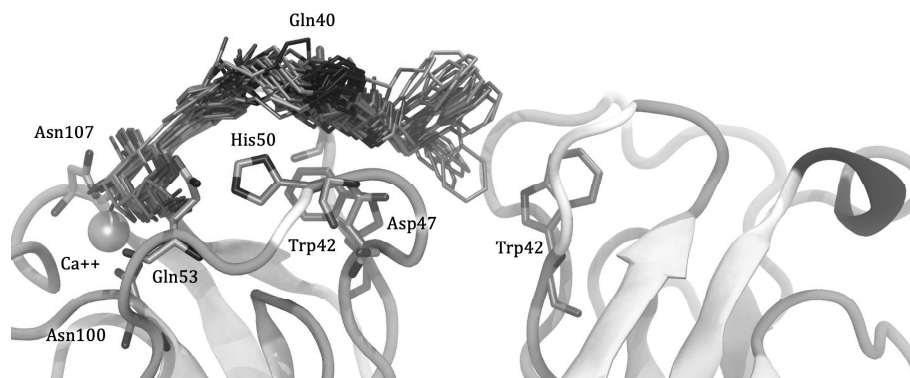


Figure 2. Molecular dynamics simulation of **1** in complex with LecA. Several snapshots of ligand **1** from the 240 ns molecular dynamic trajectory are depicted as cyan sticks. The phenoxy group of **1** is partially entering the central pocket of LecA and interacts with the Trp42 of both monomers (depicted as pink sticks). The galactosyl residue firmly coordinates the calcium ion (yellow sphere) in the carbohydrate binding site, and its β -phenyl aglycon binds to His50. Residues involved in hydrogen bonds to the galactosyl moiety are depicted by green sticks.

Synthesis

Galactose pentaacetate **S1** was reacted with thiol **S2** in presence of a Lewis acid to give β -thioglycoside **S3** in 95% yield (Scheme 1). Then, the methyl ester in **S3** was cleaved under S_N2 conditions using LiI to leave the acetates unchanged. The corresponding acid intermediate was then activated with EDC/HOBt and coupled with amine **S4** to give amide **S5** (57%, 2 steps). The latter was subsequently deprotected under Zemplén conditions to yield azide **S6**.

L-Propargylglycine **S7** was first transformed into ester **S8** using thionylchloride in MeOH and then stirred in neat ethanolamine to give amide **S9**. In the next step, the differently substituted acetic acids bearing the central pocket-targeting pharmacophores were introduced by amide coupling with alkynyl amine **S9** into amides **S10a–i** (36–71%). These alkynyl amides were then linked to azide **S6** in a copper(I)-catalyzed azide-alkyne cycloaddition reaction resulting in the final 9 triazoles **1–9** with varying side chains in high yields (50–92%).

Biophysical evaluation

All nine synthesized LecA ligands were then tested for LecA inhibition in a competitive binding assay based on fluorescence polarization^[25] and azide **S6** was included as a control devoid of the side chain targeting the central pocket. In addition, the two positive controls methyl α -D-galactoside (Me- α -D-Gal) and *para*-nitrophenyl β -D-galactoside (pNPG) were included (Figure 3A). Unsubstituted azide **S6** ($IC_{50} = 50.3 \pm 5.4 \mu M$) showed a fourfold higher activity than Me- α -D-Gal ($IC_{50} = 196 \pm 7.8 \mu M$). Phenoxy acetate **1** ($IC_{50} = 49 \pm 4.5 \mu M$), the closest derivative of the original divalent ligand, was as active as unsubstituted **S6**. Compounds **6** and **7** carrying electron donating or withdrawing groups in *para* position of the phenyl ring, as well as naphthyl **8** and phenol **9** did not show increased potency compared to **1** in this assay. Derivatives **3**, **4** and **5** with increased spacer length

by two to four methylene groups showed highest inhibition of LecA with IC_{50} s between 39 and 43 μM .

To further evaluate the affinity of the ligands towards LecA in a direct binding experiment, we performed surface plasmon resonance (SPR) analysis in which LecA was immobilized and compounds **1–9** were injected. Dissociation constants at steady state were calculated from multiple cycle analysis after injection of the tested compounds in 0–200 μM range (Figure 3C). The kinetic rate constants k_{on} and k_{off} could not be reliably determined due to fast association and dissociation of the compounds and LecA. All compounds showed affinities towards LecA with little variation between derivatives in the μM range (6.0 ± 0.2 to $10.3 \pm 1.3 \mu M$, Table 1) which is comparable to previously reported aromatic galactosides. The similar affinities observed for compounds carrying the additional aryloxy pharmacophores and that of **S6** suggest that the central pocket targeting pharmacophores in **1–9** either do not reach the intended site or the entropic penalties upon binding are compromising attractive interactions.

Isothermal titration calorimetry was used to obtain thermodynamic data on the LecA – inhibitor interaction with both binding partners in solution. Supporting the competitive binding assay and SPR, all tested compounds showed affinities in the low micromolar range in ITC (Table 1). Surprisingly, the unsubstituted ligand **S6** was the most active compound with a K_d of 4.8 μM . The triazole derivatives had K_d values between 6.3 and 9.4 μM , and binding enthalpies between -38.7 and -50.3 kJ/mol. Compound **2**, with an increased linker length of 1 methylene group compared to **1**, showed the highest affinity (K_d (**6**) = 6.3 μM) and a maximal enthalpy gain (ΔH (**2**) = -50.3 kJ mol $^{-1}$), followed by compound **8** with the longest linker of 4 methylene groups (K_d (**5**) = 6.7 μM , ΔH (**5**) = -46.3 kJ mol $^{-1}$). When comparing compounds **3**, **4** and **5**, that carry 2, 3 or 4 methylene groups respectively, a steady increase in binding enthalpy is observed which was counterbalanced by an unfavorable trend in entropy. Substitution in *para/ortho* position of the phenyl ring (**6**, **7**, **9**) or its replacement (**8**) did not lead to significant changes in binding affinity, except for

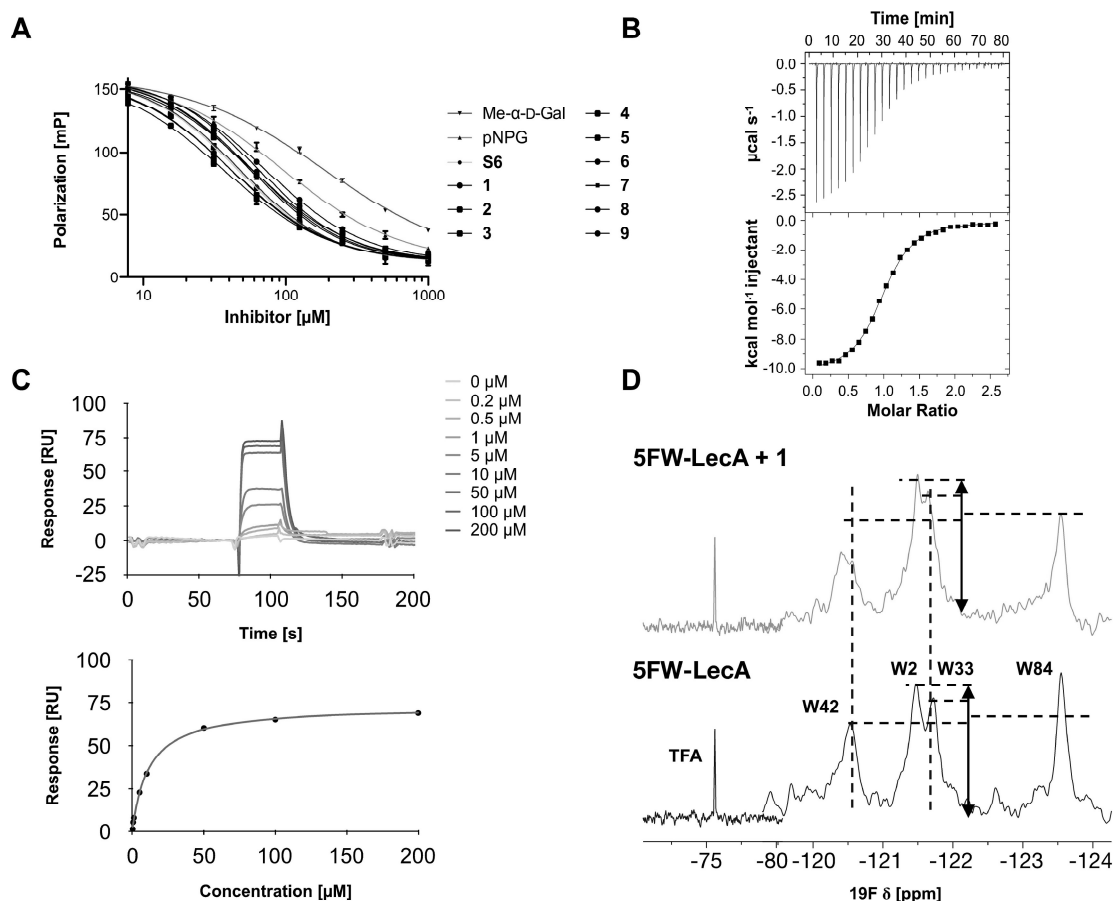


Figure 3. Analysis of synthetic inhibitor 1 interacting with LecA using (A) a competitive binding assay based on fluorescence polarization for all nine compounds including azide S6 and controls methyl α -D-galactoside and *para*-nitrophenyl β -D-galactopyranoside, (B) isothermal titration calorimetry sensorgram (top panel) obtained by titration of 1 to LecA with integration of peaks and fit (bottom panel), (C) surface plasmon resonance using multi-cycle kinetic studies (data shown for 1), top: sensorgram, bottom: affinity analysis, and (D) 19 F-protein-observed fluorine (PrOF) NMR demonstrates the impact of binding of 1 on the NMR resonances of Trp.

tolyl 6 that was the most active (K_d (6) = 6.8 μ M), with highest binding enthalpy (ΔH (6) = -45.1 kJ mol $^{-1}$) at the cost of the highest entropic penalty ($-\Delta S$ (6) = 14.6 kJ mol $^{-1}$) among this subset of compounds 6–9.

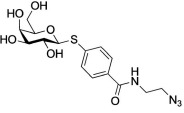
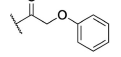
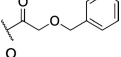
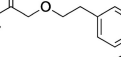
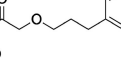
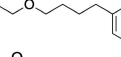
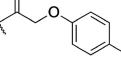
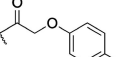
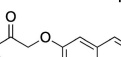
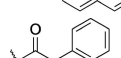
Protein observed 19 F (PrOF) NMR allows the detection of weak binding interactions with ligands. However, it requires the introduction of fluorine nuclei as sensitive NMR probes into the protein using 19 F-labelled amino acids or its precursors. Given the presence of tryptophans inside of the central pocket, we aimed to use PrOF NMR for detection of binders to this pocket.

To this end, all four tryptophans of LecA were simultaneously metabolically labeled at position 5 on the indole rings with fluorine and assignment was previously done by site-directed mutagenesis.^[31] Compounds 1, 2 or 8 were added to the labeled protein and PrOF NMR spectra were recorded and compared to those of the protein in absence of ligands (see Figure 3D). Trp42 is located in the carbohydrate binding site and its resonance was therefore affected by all tested ligands as observed by line broadening and chemical shift perturbation

(CSP₍₁₎ = 0.13 ppm, CSP₍₂₎ = 0.14 ppm and CSP₍₈₎ = 0.14 ppm respectively). The signals for Trp2 and Trp84 were not affected in any case while the signal intensity of the tryptophan located in the central pocket, Trp33, was shifted in presence of 1 (CSP₍₁₎ = 0.08 ppm), 2 (CSP₍₂₎ = 0.06 ppm) and 8 (CSP₍₈₎ = 0.07 ppm) (Figure S2), supporting a contact of the arylether moieties with the central pocket of LecA.

To analyze the interaction of ligands with the central pocket at atomic resolution, we crystallized and analyzed LecA in complex with 1. The resulting structure was determined at 1.53 Å resolution in space group $P2_12_12_1$ (Figure 1A, Table S3). The asymmetric unit consisted of a homotetramer of LecA, in agreement with previously reported LecA structures. One Ca $^{2+}$ -ion and one molecule of 1 could be located in each monomer, although the completeness of electron density of the ligand beyond the triazole ring varies among the four sites (Figure 4A). The galactoside moiety of 1 interacts with LecA in the same manner as previously reported, with hydrogen bonds to residues His50, Gln53, Asp100, Thr104, Asn107 and Asn108, and

Table 1. Analysis of LecA inhibitors **56** and **1–9** in direct binding (ITC, SPR) and competitive binding (FP) biophysical assays.

Structure	Name	ITC K_D [μM]	ΔH [kJ/mol]	$-T \Delta S$ [kJ/mol]	n	FP IC_{50} [μM]	SPR K_D [μM]
	56	4.8 ± 0.3	-48.9 ± 1.6	18.6 ± 0.5	0.9 ± 0.0	50.3 ± 5.4	6.7 ± 0.3
	1	9.4 ± 4.3	-46.2 ± 3.5	17.3 ± 1.5	1.0 ± 0.1	49.0 ± 4.5	10.3 ± 1.3
	2	6.3 ± 0.3	-50.3 ± 1.1	20.7 ± 0.3	1.0 ± 0.1	60.7 ± 2.5	6.0 ± 0.2
	3	8.3 ± 0.3	-41.0 ± 1.1	12.0 ± 0.4	1.0 ± 0.1	43.3 ± 3.2	6.9 ± 0.4
	4	8.6 ± 2.2	-43.9 ± 2.6	14.9 ± 0.8	1.0 ± 0.0	41.7 ± 3.3	7.2 ± 0.2
	5	6.7 ± 2.4	-46.3 ± 3.0	16.6 ± 1.0	0.8 ± 0.1	39.3 ± 2.6	6.7 ± 0.3
	6	6.8 ± 0.9	-45.1 ± 1.7	15.6 ± 0.6	0.8 ± 0.1	60.7 ± 6.8	7.7 ± 0.4
	7	8.6 ± 0.6	-43.5 ± 2.0	14.6 ± 0.7	1.0 ± 0.2	58.3 ± 4.0	7.8 ± 0.4
	8	7.6 ± 1.7	-38.7 ± 2.8	9.4 ± 0.8	1.1 ± 0.1	67.7 ± 4.5	6.4 ± 0.2
	9	8.8 ± 0.7	-41.4 ± 0.6	12.2 ± 0.1	0.9 ± 0.1	75.3 ± 3.4	9.8 ± 0.3

FP-data for controls: IC_{50} Me- α -D-Gal = $196 \pm 7.8 \mu\text{M}$; IC_{50} pNPG = $103 \pm 6.1 \mu\text{M}$.

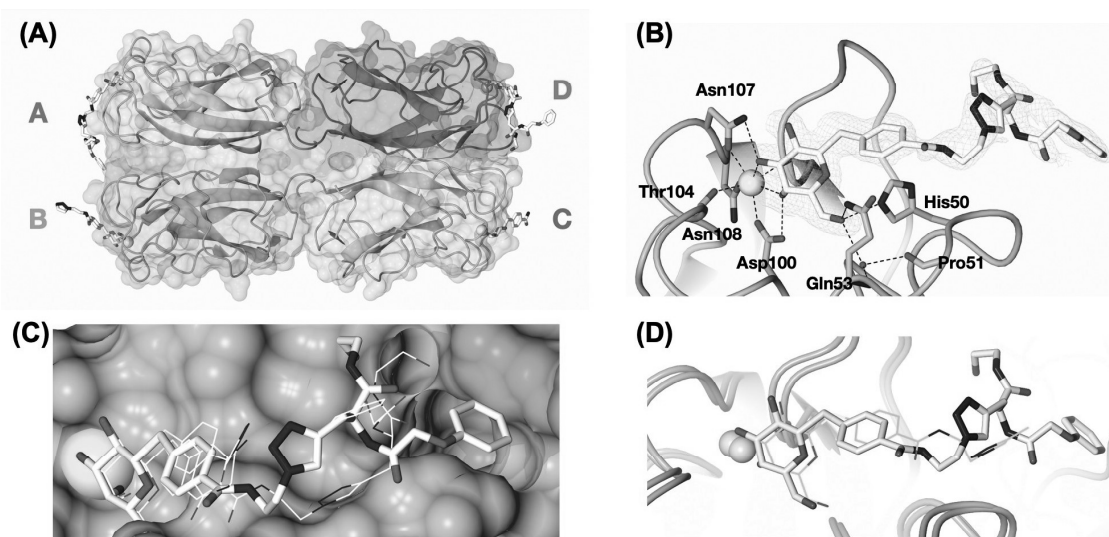


Figure 4. Crystal structure of LecA in complex with **1**. (A) The overall structure with **1** bound in all four monomers. (B) Binding pose of ligand **1** in site A with electron density displayed for the ligand. (C) Superposition of ligand conformations from all 4 binding sites. Ligand in site A is shown as sticks, whilst the others are shown as lines. (D) superposition of LecA structure in complex with **1** (protomer A) and structure in complex with peptide-based divalent ligand (PDB: 4CP9) reported by Winssinger *et al.*^[29] **1** is shown in white, thick stick and the visible part of the peptide-based ligand reported by Winssinger *et al.*^[29] in pink line.

coordination with the calcium ion.^[27] The aglycon forms CH- π interaction with His50 (Figure 4B) in the same way as previously reported for aromatic galactosides and LecA.^[21,32] However, superposition of the ligand in all four sites showed that the position of the phenyl ring varied slightly (Figure 4C).

The visible triazole rings in site A, B and D are surprisingly found in different positions and do not form productive contacts with the protein or water molecules (Figure 4B, C and Figure S4), which is in contrast to the previously reported divalent ligand with the similar phenyl aglycon and triazole design (Figure 1C),^[29] where the triazole ring interacted with the side chains of His50, Tyr36 and Asp47 through a water bridge. Superposition of LecA in complex with **1** and that in complex with the peptide-based divalent ligand reported by Winssinger *et al.*^[29] shows that the triazole ring oriented differently among the two ligands (Figure 4D). However, the conformation of our ligand **1** in protomer A is supporting our working hypothesis, where the terminal phenyl ring reaches the targeted central pocket. This data for the shortest derivative **1** also shows that an extension of the linker as implemented in molecules **2–6** could enable a better interaction with this pocket, which is consistently supported by our SPR and FP data for the extended molecules **3–5**.

Conclusions

We identified a central pocket localized between two monomers of the bacterial lectin LecA as a new target site for additional pharmacophores introduced into galactose-based inhibitors. Inhibitors were designed to bind simultaneously to the carbohydrate binding site and the central pocket. The nine synthesized inhibitors were analyzed for their interaction with LecA in the complementary biophysical assays such as ITC, SPR and a competitive binding assay. All compounds bound to LecA in the low micromolar range. Protein NMR data supported the interaction between the phenoxyacetate moiety with the central pocket as deduced from the CSP on the central pocket localized Trp33. Finally, the crystal structure of **1** in complex with LecA revealed that the additional aromatic moiety is oriented towards the central pocket. This is further supported by the increased affinity and binding enthalpy for the extended derivative **2** and the slight increase in binding affinity for the longer spacer containing compounds **3–5** suggests their interaction with the central pocket.

One fundamental question is raised by the fact that such a galactose-based inhibitor targeting the central pocket may block the single central pocket between two C2-symmetry related galactose binding sites and prevent efficient binding of a second copy of the ligand bound to the other galactose site. This question may be solved by future bivalent galactosides carrying an appropriately positioned single pharmacophore addressing the central pocket.

Experimental Section

In silico crystal structure analysis: Twenty three crystallographic structures were retrieved from the Protein Data Bank^[30] (PDB) (Table S1) and checked for the presence of the cavity termed here as central pocket. Cavity detection was performed using VolSite.^[33] VolSite detects the shape and pharmacophore properties e.g. H-Bond donor/acceptor, hydrophobic and aromatic points of protein cavities and gives details about its buriedness, volume and a drugability score. The central pocket cavity was detected in 25 out of the 31 dimers that are available as crystallographic structures on PDB (Table S2).

Docking of **1 and molecular dynamics simulation:** We used two crystal structures for docking. The structure of Winssinger's divalent LecA ligand (PDB code: 4cp9) co-crystallized with LecA contains parts of the ligand **1** and was used to guide the docking of the whole ligand **1**. It was further redocked on a LecA dimer structure with the central pocket in a more open conformation (PDB code: 4lke). Hydrogens were added to the protein structures using PROTONS.^[34] A 3D model of **1** carrying the phenoxy acetate side chain was prepared using CORINA.^[35–37] Compound **1** was docked with the placed fragments method from Surflex.^[38]

The selected docked conformation for the complex was parameterized according to the General Amber Force Field 2 (GAFF2)^[39] using AmberTools 16.^[40] RESP charges were fitted from the electrostatic potential and optimized at the HF/6-31G* level of theory by using Gaussian 09.^[41] The protein was prepared according to the ff14SB forcefield,^[42] using *tleap*. Ions (K^+ , Cl^- and Ca^{2+}) were used to neutralize the complex. PBradii was set to mbondi3. The system was solvated with TIP3P^[43] by using water molecules in an octahedral box, under periodic boundary conditions, with a distance of 12.0 Å between the protein and each face of the box. The systems were minimized in three steps by using the steepest descent, followed by a conjugate gradient after 2000 cycles: (i) position restraints for all heavy atoms of the complex (weight of 10 kcal/mol-Å²) (ii) position restraints for the atoms of the backbone of LecA, (weight of 10 kcal/mol-Å²); (iii) weak position restraints for the backbone atoms (weight of 5 kcal/mol-Å²). After minimization, the system equilibration was performed on the solvent/ions in three steps: (i) the system was gradually heated from 30 to 100 K for 50 ps under the NVT ensemble with weak position restraints for the heavy atoms of LecA, by using a weight of 5 kcal/mol-Å²; (ii) heating from 100 to 298.15 K for 250 ps under the NPT ensemble with weak position restraints for the heavy atoms of the protein using a weight of 5 kcal/mol-Å²; (iii) MD under the NPT ensemble at constant temperature of 298.15 K for 1 ns with very weak position restraints for the protein heavy atoms, using a weight of 0.1 kcal/mol-Å².

The temperature was set to at 300 K using the Langevin dynamics with a collision frequency of 2 ps⁻¹. The pressure was kept constant using the Monte Carlo barostat with a relaxation time of 2 ps. SHAKE^[44] was used to control bonds length involving hydrogen atoms. The production runs, under the NPT ensemble, consisted in one replica of ~240 ns and three independent replicas of 30 ns each, to verify the influence of sampling on ligand **1** conformations.

Visualization and hydrogen bond calculations for the trajectories were done using VMD.^[45] Hydrophobic contacts were calculated using the cpptraj module available with AmberTools with the native contacts routine using a distance cutoff of 5 Å and a filter selection on non-polar amino acids. The short trajectories and the long MD yielded similar results regarding the mentioned analyzes. Images were rendered using VMD and UCSF-Chimera.^[46]

Recombinant expression and purification of LecA: LecA expression and purification was performed as previously described.^[20] The expression strain *E. coli* BL21(DE3) carrying the pET25pa11^[47] plasmid

was grown in 4 L LB containing ampicillin (100 $\mu\text{g/mL}$) at 37 °C and 180 rpm until an $\text{OD}_{600} \approx 0.5$ was reached. For induction of protein expression IPTG (0.25 mM) was added, and the culture was grown for additional 4 h at 30 °C. Bacterial cells were harvested by centrifugation (3000 g, 10 min, 4 °C) and the cell pellet was resuspended in TBS/ Ca^{2+} -buffer (1 mM CaCl_2 , 150 mM NaCl, 20 mM TRIS, 2.5 mM KCl, pH = 7.4) and PMSF (1 mM) and lysozyme (0.4 mg/mL) were added. Cell lysis was performed by 5 cycles in a homogenizer (M-110P, Microfluidics, USA), debris was removed by centrifugation (60 min, 10000 g, 4 °C), and LecA from the supernatant was purified on a galactosylated Sepharose CL-6B column using an Äkta start chromatography device. After washing with TBS/ Ca^{2+} -buffer, bound LecA was eluted with 100 mM D-galactose in buffer and then dialyzed against fresh TBS/ Ca^{2+} -buffer every day for 7 days to remove galactose. Protein concentration was determined by UV spectroscopy ($\epsilon = 27960 \text{ M}^{-1} \text{ cm}^{-1}$, molecular weight 12893 g mol^{-1} , ExPASy ProtParam). LecA used for ITC was dialyzed against ddH₂O for 7 days and lyophilized afterwards prior to dissolving in the ITC buffer.

Fluorescence polarization assay: This assay was performed in analogy to Joachim *et al.*^[20] using SulfoCy5Gal, an adapted Cy5 dye.^[25] 10 μL of a solution containing SulfoCy5Gal (20 nM) and LecA (40 μM) in TBS/ Ca^{2+} -buffer were added to 10 μL of inhibitor in the same buffer at concentrations from 2000–15.6 μM and containing 2 % DMSO in a black 384-well plate (cat no 781900, Greiner Bio-One, Germany) in technical triplicates. Methyl α -D-galactopyranoside and 4-nitrophenyl β -D-galactopyranoside were included as positive controls. The plate was sealed (EASYseal, cat no 676001, Greiner Bio-One), centrifuged (1500 \times g, 1 min, 25 °C) and incubated in a dark chamber under shaking conditions for 16 h at r.t. The foil was removed and the fluorescence intensity was measured with a PheraStar FS microplate reader (BMG Labtech GmbH, Germany) at ex. 590 nm and em. 675 nm. Data were analyzed using MARS Data Analysis Software (BMG Labtech GmbH, Germany) after subtracting blank values (LecA in TBS/ Ca^{2+} -buffer with 1 % DMSO) from the samples. Fluorescence polarization was calculated and the data were fitted according to the four-parameter variable slope model. This experiment was independently repeated three times and data were averaged and visualized using GraphPad PRISM version 5.

Isothermal titration calorimetry (ITC): Galactoside (1.0 or 1.5 mM) in TBS/ Ca^{2+} -buffer was titrated into a stirred (700 rpm) LecA solution (75–232 μM) dissolved in the same buffer using a MicroCal ITC₂₀₀ (Malvern, United Kingdom) instrument at 25 °C. The reference power was set to 5 $\mu\text{cal s}^{-1}$, the filter period to 5 s, and 20–39 injections (0.5–2 μL per injection) with an injection duration of 1 s were performed per experiment with a spacing of 240 s between each injection. For some titrations (in case of compounds 2 (1x), 4 (2x) and 8 (2x)), the syringe was refilled after the first titration ended, and the experiment was continued with the same sample cell contents to reach saturation, and the resulting files were merged using the MicroCal Concat ITC software. The first injection of each experiment was discarded, and the data were analyzed with the MicroCal Origin software using the one-site binding model. Individual titrations are depicted in Figure S1.

Protein-observed ^{19}F (ProF) NMR: Labelled 5FW-LecA for ProF NMR studies was produced recombinantly as reported previously.^[31] NMR experiments were conducted on a Bruker AscendTM700 (AvanceIII HD) spectrometer equipped with a 5 mm TCI700 CryoProbe in 3 mm tubes (Norell S-3-800-7). ProF NMR was recorded using 200 μM 5FW-LecA in 20 mM Tris-HCl pH 7.8 with 150 mM NaCl, 10 % D₂O and 100 μM TFA at 310 K. Changes in the spectra in presence of DMSO and upon addition of 3 mM ligand were processed and referenced to trifluoroacetic acid (TFA) as internal reference at -75.6 ppm . We considered only changes in chemical shift perturbation (CSP) upon ligand addition being two-fold greater than standard deviation (2x std. dev., 0.03 ppm) of the fluorine resonance.

Surface plasmon resonance (SPR): All experiments were performed on a BIACORE X100 at 25 °C. For activation and immobilization of LecA, buffer A (10 mM phosphate buffer + 2.7 mM KCl, 137 mM NaCl, 0.05 % Tween 20 + 100 μM CaCl_2) was used as a running buffer. A CM5 BIACORE chip was activated by 3 injections of 1:1 NHS/EDC mixture (contact time = 540 s, flow rate = 10 $\mu\text{L/min}$) on channel 1 and 2, followed by injections of LecA (100 $\mu\text{g/mL}$ in 10 mM sodium acetate buffer pH 4.5) on channel 2 only until the final binding response reached 2939 RU. Compounds were dissolved in DMSO to the final concentration of 100 mM and subsequently diluted to 5 mM in buffer A. The compounds were then prepared to the required analytic concentrations (0.2–200 μM) in buffer A supplemented with 5 % DMSO. For multi-cycle kinetic analysis of the compounds, buffer A supplemented with 5 % DMSO was used as a running buffer. The compounds were analyzed by injections of multiple concentrations on the immobilized LecA (contact time = 30 s, dissociation time = 60 s, flow rate 30 $\mu\text{L/min}$) and affinity analysis was performed using BIACORE evaluation software, which plots the binding response at the steady state against the analyzed concentrations and fits a non-linear curve to obtain K_d values. Individual experiments are depicted in Figure S3.

X-ray crystallography: Lyophilized LecA was dissolved to saturation in PBS at pH 7.4, containing 100 μM CaCl_2 . 1 μL of the protein solution was mixed with 1 μL of the reservoir solution (20 % PEG6000, 1 M LiCl, 100 mM sodium acetate pH 4.2, 5 % DMSO containing 1 mM of compound 1 and the mixture was deposited on a siliconized glass circle cover slide (22 mm, Hampton research). Crystallization was performed by hanging drop vapor diffusion at 19 °C. Protein crystals were cryo-protected in 20 % PEG6000, 1 M LiCl, 100 mM sodium acetate pH 4.2 supplemented with 20 % ethylene glycol and flash frozen in liquid nitrogen. Data collection was conducted at SOLEIL PROXIMA 1 beamline (Saint Aubin, France). The recorded data were indexed, integrated and scaled at SOLEIL using XDS,^[48] and merged using AIMLESS.^[49] The structure was solved by molecular replacement in PHASER^[50] using 1OKO as a searching template. The model was improved by manual rebuilding in COOT^[51] and refinement in REFMAC5.^[52] The model of compound 1 was manually built in AceDRG^[53] and manually placed in the protein model. The final model was validated with MolProbity,^[54] PDB-redo^[55] and wwPDB validation server (<http://validate-rcsb-1.wwpdb.org>). Structural figures were prepared using CCP4MG.^[56] Data merging, phasing and model re-building and refinement were performed through CCP4i2^[57] graphical interface.

Supporting Information

The supporting information contains the synthesis of LecA inhibitors and ^1H - and ^{13}C -NMR spectra of new compounds, the tables of X-ray structures used for the analysis and description of the carbohydrate binding site and the central pocket, ITC and ProF NMR raw data, the sensorgrams and affinity analyses of the SPR data as well as the X-ray structure of compound 1 with focus of each binding center of the LecA tetramer.

Acknowledgements

The authors acknowledge the financial support of the French-German ANR/DFG project (ANR-AAPG-2017) funded by the Agence Nationale de la Recherche (grant no. ANR-17-CE11-0048) and Deutsche Forschungsgemeinschaft (grant no. Ti756/5-1 and RA1944/

Discovery of *N*- β -L-Fucosyl Amides as High-Affinity Ligands for the *Pseudomonas aeruginosa* Lectin LecB

Patrycja Mała^{#1,2}, Eike Siebs^{#1,3,4}, Joscha Meiers^{1,3,4}, Katharina Rox^{4,5}, Annabelle Varrot⁶, Anne Imberty⁶, Alexander Titz^{1,3,4*}

¹Chemical Biology of Carbohydrates (CBCH), Helmholtz-Institute for Pharmaceutical Research Saarland (HIPS), Helmholtz Centre for Infection Research, 66123 Saarbrücken, Germany;

²Faculty of Chemistry, Adam Mickiewicz University, 61-614 Poznań, Poland;

³Department of Chemistry, Saarland University, 66123 Saarbrücken, Germany;

⁴Deutsches Zentrum für Infektionsforschung (DZIF), Standort 38124 Braunschweig, Germany;

⁵Chemical Biology (CBIO), Helmholtz Centre for Infection Research, 38124 Braunschweig, Germany;

⁶Univ. Grenoble Alpes, CNRS, CERMAV, 38000 Grenoble, France.

[#]These authors contributed equally to this work.

*corresponding author e-mail: alexander.titz@helmholtz-hzi.de

Keywords: Lectin, carbohydrates, fucose, glycomimetics, LecB

Abstract

Pseudomonas aeruginosa is listed by the World Health Organization (WHO) as the most critical Gram-negative pathogen resisting antimicrobial treatment. It causes severe infections mainly in immuno-compromised or cystic fibrosis patients, which are difficult to treat due to its ability to form biofilms that protect the bacteria from antibiotics. One of the key players in bacterial adhesion to the host and biofilm formation is the lectin LecB, an extracellular protein that stabilizes the biofilm matrix. For the inhibition of LecB, we designed and synthesized a set of fucosyl amides, sulfonamides and thiourea derivatives. Then, we analyzed their binding to LecB in competitive and direct binding assays. We identified β -fucosyl amides as high-affinity ligands with unprecedented affinity for LecB in the two-digit nanomolar range. The molecules further showed good stability in murine and human blood plasma and hepatic metabolism, providing a basis for future development into antibacterial drugs. Finally, X-ray crystallography of an α - and a β -anomer of *N*-fucosyl amides in complex with LecB revealed the interactions responsible for the high affinity of the β -anomer at atomic level.

Introduction

Antimicrobial resistance is a rapidly developing threat to humanity.^[114] The Gram-negative bacterium *P. aeruginosa* belongs to the problematic ESKAPE panel and is listed as the most critical drug-resistant bacterial pathogen by the WHO.^[69] It is a threat to people with cystic fibrosis (CF) or chronic obstructive pulmonary disease (COPD) and to hospitalized immunocompromised patients.^[39] This bacterium can form biofilms which renders standard-of-care antibiotics orders of magnitude less effective.^[74] Moreover, many *P. aeruginosa* strains have become multi-drug resistant^[16,115] and several approaches to combat this problem are in the pipeline.^[39] An alternative strategy to antibiotics are antivirulence agents or pathoblockers, that instead of killing aim at disarming the bacteria in order to neutralize bacterial virulence and thereby provide protection to the host.^[39,116]

As the major resistance mechanism, the biofilm matrix hinders penetration of antibiotics and in addition, the embedded bacteria also reduce their metabolic activity rendering them persistent to treatment. One promising therapeutic option is thus to interfere with biofilm formation to restore antibiotic efficacy and also to provide access

to the bacteria for the immune system. For establishing the biofilm, *P. aeruginosa* expresses two extracellular lectins LecA and LecB. Both proteins are crucial for initial cell adhesion. They are also essential constituents of the biofilm matrix^[49,50] and they establish interactions by binding to carbohydrate epitopes of the bacterial exopolysaccharides as well as to the bacterial and the host glycocalyx.^[47,54,117] The sequences of LecA and LecB have been analyzed for various clinical and environmental isolates, demonstrating the lectins' functional conservation.^[55,60]

The binding of LecB to the biofilm matrix exopolysaccharide Psl^[54] and its requirement for mature biofilm formation was demonstrated.^[50] Furthermore, LecA is involved in host cell invasion of *P. aeruginosa* by binding to the glycosphingolipid Gb3.^[118] In that process, LecA induces phosphorylation of the adaptor protein CrkII that mediates signaling across the host's plasma membrane which most likely assists the membrane engulfment.^[75] Infection experiments using *lecA* or *lecB* knockout strains revealed improved lung bacterial clearance in mice and better epithelial wound healing when compared to the corresponding wildtype bacteria.^[117,119]

A study on cystic fibrosis patients with chronic *P. aeruginosa* infections showed that inhalation of a L-fucose/ D-galactose solution reduced the amount of bacteria in sputum.^[40] In another study on *P. aeruginosa* lung infected mice, it could be demonstrated that administration of carbohydrates in combinations with antibiotics reduced the bacterial burden more efficiently than single treatments of antibiotics.^[77] Therefore, the synthesis of LecA and LecB inhibitors as novel anti-infectives is an active field of research.^[39,81,120]

LecA forms homotetramers and binds to D-galactosides via a calcium ion in its carbohydrate binding site.^[121] Several monovalent galactosides were synthesized as LecA inhibitors reaching moderate binding in the micromolar range.^[102,103,122] In contrast, divalent galactoside inhibitors with simultaneous binding to two adjacent carbohydrate binding sites of the LecA tetramer gave low nanomolar inhibitors.^[81,85,87,123] Novel concepts were also reported such as the development of covalent lectin inhibitors^[89], addressing a subpocket between the two adjacent carbohydrate binding sites^[90], or the development of non-carbohydrate glycomimetics^[91,92].

LecB also forms homotetramers and possesses two calcium ions per carbohydrate binding site mediating binding to its fucoside or mannoside ligands.^[124] The affinity of

fucosides is increased compared to mannosides due to an additional lipophilic interaction of the fucose C6 methyl group with the protein at Thr45, resulting in sub-micromolar binding (**Figure 6A**, K_d of Me- α -L-Fuc = 0.43 μ M^[53], K_d of Me- α -D-Man = 71 μ M^[53]). Furthermore, it was demonstrated that the CH₂OH group of mannosides adopts a sterically hindered position when bound to LecB.^[125] However, the co-crystal structure of fucose in complex with LecB^[124] revealed a subpocket next to the anomeric center that was subsequently addressed in our program for LecB inhibitors (**Figure 6A**, compound **II**).^[78,80,126,127] A small cleft between the carbohydrate binding site and said subpocket is surrounded by amino acids Ser22 and Asp96. The subpocket itself differs slightly between the two *P. aeruginosa* strains.^[60]

Consequently, we have designed fucose-mannose hybrid glycomimetics^[79,80,125] combining the properties of our first set of D-mannose-derived inhibitors^[78] with L-fucose resulting in sub-micromolar affinities (K_d (PAO1) (**I**) = 0.83 μ M^[79], K_d (PA14) (**I**) = 0.29 μ M^[79], **Figure 6B**), oral bioavailability in mice and good antibiofilm activity *in vitro*. Our initial mannosyl amides and sulfonamides as well as fucose-mannose hybrid molecules targeted this subpocket and the interactions could be explained on the basis of several X-ray LecB structures for both strains.^[78–80] We demonstrated that the mannose sulfonamides inhibited LecB better than the corresponding amides (IC_{50} (PAO1) (**III**) = 110 μ M, IC_{50} (PA14) (**III**) = 42 μ M, IC_{50} (PAO1) (**IV**) = 16 μ M, IC_{50} (PA14) (**IV**) = 3.3 μ M).^[78,80] Due to the different geometry in the sulfonamide moiety, these molecules can circumvent a steric clash of the amides with Ser97 on LecB_{PA14}.^[79] However, the fucose-mannose hybrid amides were as active as the mannose sulfonamides (IC_{50} (PAO1) (**V**) = 8.7 μ M, IC_{50} (PA14) (**V**) = 3.5 μ M) due to hydrophobic interactions of their methyl groups (C6) with Thr45.^[79] Shifting the amide/sulfonamide linker function further towards this subpocket by using an elongated heptose derivative resulted in a loss in affinity for LecB_{PAO1} ($IC_{50} \geq 82 \mu$ M).^[126]

In the present work, we further assessed the positioning and nature of the linking unit between the carbohydrate and the pharmacophore targeting the additional subpocket in LecB. To this end, we shortened the linking function in those molecules by removing a methylene group and combined the fucose pharmacophore with amide and sulfonamide functions, and furthermore, replaced the previously reported linkers in the hybrid-type molecules with thioureas.

Results and Discussion

Design of Fucosyl Amides and Sulfonamides, Fucosylmethyl Thioureas

To study the influence of the linking units between fucose and aromatic pharmacophores targeting the subpocket in LecB, we replaced the known amide and sulfonamide linkers of previous hybrid-type molecules **I** and **V**. First, we modified the chemical nature of the linker and introduced thioureas providing hydrogen-bond donor/acceptor properties resulting in molecules such as **11c** (**Figure 6C**) with altered geometry of the linking unit and elongation. Then, we shortened the molecules into analogs devoid of the methylene group in compounds **I** and **V** and designed *N*-fucosides of amides **4** and sulfonamides **8** to assess hydrogen bond formation with the amino acids such as Ser22 (**Figure S5**).

These molecules were first docked *in silico* into LecB. For fucosylmethyl *p*-tolylthiourea (**11c**), the carbohydrate moiety superimposes with fucose bound to LecB (**Figure S5**), while the thiourea forms hydrogen bonds with the backbone of Asp96 and with its sidechain. Interestingly, the software docked the thiourea function in its tautomeric thiol form, although the thione is dominating in aqueous solutions.^[128] The docked binding pose of β -fucosyl benzamide (**4a**) showed again an identical orientation of the fucose, while its amide serves as hydrogen bond acceptor for the sidechain of Ser22. In addition, the aromatic ring forms lipophilic contacts with Gly24 and Val69 of the adjacent pocket (**Figure S5**).

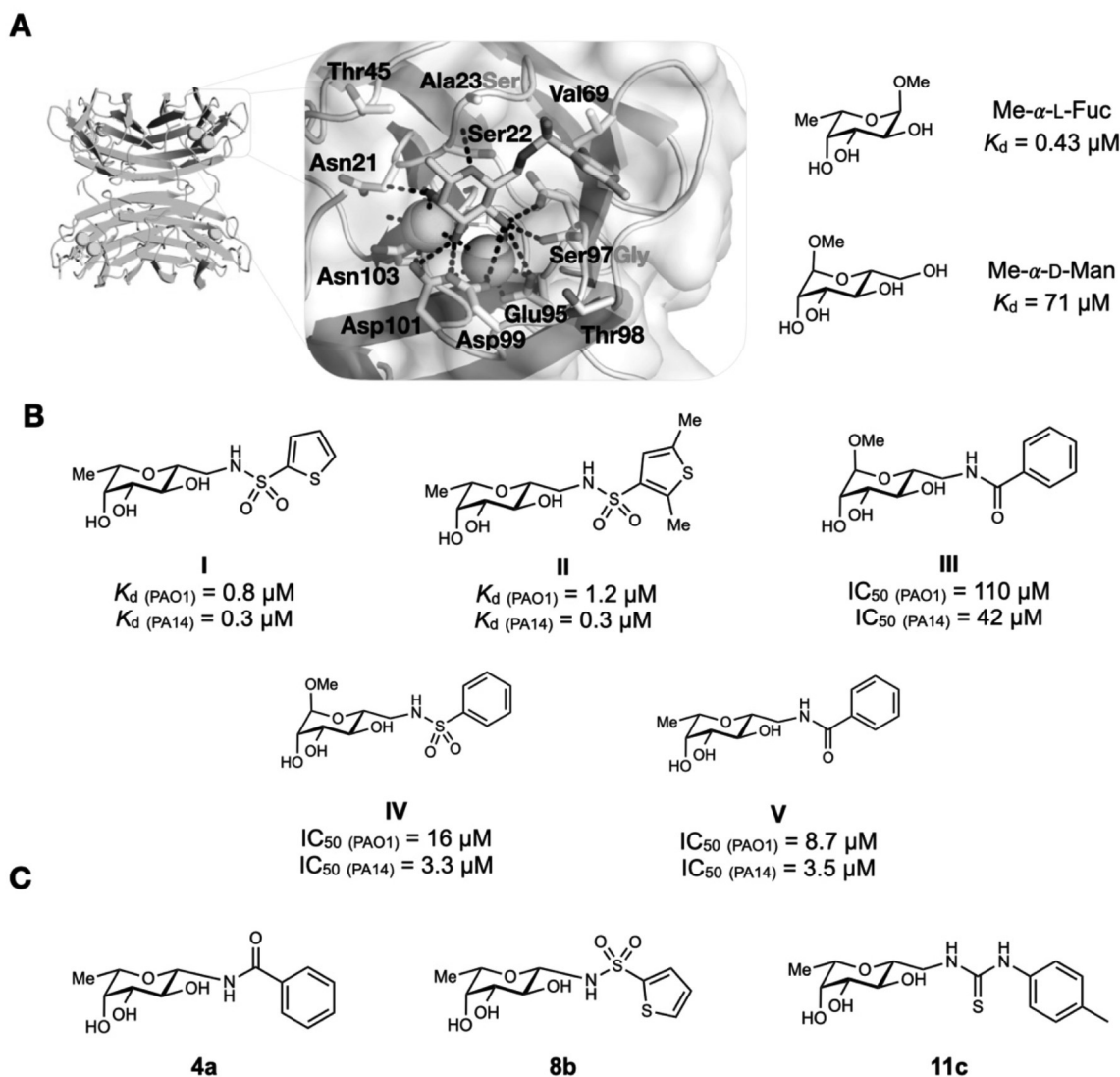


Figure 6: **A)** Crystal structure of LecB_{PA14} in complex with hybrid **II** (pdb: 5MAZ^[79]) with overall tetramer and zoom into the carbohydrate binding site (green spheres: calcium ions, red spheres: waters, red: oxygens, blue: nitrogen, pink: amino acid variations in LecB_{PA01}); **B)** Previous low-molecular weight LecB inhibitors **I**^[80], **II**, **V**^[79], and **III–IV**^[78,80]. **C)** Designed LecB inhibitors *N*-fucosyl amides (e.g. **4a**), *N*-fucosyl sulfonamides (e.g. **8b**) and fucosylmethyl thioureas (e.g. **11c**) for this work.

Synthesis of Amides, Sulfonamides and Thioureas

β -Fucosyl amides were obtained in a three-step synthesis from fucose tetraacetate **1** (**Scheme 1**). Transformation of the tetraacetate into azide **2** was achieved in good yield (77%). This azide was further converted in a STAUDINGER reduction followed by acylation with acyl chlorides to the corresponding protected β -fucosyl amides **3a–n** in yields of 12–71%. The target amides **4a–n** were then obtained after deacetylation under ZEMPLÉN conditions (47–99%). Furthermore, one α -fucosyl amide, benzamide **6**,

was also synthesized to serve as a control molecule. **6** was obtained over two steps commencing from azide **2** by an activation with triphenylphosphine under reflux to form the α -oxazoline intermediate, which was coupled with a thiopyridyl ester of benzoic acid using DE SHONG^[129] conditions to form the protected α -fucosyl amide **5** in 27% yield. The latter compound was finally deprotected to give the α -anomer **6** in 98% yield.

Two representative examples of fucosyl sulfonamides **8a–b** were synthesized by *N*-glycosylation of the respective sulfonamides. To this end, the sulfonamide acceptors were treated with tetraacetate **1** as donor under Lewis acid catalysis and the pure β -glycosides **7a–b** were obtained in good yields. Unfortunately, the subsequent deprotection step using sodium methoxide inevitably resulted in anomerization and the test compounds were obtained as anomeric mixtures.

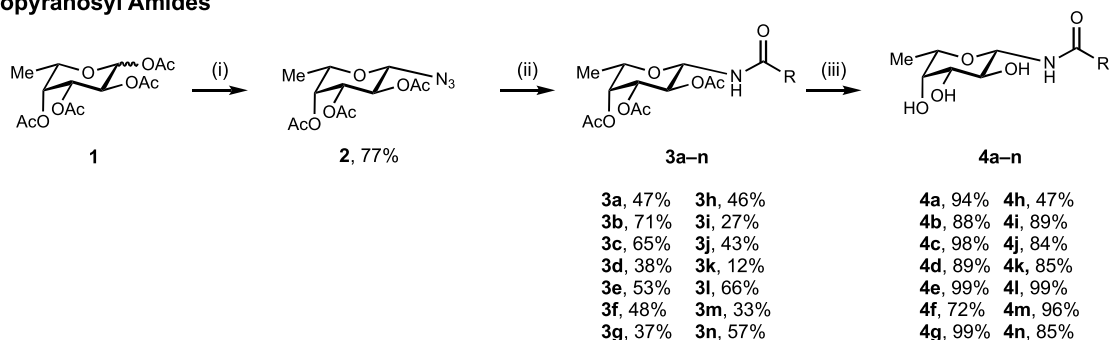
The β -fucosylmethyl thioureas **11c–e, o, p** were obtained in a three step synthesis from L-fucose. Optimized HENRY^[130] reaction conditions of fucose and nitromethane were used to obtain condensation product **9**. After reduction of nitro **9** to the amine **10**, the latter was reacted with various isothiocyanates to give the thioureas **11c–e, o, p** in good yields (71–80%).

In total, 22 fucose derivatives were synthesized (**Scheme 1**), among which are 14 β -fucosyl amides, one α -fucosyl amide, two fucosyl sulfonamides obtained as α/β mixtures, and five β -fucosylmethyl thioureas.

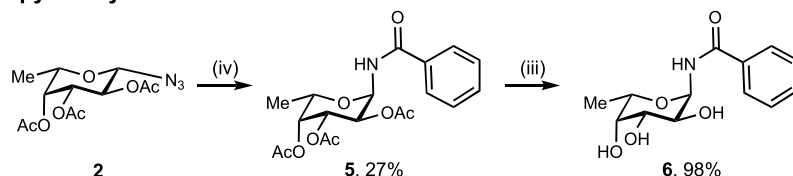
Evaluation of LecB Binding in Biophysical Assays

All ligands were then tested in a competitive binding assay^[78] based on fluorescence polarization for dose-dependent inhibition of LecB_{PAO1}. Due to obtained high affinities of the β -fucosyl amides, the assay was slightly modified and a lower LecB concentrations of 75 nM was used (**Table 4**). Therefore, the obtained data for L-fucose was also slightly lower with an $IC_{50} = 1.35 \pm 0.04 \mu M$ compared to the reported value of $2.74 \mu M$ ^[78].

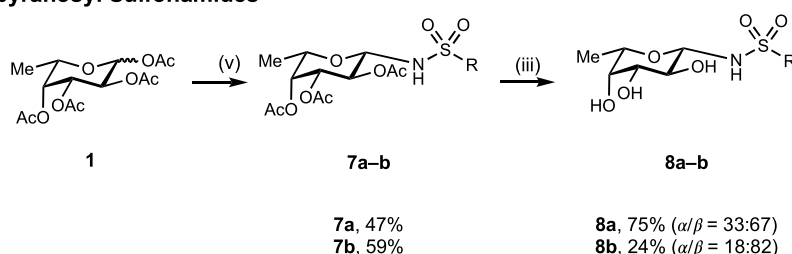
β -Fucopyranosyl Amides



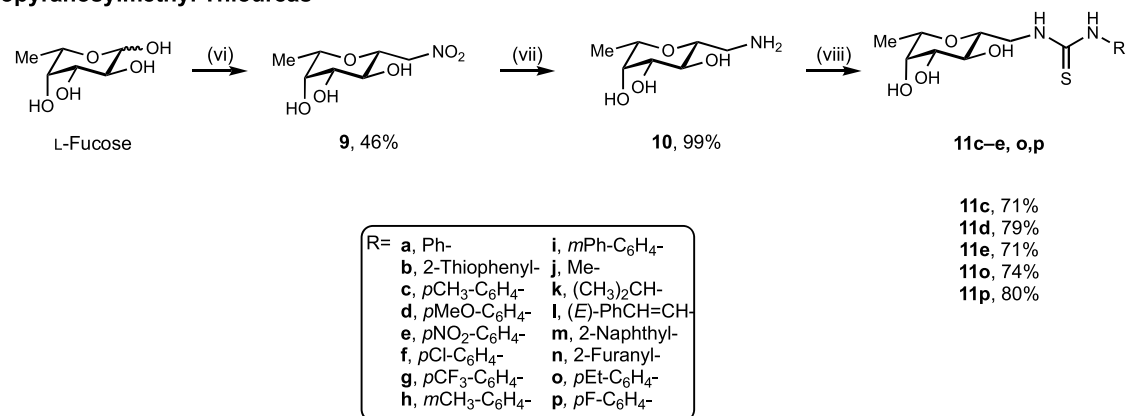
α -Fucopyranosyl Amide



Fucopyranosyl Sulfonamides



β -Fucopyranosylmethyl Thioureas



Scheme 1: Synthesis of LecB ligands β -fucosyl amides (**4a–n**), α -fucosyl amide (**6**), fucosyl sulfonamides (**8a–b**) and β -fucosylmethyl thioureas (**11c,d,e,o,p**). Reagents and conditions: (i) Me₃SiN₃, SnCl₄, CH₂Cl₂, 25 °C, 1.5 h; (ii) RCOCl, PPh₃, Et₃N, CH₂Cl₂, 0–25 °C, o.n.; (iii) NaOMe, MeOH, -25 to -15 °C or 0 °C, o.n. or 1.5 h; (iv) 1. PPh₃, MeNO₂, 4 Å molecular sieves, reflux, 24 h; 2. S-(pyridin-2-yl) benzothioate, CuCl₂·H₂O; (v) phenylsulfonamide or thiophene-2-yl-sulfonamide, BF₃·OEt₂, MeCN, 25 °C, 24 h; (vi) 1. MeNO₂, NaOMe (cat.), DMSO, 25 °C, 6 h; 2. HCl (1 M, pH = 4), H₂O, reflux, o.n.; (vii) Pt/C, H₂, MeOH, 25 °C, 48 h; (viii) isothiocyanates, MeOH, 0–25 °C, o.n.

In this assay, all tested β -fucosyl amides inhibited LecB in the nanomolar range. Acetamide **4j** showed the weakest inhibition among the series giving an IC₅₀ of 902 ±

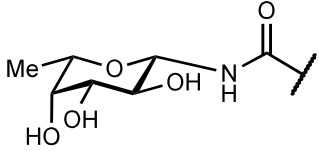
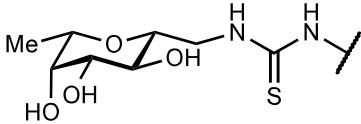
69 nM. The affinity increased among this compound class when the acetamide was replaced with larger substituents: replacing its methyl with aromatic rings such as 2-furanosyl (IC_{50} (**11**) = 272 ± 26 nM), 2-thiophenyl (IC_{50} (**4b**) = 122 ± 21 nM) or a phenyl ring increased the affinity towards LecB up to tenfold into the two-digit nanomolar range (IC_{50} (**4a**) = 88 ± 12 nM), which rendered fucosyl benzamide (**4a**) 15-fold more potent than L-fucose. Introducing electron-donating substituents on the phenyl ring (e.g. in **4c**, **4d**, **4f**, and **4h**) had a negligible effect on affinity, whereas the strongly electron-withdrawing substituents in **4e** and **4g** reduced the affinity by a factor 2. Extension of the ring system into a naphthyl (IC_{50} (**4m**) = 92 ± 13 nM) or a biphenyl (IC_{50} (**4i**) = 85 ± 16 nM) residue resulted in similarly high affinity as the benzamide **4a**. However, changing the configuration at the anomeric center from β -glycoside **4a** to its isomeric α -anomer **6** resulted in a 26-fold drop in affinity (IC_{50} (**6**) = 2324 ± 432 nM).

The fucosyl sulfonamides **8a** and **8b** were tested as anomeric mixtures and showed an affinity comparable to L-fucose (IC_{50} (**8a**) = 1496 ± 512 nM and IC_{50} (**8b**) = 1144 ± 247 nM). Since the β -anomer was the major anomer present in both cases (67% and 82%), we assigned a reduced LecB binding potency compared to the synthesized fucosyl analogs. When comparing these results to our previous β -fucosylmethyl sulfonamides for LecB_{PAO1} (IC_{50} = 0.97–1.80 μ M^[80]) we conclude a comparable activity of the sulfonamides independent of the presence of the methylene group, e.g. 2-thiophene **8b** lacking the methylene group is only slightly more active (factor 1.6) than its β -fucosylmethyl homolog (IC_{50} = 1.8 μ M^[79]).

The β -fucosylmethyl thioureas were also tested and proved 100-fold less active than the best carboxamides reported here (IC_{50} (**11c–e**, **11o**, **11p**) = 7.3–8.8 μ M). We observed a strong decrease in affinity of the thioureas compared to the previously reported sulfonamides **I** or **II**, although they were as potent as the fucosylmethyl carboxamides **V**.

The strong decrease in affinity for LecB observed here between the very active fucosyl amides and the less active sulfonamides could possibly result from two factors. First, the increased acidity of the sulfonamides compared to the carboxamides could impact their hydrogen-bonding properties. Second, the geometry of the substituents of an amide or a sulfonamide differ from planar trans to staggered gauche, which results in an altered orientation of the linked pharmacophores.

Table 4: Inhibition of LecB_{PAO1} by fucosyl amides, sulfonamides and fucosylmethyl thioureas in a competitive binding assay^[103]. IC₅₀s and std. dev. determined from three independent experiments.

No.	Structure	IC ₅₀ [nM]	No.	Structure	IC ₅₀ [nM]
					
4a	Ph-	88 ± 12	4h	<i>m</i> CH ₃ -C ₆ H ₄ -	120 ± 13
4b	2-Thiophenyl-	122 ± 21	4i	<i>m</i> Ph-C ₆ H ₄ -	85 ± 16
4c	<i>p</i> CH ₃ -C ₆ H ₄ -	110 ± 17	4j	CH ₃ -	902 ± 69
4d	<i>p</i> CH ₃ O-C ₆ H ₄ -	138 ± 20	4k	(CH ₃) ₂ CH-	155 ± 18
4e	<i>p</i> NO ₂ -C ₆ H ₄ -	204 ± 44	4l	(<i>E</i>)-PhCH=CH-	302 ± 43
4f	<i>p</i> Cl-C ₆ H ₄ -	130 ± 19	4m	2-Naphthyl-	92 ± 13
4g	<i>p</i> CF ₃ -C ₆ H ₄ -	211 ± 41	4n	2-Furanyl-	272 ± 26
6	α -Fucosyl benzamide	2324 ± 432			
8a	Fucosyl benzene-sulfonamide	1496 ± 512 (α/β = 33:67)	8b	Fucosyl thiophene-2-sulfonamide	1144 ± 247 (α/β = 18:82)
					
11c	<i>p</i> CH ₃ -C ₆ H ₄ -	7387 ± 473	11o	<i>p</i> CH ₃ CH ₂ -C ₆ H ₄ -	8112 ± 374
11d	<i>p</i> CH ₃ O-C ₆ H ₄ -	7658 ± 288	11p	<i>p</i> F-C ₆ H ₄ -	8831 ± 103
11e	<i>p</i> NO ₂ -C ₆ H ₄ -	5315 ± 292			

To validate the high affinity of the described LecB inhibitors in an orthogonal assay, we analyzed the two anomers of fucosyl benzamides, α -**6** and β -**4a**, by isothermal titration calorimetry (**Table 5**). The K_d of β -fucosyl amide **4a** obtained by ITC was 195 ± 97 nM, and thus somewhat higher than the IC₅₀ of 88 ± 12 nM obtained in the competitive binding assay. The binding enthalpy of $\Delta H = -29.4 \pm 1.3$ kJ mol⁻¹ was similar to the one reported for L-fucose ($\Delta H = -31.2$ kJ mol⁻¹^[61]), but the favorable entropic contribution $T\Delta S = 9.2$ kJ mol⁻¹ was much higher (L-Fuc = 0.3 kJ mol⁻¹^[61]) explaining the increase in binding affinity. α -Fucosyl benzamide (**6**) showed a K_d of 2.3 ± 0.4 μ M corresponding to the one previously reported for other types of α -fucosyl amides and LecB binding ($K_d = 1.2$ – 2.1 μ M^[131]). The enthalpy of LecB-binding of the α -anomer **6** ($\Delta H = -28.0$ kJ mol⁻¹) was nearly identical to the one observed for the β -anomer, while the entropic contribution was much lower for the α -anomer ($T\Delta S = 4.2$ kJ mol⁻¹). In general,

α -linked substituents at the anomeric center of fucose point towards the solvent when bound to LecB, while β -linked substituents are oriented towards the protein surface, and thus the observed difference in binding entropy possibly resulted from displaced protein-bound water molecules for β -benzamide **4**. Favorable entropy of binding is very unusual in protein-carbohydrate interaction, and optimizing this thermodynamic contribution appears here as a valuable strategy.

Table 5: Isothermal titration calorimetry of LecB with fucosyl benzamides β -**4a** and α -**6**; means and std. dev. were calculated from three independent titrations.

Compound	ΔG [kJ mol ⁻¹]	K_d [nM]	n	ΔH [kJ mol ⁻¹]	$T\Delta S$ [kJ mol ⁻¹]
4a	-38.6	195 ± 97	1.0 ± 0.1	-29.4 ± 1.3	9.2 ± 0.4
6	-32.2	2310 ± 350	0.9 ± 0.1	-28.0 ± 1.7	4.2 ± 1.3

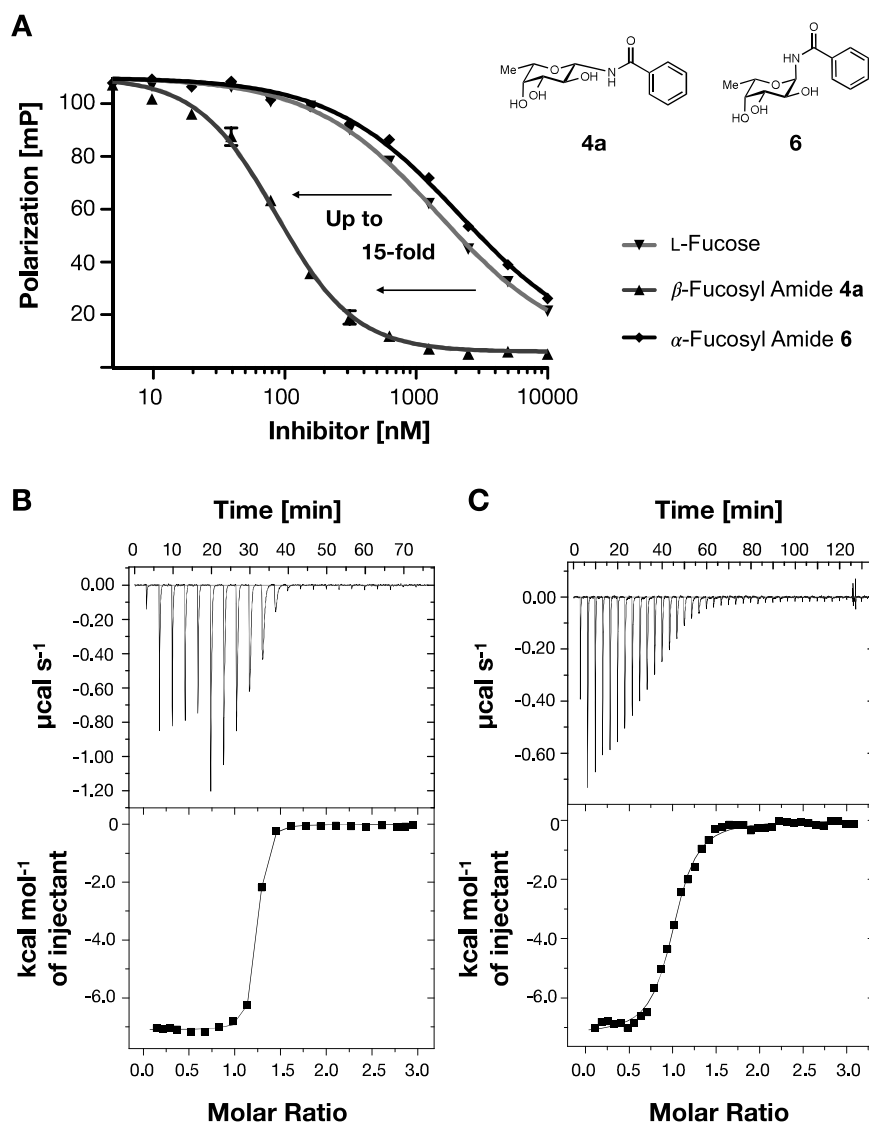


Figure 7: Biophysical analysis of fucosyl benzamides β -**4a** and α -**6** with LecB_{PAO1}: **A)** competitive binding assay based on fluorescence polarization shows a 15-fold increase of LecB inhibition for β -**4a** compared to its α -anomer **6**; **B)** isothermal titration calorimetry of β -fucosyl benzamide **4a** and of **C)** α -fucosyl benzamide **6** against LecB. Both experiments show one respective titration of each fucosyl benzamide from one of the independent replicates, for **A** error bars correspond to those from technical triplicates of one independent replicate.

X-Ray Crystallography of α - and β -Fucosyl Amides in Complex with LecB

To analyze the interactions of the fucosyl amides at the atomic level, we co-crystallized compounds **4a**, **4i** and **6** with LecB_{PAO1}. Co-crystals were obtained for all compounds, but **4a** in complex with LecB diffracted to lower resolution. One dataset was collected for β -benzamide **4a** in complex with LecB and was solved at 2.5 Å resolution with two tetramers in the asymmetric unit. The electron density was poor for one of the

tetramers and in some binding sites and not always well defined for the aglycon, so we decided not to refine it. The datasets obtained for β -biphenyl **4i** and α -benzamide **6** in complex with LecB were of high quality and the structures were solved at high resolution (**4i**: 1.55 Å, **6**: 1.50 Å) (**Figure 8**, **Table S7** for data quality). In both cases, the four carbohydrate binding sites of LecB were occupied by the ligands **4i** or **6**, and the fucose moiety of both ligands was firmly bound to the calcium-ions inside the carbohydrate binding site as reported previously for the fucose/LecB complex (pdb: 1GZT^[124]). In both ligands, the three hydroxy groups OH2, OH3 and OH4 bind to the calcium-ions and their C6 methyl group interacts with Thr45 and Ser23 *via* hydrophobic contacts (**Figure S6** and **Figure S7**, for individual protomers).

In the complex of α -fucosyl benzamide **6** (**Figure 8A**, **Figure S6**), the amide nitrogen points away from the surface of the protein and is integrated in the hydrate layer where it loosely binds to Thr98(NH) *via* one water molecule. Its carbonyl oxygen atom points towards the binding site and forms one hydrogen bond with Ser23 (2.47 – 3.25 Å). Additionally, a hydrogen bond *via* a water to Asp96 can be observed. The aromatic ring of the benzamide is rotated out of the plane of the amide bond. Its orientation depends on a crystal contact based on an edge-to-face interaction with another benzamide. Further, the benzamide ring forms loose lipophilic interactions with Gly97 (3.76 Å – 4.76 Å), that slightly depend on the rotation of the aromatic ring. Thus, the aromatic ring of α -fucosyl benzamide **6** only weakly contributes to the compounds' binding affinity towards LecB.

In the LecB complex of β -fucosyl biphenylamide **4i** (**Figure 8C**, **Figure S7**), the amide function occupies similar orientations in protomers B–D and differs slightly in protomer A. In the latter, the amide NH of **4i** serves as a hydrogen bond donor for Ser22(OH) that makes an ion-dipole interaction with Asp96(COO⁻). In protomers B–D there is a rotation of the carbonyl oxygen by approximately 60°. This allows the carbonyl oxygen to establish a hydrogen bond with Ser23(OH) facing toward the binding site (**Figure 8C**). The rotation further brings the NH in a dipole-ion interaction with Asp96, which could be an explanation for the high binding affinity of the β -fucosyl amides. The directly attached phenyl ring forms lipophilic contacts with Gly24 and Val69. The distal ring is largely solvent exposed but also forms hydrophobic contacts with Asn70. This rather small additional contact area could explain the only very small affinity increase of biphenyl **4i** compared to the benzamide **4a** in the competitive binding assay.

To better understand the high affinity of fucosyl biphenylamide **4i**, we compared the two co-crystal structures of *manno*-cinnamide **S1** (pdb: 5A3O, $K_d = 18.5 \mu\text{M}$ against LecB_{PAO1}^[127]) or fucose-mannose hybrid **II** (pdb: 5MAZ, LecB_{PA14}, $K_d = 0.8 \pm 0.1 \mu\text{M}$ against LecB_{PAO1}^[80]) with biphenyl fucosylamide **4i** in complex with LecB_{PAO1} (**Figure S8**, **Figure S9**).

Superimposition of the complexes LecB_{PAO1}-**4i** and LecB_{PA14}-**S1** clearly shows the effect of the additional methylene group of **S1**. The amide NH of **4i** points in between Ser22 and Asp96 to form a hydrogen bond or dipole-ion interaction. On the other hand, the NH of **S1** is shifted by the methylene group and points beyond the carboxylic acid of Asp96 (**Figure S8**), therefore unable to neither form a hydrogen bond with Ser22, nor an efficient dipole-ion interaction with Asp96.

Also, for the fucose-mannose hybrid **II** the additional CH₂-spacer prevents an interaction of the sulfonamide NH with Ser22 and Asp96 (**Figure S9**). But in contrast to the amides, the nitrogen atom in the sulfonamide function is sp³ hybridized resulting in a conformationally distinct and more flexible linker that enables lipophilic interactions of the thiophene residue with Ser97, Gly24, Val69 and the CH₂ of Asp96, while the large biphenyl residues of **4i** point away from this shallow binding pocket. Interestingly, this orientation of **4i** is in alignment with the previously observed binding pose of **II** with crystal-contacts. The superimposition of **4i** and **II** suggests that introduction of a CH₂-linker next to the carbonyl of the amide functionality of **4i** could allow a certain flexibility of the aromatic substituent and perhaps restore the lipophilic interactions in **4i** as observed for dimethylthiophene **II**.

In summary, β -fucosyl biphenylamide **4i** aligns optimally with the surface of LecB because its amide nitrogen atom is ideally positioned to form a hydrogen bond with Ser22 or an ion-dipole interaction with Asp96, its amide carbonyl oxygen interacts with Ser23 and its proximal phenyl ring forms hydrophobic interactions in the adjacent pocket with Gly24 and Val69.

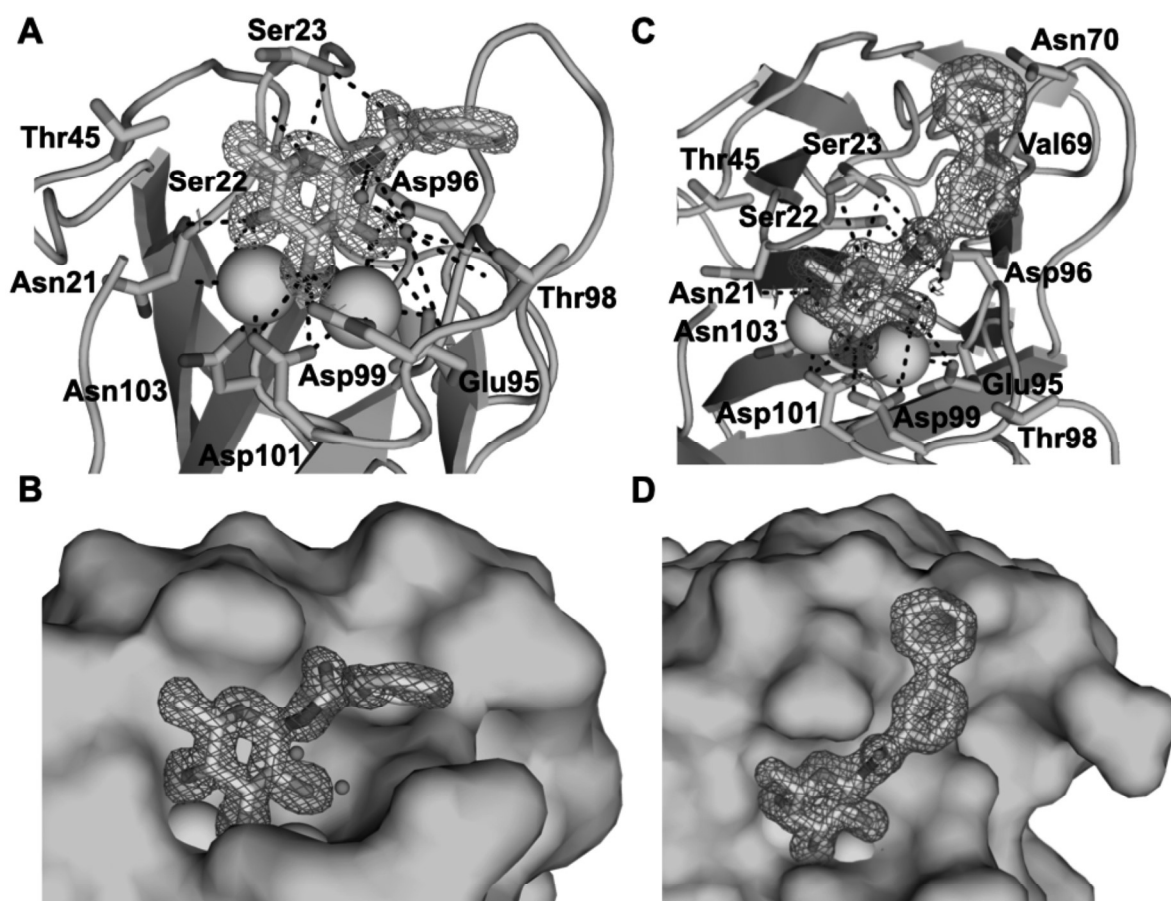


Figure 8: Co-crystal structures of α -fucosyl amide **6** (**A**, **B**, protomer B) and β -fucosyl amide **4i** (**C**, **D**, protomer D) in complex with LecB (pdb: 8AIY for **4i**, 8AIJ for **6**). Electron density is displayed at 1σ , ligands and amino acid residues in the binding site are shown as sticks, water molecules in red and Ca^{2+} -ions as green spheres. Dashed lines indicate hydrogen-bonding interactions of the specific ligand with the protein.

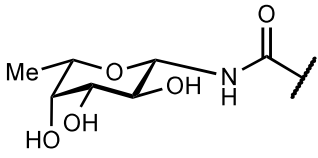
Metabolic Stability, Plasma Protein Binding, and Cytotoxicity

Next, we evaluated the compounds' metabolic stability using murine/human plasma and liver microsomes and analyzed their plasma protein binding (PPB) capacity (**Table 6** and **Table 7**). In mouse plasma, seven out of 15 compounds showed good stability with half-lives between 85–145 min and 6 further compounds were fully stable: α -fucosyl benzamide (**6**), simple and aromatic β -fucosyl amides (**4b**, **4d**, **4k**, **4g**) and sulfonamide **8a**. Only compound **4f** with a *para*-chlorophenyl was less stable ($t_{1/2}$ = 23 min). In human plasma, all compounds were very stable except fucosyl sulfonamide **8b** that degraded slowly ($t_{1/2}$ = 112 min). Exceptions were acetamide **4j** and 2-furanoyl amide **4n** which were not detectable in human and murine plasma, suggesting fast degradation.

For mouse plasma, the lowest protein binding of 21% was obtained for isopropyl amide **4k**. Bulky β -fucosyl amide **4i** as well as 2-thiophenyl sulfonamide **8b** showed 100% binding, while other fucosyl amides showed PPB between 24–76%. *meta*-Methylbenzamide **4h** showed only 45% binding in mouse plasma compared to its *para*-isomer **4c** with non-detectable binding due to degradation processes in mouse plasma. In human plasma, most of the compounds showed a similar binding profile as in mouse plasma. However, especially **8b** exhibited differences: whereas **8b** showed much lower binding in human plasma, full binding in mouse plasma was observed. For compound **8b** these differences might be an artefact as **8b** was less stable in plasma *per se*, with better signals in human plasma. Biphenyl amide **4i** exhibited the highest plasma protein binding in both species compared to all other tested fucosyl amides.

Table 6: Mouse & human plasma stability and plasma protein binding of α -/ β -fucosyl amides and sulfonamides.

No.	Structure	Plasma		PPB	
		Mouse	Human	Mouse	Human
		$t_{1/2}$ [min]	$t_{1/2}$ [min]	[%]	[%]

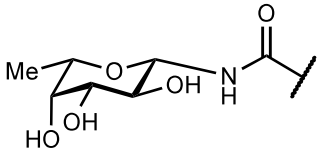


4a	Ph-	n.d.	n.d.	38 ± 5	49 ± 16
4b	2-Thiophenyl-	stable	stable	28 ± 10	70 ± 26
4c	<i>p</i> CH ₃ -C ₆ H ₄ -	130	stable	n.d.	31 ± 9
4d	<i>p</i> CH ₃ O-C ₆ H ₄ -	stable	stable	36 ± 14	12 ± 3
4e	<i>p</i> NO ₂ -C ₆ H ₄ -	144	stable	46 ± 3	57 ± 16
4f	<i>p</i> Cl-C ₆ H ₄ -	24	stable	n.d.	55 ± 18
4g	<i>p</i> CF ₃ -C ₆ H ₄ -	stable	stable	58 ± 19	43 ± 29
4h	<i>m</i> CH ₃ -C ₆ H ₄ -	85	stable	45 ± 20	35 ± 14
4i	<i>m</i> Ph-C ₆ H ₄ -	119	stable	100 ± 0	90 ± 0
4j	Me	n.d.	n.d.	-	-
4k	(CH ₃) ₂ CH-	stable	207	21 ± 13	12 ± 9
4l	(<i>E</i>)-PhCH=CH-	145	stable	18 ± 6	33 ± 14
4m	2-Naphthyl-	131	stable	76 ± 21	68 ± 6
4n	2-Furanoyl	n.d.	n.d.	-	-
6	α -Fucosyl benzamide	stable	stable	24 ± 8	78 ± 20

8a	Fucosyl benzene-sulfonamide	stable	stable	33 ± 7	50 ± 7
8b	Fucosyl thiophene-2-sulfonamide	123	112	100 ± 0	15 ± 9

n.d. = not detected; stable ≥ 240 min.

Table 7: Metabolic stability of α/β -fucosyl amides and sulfonamides in mouse or human liver microsomes.

No.	Structure	Mouse		Human	
		$t_{1/2}$ [min]	Cl_{int} [$\mu\text{L min}^{-1}$ mg^{-1} [protein]]	$t_{1/2}$ [min]	Cl_{int} [$\mu\text{L min}^{-1}$ mg^{-1} [protein]]
					
4a	Ph-	27.7	179.3	23.7	58.4
4b	2-Thiophenyl-	> 60	< 23	34.8	39.8
4c	<i>p</i> CH ₃ -C ₆ H ₄ -	22.4	61.8	11.5	120.4
4d	<i>p</i> MeO-C ₆ H ₄ -	> 60	< 23	161.6	8.6
4e	<i>p</i> NO ₂ -C ₆ H ₄ -	50.2	27.5	32.6	42.5
4f	<i>p</i> Cl-C ₆ H ₄ -	69.4	20.0	119.1	11.6
4g	<i>p</i> CF ₃ -C ₆ H ₄ -	161.8	8.6	229.7	6.0
4h	<i>m</i> CH ₃ -C ₆ H ₄ -	47.0	29.5	146.4	9.5
4i	<i>m</i> Ph-C ₆ H ₄ -	75.3	18.4	> 60	< 23
4j	CH ₃ -	> 60	< 23	86.4	16.0
4k	(CH ₃) ₂ CH-	> 60	< 23	25.2	55.0
4l	(<i>E</i>)-PhCH=CH-	> 60	< 23	247.2	5.6
4m	2-Naphthyl-	27.6	50.2	65.7	21.1
4n	2-Furanyl-	104.4	13.3	59.6	23.2
6	α -Fucosyl benzamide	89.9	15.4	83.7	16.5
8a	α/β -Fucosyl benzenesulfonamide	> 60	< 23	> 60	< 23
8b	α/β -Fucosyl thiophene-2-sulfonamide	56.1	24.7	43.9	31.5

Then, the compounds were tested for their stability in presence of mouse or human liver microsomes (**Table 7**). Most compounds exhibited a rather low clearance, both in the presence of mouse microsomes (11/17 compounds $CL_{int}(m) < 23 \mu\text{L min}^{-1} \text{mg}^{-1}$ [protein]) and human liver microsomes (10/17 compounds $CL_{int}(h) < 23 \mu\text{L min}^{-1} \text{mg}^{-1}$

[protein]). *meta*-Methylbenzamide **4h** gave a moderate clearance with mouse microsomes ($CL_{int} (m, \mathbf{4h}) = 30 \mu\text{L min}^{-1} \text{mg}^{-1} [\text{protein}]$), and low clearance with human liver microsomes ($CL_{int} (h, \mathbf{4h}) = 10 \mu\text{L min}^{-1} \text{mg}^{-1} [\text{protein}]$). In contrast, its *para*-methyl isomer **4c** showed high clearance with microsomes of both species ($CL_{int} (m, \mathbf{4c}) = 61 \mu\text{L min}^{-1} \text{mg}^{-1} [\text{protein}]$, $CL_{int} (h, \mathbf{4c}) = 120 \mu\text{L min}^{-1} \text{mg}^{-1} [\text{protein}]$). Interestingly, the two most potent inhibitors **4a** and **4i** behaved completely differently. While benzamide **4a** resulted in the highest intrinsic clearance with mouse and the second highest clearance with human liver microsomes ($CL_{int} (\text{mouse}, \mathbf{4a}) = 179 \mu\text{L min}^{-1} \text{mg}^{-1} [\text{protein}]$; $CL_{int} (\text{human}, \mathbf{4a}) = 58.4 \mu\text{L min}^{-1} \text{mg}^{-1} [\text{protein}]$), biphenyl **4i** showed low clearance ($CL_{int} (\mathbf{4i}) < 23 \mu\text{L min}^{-1} \text{mg}^{-1} [\text{protein}]$) in presence of microsomes from both species. In general, several tested compounds had a good metabolic stability in the presence of mouse or human microsomes *in vitro* with $t_{1/2} > 60$ min, except for the *para/meta*-methyl benzamides **4c** and **4h**, unsubstituted benzamide **4a**, and naphthyl **4m**. Nitrophenyl **4e**, thiophenyl **4b** and furanyl **4n** were moderately stable with half-lives between $t_{1/2} = 47\text{--}56$ min in mouse and $t_{1/2} = 25\text{--}60$ min in human liver microsomes.

Furthermore, cytotoxicity of α -/ β -fucosyl amides and sulfonamides was assessed *in vitro* using three different cell lines, i.e., epithelial lung cell line (A549), Chinese hamster ovary cells (CHO) and epithelial liver cell line (HepG2) (**Figure 9**). All tested LecB ligands displayed no toxicity against A549 cells at 10 nM and 1000 nM. One exception was obtained for the amide derivative **4j** with slightly reduced viability at 1 μM . Testing against Chinese hamster ovary cells (CHO) also revealed no toxicity for most of the tested compounds. However, a noticeable dose-dependent toxicity was observed in the case of the two amide derivatives **4d** and **4j**. Finally, the testing against HepG2 liver epithelial cells resulted in heterogeneous cytotoxicity across our compounds. Eleven out of 15 fucosyl amides showed detectable cytotoxicity with in part strong reduction of cellular viability. Four exceptions devoid of detectable cytotoxicity against HepG2 cells were *para*-nitrophenyl amide **4e**, *meta*-methylphenyl amide **4h**, 2-furanoyl amide **4n** as well as biphenyl amide **4j**, a compound that was however somewhat toxic against the other two cell lines, A549 and CHO cells. No cytotoxicity against HepG2 cells was detected for sulfonamide **8a** while sulfonamide **8b** displayed significant cytotoxicity at 1000 nM.

In summary, all fucosyl amides were sufficiently stable in human plasma and most withstood degradation by human liver microsomes. Their intrinsic clearance was generally good except for benzamide **4a**, *para*-methylbenzamide **4c**, and for 2-

naphthyl **4m**. With regard to cytotoxicity, a large variation was observed between the three tested cell lines and the 15 tested amides. Numerous compounds showed significant cytotoxicity especially against HepG2 cells, however some compounds did not show acute cytotoxicity in all three cell lines.

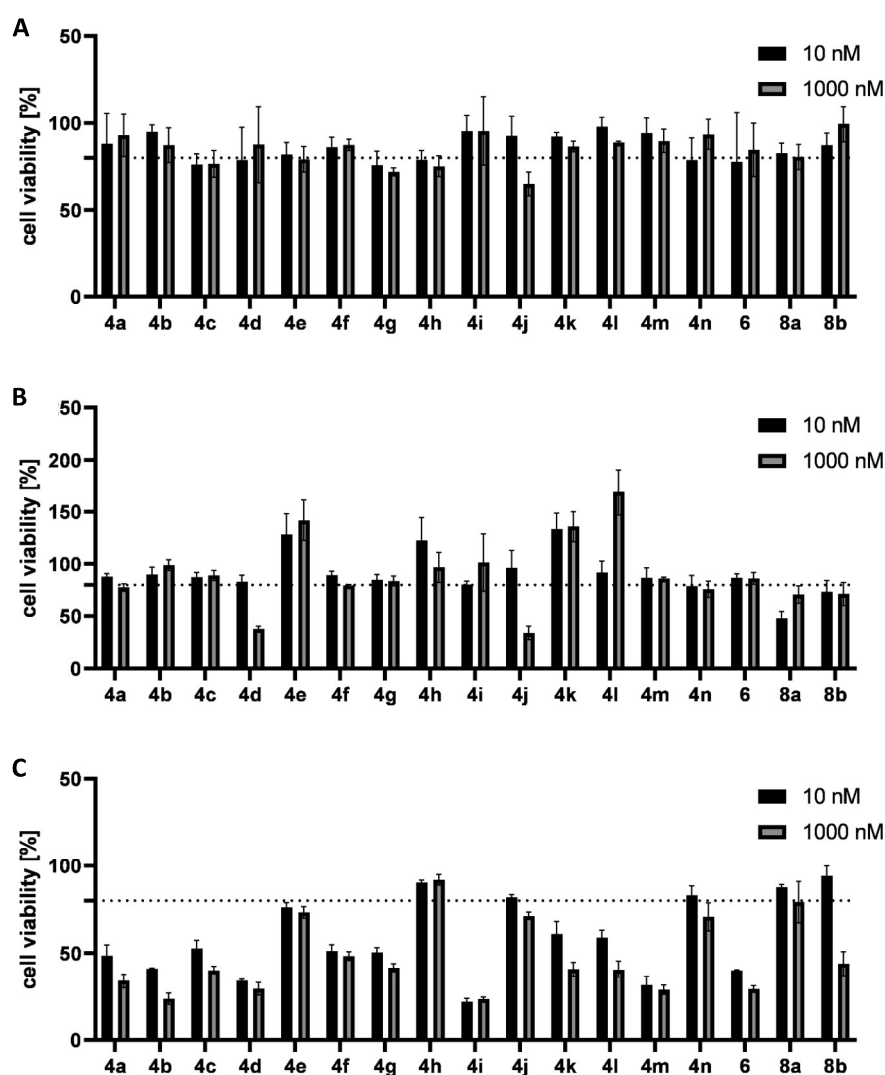


Figure 9: Cytotoxicity of α/β -fucosyl amides and sulfonamides against the lung epithelial cells A549 (**A**), Chinese hamster ovary cells (CHO) (**B**) and liver epithelial cells (HepG2) (**C**) at 10 nM and 1000 nM concentration of LecB ligands. Cells treated with vehicle only (DMSO diluted in PBS, final DMSO concentration in the cell assay: 0.1 %) served as a negative control indicated as the dashed line. Pure medium (DMEM + 10 % FCS) and completely damaged cells served as positive controls. The error bars show the standard deviation of minimum three independent experiments.

Conclusion

In our search for novel LecB inhibitors, we have designed and synthesized fucosylmethyl thioureas as well as shortened molecules, fucosyl amides and sulfonamides, lacking the methylene group in order to analyze the altered linker position on LecB binding. Surprisingly, the fucosylmethyl thioureas only showed moderate binding in the micromolar range which was also observed for the *N*-linked fucosyl sulfonamides. On the other hand, our β -fucosyl amides devoid of the methylene bridge constitute the first monovalent two-digit nanomolar LecB inhibitors with IC₅₀ of 88 nM for benzamide **4a** and 85 nM for biphenyl derivative **4i**. Noteworthy, α -fucosyl amides had been studied before with LecB and these molecules showed moderate micromolar affinities^[131], which was confirmed by our control molecule α -benzamide **6**.

In the co-crystal structure of β -linked **4i** in complex with LecB, we demonstrated that the amide function is crucial for binding by forming a hydrogen bond with Ser22 which is located between the carbohydrate binding site and the additional subpocket, that is occupied by the proximal aromatic ring which hydrophobically interacts with Gly24 and Val69. These interactions are absent in the crystal structure of the α -anomer **6** in complex with LecB. Both crystal structures provide a basis for the interpretation of the microcalorimetric titration data for LecB with **4a** or **4i**. Compounds have also been analyzed for *in vitro* early ADME/Tox. In general, satisfying properties could be identified in all assays, except for significant cytotoxicity against one out of the three tested cell lines. Despite the fact that numerous compounds proved toxic, the presence of several derivatives without detectable cytotoxicity clearly underlines their potential for future optimization.

During the writing of this manuscript, a set of β -fucosyl amides has been reported as weak inhibitors of the N-terminal domain of BC2L-C, a lectin from *Burkholderia cenocepacia*^[132,133], that also binds fucose but displays different sequence, structure and binding site architecture compared to LecB. Despite this low affinity, further optimization could open a possibility for a molecule that potently inhibits both lectins, which could be of interest since *P. aeruginosa* and *B. cenocepacia* often co-infect cystic fibrosis patients. Thus, these β -fucosyl amides constitute a promising new class of LecB inhibitors for future use as pathoblockers against infections with *P. aeruginosa* and beyond.

Material and Methods

General experimental details and syntheses are described in the supporting information.

Docking

The crystal structure coordinates of LecB in complex with fucose (pdb: 1OXC) was adjusted for docking in MOE (Molecular Operating Environment, Chemical Computing Group, Canada version: 2014.0901) by removing all ligands and water molecules and keeping one monomer. The coordinates of the carbohydrate binding site were determined in AutoDockTools^[134] and added into the docking-file. Four amino acid Asn21, Glu95, Asp101 and Asn103 were kept flexible during the docking-run. In parallel, ligands for docking were drawn in ChemDraw, exported as SMILES code and translated into a pdb-file using the online SMILES Translator.^[135] The ligand pdb-files were processed in MOE, bonds were set to the lowest energy level, and exported as a mol2-file and added to the docking file. The docking was performed with Plants1.1^[136] using a binding site radius of 13 Å. For validation of the protocol, docking was performed first with α -L-fucose and the resulting pose was then superimposed with its co-crystal structure in the Molecular Operating Environment (MOE), which confirmed the same interactions. Afterwards, the two compounds **4a** and **11c**, were docked. The results were visualized and analyzed in MOE.

Recombinant Expression and Purification of LecB

LecB from *P. aeruginosa* PAO1 was expressed and purified from *Escherichia coli* BL21 (DE3) and the plasmid pET25pa2l^[137] as described before.^[78]

Competitive Binding Assay

The competitive binding assay based on fluorescence polarization was performed with small modifications in analogy to the one described before.^[60] To a 10 μ L solution of LecB_{PAO1} (150 nM) and the fluorescent reporter ligand (*N*-(fluorescein-5-yl)-*N'*-(α -L-fucosyl ethylene)-thiocarbamide, 20 nM) in TBS/Ca²⁺-buffer containing 0.02% DMSO (20 mM Tris*HCl, 137 mM NaCl, 2.6 mM KCl, 1 mM CaCl₂, pH = 7.4) in a black 384-well plate (Greiner Bio-One, Germany, cat. no. 781900), a 10 μ L serial dilution of the inhibitor in the same buffer (10–0.005 μ M, dilution factor 2) was added (technical triplicates). The plate was sealed (EASYseal, Greiner Bio-One, cat. no. 676001), centrifuged (1500 x g, r.t., 1 min) and incubated in a dark chamber under shaking conditions for 24 h. Afterwards, the seal was removed and fluorescence was measured

on a PHERAstar FS (BMG Labtech, Germany, filter ex.: 485 nm em.: 535 nm). After blank value (TBS/Ca²⁺-buffer with LecB) subtraction from the measured fluorescence intensities, polarization was calculated and the data were analyzed using the four-parameter variable slope model in MARS software (BMG Labtech). Then, top and bottom plateaus were redefined using the full inhibition value in presence of the positive control L-fucose and full binding value for LecB and reporter ligand in the absence of inhibitor as a negative control. The experiments were performed in three independent experiments, the data were averaged and visualized using GraphPad PRISM (version 5). Fucosylmethyl thioureas were tested at a final LecB concentration of 150 nM, a dilution series 100–0.78 μ M and 0.1% DMSO in the TBS/Ca²⁺-buffer.

Isothermal Titration Calorimetry

Compounds **4a**, **6** and LecB_{PAO1} were separately dissolved in the same TBS/Ca²⁺-buffer and the concentration for LecB was determined by UV absorbance (ϵ = 6990 M⁻¹ cm⁻¹[138]). Experiments were performed on a MicroCal iTC₂₀₀ (Malvern Panalytical, United Kingdom) instrument by titrating the ligand (500–1500 μ M) into LecB solution (100 μ M) with stirring (700 rpm) at 25 °C. The reference power was set to 5 μ cal s⁻¹, the filter period to 5 s, and 19–39 injections (0.5–2 μ L per injection) with an injection duration of 1 s and a spacing of 240 s between each injection were performed per experiment. In case of titrations with lower LecB concentrations (500 μ M), the syringe was refilled after the first run ended, and the experiments were continued with the same sample cell contents to reach saturation. The resulting data files were merged with the MicroCal Concat ITC software. The first injection of every titration was discarded and the data were analyzed with the MicroCal Origin software using the one-site binding model. ITC data are depicted in **Figure 7** and **Table S6**.

X-ray Crystallography of LecB_{PAO1} in complex with 6 or 4i

LecB_{PAO1} at 10.8 mg mL⁻¹ in water with 1 mM CaCl₂ (**4a** and **4i**) or at 9.3 mg mL⁻¹ in 20 mM Hepes pH 7.5, 100 mM NaCl and 100 μ M CaCl₂ (**6**) was incubated in a 19:1 ratio with 2.5 mM of compound for 30 min to 1 h prior to crystallization. Stock solution at 50 mM of compound were made in water for **4a** and in 100% DMSO for **4i** and **6**. Hanging drop vapor diffusion method using 1 μ L of protein-ligand mixture with 1 μ L of reservoir solution at 19 °C in a 24-well plate yielded crystals after 1–3 days. The crystals for the LecB_{PAO1}-**4a** or LecB_{PAO1}-**4i** complexes were obtained with 30 and 28% PEG 8K, 200 mM (NH₄)₂SO₄ and 100 mM Tris pH = 8.5, respectively and those for LecB_{PAO1}-**6** with 24% Peg8K, 1 M LiCl and 100 mM sodium acetate pH 4.4. All solutions were

cryoprotected and the crystal was directly mounted in a cryoloop and flash-frozen in liquid nitrogen. Diffraction data were collected at 100 K at Synchrotron SOLEIL (Paris, France) on beamline Proxima 1 using an EIGER X 16M area detector for LecB_{PAO1}-**4a** and LecB_{PAO1}-**4i** or Proxima 2 using a EIGER X 9M area detector for LecB_{PAO1}-**6**. The data were processed using XDS and XDSme.^[107,139] All further computing was performed using the CCP4 suite.^[140] Five percent of the observations were set aside for cross validation analysis and hydrogen atoms were added in their riding positions, and used for geometry and structure-factor calculations. The structure was solved by molecular replacement using PHASER.^[141] For complexes with **4i** and **6**, the coordinates of the 5A3O tetramer were used as a search model to search for one tetramer in the asymmetric unit. The structures were refined with restrained maximum likelihood refinement using REFMAC 5.8^[111], iterated with manual rebuilding in Coot.^[142] Ligand libraries were created using JLigand. The ligands were introduced after inspection of the 2Fo – DFc weighted maps. Water molecules, introduced automatically using Coot, were inspected manually. The stereochemical quality of the models was assessed with the PDB Validation Server. The structure of LecB in complex with **4a** was solved by molecular replacement at 2.5 Å using Phaser and a search for one tetramer and 2 dimers from model 1GZT. The low resolution did not give suitable electron density and we decided not to refine it. Data quality and refinement statistics are summarized in Supporting Information, **Table S7**. All structural figures were created using PyMOL.

Plasma Stability Assay

Each compound, dissolved in DMSO (10 mM), was added to mouse plasma (pH = 7.4, 37 °C) or to human plasma (pH = 7.4, 37 °C) to yield a final concentration of 1 µM. In addition, procaine and procainamide (dissolved in DMSO) were added to mouse plasma or to human plasma (pH = 7.4, 37 °C) to yield a final concentration of 1 µM. Procaine served as a positive control as it is unstable in mouse plasma. Procainamide served as a negative control as it is stable in mouse plasma. The samples were incubated for 0, 15, 30, 60, 90, 120, and 240 min at 37 °C. At each time point, 10 µL of the respective sample was extracted with 90 µL acetonitrile and 12.5 ng/mL of caffeine as an internal standard for 5 min at 2000 rpm on a MixMate vortex mixer (Eppendorf). Acetonitrile and caffeine were dispensed using a Mantis Formulatrix. Then samples were centrifuged for 20 min at 2270 x g at 4 °C and the supernatants were transferred to 96-well Greiner V-bottom plates. Peak areas of each compound

and of the internal standard were analyzed using the MultiQuant 3.0 software (AB Sciex). Peak areas of the respective compound were normalized to the internal standard peak area and to the respective peak areas at time point 0 min with equation (4):

$$Norm_{peak\ area} = \frac{C * D^{-1}}{A * B^{-1}} \quad (4)$$

A: peak area of the compound at the time point 0 min, B: peak area of the internal standard at time point 0 min, C: peak area of the compound at the respective time point, D: peak area of the internal standard at the respective time point.

In vitro Metabolic Stability Assay

Liver microsomes (mouse and human, Thermo Fisher) were thawed slowly on ice. 5 mg mL⁻¹ of microsomes, 2 µL of a 100 µM solution of every compound and 183 µL of 100 mM phosphate buffer were incubated 5 min at 37 °C in a water bath. Reactions were initiated using 10 µL of 20 mM NADPH (Carl Roth, Germany). Samples were incubated in three replicates at 37 °C under gentle agitation at 150 rpm. At 0, 5, 15, 30, and 60 min, reactions were terminated by the addition of 180 µL acetonitrile. Samples were vortexed for 5 min and then centrifuged at 2270 x g for 20 min at 4 °C. The supernatants were transferred to 96-well Greiner V-bottom plates, sealed with WebSeal non-sterile mats (Thermo Fisher) and analyzed according to the section HPLC-MS analysis. Peak areas of the respective time point of the compounds were normalized to the peak area at time point 0 min. Then half-life was calculated using linear regression. CL_{int} [µL/min/mg protein] was calculated using the following equation (5):

$$CL_{int} [\mu L \min^{-1} mg_{(protein)}^{-1}] = \frac{0.693}{0.005 * t_{1/2}} \quad (5)$$

Plasma Protein Binding

Plasma protein binding was assessed using the rapid equilibrium device (RED) system from ThermoFisher. Compounds were dissolved in DMSO to a concentration of 10 mM. Naproxene served as control as it shows high plasma protein binding. Compounds were diluted in murine plasma (from CD-1 mice, pooled, Biomol GmbH) or in human plasma (human donors, both genders, pooled, antibodies-online GmbH) to a final concentration of 1 µM. Dialysis buffer and plasma samples were added to the respective chambers according the manufacturer's protocol. The RED plate was

sealed with a tape and incubated at 37 °C for 2 hours at 800 rpm on an Eppendorf MixMate vortex-mixer. Then samples (dialysis and plasma samples) were withdrawn from the respective chambers. To 25 µL of each dialysis sample, 25 µL of plasma and to 25 µL of plasma sample, 25 µL of dialysis buffer was added. Then, 150 µL ice-cold extraction solvent (MeCN/H₂O (90:10) containing 12.5 ng mL⁻¹ caffeine as internal standard) was added. Samples were incubated for 30 min on ice. Then, samples were centrifuged at 4 °C at 2270 x g for 10 min. Supernatants were transferred to Greiner V-bottom 96-well plates and sealed with a tape. The percentage of bound compound was calculated with the equations (6) and (7):

$$\%_{free} = \frac{C_{buffer\ chamber}}{C_{plasma\ chamber}} * 100 \quad (6)$$

$$\%_{bound} = 100\% - \%_{free} \quad (7)$$

HPLC-MS Analysis

Samples were analyzed using an Agilent 1290 Infinity II HPLC system coupled to an AB Sciex QTrap 6500plus mass spectrometer. LC conditions were as follows: column: Agilent Zorbax Eclipse Plus C18, 50x2.1 mm, 1.8 µm; temperature: 30 °C; injection volume: 5 µL per sample; flow rate: 700 µL min⁻¹. Samples were run under acidic and buffered conditions. Solvents for acidic conditions: A1: water + 0.1% formic acid; solvent B1: 95% MeCN/5% H₂O + 0.1% formic acid; solvents for buffered conditions: A2: 95% H₂O + 5% MeCN + 5 mM ammonium acetate + 40 µL L⁻¹ acetic acid; B2: 95% MeCN + 5 % H₂O + 5 mM ammonium acetate + 40 µL L⁻¹ acetic acid. The same gradient was applied for acidic and buffered conditions: 99% A at 0 min, 99% A until 1 min, 99–0% A from 1.0 min to 4.0 min, 0% A until 5.0 min. Mass transitions for controls and compounds are depicted in **Table S8**.

Cytotoxicity

The epithelial liver cell line HepG2 (ATCC HB-8065TM) and the epithelial lung cell line A549 (ATCC CCL-185) were cultivated in Dulbecco's modified Eagle's medium (DMEM) with 10% heat-inactivated fetal calf serum (FCS) at 37°C and 5% CO₂. CHO cells (ATCC CCL-61) were cultivated in Gibco Ham's F-12K medium supplemented with 10% FCS. Cells were seeded into a 96-well plate (Nunc, Roskilde, Denmark) and grown to 75% confluency. The following compounds were tested in the cell assay: **4a-4n**, **6**, **8a** and **8b**. Every compound was dissolved in DMSO and diluted in PBS (final

DMSO concentration in the cell assay: 0.1 %). Cells were incubated for 24 h at 37°C and 5% CO₂ with the respective compound at two different concentrations (10 nM and 1 µM) allowing for a rapid screen. Cells treated with vehicle only (DMSO diluted in PBS, final DMSO concentration in the cell assay: 0.1 %) served as a negative control. Furthermore, pure medium (DMEM + 10 % FCS) and completely damaged cells served as positive controls. To damage cells, cells were treated with 0.5% Triton X-100 1 h prior to addition of MTT (Sigma). After 24 h cells were washed twice with DMEM + 10% FCS (A549 and HepG2 cells) or Ham's F-12K + 10% FCS (CHO-cells). MTT diluted in PBS (stock solution 5 mg/mL) was added to the wells at a final concentration of 1 mg/mL. The cells were incubated for 3 h at 37°C and 5% CO₂. Medium was removed and 0.04 M HCl in 2-propanol was added. The cells were incubated at room temperature for 15 min. Then the supernatant was transferred to a 96-well plate. UV absorbance of the samples was measured at 560 nm and at 670 nm as a reference wavelength on a Tecan Sunrise ELISA reader using Magellan software. Data was normalized using the following formula: (A-B)/(C-B) with 'A' as the respective data point, 'B' as the value of the Triton X-100-treated control and 'C' as the vehicle control. The experiment was repeated at least three times. The error bars indicate the standard deviation.

Supporting Information

The supporting information contain purity of key compounds by LCMS, protein-ligand interaction maps of the docked compounds, β -fucosyl benzamide **4a** and β -fucosylmethyl thiourea **11c**, ITC data of all replicates for titrations of LecB with β -/ α -fucosyl benzamides **4a** and **6**, X-ray data collection and refinement statistics of LecB complexed structures and a zoom in the electron density and interactions of the ligand with each LecB protomers, and m/z search window for plasma stability and mass transitions of the tested compounds from the ADME studies.

Orcid

Patrycja Mała: 0000-0002-1259-6484

Eike Siebs: 0000-0002-7349-9872

Joscha Meiers: 0000-0003-0869-0088

Katharina Rox: 0000-0002-8020-1384

Annabelle Varrot: 0000-0001-6667-8162

Anne Imberty: 0000-0001-6825-9527

Alexander Titz: 0000-0001-7408-5084

Notes

The authors declare no competing financial interest.

Acknowledgements

A.T. and A.I. acknowledge by financial support of the French-German ANR/DFG project (ANR-AAPG-2017) funded by the Agence Nationale de la Recherche (grant no. ANR-17-CE11-0048) and Deutsche Forschungsgemeinschaft (grant no. Ti756/5-1). A.I and A.V. acknowledge support from ANR project Glyco@Alps (ANR-15-IDEX-02), Labex ARCANÉ, and CBH-EUR-GS (ANR-17-EURE-0003). The authors also thank DAAD for a scholarship to P.M., acknowledge grant no. POWR.03.02.00-00-I026/16 co-financed by the European Union through the European Social Fund under the Operational Program Knowledge Education Development, for her support, the Ambassade de France en Allemagne for a PROCOPE-Mobility scholarship to J.M., and the European Research Council for an ERC Starting Grant (Sweetbullets, grant no 716311) to A.T. KR receives support from the German Centre for Infection Research (DZIF, TTU 09.719) for operating the PK/PD unit. We acknowledge the synchrotron SOLEIL (Saint Aubin, France) for access to beamlines Proxima 1 and 2 (Proposal Number 20191314) and for the technical support of Pierre Montaville and Martin Savko. We also acknowledge Yasmina Grimoire, Emilie Gillon (both CERMAV), Dirk Hauck (HIPS), Janine Schreiber and Kimberley Vivien Sander (both HZI) for excellent technical support.

Ligand-Based Design of *Burkholderia ambifaria* Lectin BambL Inhibitors

Introduction

Burkholderia ambifaria is a Gram-negative bacterium that belongs to the *Burkholderia cepacia* complex (Bcc). This group currently consists of 20 closely related bacterial species, which were previously subordinated to *Pseudomonas aeruginosa*.^[32,35,143] Some of the *cepacia* species play a crucial role in cystic fibrosis (CF) and cause sudden death in 20% of the infected patients within weeks/month, which is termed “*cepacia* syndrome”.^[64,144] Most of the Bccs in CF are multi-drug resistant, which makes it important to find new strategies to kill those bacteria.^[145]

CF lungs show increased amounts of fucosylated cell-surface carbohydrates on nasal and trachea membrane glycoproteins/glycopeptides compared to healthy people^[146], which suggest that lectins may be involved in the attachment and recognition of host cell by binding to its glycocalyx as proposed for the *P. aeruginosa* lectins, LecA and LecB.^[117] So far, different lectins were isolated from the Bcc such as the *B. cenocepacia* lectins BC2L-A^[147], BC2L-B^[148] (uncharacterized^[149]), and BC2L-C^[132] which contain a LecB-like domain, and BP39L^[150], a seven-bladed β -propeller folded lectin, expressed by *B. pseudomallei*. Furthermore, *B. ambifaria* expresses BambL as a carbohydrate binding protein that binds to fucosylated glycans such as the human histo blood group H-type 2 and Lewis^Y and showed sub-micromolar affinities for the blood group A- and H-type 5 trisaccharides.^[63] It was also shown that it is able to bind lung tissues and human saliva.^[63]

Recently, it was reported that BambL acts as a B cell superantigen that leads to an activation-induced cell death by impeding the adaptive immune response at high concentrations and favors infections.^[151] This lectin interacts with the B cell surface glycans activating the expression of the cell surface molecules CD54, CD69, and CD86. This adhesion/activation by BambL can be blocked by using either B cell receptor pathway inhibitors or an excess of fucose.

BambL is a six bladed homotrimer with one fucose binding site per protomer and three additional binding sites at the interface of the monomers (**Figure 10A**). The carbohydrate binding sites (CBS) of the monomers and of their interfaces are quite similar, but differ slightly in their fucose binding mode (*intra* pocket: Trp74 vs *inter* pocket: Tyr29 and amino acids Tyr35/Thr80, Ser13/His58 in the vicinity). Fucose binds

to BamBL and forms six hydrogen bonds per pocket (*intra*: Arg15, Glu26, Trp79, Ala38; *inter* Trp34, Arg60, Glu71, Ala83) and strong hydrophobic interactions with the aromatic residues (Tyr29 or Trp74).^[63] The interactions lead to affinities in the low micromolar range for small monovalent ligands ($K_d(\text{Me-}\alpha\text{-L-Fuc}) = 1.02 \mu\text{M}$, $n = 1.9$).^[68] So far, only a few compounds were tested on BamBL, fucosylated glycans with preferred binding to blood group H and Lewis^Y oligosaccharides^[63], monovalent $\alpha\text{-L-fucopyranose}$ compound **I** (**Figure 10B**) showing two binding modes, one with a high and one with a moderate affinity ($K_d(\text{I}) = 0.24/23 \mu\text{M}$ ^[68]), and mannose/phosphate centered fucosylated multivalent glycoclusters e.g. **II** ($K_d(\text{II}) = 43 \pm 0.2 \text{ nM}$ ^[152]).

Here, we focused on rational ligand-based design of compounds derived from an *in-house* library screening for targeting the CBSs of the *B. ambifaria* lectin BamBL. Afterwards, we synthesized inhibitors and studied their affinities *in vitro* and co-crystallized promising hits with BamBL.

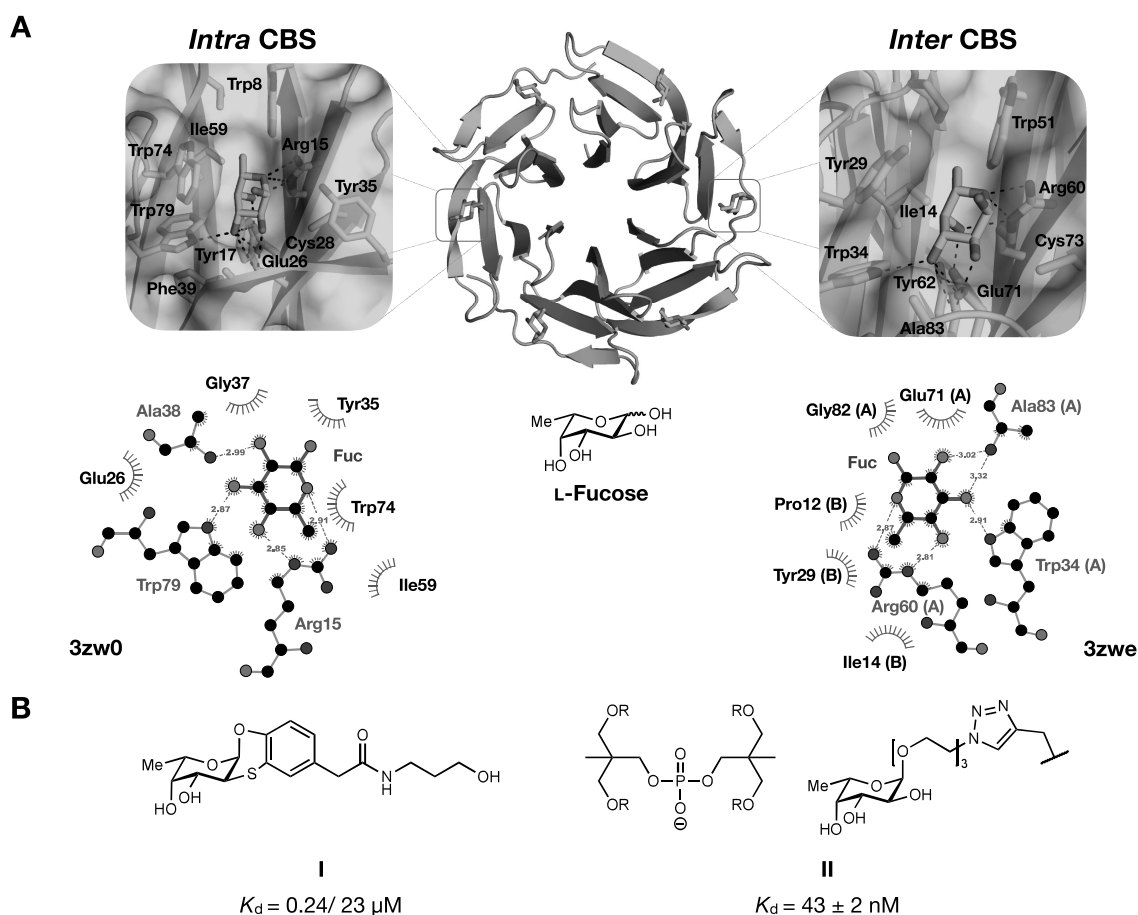


Figure 10: A) Co-crystal structure of L-fucose in complex with BamBL showing the interactions at the *intra* and *inter* CBSs (pdb: 3zw0, *intra*; pdb: 3zwe, *inter*, visualized in PyMOL and LigPlot; color in 2D interaction map: ligand bond (purple), non-ligand bond (brown), hydrogen bond and its length (green dashes), non-ligand residues involved in hydrophobic contract(s) (reddish brown), corresponding atoms involved in

hydrophobic contact(s) (reddish brown dashes around atom). **B)** Reported synthesized BambL inhibitors. Left: monovalent α -L-fucopyranose **I**.^[68] Right: fucosylated tetravalent glycocluster **II**.^[152]

Results and Discussion

Biophysical Screening of Fucopyranosides against BambL

For the development of new BambL inhibitors, we selected 11 compounds from our *in-house* library carrying changes in different positions on the core structure of L-fucose to cover a broad spectrum in modifications to reveal positions that are tolerated for later extensions to further improve the binding affinity towards this lectin. The compounds were tested in the competitive binding assay based on fluorescence polarization against BambL with Me- α -L-Fuc as a positive control as reported before (**Table 8** and visualized in **Figure S10**).^[153]

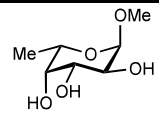
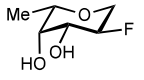
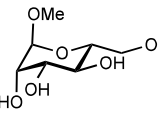
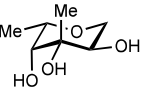
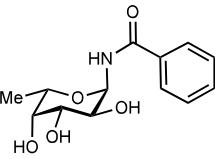
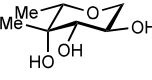
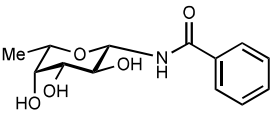
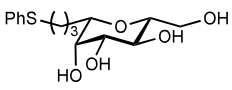
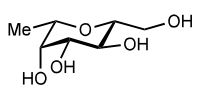
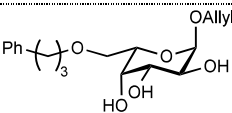
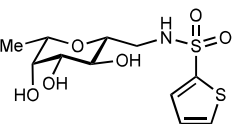
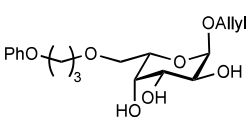
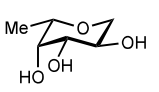
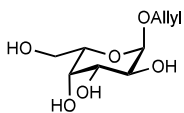
In these experiments, Me- α -L-Fuc (**1**) ($IC_{50} = 3.1 \pm 0.6 \mu M$) showed a comparable BambL affinity to the reported value ($IC_{50} = 2.4 \mu M$ ^[153]). Interestingly, Me- α -D-Man (**2**) did not inhibit BambL under those FP assay conditions although fucose specific lectins often also bind to mannosides with lower affinities or *vice versa* (e.g.: bacterial lectin LecB or C-type lectins with a carbohydrate recognition domain Glu-Pro-Asn (Langerin))^[154]. These findings confirm the importance of the equatorial methyl group in position five of L-fucose to form the CH $\cdots\pi$ -interaction with the protein.

Modifications at the anomeric position (C1) were tolerated, and the α -fucosyl benzamide (**3**) (IC_{50} (**3**) = $10.5 \pm 0.8 \mu M$) was clearly favored over the β -analog **4** (IC_{50} (**4**) = $136.9 \pm 10.7 \mu M$) and was nearly as active as Me- α -L-Fuc (**1**). Replacement of the bulky phenyl ring in β -**4** with a small flexible hydroxymethyl group (Fuc-Man hybrid **5**, $IC_{50} = 9.3 \pm 1.3 \mu M$), led to higher values than Me- α -L-Fuc (**1**) comparable to the α -benzamide **3**. Extension of the aglycon (fucosylmethyl sulfonamide **6** (IC_{50} (**6**) = $4.1 \pm 0.6 \mu M$) showed an affinity similar to Me- α -L-Fuc (**1**). Removal of the hydroxymethyl group from C1 in **5** was tolerated (IC_{50} (**7**) = $7.4 \pm 1.2 \mu M$), which can be explained by the x-ray co-crystal structure of fucose with BambL showing that only the hydroxy groups OH2, OH3 and OH4 are involved in binding (**Figure 10A**). In contrast, modifying C2 by replacing the hydroxy group with a fluorine atom led to an affinity decrease by 3.5-fold. Implementing different substituents on C3 or C4 of L-fucose while keeping the hydroxy groups for binding led to significant affinity decrease (IC_{50} (**9**) >340 μM and

IC₅₀ (**10**) >300 μM). This can be explained by the CBSs of BambL that only provides space for unsubstituted L-fucose on C3 and C4 to align and interact with the aromatic residues Tyr29 or Trp74. Interestingly, extending the methyl group of C5 by an alkyl chain as shown for 6-phenylthio propyl fucoside **11** did not affect the binding affinity, which revealed C6 as an excellent position for further modification.

Taken together, it is possible to perform substitutions in α-position and C6 of L-fucose. This observation corresponded to interactions shown in the co-crystal structures of fucose with BambL (pdb: 3zwe, 3zw0) where the three conserved hydroxy groups (OH2, OH3 and OH4), and the hydrophobic backbone together with the methyl group in position C6 interact with the protein. These information were then used for the design of new BambL inhibitors.

Table 8: Competitive binding results of *in-house* library fucosides (**1–11**) and synthesized galactoside inhibitors (**16a–b**, **17**) against BambL. IC₅₀s and std. dev. determined from three independent experiments.

No.	Structure	IC ₅₀ [μM]	No.	Structure	IC ₅₀ [μM]
1		3.1 ± 0.6	8		26.0 ± 1.0
2		n.i.	9		>340*
3		10.5 ± 0.8	10		>300*
4		136.9 ± 10.7	11		5.6 ± 0.4
5		9.3 ± 1.3	16a		4.6 ± 0.5
6		4.1 ± 0.6	16b		3.5 ± 0.4
7		7.4 ± 1.2	17		7.8 ± 0.9

[*] = tested only ones.

In Silico Design of New BambL Inhibitors

The biophysical validation revealed that L-fucose can be modified in α -position on C1, and in equatorial direction on C5, whereupon molecule **12** was designed that was extended on C5 with an alkyl phenyl chain (**Figure 11A**). To improve synthetic accessibility, an ether function was introduced that might form a hydrogen bond with Arg15 (**Figure S11**). At the α -position, we introduced an instable hemiacetal function that had to be protected during the synthesis but the screening results showed that introducing different groups in α -position including bulky benzamides did not change the affinity significantly. The designed L-galactose compound **12** was then docked with BambL. The top ranked molecule pose was discarded, since the phenyl ring was placed in the CBS and the sugar moiety was randomly attached to the surface (**Figure 11B**, beige). The second ranked molecule showed the sugar part in the correct place, however, the phenyl ring did not interact with the surface of BambL (**Figure 11B**, yellow; **Figure S12**). The third ranked pose showed the fucose moiety in the CBS, the alkyl chain was embedded in a small cleft along the surface and the phenyl ring reached a second pocket where it might interact with Trp8 to form a $\text{NH}\cdots\pi$ interaction, which is known to be slightly stronger than a $\text{CH}\cdots\pi$ interaction (**Figure 11C**).^[155] The rest of the phenyl ring showed only minor interactions with the protein surface. Compound **12** was then superimposed with the co-crystal structure of fucose with BambL where the sugar moiety showed the same orientation and interactions although the two structures were not perfectly aligned (**Figure 11D**).

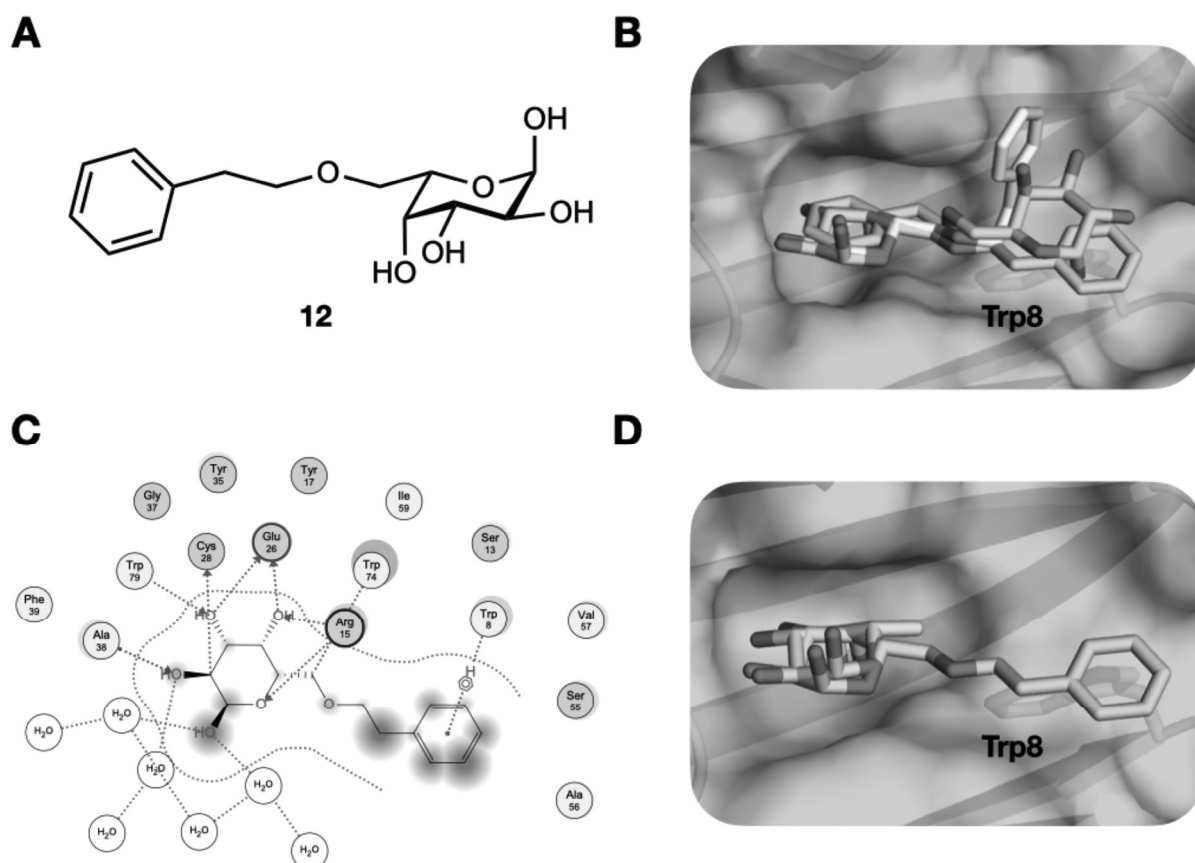


Figure 11: **A)** Designed L-galactose compound **12**. **B)** The top three ranked poses in complex with BambL obtained *via* docking – beige: the phenyl ring bound into the carbohydrate binding site, yellow: second best binding pose of **12** showed tilted phenyl ring without interacting with the surface and the third pose (green) shows the most reasonable orientation of the phenyl residue. **C)** Protein-ligand interaction 2D map of **12** in complex with BambL. The phenyl ring formed a $\text{NH}\cdots\pi$ interaction with Trp8 (purple = polar, purple red = acidic, purple blue = basic, green = hydrophobic, blue arrow = backbone acceptor/ donor, green arrow = sidechain acceptor/ donor, blue = ligand exposure, light blue = receptor exposure, grey line = solvent contract). **D)** Compound **12** (green) docked with BambL and superimposed with fucose (blue) from the co-crystal structure (pdb: 3zw0).

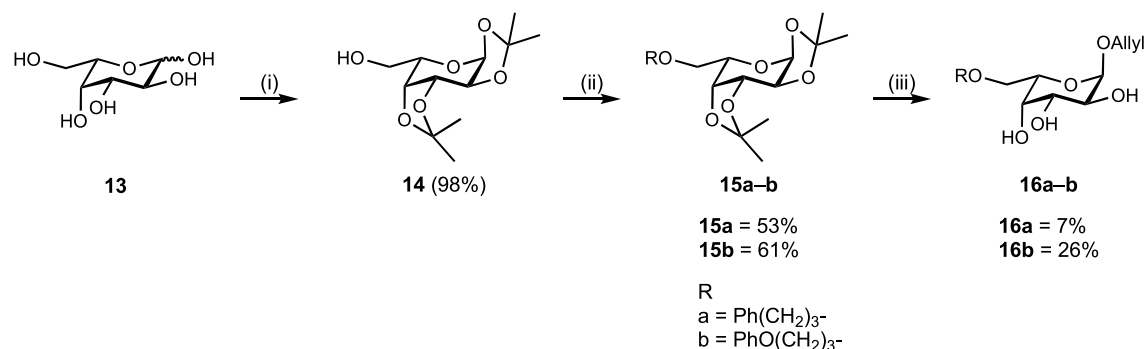
Synthesis Towards Designed Compound 12

In the next step, the designed L-galactose compound **12** was synthesized to confirm the $\text{NH}\cdots\pi$ interaction of the phenyl ring and Trp8. The hemiacetal was protected in the anomeric position with an allyl functional group, which could be either deprotected or modified and extended later. First, the synthesis was established on D-galactose due to a significant cost difference (D-Gal: 1kg/80€, L-Gal: 1 g/260 €, Carbosynth 02/2022) by performing the allylation of D-galactose in the beginning, followed by a derivatization in the second step, but, besides changing the bases, amounts of reagents, adding

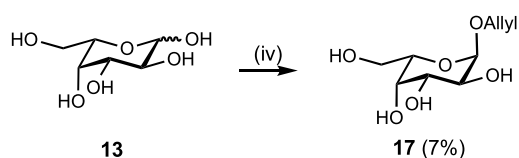
TBAI for a halide exchange and changing the global reaction conditions (heat, time), the reaction did not lead to sufficient yields and was not further used (overall: 0-6% on D-Gal, data not shown). Then, the method from Koch *et. al.*^[156] was tested on D-galactose and later used for the synthesis of two additional L-galactose derivatives. They were designed based on compound **12** but had one/two atoms more in their alkyl chains between O6 and the phenyl ring. Due to a lack of time, the synthesis of the *in silico* compound **12** was not finished. The synthesis of the two derivatives was performed from L-galactose **13** by protecting the *vicinal* hydroxy groups with acetone to afford compound **14** with an unprotected primary hydroxy group in C6. Then, the alkylation of the free hydroxy group was performed under basic conditions leading to the intermediates **15a** and **15b**, which were then deprotected with DOWEX and allylated with allyl alcohol to give the desired L-galactose derivatives **16a** and **16b**.

Allyl- α -L-Gal (**17**) was synthesized by FISCHER glycosylation with allyl alcohol from L-galactose **13** to serve as a control in the biophysical assay. The yield of **17** was low, despite showing a high turnover during the reaction, due to material losses caused by silica purification but mainly in the crystallization and recrystallization step from EtOH afterwards.

Allyl 6-O-Alkyl- α -L-Galactosides



Allyl- α -L-Galactoside



Scheme 2: Synthesis of **allyl 6-O-alkyl- α -L-galactopyranosides (16a-b)**: (i) CuSO₄, H₂SO₄, acetone, 25 °C, 19 h; (ii) KOH, RBr, toluene, 100 °C, 2 d; (iii) Allyl alcohol, DOWEX 50WX4, reflux. 2 d; and **allyl- α -L-galactopyranoside (17)**: (iv) Allyl alcohol, CH₃COCl, 0–70 °C, 16 h.

Biophysical Binding Validation of Synthesized BambL Compounds

The biophysical validation of the L-galactose derivatives on BambL was performed in the competitive binding assay^[153]. Allyl-L-galactopyranoside (**17**) inhibited BambL in the low micromolar range (IC_{50} (**17**) = $7.8 \pm 0.9 \mu M$) showing that extending L-fucose in α -position has no negative impact on the affinity which correlates to the blood groups: A-trisaccharide, H type-2 trisaccharide, and H disaccharide with either $Fuc\alpha(1-2)Gal$ or $Fuc\alpha(1-3)Gal$ linkages, the $Fuc\alpha(1-6)GlcNAc$, $GlcNAc\beta(1-4)[Fuc\alpha(1-6)]GlcNAc$ -1-OMe (IC_{50} = $5.5-10.3 \mu M$)^[63,153] or our *in-house* library compound α -fucopyranosyl benzamide (**3**). The slight affinity loss compared to Me- α -L-Fuc (**1**) could be attributed to the allyl group in comparison to the smaller methoxy group of **1**. The 6-O-phenylpropyl L-galactoside **16a** (IC_{50} (**16a**) = $4.6 \pm 0.5 \mu M$) and 6-O-phenoxypropyl L-galactoside **16b** (IC_{50} (**16b**) = $3.5 \pm 0.4 \mu M$) showed the same activities as the positive control Me- α -L-Fuc (**1**) (IC_{50} (**1**) = $3.1 \pm 0.6 \mu M$) suggesting that the alkyl chain with the phenyl ring of the L-galactose derivatives bind to the surface resulting in lower IC_{50} values compared to allyl α -L-galactoside (**17**). One possibility could be that the additional ether function forms a hydrogen bond with Arg15 or the compounds form a $NH\cdots\pi$ interaction with Trp8, however, in that case, we would expect a higher affinity boost.

Crystallographic Studies on Different BambL Inhibitors

To analyze the interactions of the fucosylated derivatives from our *in-house* library and the synthesized L-galactosides with BambL at atomic level, we co-crystallized BambL with five compounds: α -fucosyl benzamide (**3**), β -fucosylmethyl 2-thiophene-sulfonamide (**6**), 1-C-(3-phenylthio)-propyl- β -D-mannoside (**11**), 6-O-phenoxy- α -allyl-L-galactoside (**16b**), and allyl- α -L-galactoside (**17**). In all cases, the collected diffraction data were in good quality with resolutions and space groups as stated: **3**: 2.0 Å, C 2 2 2₁, **6**: 1.8 Å, P 1 2₁ 1, **11**: 1.7 Å, C 1 2 1 **16b**: 1.4 Å, P 2₁ 2₁ 2₁, and **17**: 1.4 Å, P 2₁ 2₁ 2₁. X-ray data collections and refinement statistics are summarized in **Table S9**.

In case of α -fucopyranosyl benzamide **3**, all six CBSs were occupied and the fucose moiety bound to the *intra/inter* CBSs as reported before.^[63] In the *intra* pocket the aromatic part of ligand **3** stacked to Trp74 in a Y-shaped conformation^[157] where two H atoms of Trp74 are directed towards the aromatic ring (**Figure 12A**). The amide function on the other hand did not interact with the surface of BambL and confirmed

the results from the competitive binding assay that growing the fucosides in α -position is possible but did not improve the affinity towards BambL. The additional stacking interactions were too weak to improve the affinity. In one of the *inter* CBSs, the aromatic ring showed hydrophobic interactions with Ser32 (**Table S10**) but they were not strong enough to stabilize the ring in that positions on BambL's surface due to variations of that part in other CBSs.

In case of the fucosylmethyl sulfonamide **6** in complex with BambL, the asymmetric unit showed two BambL trimers. In the following, only one of the two identical structures was described (**Figure 12A**). All six CBSs (*intra/inter* pockets) were occupied by one molecule of **6** where the fucose moiety showed the same orientations and interactions as previously described.^[63] The ligand was resolved from the electron density for the *intra* CBSs in all cases while two structures from the *inter* CBSs were only solved up to the sulfonamide functional group. In the *intra* CBSs, the thiophene moiety interacted in a parallel offset geometry^[158] with Tyr35 and the nitrogen of the sulfonamide function formed a hydrogen bond with the hydroxy group of Tyr35. On the other hand, the thiophene moiety of **6** from the *inter* CBS formed hydrophobic contacts with Asp75 (**Table S10**). This aromatic interactions of the thiophenes with BambL may explain the good affinity (IC_{50} (**6**) = $4.1 \pm 0.6 \mu M$) that was similar to the value of Me- α -L-Fuc in the competitive binding assay.

In the co-crystal structure of β -C-mannoside **11** in complex with BambL, all six CBSs were occupied and the fucose moiety bound to BambL as reported previously^[63], yet, in one *intra* CBS, the structure was not fully resolved (phenyl ring missing). The alkyl chains of the two resolved ligands **11** from the *intra* pocket showed hydrophobic interactions with Trp74 and the aromatic ring was stabilized by Trp8, Ser13, and Asp30. In the *inter* pocket, the alkyl chain showed contacts with BambL (Ile14, Pro12, Try29 and Trp51) and the phenyl ring showed hydrophobic interactions with His58 and Asp75. The sulfur atom did not interact with the surface in both cases.

Allyl α -L-galactopyranoside (**17**) bound only to the *intra* CBSs of BambL. The allyl ether function pointed from the surface and the oxygen atom was not involved in binding. The vinyl part of the flexible allyl-function was not properly represented by the electron density and showed slightly different orientations for the three ligands but they form π - π -stackings with Trp74. The three hydroxyl groups OH2, OH3 and OH4 and the protons in position H3, H4 and H5, representing the same stereochemical orientations of the

same atoms in fucose which is 6-deoxy-L-galactose, bound in both cases to the same amino acid residues in *intra* CBSs. O6 of **17** was imbedded in the solvate layer and bound to Trp74 *via* a water bridge and was too far ($>3.3\text{\AA}^{[159]}$) from Arg15 to form a hydrogen bond.

The co-crystal structure of α -allyl-6-O-phenoxy-L-galactoside (**16b**) in complex with BambL showed the ligand only in one *intra* CBS. In this structure, the vinyl part of the allyl group showed interactions with Tyr35 instead of Trp74 as shown for ally α -L-galactoside (**17**) representing the flexibility and weak affinity of the allyl group in both cases. The core structure of this molecule showed the same interactions as **17**, the oxygen atom from the ether function was not involved in binding, the flexible alkyl chain showed hydrophobic interactions with Val57 and its phenyl ring had hydrophobic interactions with Ser13 and Asp30.

In summary, the co-crystal structures confirmed that modifications on L-fucose in α -position and on its methyl group in position six are possible, however, the additional interactions of the studied ligands from the *in-house* library with BambL did not contribute to a sufficient affinity increase. The L-galactosides showed binding to the CBSs with the same orientations as L-fucose but the two flexible residues, the allyl group and the alkyl chain, hinder the molecules to form attractive interactions with the surface. The $\text{NH}\cdots\pi$ interaction of the L-galactose series with Trp8 of BambL's *intra* CBS and their phenyl rings was not achieved due to the additional atoms in the linkers in **16a** and **16b** compared to the designed compound **12** and could be adjusted with the established synthesis procedure in the future.

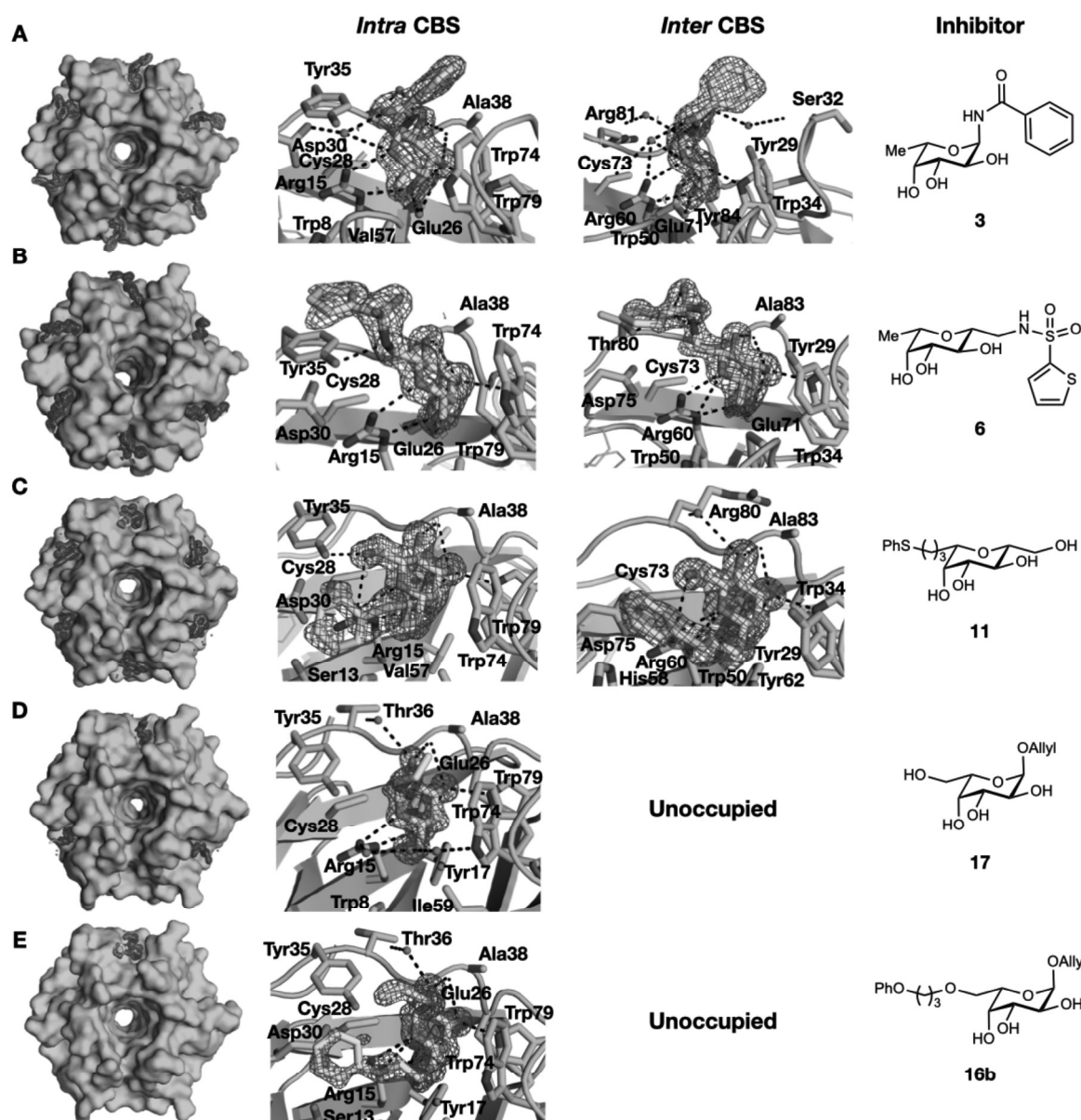


Figure 12: Overall complexes of BambL inhibitors in the *intra/inter* CBSs. The individual interaction of the ligands with BambL are shown for **A)** α -fucosyl benzamide (**3**) **B)** β -fucosylmethyl 2-thiophene-sulfonamide (**6**) **C)** 1-C-(3-phenylthio)-propyl- β -D-mannoside (**11**) **D)** allyl α -L-galactoside (**17**) and **E)** α -allyl-6-O-phenoxy-L-galactoside (**16b**). Electron density is displayed at 1σ , ligands (cyan) and amino acid residues in the binding site are shown as sticks, water molecules in red and calcium ions as green spheres. Dashed lines indicate hydrogen-bonding interactions of the specific ligand with the protein.

Conclusion and Outlook

In this work, we aimed for new *Burkholderia ambifaria* lectin BambL inhibitors by a ligand-based-design approach. For that, we screened our *in-house* library of fucosides with modifications in different positions against BambL resulting in two possible growth vectors on L-fucose, the α -position and its equatorial methyl group (C6). Afterwards,

we designed and studied 6-O-phenylethyl- α -L-galactoside (**12**) *in silico*. The phenyl ring showed a NH $\cdots\pi$ -interaction with Trp8 in the *intra* CBS and the ether function could form a hydrogen bond with Arg15. Two L-galactosides were synthesized with varying linker lengths between the oxygen and the phenyl ring (3 and 4 atoms) and the unstable hemiacetal in **12** was protected by an allyl group which could be either deprotected or modified and extended later. The designed compound **12** was not synthesized due to a lack of time. In addition, we synthesized allyl α -L-galactoside (**17**) as a control that showed a two-fold lower affinity in binding than the positive control Me- α -L-Fuc ($IC_{50(15)} = 3.5 \pm 0.4 \mu M$). The galactosides **16a** and **16b** were as affine as the positive control compensating the affinity loss of the α -allyl functional group by forming hydrophobic contacts with their phenyl residues and the surface of BamBL. Crystallography studies showed hydrophobic interactions of the phenyl residue of **16b** with the surface, however, the alkyl spacer was too flexible and too long with its additional atoms compared to the design compound **12** to form the expected NH $\cdots\pi$ -interaction of its phenyl ring with Trp8 as shown in **Figure 11**. In future, the designed compound **12** has to be synthesized to address the NH $\cdots\pi$ interaction of Trp8 with its phenyl ring. Further, crystallography studies of different fucosides with BamBL confirmed our structure activity relationship from the *in vitro* studies that it is possible to grow fucose in α -position and its equatorial methyl group (C6) to achieve promising BamBL inhibitors.

Material and Methods

Docking

The crystal structure coordinates of BamBL in complex with fucose (pdb: 3zw0) was adjusted for docking in MOE (Molecular Operating Environment, Chemical Computing Group, Canada, version: 2014.0901) by removing all ligands and water molecules by keeping one protomer. The coordinates of the CBS were determined in AutoDockTool^[134] and added to the docking-file. Five amino acids (Arg15, Tyr17, Glu26, Ala38, Trp74, and Trp79) were kept flexible during the docking-run. In parallel, the compounds (α -L-fucose and **12**) for docking were prepared in ChemDraw, saved as a SMILES code and translated into a pdb-file with the online SMILES Translator^[135]. The received pdb-file of the compound was opened in MOE, the bonds were set to the lowest energy level and saved as a mol2-file and added to the docking file. The docking

was performed with Plants1.1^[136] with a binding site radius of 13 Å. To validate the protocol, fucose was docked with BambL and superimposed with its co-crystal structure showing the same interactions with the protein. The results were visualized and analyzed by MOE.

Expression of BambL

BambL from *B. ambifaria* was expressed and purified from *Escherichia coli* BL21 (DE3) and the plasmid pET25b in Anne Imberty's laboratory as described before^[63] that was then provided by Sue Kuhaudomlarp and Elena Shanina.

Competitive Binding Assay

Solid BambL was dissolved in TBS/Ca²⁺-buffer (20 mM Tris, 137 mM NaCl, 2.6 mM KCl, 1 mM CaCl₂, pH = 7.4) and the concentration was determined ($M = 9381.21$ g/mol; $\varepsilon = 40575$ M⁻¹ cm⁻¹^[138] with NanoDrop2000c (ThermoFisher Scientific).

The competitive binding assay based on fluorescence polarization was performed in analogy to the described procedure before.^[60] To a solution (10 µL) containing BambL (300 nM), a fluorescent reporter ligand (*N*-(fluorescein-5-yl)-*N'*-(α -L-fucopyranosyl ethylene)-thiocarbamide, 20 nM) in PBS-buffer (140 mM NaCl, 2.7 mM KCl, 1.8 mM KH₂PO₄, 10.0 mM Na₂HPO₄·7 H₂O) was added to a 10 µL serial dilution of the compounds in the same buffer (100 – 0.78 µM, dilution factor 2) in triplicates in a 384-black well plate (Greiner Bio-One, Germany, cat. no. 781900). The plate was sealed (EASYseal, Greiner Bio-One, cat. no. 676001), centrifuged (3000 rpm, r.t., 1 min) and incubated in a dark chamber for 1 h and 16 h. Afterwards, the fluorescence intensities were measured and the polarization was calculated with a PHERAstar FS (BMG Labtech, Germany, filter ex.: 485 nm em.: 535 nm). The data were analyzed using the four parameter variable slope model in MARS software (BMG Labtech). The control data (PBS-buffer with BambL) were subtracted from the measured values and the polarization plateaus were set according to the positive control, Me- α -L-Fuc, and the negative control (PBS-buffer with BambL plus reporter ligand). The graphs were visualized with Prism. The experiments were performed in three independent experiments.

X-ray Crystallography

BambL (9–10 mg mL⁻¹ in 20 mM tris-HCl pH = 7.5, 150 mM NaCl) was incubated in a 19:1 ratio with 2.5 mM of **3**, **6**, **11**, **16b** or **17** and additional DMSO in case of **3** (5%), **16b** (5%), and **17** (1.25%) 1h prior to crystallization. Hanging drop vapor diffusion

method using 1 μ L of protein-ligand mixture with 1 μ L of reservoir solution at 19 °C in a 24-well plate yielded crystals after 1–3 days. For the Bambl complexes, data were collected from plate crystals obtained under different conditions (Bambl-**3**: 4.3 M NaCl and 100 mM HEPES pH = 7.5 (SG1 screening kit, condition 2-48, Molecular Dimensions); Bambl-**6**: 24% PEG 4K and 10% *i*PrOH (SG1 screening kit, condition 1-6, Molecular Dimensions); Bambl-**12**: 30% PEG 2K MME, 200 mM (NH₄)₂SO₄ and 100 mM NaOAc pH 4.6 (SG1 screening kit, condition 1–45, Molecular Dimensions); Bambl-**16b**: 20% PEG 10K and 100 mM HEPES pH = 7.5; Bambl-**17**: 26% PEG 2K MME and 100 mM NaOAc pH = 4.6). The crystals were mounted in a cryoloop and flash-frozen in liquid nitrogen. Diffraction data was collected at 100 K at Synchrotron SOLEIL (Paris, France) on beamline Proxima 2 using a EIGER X 9M area detector. The data were processed using XDS.^[107] All further computing was performed using the CCP4 suite.^[140] Five percent of the observations were set aside for crossvalidation analysis, and hydrogen atoms were added in their riding positions and used for geometry and structure-factor calculations. The structure was solved by molecular replacement using PHASER.^[141] For the complexes Bambl-**3**, Bambl-**6**, Bambl-**11**, Bambl-**16b**, and Bambl-**17**, the coordinates of the 3zw0 trimer were used as a search model to search for one trimer in the asymmetric unit. The structures were refined with restrained maximum likelihood refinement using REFMAC 5.8^[160] iterated with manual rebuilding in Coot.^[142] Ligand libraries were created using JLigand. The ligands were introduced after inspection of the 2Fo–DFc weighted maps. Water molecules, introduced automatically using Coot, were inspected manually. The stereochemical quality of the models was assessed with the PDB Validation Server. Data quality and refinement statistics are summarized in Supporting Information, **Table S9**. The complexes were visualized with PyMOL and analyzed there and in LigPlot^[161].

Supporting Information

The supporting information contains the synthetic procedure of the L-galactose derivatives including their analyses and NMR spectra, the dose-dependent validation of the *in-house* library compounds and the L-galactose derivatives in the competitive binding assay based on fluorescence polarization, the interactions of the second best pose from the *in silico* studies of **12** in complex with Bambl, and the x-ray data collection and refinement statistics of the individual complexes and their 2D presentations.

Conclusion and Outlook

The biofilm is a great defense mechanism of bacteria against cell eradications by external threats. In addition to limited penetration of antibiotics through the biofilm, compounds also encounter persister cells that hinder the drug uptake into the cells. These limitations have to be overcome especially for immunocompromised or cystic fibrosis patients who suffer from chronic wound infections, pneumonia or sinus infections caused by *Pseudomonas aeruginosa* or bacteria belonging to the *Burkholderia cepacia* complex.

This work focused on the inhibition of lectins that play an important role in the biofilm formation and might reduce the biofilm growth. This idea has already been used in our group, which led to the development of potent fucose-mannose hybrid glycomimetic inhibitors for the *P. aeruginosa* lectin LecB by rational ligand-based design with promising results in pharmacokinetics and reduction of the biofilm thickness.^[80]

In this work, different approaches for finding new inhibitors for the extracellular *P. aeruginosa* lectins, LecA and LecB, and for the *B. ambifaria* lectin, BamBL were investigated. *In silico* screening and biophysical validations were performed to identify structural motifs of new non-carbohydrate inhibitors against LecA that can be further optimized to overcome chemical limitations. A new binding pocket of LecA, the central pocket, was identified, analyzed and addressed with new ligands. In addition, rational-ligand-based-design was applied for the identification of new LecB and BamBL inhibitors.

For LecA, catechols were identified as the first non-carbohydrate lectin inhibitor class with affinities in the low millimolar range. They are a general binding motif to inhibit additional Ca^{2+} -ion dependent lectins such as langerin and an alternative to carbohydrate based glycomimetics. Although, catechols are considered as PAINS, a careful validation in different biophysical experiments and later by co-crystal studies showed that electron-deficient catechols are stable in presence of LecA. Further optimization of this compound class improved the binding affinity to micromolar affinities as shown for tolcapone ($K_d = 181 \pm 6 \mu\text{M}$). An *in vitro* screening of 3564 catechol-like compounds revealed 48 tolcapone-like hits with peak affinities in the range of pNPG ($K_d = 14.1 \mu\text{M}$ ^[102]). The compounds will be further analyzed in different biophysical assays and later tested against the *P. aeruginosa* biofilm.

In addition, the central pocket between parallel LecA monomers was analyzed and investigated as a new target. Different galactosides carrying an additional aromatic ring were synthesized to reach the central pocket of which one compound was co-crystallized with LecA. It was shown that the aromatic part was in proximity to the central pocket, however, not inside as hypothesized. Compounds with longer spacers between the core structure and the aromatic residue improved the binding affinity, suggesting an increased interaction with the pocket. This has to be confirmed by additional crystallography studies in the future.

Rational ligand-based design led to highly potent monovalent LecB inhibitors in the two digit nanomolar range ($IC_{50} = 85 \pm 16$ nM). The micromolar affinity fucose-mannose hybrids served as models that showed reduction of the biofilm growth. The original sugar moiety was replaced by fucose, the more potent natural LecB ligand, and the linker between the carbohydrate and their aglycons was removed. The resulting *N*-fucosyl amide class showed binding of their amide nitrogen to Ser22 *via* a hydrogen bond or Asp96 *via* an ion-dipole interaction and hydrophobic contacts with their aromatic residues with Gly24 and Val69 from the surface. In addition, most compounds were stable in murine and human blood plasma and in murine and human liver microsomes. Those *N*-fucosyl amides are a promising new LecB inhibitor class against *P. aeruginosa* biofilms.

For BamBL, ligand-based design led to new inhibitors that were as potent as the natural ligand (α MeFuc). Structure-activity-relationship studies of fucosides, modified in different positions and tested in a competitive binding assay based on fluorescence polarization, revealed the possibility to modify fucose in α -position and on its equatorial methyl group (C6). Further *in silico* studies demonstrated that an aromatic aglycon with an ether function in its linker was favored. Thereupon, L-galactosides (fucosides with an additional hydroxy group on C6) were synthesized and afterwards tested *in vitro* resulting in affinities in the single-digit micromolar range ($IC_{50} = 4\text{--}8$ μ M). Crystallography revealed that those galactosides bound to BamBL with the same interactions as fucose, however, the aromatic aglycons bound only loosely to the proteins' surface.

All these pharmacophore studies on the *P. aeruginosa* lectins, LecA, LecB and *B. ambifaria* lectin BamBL with monovalent ligands that are comparably easy to access, helped to understand how the ligands interact at the atomic level. Promising

ligands such as the *N*-fucosyl amides for LecB or the potent non-carbohydrate tolcapone-like LecA inhibitors will be further improved to lead structures. Later, the ligands can be modified to become divalent or even multivalent inhibitors connecting several binding sites which will boost their affinities up to three orders in magnitude.

Overall, pathoblocker approaches where an inhibitor competes with the natural ligands over the active binding sites of lectins are very promising. This inhibition might not only reduce the amount of biofilm formation by interfering with bacterial adhesion, but it might also hamper the formation of persister cells since the biofilm is smaller and starvation of planktonic cells is prevented. There are already approaches of antibiotic-carbohydrate conjugates where the carbohydrate function prevents biofilm formation and the antibiotic kills the cells.^[162] This increases the chance of killing bacteria with traditional antibiotics by treating a patient with a combination of a biofilm blocker together with antibiotics.



Supporting Information

Non-Carbohydrate Glycomimetics as Inhibitors of Calcium(II)-Binding Lectins

Sakonwan Kuhaudomlarp⁺, Eike Siebs⁺, Elena Shanina, Jérémie Topin, Ines Joachim, Priscila da Silva Figueiredo Celestino Gomes, Annabelle Varrot, Didier Rognan, Christoph Rademacher, Anne Imberty, and Alexander Titz**

ange_202013217_sm_miscellaneous_information.pdf

Supplementary information

Materials and Methods

Computational methods

Protein preparation

Several crystal structures of LecA complexed with a galactoside ligand were examined (LecA-GalBG0 PDB entry 3ZYH,^[1] LecA-Gal PDB entry 1OKO,^[2] LecA-pNPGal PDB entry 3ZYF,^[1] and LecA-GalAG0 PDB entry 3ZYG^[1]). The protein structure is highly conserved with a RMSD measured on all atoms smaller than 0.5 Å. Particularly, no deviation is observed in the binding pocket. The crystal structure of LecA complexed with a peptide-linked propyl-thiogalactoside derivative was chosen due to its high resolution (PDB code 3ZYH, 1.5 Å).^[1] The protein was prepared following a previously reported protocol^[3] using the Protein Preparation Wizard tool in the Maestro suite Version 9.5. Only chain A, the calcium ion and the ligand (ethyl thiogalactoside corresponding to the fragment with electron density) were considered while all other molecules, ions and water were removed. The ligand was kept in the active site in order to avoid any structural deviation during the minimization phase. The residues' protonation states were modelled at neutral pH using PropKa Version 2.^[4] His48 was predicted to be protonated and His50 was considered as neutral, with the proton carried by nitrogen N_ε. The structure of the complex was minimized (RMSD 0.3 Å) using impref module^[5] with OPLS2005 force field.^[6]

Ligand preparation

The ligands were taken from the National Cancer Institute (NCI) diversity set IV (<https://dtp.cancer.gov/>). All ligands were prepared using LigPrep module (Version 2.3). The OPLS2005 force field was used. Ionization states were generated using Epik at neutral pH.^[7] 32 stereoisomers were generated while retaining the chirality when specified, and varying other non-chiral centres.

Grid generation

The grid was built with the default parameters around the centre of mass of the crystallographic ligand. The final grid consists of a cube of 20.6 Å length. No scaling factor was used for protein van der Waals radii. Five constraints were defined according to the specificity of the interactions between the galactose moiety and the binding site: four hydrogen bonds involving the following atoms (Asn107 OD1, Asp100 OD1, His50 NE2 and Gln53 OE1) and coordination with the calcium ion. These constraints ensure that the ligand will be docked into the binding site with a similar conformation of the galactose. During the docking phase, the ligand should fulfil at least one of these constraints.

Docking

The ligands were docked using Glide^[8] in Simple Precision (SP) mode with default parameters. The protein was considered rigid while the ligands were flexible. Strain correction terms were applied. These corrections allow identifying ligands with a large strain in the docking site. Ligands with more than 4 kcal/mol energy between the docked and free conformations received penalties. The distance dependent dielectric constant was set to 2. Epik state penalties were added to the docking score.

Sorting the poses

The poses were ranked according their E-model score. The first 100 best ligands were selected (taking a maximum of 3 conformers per ligand). The binding free energy was further estimated using the MM-GBSA protocol available in Prime software (Schrödinger). This protocol based on the molecular mechanics (MM) OPLS-2005 force field and the Generalized Born Surface Accessible (GBSA) continuum model has already shown good results in ranking ligands for lectin.^[9] The binding free energy of each ligand $\Delta G_{\text{MM-GBSA}}$ is described as the difference in the sums of MM energy, GBSA solvation energy (including surface area energy SA) of the complex on one hand and of the free ligand and free protein on the other hand. Strain energies were calculated for both the protein and the ligand. This feature is directly implemented in the Schrödinger suite and consists of quantifying the difference in energy between the free and the complex state of, both, the ligand and the protein after energy minimization. The energy minimization of the protein was performed using Prime by allowing the sidechains within 10 Å of the ligand to move, in order to avoid large structure deviation. The ranking was eventually made according the highest absolute value of the calculated $\Delta G_{\text{MM-GBSA}}$ of binding.

Virtual screening

The protocol was tested by re-docking the four crystallographic ligands. The obtained complexes are compared with the crystallographic structure in Figure S1. In all cases the galactose moiety is found superimposed with the crystallographic reference.

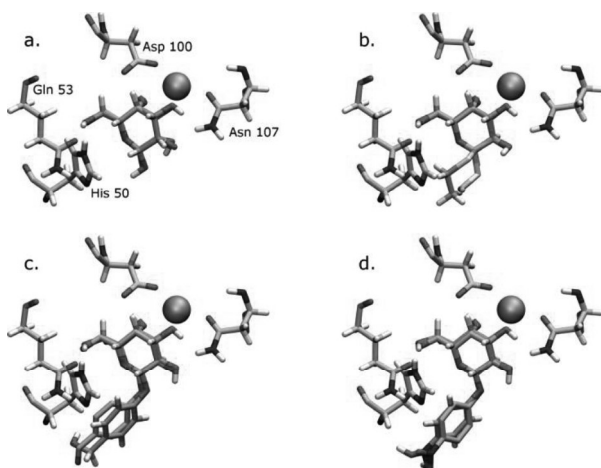


Figure S1: Superimposition of the four docked ligands with the corresponding crystallographic structure (a. D-Gal, b. GalBG0, c. GalAG0, d. pNPGal). The color codes for the ligand is as follows: carbon, hydrogen, oxygen, sulfur, nitrogen atoms in cyan, white, red, yellow and blue respectively. The carbon atoms of the crystallographic references are colored in orange, while the residues involved in hydrogen bonds have their carbon atoms colored in lime.

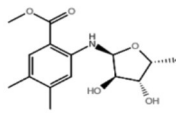
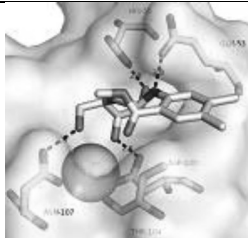
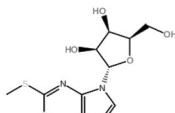
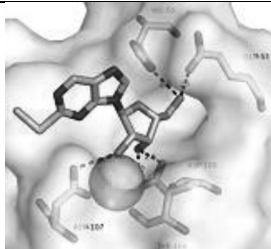
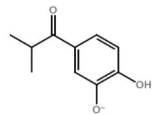
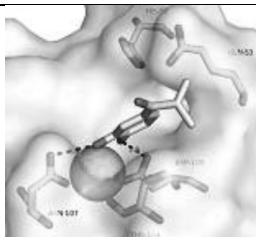
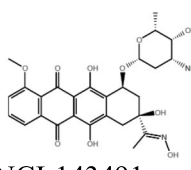
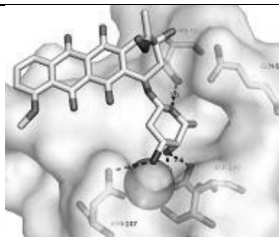
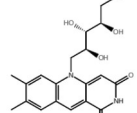
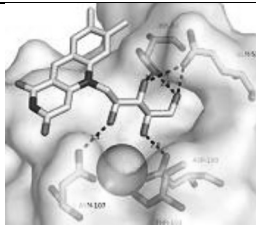
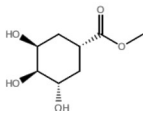
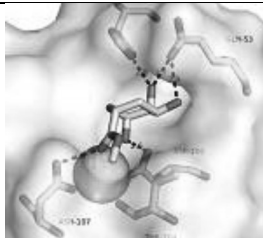
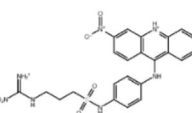

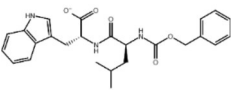
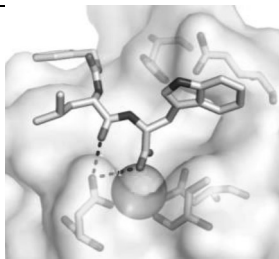
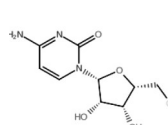
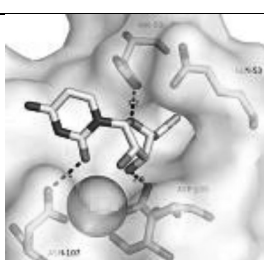
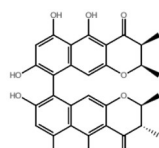
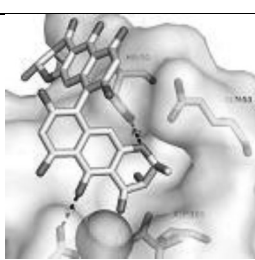
Cluster #Representative (code & Emin)	Docked	Cluster #Representative (code & Emin)	Docked
1  NCI-112203 (E = -73.0)		6  NCI-25740 (E = -62.2)	
2  NCI-27389 (E=-48.1)		7  NCI-143491 (E=-57.6)	
3  NCI-275266 (E=-64.4)		8  NCI-62511 (E=-57.5)	
4  NCI-177365 (E=-64.0)		9  NCI-335979 (E=-56.3)	
5  NCI-318799 (E=-63.7)		10  NCI-345647 (E=-53.9)	

Figure S2: Representation of the 10 lowest energy clusters. The E-model score (E) is given in kcal/mol

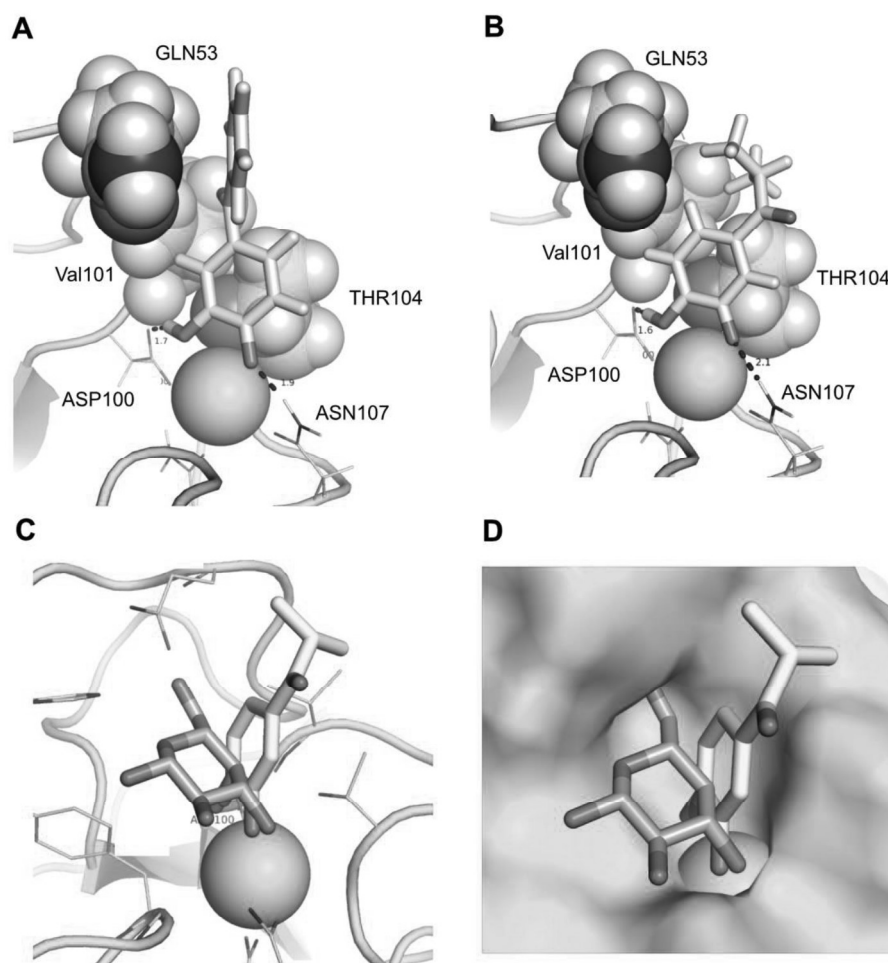


Figure S3: Docking of the two positive catechol derivative hits on LecA, A: 4-benzoylcatechol **23**, B: 4-isobutyryl catechol **21**. The coordination of the two compounds *in silico* occurs the same way: The catechol moiety binds with its two hydroxy groups to the calcium ion, with one hydroxy group to Asp100 and with the other one to Asn107. C. Superimposition of catechol **21** with galactose binding to LecA (PDB code: 1OKO). D. Surface representation of panel C.

Table S1: 60 top compounds from the virtual screening. Specific binding features are gathered (X corresponds to a positive feature while blank means an absence of interaction).

Num NCI	Formula	Test	CAS	$\Delta G_{\text{MM-GBSA}}$	Coord Ca ²⁺	H-bond					Cluster
						H50	Q53	D100	T104	Hydroph	
268251	C29H49N7O5		67642-36-8	-80.674	1	1	1	2	1	XX	
112203	C15H21NO6	YES	26909-51-3	-73.022	2	1	1	2	1	XX	1
19803	C21H20O12	YES	17912-87-7	-71.124	2		1	1	1	X	2
275266	C19H22N2O6		65176-88-7	-64.376	2	2	2	1	1	x	3
177365	C23H23N7O4S		63345-17-5	-64.024	2		1		1	X	4
318799	C9H12CIN3O4	YES	32659-31-7	-63.658	2	1		1	1	X	5
89821	C20H27CIN4O3		2499-43-6	-63.463	1		1	1	1		
25740	C11H14N4O4S	YES	27640-80-8	-62.2	2	1	1	1	1	x	6
352890	C11H14N4O4	YES	77691-03-3	-60.22	2	1	1	2	1	x	6
46385	C16H18N6O4S	YES	13153-28-1	-59.909	2	1	1	1	1	x	6
143491	C27H30N2O10	YES	34610-60-1	-57.604	2	1		2	1	x	7
62511	C8H14O5	YES	(None)	-57.488	2	1	2	1	1	x	8
335979	C25H29N3O5		53262-00-3	-56.291	2				2		9
316458	C11H13N5O3	YES	(None)	-56.222	2	1	1	1	1	x	6
370383	C11H12N4O4	YES	91296-23-0	-55.688	2	1	1	1	1	XX	1
2561	C12H16O5	YES	5329-50-0	-55.021	2	1		1	2	x	Gal
105827	C12H15N5O4S	YES	22242-90-6	-54.089	2	1	1	1	1	XX	1
345647	C30H26O10		(None)	-53.93	2	1			1		10
349156	C14H15NO8	YES	96203-70-2	-53.569	2	1	1	1	2	x	7
70307	C12H17N3O4	YES	(None)	-53.088	1				1		
5856	C23H20N4O2S2	YES	(None)	-53.028	1				1		
107022	C12H6O8	YES	568-80-9	-52.664	2			2	1	XX	2
121182	C10H14N6O3	YES	24807-84-9	-51.72	2	1	1	1	1	x	6
215718	C18H13BrN2O2S		(None)	-51.673	1				2	XX	
638080	C10H8N2O4		(None)	-50.95	2			2	1	XX	2
128606	C27H17NO5		27021-99-4	-50.553	1				1		
45383	C25H22N4O8	YES	3930-19-6	-50.235	1				1	x	
176324	C5H7CIN2O4	YES	(None)	-49.5	2	1		1	1		11
95204	C10H19N4O3PS	YES	13259-52-4	-49.034	1		2	1	1		
244387	C17H28O6		469-77-2	-48.776	2		1		1	X	9
265372	C12H20N6O		(None)	-48.193	1			1	2		
27389	C10H12O3	YES	5466-89-7	-48.074	2			1	1	XX	2
98026	C17H16O5		(None)	-47.693	1				1	X	

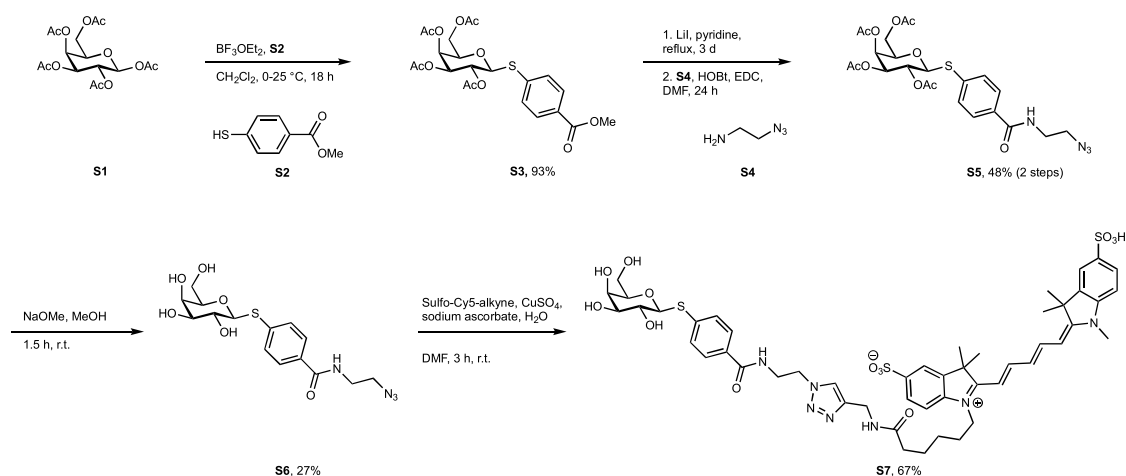
76350	C21H22N2O6	(None)	-47.452	2			1	XX	9
270063	C15H14ClNO4	35071-94-4	-47.413	2			2	XX	9
133075	C15H13N5O3S	28558-60-3	-46.41	1		1	1		
36586	C15H10O5	YES 446-72-0	-46.128	1			1	2	x
241998	C14H18N2O5	36200-99-4	-45.698	1	1		1	1	X
89759	C15H14O4	YES 487-49-0	-45.482	1			1		XX
3001	C11H18O4	YES 5345-12-0	-45.057	1			1	1	XX
73735	C28H24N2O6	YES 52962-95-5	-44.551	2				1	XX 9
91529	C25H24O12	YES 1182-34-9	-44.515	2		1		1	X 9
49643	C16H20O5	YES 58661-24-8	-44.507	1		1			X
372221	C12H14N4O	74526-21-9	-43.633	1		1	1	1	
83950	C25H22N4O8	YES 3930-19-6	-43.614	1					XX
25368	C11H20O4	YES 06/04/7248	-43.496	2			1	1	XX 12
82560	C10H11N7O3S	2877-68-1	-43.485	1	1	1	1	1	
144982	C12H18O4	(None)	-42.782	1	1			1	XX
37408	C8H6O4	YES 6272-27-1	-42.768	2			1	1	XX 2
194308	C12H22N2O3S2	90379-42-3	-42.749	2		1	1	1	XX 13
120913	C14H18N2O3	YES (None)	-42.437	1	1	1		1	XX
303603	C7H8N2O5	19561-70-7	-42.181	1			1	1	X 3
22939	C12H20N2O2	YES (None)	-41.596	1				2	XX
127133	C27H18N2O4	(None)	-41.243	2				2	XX 14
32838	C11H10N2O2	YES 6252-98-8	-41.2	1		1	1		X
107679	C13H18ClN3O4S2	YES 742-20-1	-40.894	2		1	1	1	X 15
13434	C12H15NO6	5450-16-8	-40.825	2	1			1	XX 14
60548	C5H5NO4S	YES 6628-19-9	-40.666	2			1	1	11
303294	C15H13N3O3S2	63018-28-0	-40.607	1	1			2	X
22847	C8H13NO3	YES 4854-46-0	-40.473	2				1	XX 9

Experimental methods

Synthesis of fluorescent reporter ligand Sulfo-Cy5-Gal S7

Chemical synthesis

Synthesis of the Cy5-linked galactoside **S7** as a primary fluorescent ligand for the competitive binding assay is described in Scheme S1. The interaction of **S7** with LecA in a direct titration experiment is shown in Figure S4.



Scheme S1: Synthesis of red-shifted reporter ligand **S7**.

Commercial chemicals and solvents were bought from TCI Chemicals, Sigma Aldrich, CarlRoth or Lumiprobe and used without further purification. Deuterated solvents were purchased from Eurisotop (Saarbrücken, Germany). The reactions were followed by thin layer chromatography (TLC) by using silica gel 60 Å coated on an aluminum plate (Macherey & Nagel, Düren, Germany), visualized by UV light (254 nm) and stained with either a molybdate solution (0.02 M of ammonium cerium sulfate dihydrate and ammonium molybdate tetrahydrate in aq. 10% H_2SO_4) or a KMnO_4 -solution (1.5 g KMnO_4 , 10 g K_2CO_3 and 1.25 mL 10% NaOH in 200 mL water). Furthermore, the reaction progress was controlled by high pressure liquid chromatography (HPLC) with mass spectroscopy on a Thermo Dionex Ultimate 3000 HPLC coupled to a Bruker amazon SL mass spectrometer, with UV detection at 254 nm using a RP-18 column (100/2 Nucleoshell RP18plus, 2.7 μM from Macherey-Nagel, Germany)

as stationary phase. Purification was performed by flash chromatography (TeledyneIsco Combiflash Rf200) on silica gel 60 Å (400 mesh particle size, Fluka). Products **S6** and **S7** were purified by the preparative HPLC using a 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). Nuclear magnetic resonance (NMR) spectroscopy was performed on a Bruker Avance III 500 UltraShield spectrometer at 500 MHz for ¹H-NMR and 126 MHz for ¹³C-NMR spectra. Chemical shifts given in ppm were calibrated on the residual solvent signals according to literature:^[10] CDCl₃ = 7.26 ppm and 77.0 ppm, DMSO-d₆ = 2.50 ppm and 39.52 ppm. Multiplicity is given as s singlet, d doublet, t triplet, q quartet and m multiplet. High resolution mass spectra (HRMS) were measured on an Ultimate 3000 UPLC system coupled to a Q exactive Focus Orbitrap system with HESI source (Thermo Fisher, Dreieich, Germany).

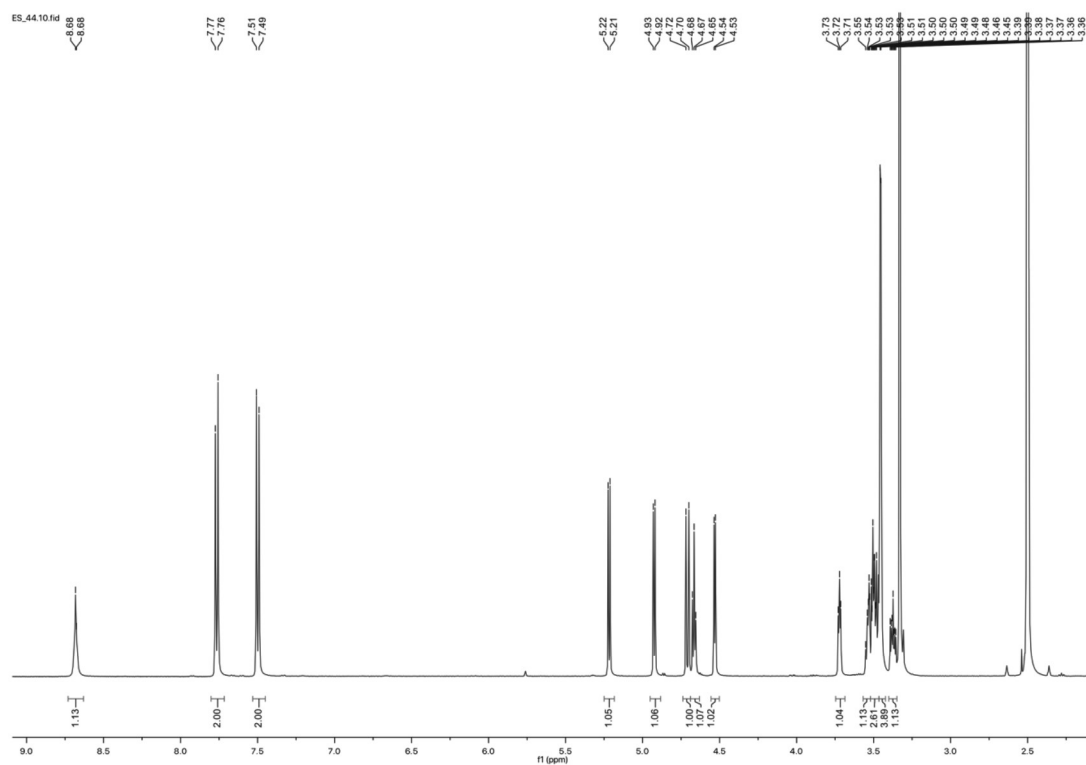
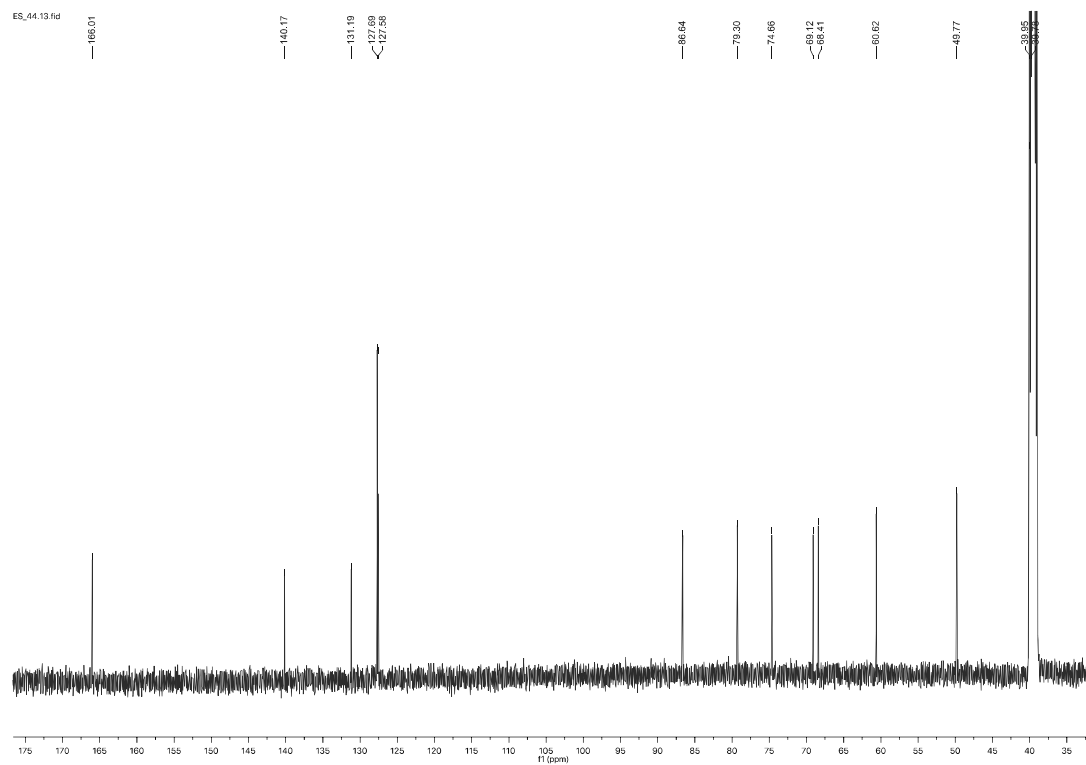
4-Methyloxycarbonylphenyl 2,3,4,6-O-acetyl 1-thio-β-D-galactopyranoside (S3). To galactose pentaacetate (**S1**, 1.50 g, 1.0 eq., 0.25 M in CH₂Cl₂), BF₃·OEt₂ (1.95 mL, 15.3 mmol, 4.0 eq.) dissolved in CH₂Cl₂ (24 mL) was added methyl-4-sulfanylbenzoate (**S2**, 3.0 eq., 8.91 mmol, 0.4 M in CH₂Cl₂). **S3** was synthesized following the procedure from Winssinger *et al.*^[2] 1.37 g (2.75 mmol, 93%) **S3** was obtained (*R*_f = 0.4 in EtOAc/Toluene 1/5). ¹H NMR (500 MHz, CDCl₃) δ 7.96 (d, *J* = 8.5 Hz, 2H, 2x ArH), 7.52 (d, *J* = 8.5 Hz, 2H, 2x ArH), 5.45 (dd, *J* = 3.4, 1.1 Hz, 1H, H-4), 5.28 (t, *J* = 10.0 Hz, 1H, H-2), 5.08 (dd, *J* = 9.9, 3.4 Hz, 1H, H-3), 4.83 (d, *J* = 10.0 Hz, 1H, H-1), 4.23 – 4.10 (m, 2H, H-6), 4.00 (ddd, *J* = 7.0, 5.8, 1.2 Hz, 1H, H-5), 3.91 (s, 3H, OCH₃), 2.15 (s, 3H, CH₃), 2.08 (s, 3H, CH₃), 2.06 (s, 3H, CH₃), 1.98 (s, 3H, CH₃). ¹³C NMR (126 MHz, CDCl₃) δ 170.53 (C=O), 170.31 (C=O), 170.17 (C=O), 169.56 (C=O), 166.66 (CO₂CH₃), 139.44 (ArC), 130.47 (2x ArCH), 130.11 (2x ArCH), 129.30 (ArC), 85.67 (C-1), 74.81 (C-5), 72.02 (C-3), 67.31 (C-4), 67.12 (C-2), 61.87 (C-6), 52.40 (OCH₃), 20.94 (CH₃), 20.85 (CH₃), 20.81 (CH₃), 20.72 (CH₃). HR-MS (ESI): calcd for C₂₂H₂₆O₁₁S [C₂₂H₂₆O₁₁S - H]⁻ 497.1112; found 497.1118.

2-Azido ethylamine hydrobromide (S4). 2-Azidoethylamine was synthesized following the procedure of Winssinger *et al.*^[11] Crude **S4** was used without further purification. ¹H NMR matches the reported data. ¹³C NMR (126 MHz, DMSO-*d*₆) δ 48.53 (NH₃CH₂CH₂N₃), 38.52 (NH₃CH₂CH₂N₃).

Azidoethyl derivative S5. **S5** was obtained following the procedure from Winssinger *et al.*^[2] Deprotection of **S3** (450 mg, 1.0 eq., 0.3 M in pyridine) using LiI (3.0 eq.) was achieved by

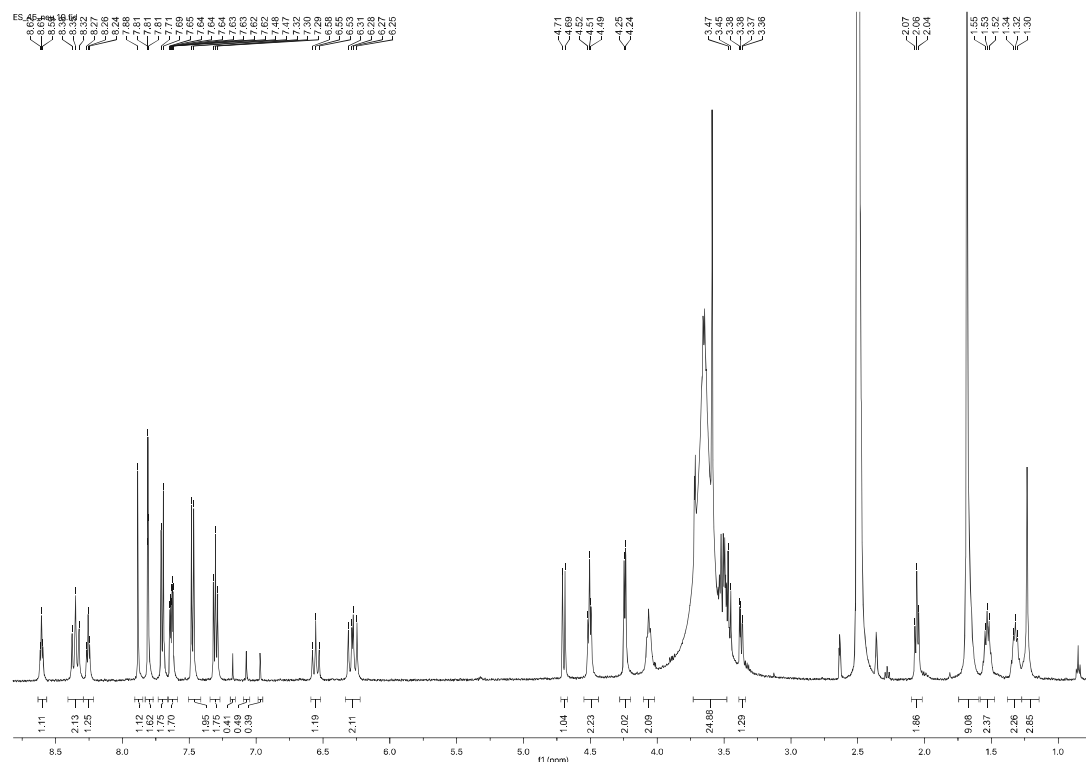
heating to reflux for 3 d. After workup, the crude acid (HR-MS (ESI): calcd for acid intermediate $[\text{C}_{21}\text{H}_{24}\text{O}_{11}\text{S} - \text{H}]^+$ 483.0967; found 483.0959.) was coupled with 2-azidoethylamine hydrobromide (**S4**, 4 eq) using EDC*HCl/HOBt (3.0 eq.). Purification by silica gel chromatography (PE/EtOAc 2/3, $R_f = 0.40$) gave 240 mg of pure **S5** after lyophilization (48% yield over 2 steps). ^1H NMR and ^{13}C NMR match the reported data. HR-MS (ESI): calcd for **S5** $[\text{C}_{23}\text{H}_{28}\text{N}_4\text{O}_{10}\text{S} + \text{H}]^+$ 553.1599; found 553.1592.

Deprotected azidoethyl galactoside S6. Compound **S5** (30 mg, 0.054 mmol, 1.0 eq.) was dissolved in MeOH (4 mL, $c = 0.01$ M), NaOMe (2 μL , 5.33 M in MeOH) was added and the reaction was stirred for 3 h at r.t.. The solution was neutralized with Amberlite (IR120/ H^+), filtered and the solvent was evaporated. The crude product was purified by preparative HPLC to give the title compound (5.7 mg, 0.015 mmol, 27%). ^1H NMR (500 MHz, $\text{DMSO}-d_6$) δ 8.72-8.65 (m, 1H, NH), 7.77 (d, $J = 8.5$ Hz, 2H, 2x ArH), 7.50 (d, $J = 8.5$ Hz, 2H, 2x ArH), 5.22 (d, $J = 6.0$ Hz, 1H, OH-2), 4.93 (d, $J = 5.6$ Hz, 1H, OH-3), 4.71 (d, $J = 9.6$ Hz, 1H, H-1), 4.67 (t, $J = 5.4$ Hz, 1H, OH-6), 4.54 (d, $J = 4.4$ Hz, 1H, OH-4), 3.72 (t, $J = 3.9$ Hz, 1H, H-4), 3.56 – 3.47 (m, 4H, H-6, H-2, H-5), 3.47-3.43 (m, 4H, $\text{NHCH}_2\text{CH}_2\text{N}_3$), 3.37 (ddd, $J = 9.0, 5.6, 3.3$ Hz, 2H, H-3). ^{13}C NMR (126 MHz, DMSO) δ 166.02 ($\text{C}=\text{O}$), 140.17 (ArC), 131.19 (ArC), 127.70 (2x ArCH), 127.59 (2x ArCH), 86.64 (C-1), 79.30 (C-5), 74.66 (C-3), 69.12 (C-2), 68.41 (C-4), 60.63 (C-6), 49.78 ($\text{NHCH}_2\text{CH}_2\text{N}_3$), 39.78 ($\text{NHCH}_2\text{CH}_2\text{N}_3$). HR-MS (ESI): calcd for $[\text{C}_{15}\text{H}_{20}\text{N}_4\text{O}_6\text{S} + \text{Na}]^+$ 407.0996; found 407.0989.

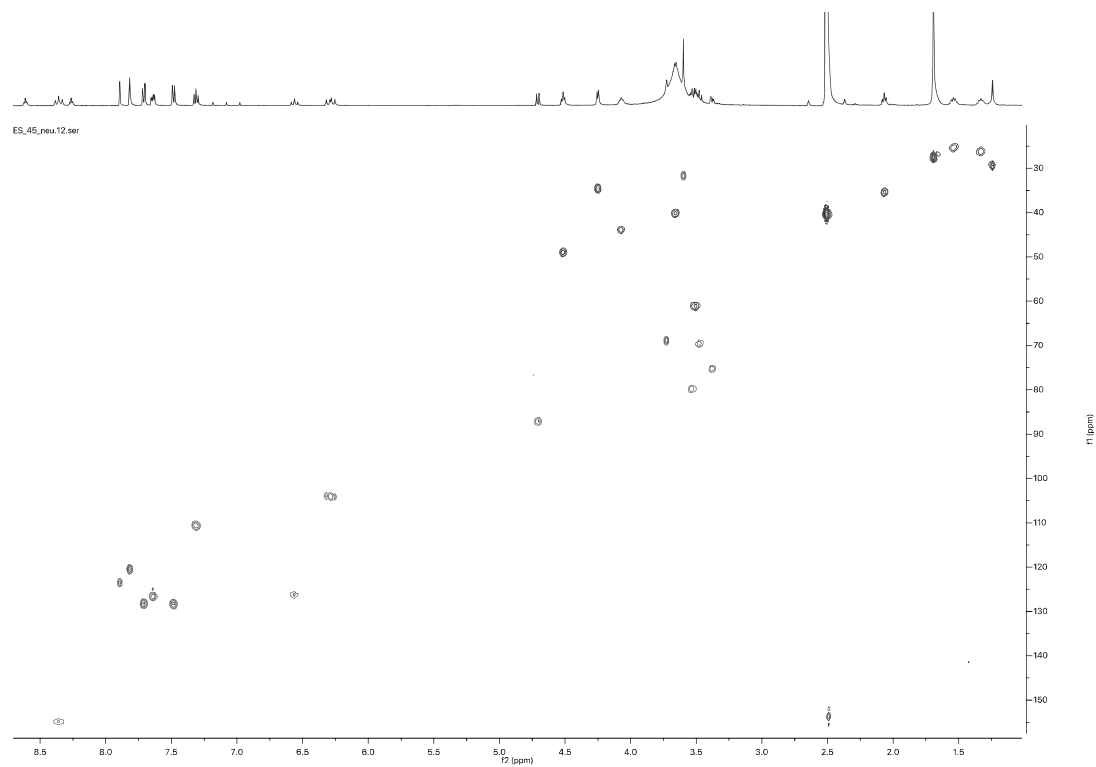
¹H NMR of **S6**¹³C NMR of **S6**

Sulfo-Cy5-galactoside S7. Azide **S6** (2.2 mg, 5.8 μmol) and Sulfo-Cyanine-5-alkyne (5.0 mg, 7 μmol , art. no. B33B0, Lumiprobe, Germany) were dissolved in dry and degassed DMF (0.4 mL). Sodium ascorbate (29 μL of a 100 mM stock solution in degassed DMF, 2.9 μmol , 0.5 eq.) and CuSO_4 (29 μL of a 100 mM stock solution in degassed DMF, 2.9 μmol) were added. The reaction was stirred at r.t. and monitored by LCMS. After 3 h, the solvent was evaporated under reduced pressure. The crude product was taken up in a 1:1 mixture of $\text{H}_2\text{O}/\text{CH}_3\text{CN}$ and purified by preparative HPLC. Fractions containing the product were pooled and lyophilized. The title compound was obtained as a blue solid (4.2 mg, 3.9 μmol , 67%). ^1H NMR (500 MHz, $\text{DMSO}-d_6$) δ 8.61 (t, J = 5.6 Hz, 1H, ArCONH), 8.40-8.30 (m, 2H, 2x indoline-H), 8.26 (t, J = 5.7 Hz, 1H, NHCO), 7.88 (s, 1H, triazole-H), 7.81 (d, J = 1.8 Hz, 1H, olefin-H), 7.70 (d, J = 8.5 Hz, 2H (SArCH), 7.63 (m, 2H, 2x olefin-H), 7.47 (d, J = 8.6 Hz, 2H, SArCH), 7.34-7.28 (m, 2H, olefin-H), 6.59-6.51 (m, 2H, 2x indoline-H), 6.33-6.21 (m, 2H, 2x indoline-H), 4.70 (d, J = 9.6 Hz, 1H, H-1), 4.51 (t, J = 6.2 Hz, 2H, $\text{NHCH}_2\text{CH}_2\text{triazole}$), 4.24 (d, J = 5.6 Hz, 2H, triazoleCH₂NH), 4.06 (d, J = 7.7 Hz, 2H, COCH₂CH₂), 3.75-3.43 (m, H-2, -4, -5, -6a, -6b, $\text{NHCH}_2\text{CH}_2\text{triazole}$, NCH₃), 3.37 (dd, J = 9.0, 3.2 Hz, 1H, H-3), 2.06 (t, J = 7.4 Hz, 2H, CH₂indoline), 1.74-1.59 (m, 11H 3x CH₃, COCH₂CH₂CH₂), 1.58-1.48 (m, 2H, CH₂CH₂indoline), 1.32 (t, J = 7.8 Hz, 2H, COCH₂CH₂CH₂), 1.28-1.20 (s, 2H, CH₂). ^{13}C NMR (126 MHz, $\text{DMSO}-d_6$, signals extracted from HSQC) δ 154.01 (2x indoline-CH), 127.47 (2x SArCH), 127.35 (2x SArCH), 125.77 (2x olefin-CH), 125.42 (2x indoline-CH), 122.68 (triazole-CH), 119.62 (olefin-CH), 109.79 (olefin-CH), 103.27 (2x indoline-CH), 86.26 (C-1), 78.94 (C-5), 74.45 (C-3), 68.78 (C-2), 68.06 (C-4), 60.31 (C-6), 48.16 ($\text{NHCH}_2\text{CH}_2\text{triazole}$), 43.08 (COCH₂CH₂), 39.33 ($\text{NHCH}_2\text{CH}_2\text{triazole}$), 34.57 (CH₂indole), 33.75 (triazoleCH₂NH), 30.86 (NCH₃), 28.61 (1x CH₂), 26.68 (3x CH₃), 26.04 (COCH₂CH₂), 25.40 (CH₂CH₂CH₂indole), 24.55 (CH₂CH₂indole). HR-MS (ESI): calcd for $[\text{C}_{50}\text{H}_{62}\text{N}_7\text{O}_{13}\text{S}_3 + \text{H}]^+$ 1064.3562; found 1064.3541.

¹H NMR of S7



¹H, ¹³C-HSQC NMR of S7



Recombinant LecA expression and purification

Expression and purification of LecA was performed as described with some modifications.^[12] The *E. coli* expression strain BL21(DE3) containing the plasmid pET25pa11 was grown in LB medium (10 g NaCl, 10 g tryptone, 5 g yeast, adjusted to 1 L and autoclaved) supplemented with ampicillin (100 µg/mL) at 37 °C and 180 rpm until an OD₆₀₀ of 0.5 - 0.6 was reached. Protein expression was induced with IPTG (final concentration: 0.25 mM) and the culture was cultivated at 30 °C and 180 rpm for 4 h. Afterwards, cells were harvested by centrifugation, the supernatant was discarded, the pellet was resuspended in TBS/Ca²⁺-buffer (80 g NaCl, 24.2 g TRIS, 1.93 g KCl, 1.47 g CaCl₂*2 H₂O, adjusted to 1 L with ddH₂O, pH = 7.4), and lysed using a Microfluidizer (Microfluidics, Benchtop High-Pressure Homogenizer). After centrifugation (24,000 x g, 30 min, 4 °C), the supernatant was filtered (0.22 µm) and purified using an Äkta Start and a melibiose-modified sepharose CL-6B column.^[9] LecA was eluted with D-Gal (100 mM) in the TBS/Ca²⁺-buffer and fractions containing protein were pooled and extensively dialyzed against TBS/Ca²⁺-buffer and protein was stored at -80 °C. The concentration was determined by UV spectroscopy at 280 nm ($\epsilon = 27385 \text{ M}^{-1} \text{ cm}^{-1}$, MW = 12893 g mol⁻¹).^[13] Alternatively, protein was dialyzed against ddH₂O and lyophilized.

Competitive Binding Assay using Fluorescence Polarization (FP-Assay)²⁷

The competitive binding assay was performed in analogy to Joachim *et al.*^[14] 10 µL of a solution containing LecA (40 µM) and sulfo-Cy5-Gal S7 (20 nM) in TBS/Ca²⁺-buffer (1 mM CaCl₂) was distributed into wells of a black 384-well microtiter plate (Greiner Bio-One, Germany, cat no 781900). Then, 10 µL of a serial dilution of test compound (8 mM to 62 µM in TBS/Ca²⁺-buffer containing 8% DMSO, with or without 8 mM 2-mercaptoethanol) was added in three technical replicates. Afterwards, the plate was sealed (EASYseal, Greiner Bio-One, cat no 676001), centrifuged (1500 x g, 1 min, 25 °C) and incubated at r.t. in the dark under shaking conditions. Fluorescence intensity was measured on a PheraStar FS microplate reader (BMG Labtech GmbH, Germany; ex. 590 nm, em. 675 nm) after 1 h and 16 h. Blank values (LecA in TBS/Ca²⁺-buffer and 4% DMSO) were subtracted from test compounds, fluorescence polarization was calculated and the data analyzed according to the four-parameter variable slope model using MARS Data Analysis Software (BMG Labtech GmbH, Germany). The top and bottom plateaus were determined from the control, Me- α -D-Gal, and data was reanalyzed with these values fixed. The experiment was repeated three times. Data were visualized using GraphPad PRISM version 6.

T_{1ρ}-relaxation NMR spectroscopy

NMR spectra were recorded on a 500 MHz Avance III (UltraShield) spectrometer (Bruker BioSpin GmbH, Rheinstetten, Germany), equipped with a 5 mm TXI cryoprobe, at 298 K. Lyophilized LecA was suspended in deuterated PBS buffer (7 mM Na₂HPO₄, 3 mM NaH₂PO₄, 137 mM NaCl, 2.7 mM KCl, D₂O, pH = 7.4). After 2 h of incubation at r.t. the mixture was centrifuged at 21,380 x g to remove insoluble material. The concentration of LecA in the supernatant was determined by measuring UV absorbance at 280 nm ($\epsilon = 27,385 \text{ M}^{-1} \text{ cm}^{-1}$). The T1ρ relaxation experiments were carried out using 50 μM LecA and 500 μM ligand, or 25 μM LecA and 1 mM of ligand. For the displacement experiments, methyl α-D-galactoside (50 mM in dPBS) was used at 500 μM to 5 mM concentration. Tetramethylammonium chloride (500 μM) was added as internal standard. All NMR spectra were recorded at spin lock times varying between 10 and 900 ms. Recorded NMR spectra were analyzed using the Bruker TopSpin 3.2 software. For evaluation, the signal intensity of the internal standard was set to 100% and the peak intensity of the signals of interest was measured and the relative decrease in signal intensity calculated, setting the intensity at 10 ms spinlock time to 100%.

Thermal Shift Assay

20 μL of a solution containing LecA (20 μM), test compound (4 mM), SyproOrange (final concentration 20x of a 5000x stock in DMSO, Sigma-Aldrich, Germany) in TBS/Ca²⁺-buffer (1 mM CaCl₂) was added to a white semi-skirted 96-well plate (Thermo Fisher) in triplicates. As controls, protein without compound as well as protein in presence of two positive controls (Me-α-D-Gal or pNPGal) were included. The melting curve measurements were performed and analyzed on a real time PCR instrument (StepOnePlus, Applied Biosystems).

Protein-observed ¹⁹F (PrOF) NMR

5FW-LecA was produced recombinantly as reported previously.^[15] The measurements 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). PrOF NMR was recorded using 200 μM 5FW-LecA in 20 mM Tris-HCl pH 7.8 with 150 mM NaCl, 10% D₂O and 100 μM TFA at 310 K. Changes in the spectra upon addition of 1.5 mM pNPGal or 1.5 mM catechols were processed, referenced and normalized to trifluoroacetic acid (TFA) as internal reference at - 75.6 ppm. We considered only changes in chemical shift perturbation (CSP) upon ligand

addition being two-fold greater than standard deviation (2x std. dev., 0.03 ppm) of fluorine resonance.

Transverse relaxation-optimized spectroscopy (^1H , ^{15}N -TROSY) NMR

^{15}N -labeled LecA was purified and all TROSY experiments were measured and processed as described previously.^[16] Following changes to the procedure have been introduced: 1) we recorded spectra of 200 μM ^{15}N -LecA alone or in presence of 1 mM compound (100 mM stock dissolved in DMSO- d_6 ; 2) a pulse sequence (trosyf3gppsi19) using 16 scans and 128 increments was applied.

Surface plasmon resonance (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 $\mu\text{L}/\text{min}$ until the response exceeded 800 RU, followed by multiple injections of LecA dissolved in 10 mM sodium acetate pH 4.5 (100 $\mu\text{g}/\text{mL}$) onto channel 2 (contact time of 540 s at a flow rate of 10 $\mu\text{L}\cdot\text{min}^{-1}$). 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 catechol compound, except for **16** and **27** which were injected at 0.2 and 0.5 mM due to limited solubility in 5% DMSO (association 30 s, dissociation 60 s, 30 $\mu\text{L}\cdot\text{min}^{-1}$ flow rate), to identify the initial binders eliciting dose-response behaviors. Injection of the positive control (0.1 mM pNPGal) was included after every 4 catechol injections to monitor the activity of immobilized LecA throughout the binding screen experiments. To account for the variation in the amount of active immobilized LecA during the analyses and the difference in the molecular mass of the catechol compounds, we normalized the binding responses into efficiency (%), which is defined as (measured binding response at 1 mM/ R_{max}) \times 100 where R_{max} for a given analyte = ($\text{MW}_{\text{analyte}}/\text{MW}_{\text{LecA}}$) \times amount of active immobilised LecA (RU), and the amount of active immobilised LecA (RU) = ($R_{\text{pNPG}}/\text{MW}_{\text{pNPG}}$) \times MW_{LecA} . MW_{pNPG} is 301.3 Da (reported K_d value = 26.1 μM ¹⁶), and MW_{LecA} as a monomer devoid of Met1 is 12762 Da. From the efficiency values, hits were defined as compounds eliciting at least a twofold response increase at the higher concentration and more than 10% efficiency (Tables 1 and 2).

The prioritized hits were subjected to multi-cycle affinity studies (60 s association, 60 s dissociation, flow rate 30 $\mu\text{L}\cdot\text{min}^{-1}$) consisting of injections of compounds at increasing concentrations onto the immobilized LecA. Binding was measured after reference subtraction of channel 1 (no immobilized LecA) and subtraction of a blank injection (zero analyte concentration). All data evaluation was performed using BIACORE X100 evaluation software (version 2.0).

X-ray crystallography

Lyophilized recombinant LecA was dissolved in MilliQ water containing 100 μM CaCl_2 to the final protein concentration of 10 $\text{mg}\cdot\text{mL}^{-1}$. 1 μL of each catechol (100 mM in DMSO) was aliquoted onto a siliconized glass circle cover slides (22 mm, Hampton research) and left to evaporate to complete dryness at r. t.. Crystallization drops of LecA containing a 1:1 mixture of the protein solution and the reservoir solution (20% PEG6000, 100 mM sodium acetate pH 4.5, 1 M LiCl, 1% DMSO) were deposited on the dried catechols and 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, 100 mM sodium acetate pH 4.5, 1 M LiCl, 1% DMSO and flash cooled in liquid nitrogen. X-ray diffraction data were collected at SOLEIL-PROXIMA1 (Saint Aubin, France) using a Pilatus 6M hybrid photon counting detector (Dectris). The recorded data were indexed, integrated and scaled using XDS^[17] and merged using AIMLESS.^[18] The structures were solved by molecular replacement using IOKO as a searching template in PHASER,^[19] followed by further iterations of manual rebuilding in COOT^[20] and restrained refinement in REFMAC5.^[21] Catechol ligands were manually built in ACEDRG^[22] in CCP4i2 suite.^[23] The final models were validated with MOLPROBITY,^[24] 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.^[25] Data processing, refinement statistics and PDB IDs of the deposited structures are provided in the Supporting Information (Table S3).

Langerin expression and purification

Langerin CRD and ECD constructs were prepared as described before.^[26]

¹H STD NMR with Langerin

^1H STD competition NMR experiments were performed with 20 μM Langerin ECD and 1 mM **3** in Tris- d_{11} 20 mM Tris with 100% D_2O pH 7.8, 150 mM NaCl and 5 mM CaCl_2 at pH 7.8 and 298 K. The experiment was repeated in absence of Langerin to exclude STD effects due to direct saturation of the compound **3**. The spectra were analyzed for changes in peak intensity in MestReNova 14.1 and referenced to the internal reference 50 μM sodium trimethylsilyl propanoate (TSP). For each spectrum 512 scans were recorded with a train of Gauss pulses at a total saturation time of 2.0 s. The on- and off- resonance irradiation frequencies were set to 0 and 80 p.p.m., respectively. The relaxation delay and the acquisition time were set to 0 s and 2 s, respectively. Resonances of **3** were assigned by comparison to its previously assigned ^1H NMR spectrum. The compound was identified as competed if the intensity was reduced upon addition of 100 mM D-mannose.

^1H - ^{15}N HSQC NMR of Langerin

^1H - ^{15}N HSQC NMR experiments were performed with 100 μM Langerin CRD containing 100 μM 4,4-dimethyl-4-silapentane-1-sulfonic acid sodium salt (DSS) as internal reference, and 10% D_2O in 20 mM HEPES pH 7.4, 150 mM NaCl, 1% DMSO and 5 mM CaCl_2 at 298 K. Next, **3** or D-mannose were added at 1 mM and 30 mM, respectively. Spectra were acquired with 128 increments and 8 scans per increment for 150 μL 3 mm tubes were recorded at 298 K. The relaxation delay and the acquisition time were set to 1.4 s and 0.09 s, respectively. NMR data was processed and analyzed as described previously.^[26]

Competitive binding assay based on fluorescence polarization

For establishing the competitive binding assay for catechols a new reporter ligand (**S7**) was synthesized to avoiding a spectral overlap of some catechols with the FITC-based probe (ex. 485, em. 520 nm) used in our lab before.^[14] The red shifted dye Cy5 (ex. 590, em. 675 nm) was chosen because most of the catechols were yellowish in TBS/Ca²⁺-buffered conditions. For that sulfo-Cy5-galatoside **S7** (Scheme S1) was synthesized.

S7 was titrated with LecA to determine its dissociation constant ($K_d = 4.12 \pm 0.39 \mu\text{M}$, Figure S4) and compound **S7** was slightly more active than the previously used fluorescein-derivative ($K_d = 7.4 \pm 2.8 \mu\text{M}$).^[14]

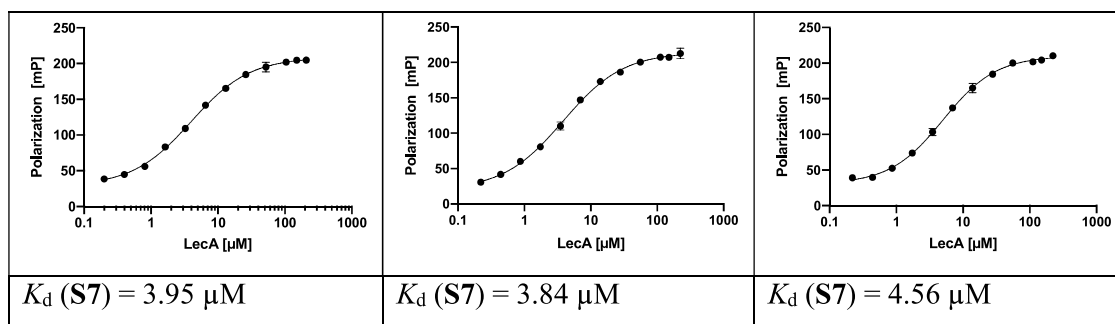


Figure S4: Three independent direct titrations of LecA against sulfo-Cy5-Gal **S7**. The K_d values were determined after fitting with MARS (BMG Labtech, Germany). **S7** was used at 10 nM, incubation time 16 h.

The newly established system was then used for the competitive binding assays. The positive control Me α -D-Gal was measured with **S7** against LecA resulting in an $\text{IC}_{50} = 150.6 \pm 11.7 \mu\text{M}$. All catechol derivatives were tested in the competitive binding assay with the Cy5-derivative **S7** in presence and absence of 2-mercaptoethanol (Figure S5).

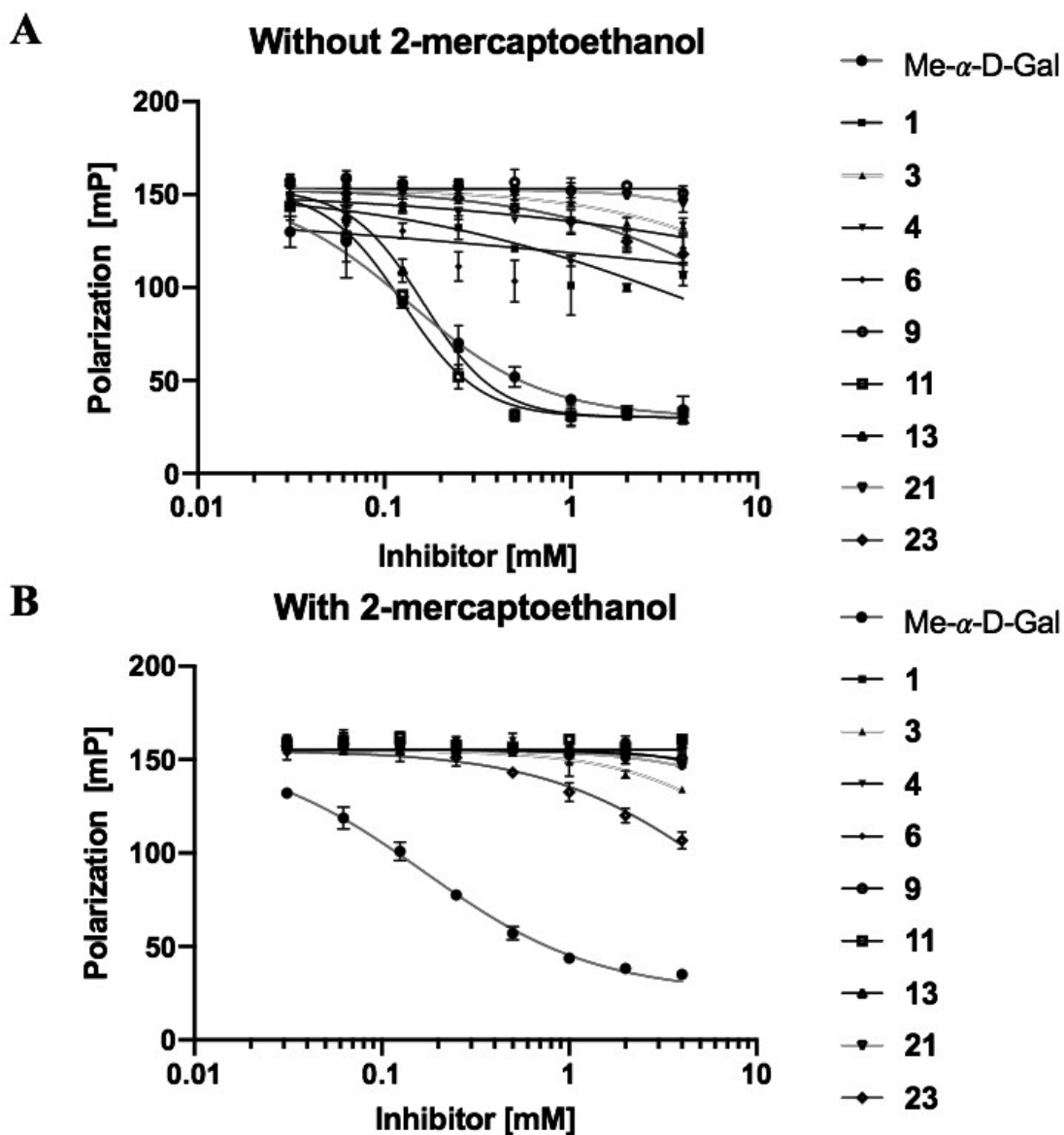
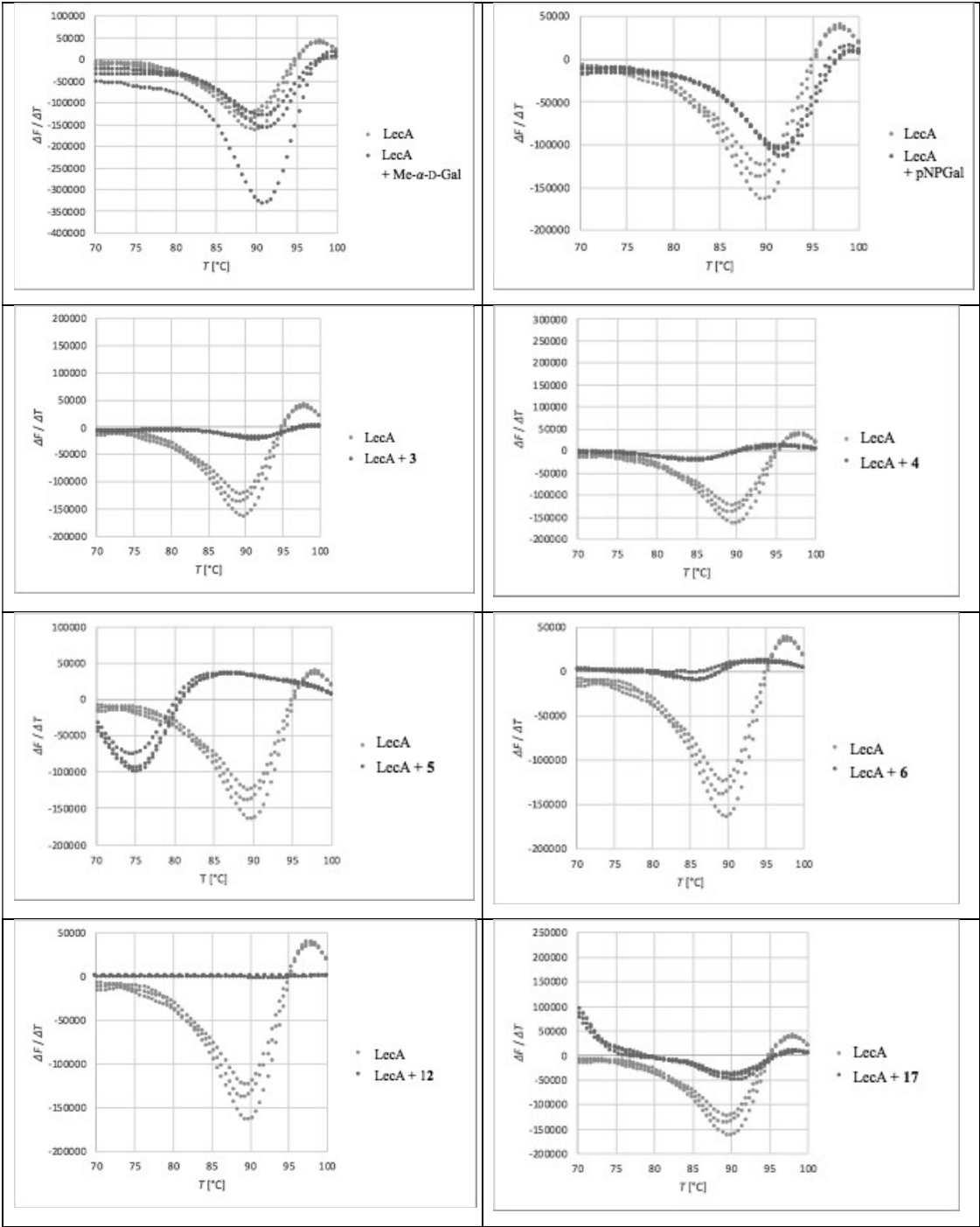


Figure S5: Evaluation of catechols in the competitive binding assay with Sulfo-Cy5-Gal **S7** in absence (A) and presence of 2-mercaptoethanol (B) measured after 16 h. Catechols carrying electron withdrawing substituents were active under both conditions (e.g. blue line). Electron-rich catechols as well as fluorinated/chlorinated catechols were inactive.

Thermal Shift Assay



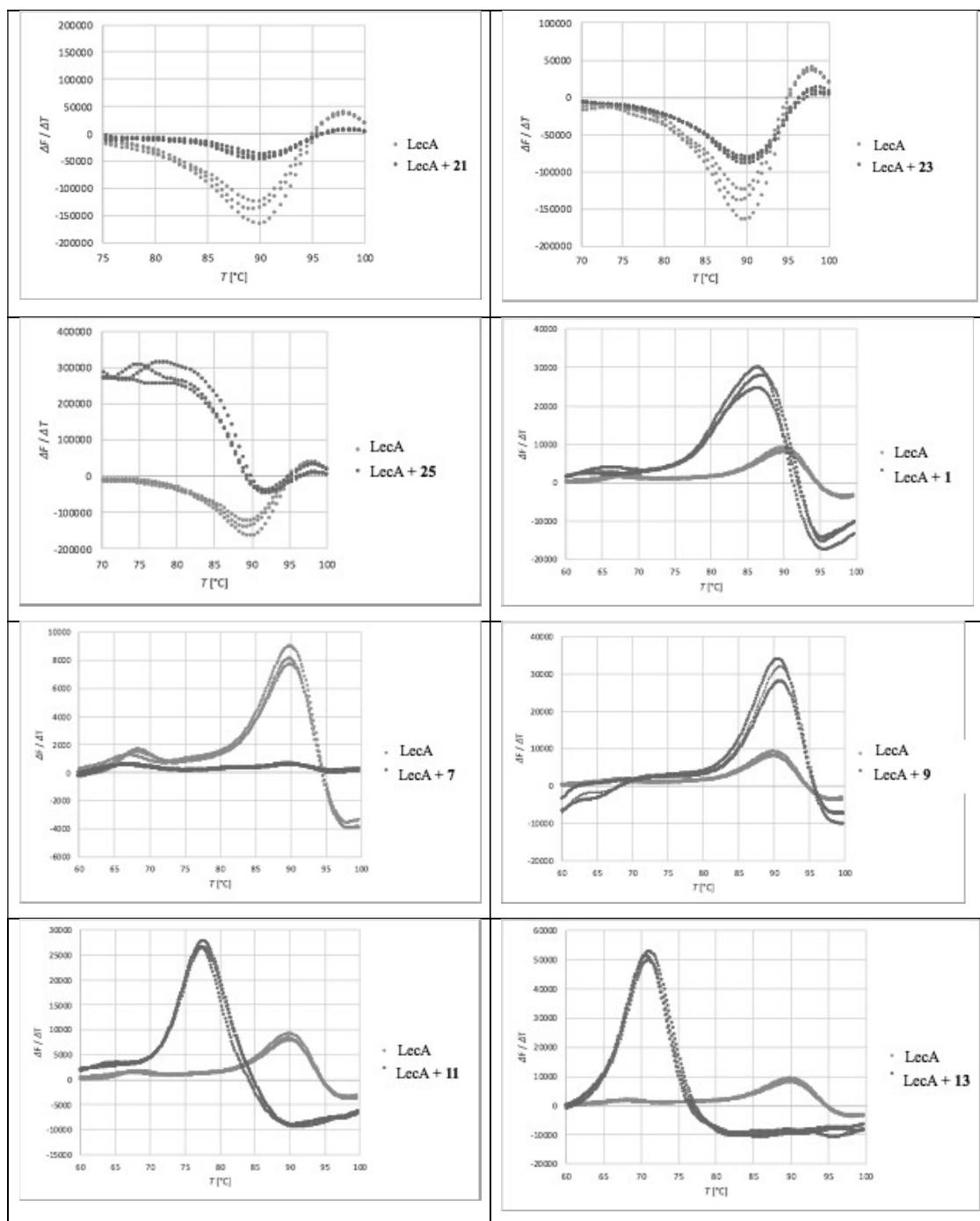


Figure S6: Thermal shift assay of a distinct set of catechols with LecA. One dataset of three technical replicates is shown in each diagram. The green line shows LecA with its melting point of 89.7 ± 0.3 °C and the red line represents LecA in presence of a catechol or a galactoside.

T_{1ρ} – NMR

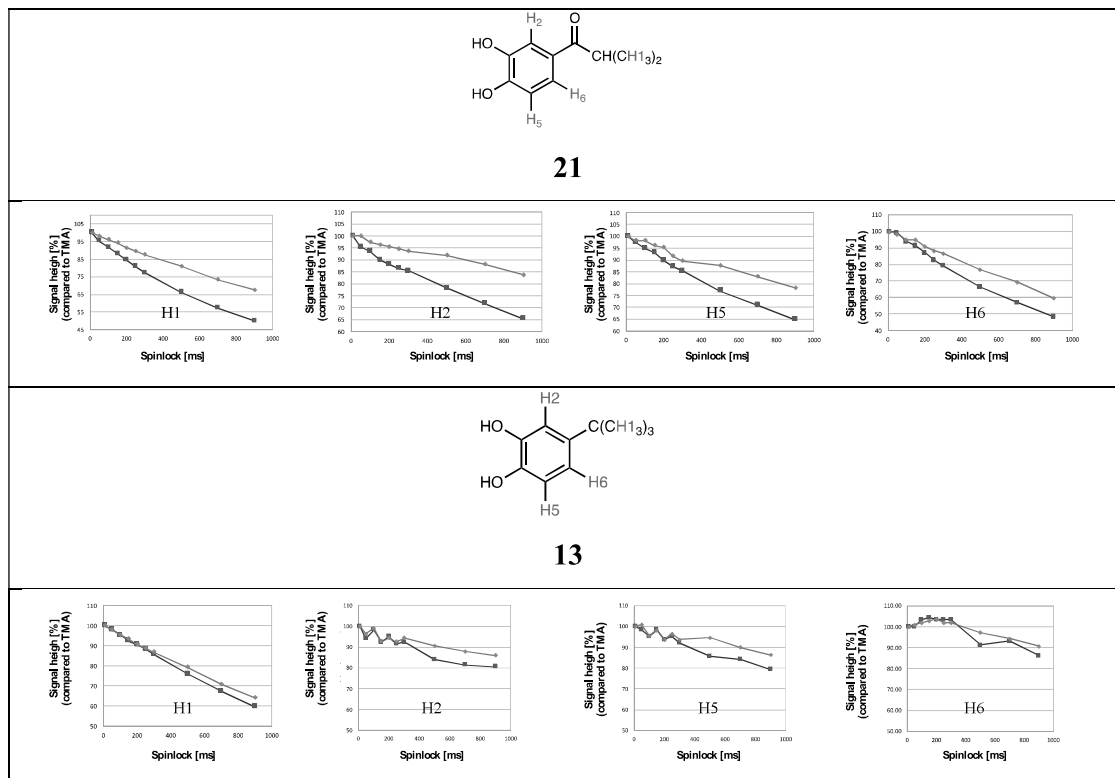


Figure S7: T_{1ρ} relaxation NMR experiment of initial hit isobutryl catechol **21** and *tert*-butylcatechol **13**. The relaxation of the individual protons is shown in each diagram. The signal intensity of each compound is shown in presence (blue line) and absence (red line) of LecA.

Surface Plasmon Resonance

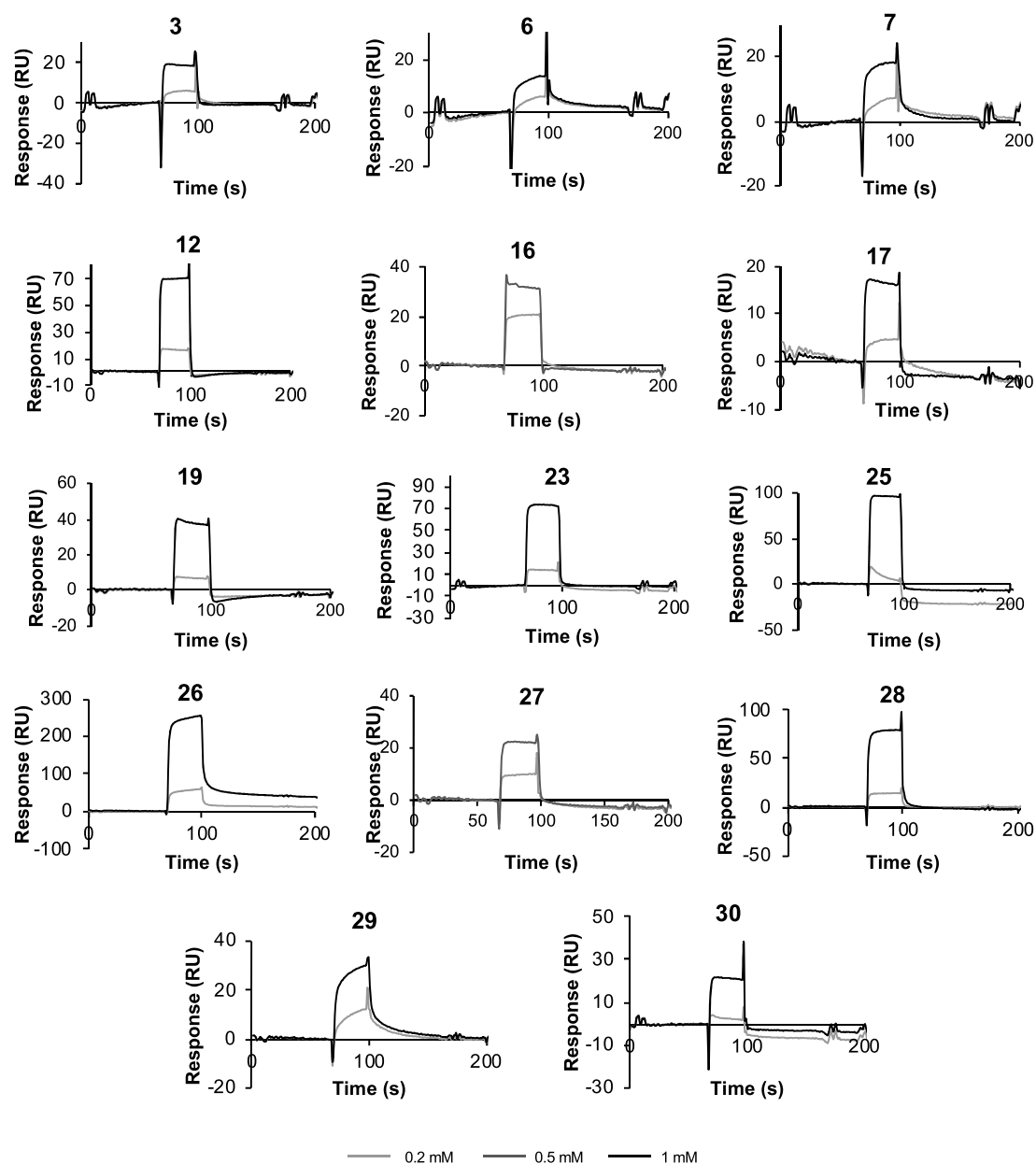


Figure S8: Sensorgrams of SPR hits which are defined as catechols that exhibited at least a twofold increase in binding response to immobilized LecA in a dose-dependent manner.

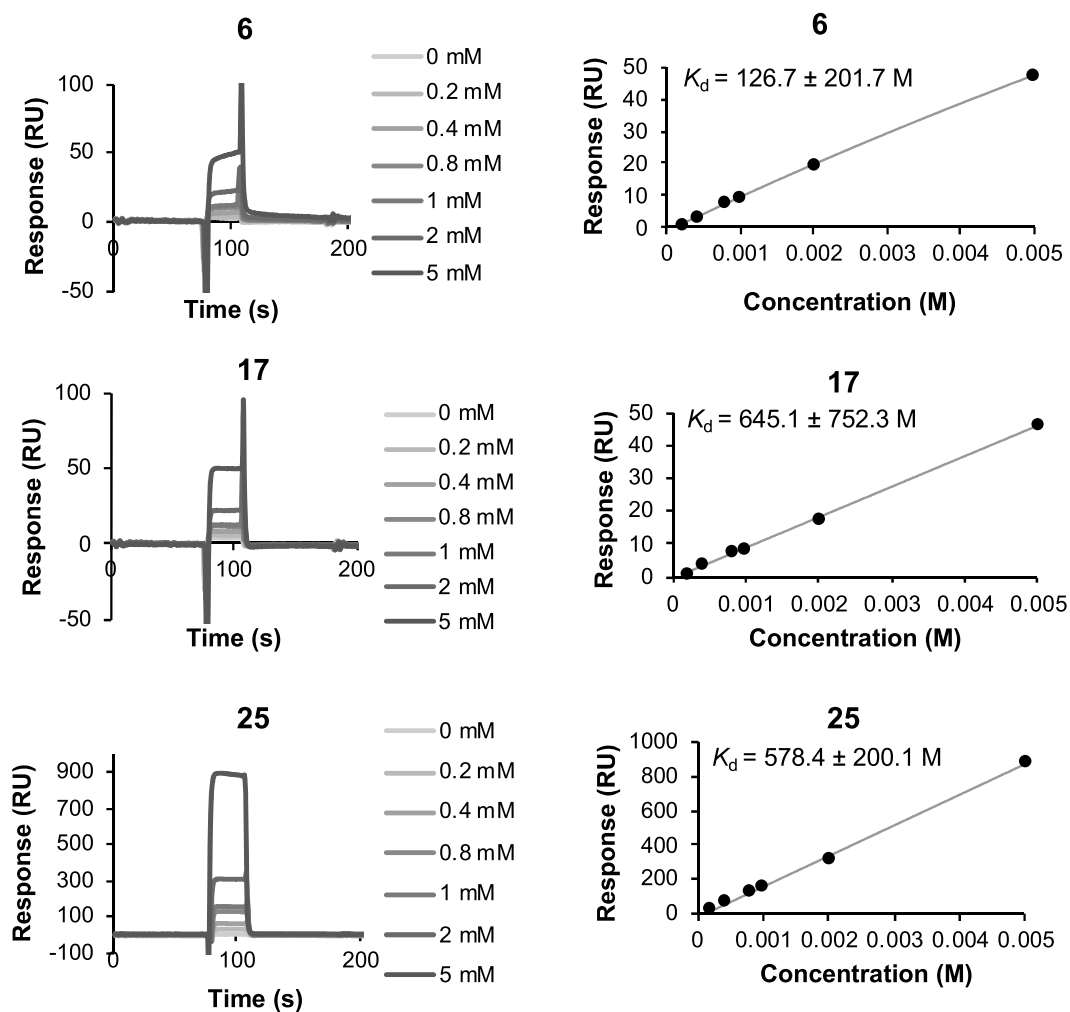


Figure S9: Multi-cycle analyses of the prioritized hits. *Left:* Sensorgrams. *Right:* affinity analyses based on the binding responses from the sensorgrams on the right. The ID of the compounds are indicated at the top of the graphs.

Table S2: Protein-observed ^{19}F (PrOF) NMR. Raw data of chemical shift perturbations of 5-fluorotryptophan residues in LecA alone or in the presence of the catechols measured in PrOF NMR experiments.

ID	^{19}F chemical shift [p.p.m.]			
	W42	W2	W33	W84
Reference	-120.525	-121.435	-121.685	-123.515
1	-120.51	-121.49	-121.7	-123.56
2	-120.53	-121.43	-121.67	-123.52
3	-120.48	-121.43	-121.7	-123.52
4	-120.51	-121.42	-121.67	-123.5
5	-120.47	-121.44	-121.72	-123.49
6	-120.49	-121.44	-121.7	-123.5
7	-120.31	-121.49	-121.69	-123.57
8	-120.51	-121.4	-121.68	-123.51
9	-120.29	-121.5	-121.7	-123.58
10	-120.52	-121.42	-121.7	-123.5
11	-120.5	-121.49	-121.69	-123.56
12	-120.46	-121.43	-121.66	-123.52
13	-120.5	-121.48	-121.7	-123.56
14	-120.53	-121.43	-121.68	-123.48
15	-120.48	-121.43	-121.69	-123.51
16	-120.5	-121.44	-121.67	-123.52
17	-120.49	-121.46	-121.7	-123.51
18	-120.53	-121.42	-121.68	-123.51
19	-120.62	-121.39	-121.74	-123.54
20	-120.52	-121.43	-121.7	-123.51
21	-120.53	-121.54	-121.7	-123.56
22	-120.46	-121.41	-121.67	-123.51
23	-120.56	-121.55	-121.79	-123.6
24	-120.34	-121.43	-121.69	-123.51
25	-120.42	-121.44	-121.68	-123.5
26	-120.4	-121.43	-121.67	-123.52
27	-120.48	-121.44	-121.66	-123.51

28	-120.49	-121.39	-121.69	-123.5
29	-120.51	-121.41	-121.69	-123.52
30	-120.47	-121.45	-121.71	-123.53
pNPGal	-120.33	-121.41	-121.74	-123.59

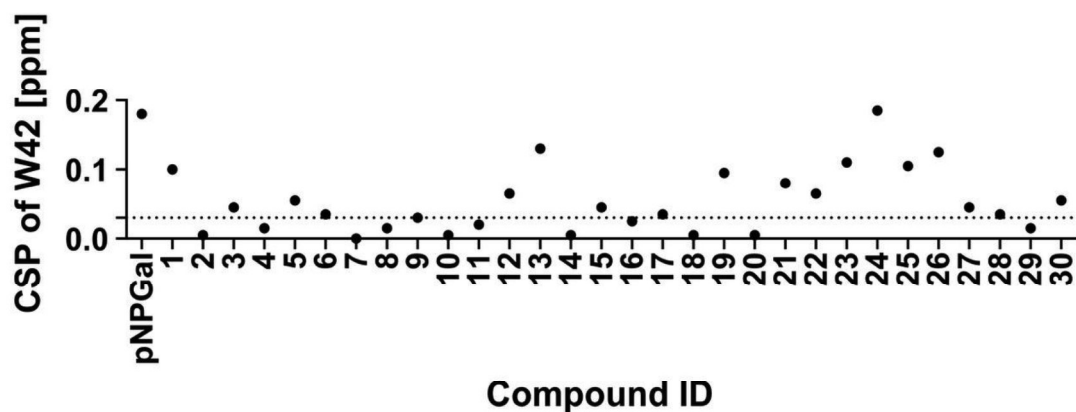


Figure S10: Interpretation of the chemical shift perturbations (CSP) data in ProOF NMR, if CSP > 0.03 ppm it was counted as a hit.

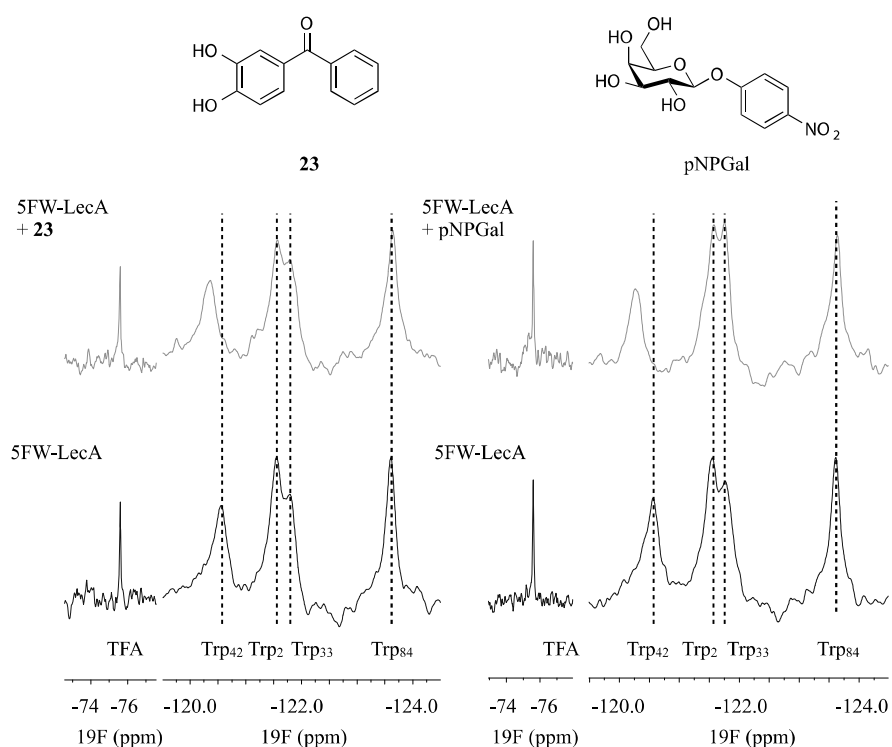


Figure S11: Protein-observed ^{19}F (ProOF) NMR spectroscopy of benzoyl catechol **23** and pNP-Gal with 5FW-LecA (W42, W2, W33, W84).

Crystallography

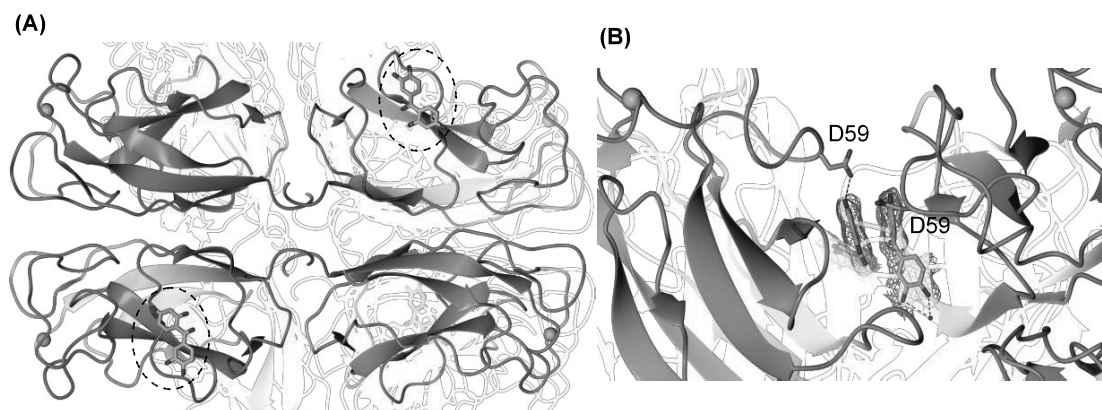


Figure S12: (A) The overall structure of LecA in complex with **25**. (PDB code: 6YOH). Dotted circles indicate the binding sites. (B) Interaction of **25** with LecA. The two molecules are created by symmetry relation. Electron density is displayed at 1σ in green mesh.

Table S3: X-ray data collection and processing of LecA-3 and LecA-25 complexes.

Data set	LecA-3	LecA-25
PDB code	6Y03	6YOH
Data Collection		
Beamline	PROXIMA1 (SOLEIL)	PROXIMA1 (SOLEIL)
Wavelength (Å)	0.9786	0.9786
Detector	Pilatus 6M	Pilatus 6M
Resolution (Å) ^a	45.46-1.88 (1.84-1.84)	48.22-1.84 (1.88-1.84)
Space Group	I 2	P 2 ₁ 2 2 ₁
a, b, c (Å)	88.94, 48.63, 128.63	48.22, 76.07, 107.15
α, β, γ (°)	90.0 95.60 90.0	90.0, 90.0, 90.0
Total observations	269513	472163
Unique reflections	47830	35137
Multiplicity ^a	5.6 (5.4)	13.4 (13.1)
Mean $I/\sigma(I)$ ^a	11.1 (2.7)	23.9 (5.4)
Completeness (%) ^a	99.6 (97.2)	99.9 (97.9)
$R_{\text{merge}}^{\text{a,b}}$	0.090 (0.520)	0.068 (0.469)
$CC_{1/2}^{\text{a,c}}$	1.0 (0.9)	1.0 (0.9)
Refinement		
Reflections: working/free ^d	47826/2432	35095/1776
$R_{\text{work}}/R_{\text{free}}^{\text{e}}$	0.171/0.206	0.193/0.225
Ramachandran plot: allowed/favoured/outliers (%) ^f	3/97/0	8/97/0
R.m.s. bond deviations (Å)	0.0104	0.0104
R.m.s. angle deviations (°)	1.559	1.545
Mean <i>B</i> -factors: protein/ligand ^f / /water (Å ²)	23/35/30	23/39/26

^a Values for the outer resolution shell are given in parentheses.

^b $R_{\text{merge}} = \sum_{\text{hkl}} \sum_i |I_i(\text{hkl}) - \langle I(\text{hkl}) \rangle| / \sum_{\text{hkl}} \sum_i I_i(\text{hkl})$.

^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: $R = \sum(|F_{\text{obs}} - F_{\text{calc}}|) / \sum |F_{\text{obs}}|$, 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 S4 : The list of resonance IDs vs assigned resonances in ^1H - ^{15}N HSQC of Langerin CRD.

ID	Assigned resonance	ID	Assigned resonance	ID	Assigned resonance	ID	Assigned resonance
1	257Lys	33	205Asn	65	221Gln	97	327Glu
2	324Val	34	245Thr	66	212Ile	98	222Phe
3	258Ala	35	221Gln	67	280Phe	99	264Trp
4	320Lys	36	218Ser	68	281Trp	100	333Leu
5	236Glu	37	302Ser	69	232Ser	101	336Gln
6	223Cys	38	220Glu	70	285Glu	102	313Lys
7	214Lys	39	316Leu	71	243Tyr	103	323Tyr
8	200Lys	40	276Gln	72	256Thr	104	345Phe
9	252Trp	41	326Ser	73	266Trp	105	275Val
10	230Leu	42	227Asn	74	306Trp	106	242Leu
11	266Trp	43	338Ser	75	210Ser	107	342His
12	321Arg	44	330Ser	76	199Trp	108	268Asp
13	267Val	45	273Asn	77	260Met	109	225Ser
14	224Val	46	304Gln	78	196Ser	110	303Leu
15	272Phe	47	229His	79	297Asn	111	334Tyr
16	208Tyr	48	228Ser	80	206Phe	112	340Trp
17	282Ile	49	235Ser	81	261Glu	113	263Asp
18	209Phe	50	276Gln	82	240Glu	114	198Gly
19	203Lys	51	277Ser	83	231Thr	115	329Gly
20	202Phe	52	332Asn	84	211Leu	116	215Thr
21	255Leu	53	197Gln	85	331Glu	117	253Ile
22	305Ala	54	344Gln	86	197Gln	118	317Phe
23	278Ala	55	227Asn	87	299Lys	119	269Asp
24	249Leu	56	221Gln	88	194Val	120	312Asp
25	226Arg	57	304Gln	89	315Phe	121	237Ser
26	300Ala	58	217Tyr	90	216Trp	122	265Ser
27	244Lys	59	279Arg	91	270Thr	123	284Gly
28	239Gln	60	273Asn	92	195Val	124	259Gly
29	251Tyr	61	304Gln	93	339Ala	125	337Gly
30	250Ile	62	332Asn	94	238Glu	126	314Thr
31	318Ile	63	197Gln	95	289Ala	127	262Gly
32	241Phe	64	227Asn	96	246Ala	128	201Tyr

129	319Cys	132	219Ala	135	234Thr	138	247Gly
130	274Lys	133	254Gly	136	296Gly	139	204Gly
131	273Asn	134	248Gly	137	332Asn		

References

- [1] R. U. Kadam, M. Bergmann, M. Hurley, D. Garg, M. Cacciarini, M. A. Swiderska, C. Nativi, M. Sattler, A. R. Smyth, P. Williams, M. Camara, A. Stocker, T. Darbre, J. L. Reymond, *Angew. Chem. Int. Ed. Engl.* **2011**, *50*, 10631-10635.
- [2] G. Cioci, E. P. Mitchell, C. Gautier, M. Wimmerova, D. Sudakevitz, S. Pérez, N. Gilboa-Garber, A. Imberty, *FEBS Lett.* **2003**, *555*, 297-301.
- [3] G. M. Sastry, M. Adzhigirey, T. Day, R. Annabhimoju, W. Sherman, *J. Comput. Aided Mol. Des.* **2013**, *27*, 221-234.
- [4] C. R. Sondergaard, M. H. Olsson, M. Rostkowski, J. H. Jensen, *J Chem Theory Comput* **2011**, *7*, 2284-2295.
- [5] J. L. Banks, H. S. Beard, Y. Cao, A. E. Cho, W. Damm, R. Farid, A. K. Felts, T. A. Halgren, D. T. Mainz, J. R. Maple, R. Murphy, D. M. Philipp, M. P. Repasky, L. Y. Zhang, B. J. Berne, R. A. Friesner, E. Gallicchio, R. M. Levy, *J. Comput. Chem.* **2005**, *26*, 1752-1780.
- [6] a) W. L. Jorgensen, J. Tirado-Rives, *J. Am. Chem. Soc.* **1988**, *110*, 1657-1666; b) G. A. Kaminski, R. A. Friesner, J. Tirado-Rives, W. L. Jorgensen, *J. Phys. Chem. A* **2001**, *105*, 6474-6487; c) D. Shivakumar, J. Williams, Y. Wu, W. Damm, J. Shelley, W. Sherman, *J Chem Theory Comput* **2010**, *6*, 1509-1519.
- [7] J. R. Greenwood, D. Calkins, A. P. Sullivan, J. C. Shelley, *J. Comput. Aided Mol. Des.* **2010**, *24*, 591-604.
- [8] T. A. Halgren, R. B. Murphy, R. A. Friesner, H. S. Beard, L. L. Frye, W. T. Pollard, J. L. Banks, *J. Med. Chem.* **2004**, *47*, 1750-1759.
- [9] J. Topin, J. Arnaud, A. Sarkar, A. Audfray, E. Gillon, S. Perez, H. Jamet, A. Varrot, A. Imberty, A. Thomas, *PloS ONE* **2013**, *8*, e71149.
- [10] H. E. Gottlieb, V. Kotlyar, A. Nudelman, *J. Org. Chem.* **1997**, *62*, 7512-7515.
- [11] A. Novoa, T. Eierhoff, J. Topin, A. Varrot, S. Barluenga, A. Imberty, W. Römer, N. Winssinger, *Angew. Chem. Int. Ed.* **2014**, *53*, 8885-8889.
- [12] B. Blanchard, A. Nurisso, E. Hollville, C. Tétaud, J. Wiels, M. Pokorná, M. Wimmerová, A. Varrot, A. Imberty, *J. Mol. Biol.* **2008**, *383*, 837-853.
- [13] E. Gasteiger, C. Hoogland, A. Gattiker, S. Duvaud, M. R. Wilkins, R. D. Appel, A. T. Bairoch, *The Proteomics Protocols Handbook Protein Identification and Analysis Tools on the ExPASy Server, The Proteo*, Humana Press, Totowa, NJ., **1999**.
- [14] I. Joachim, S. Rikker, D. Hauck, D. Ponader, S. Boden, R. Sommer, L. Hartmann, A. Titz, *Org. Biomol. Chem.* **2016**, *14*, 7933-7948.
- [15] E. Shanina, E. Siebs, H. Zhang, D. V. Silva, I. Joachim, A. Titz, C. Rademacher, *Glycobiology* **2020**.
- [16] V. Denavit, D. Lainé, C. Bouzriba, E. Shanina, E. Gillon, S. Fortin, C. Rademacher, A. Imberty, D. Giguère, *Chemistry Eur. J.* **2019**, *25*, 4478-4490.
- [17] W. Kabsch, *Acta Crystallogr. D. Biol. Crystallogr.* **2010**, *66*, 125-132.
- [18] P. R. Evans, *Acta Crystallogr. D. Biol. Crystallogr.* **2011**, *67*, 282-292.

- [19] A. J. McCoy, *Acta Crystallogr. D. Biol. Crystallogr.* **2007**, *63*, 32-41.
- [20] P. Emsley, B. Lohkamp, W. G. Scott, K. Cowtan, *Acta Crystallogr. D. Biol. Crystallogr.* **2010**, *66*, 486-501.
- [21] G. N. Murshudov, P. Skubak, A. A. Lebedev, N. S. Pannu, R. A. Steiner, R. A. Nicholls, M. D. Winn, F. Long, A. A. Vagin, *Acta Crystallogr. D. Biol. Crystallogr.* **2011**, *67*, 355-367.
- [22] F. Long, R. A. Nicholls, P. Emsley, S. Graaeulis, A. Merkys, A. Vaitkus, G. N. Murshudov, *Acta Crystallogr D Struct Biol* **2017**, *73*, 112-122.
- [23] L. Potterton, J. Agirre, C. Ballard, K. Cowtan, E. Dodson, P. R. Evans, H. T. Jenkins, R. Keegan, E. Krissinel, K. Stevenson, A. Lebedev, S. J. McNicholas, R. A. Nicholls, M. Noble, N. S. Pannu, C. Roth, G. Sheldrick, P. Skubak, J. Turkenburg, V. Uski, F. von Delft, D. Waterman, K. Wilson, M. Winn, M. Wojdyr, *Acta Crystallogr D Struct Biol* **2018**, *74*, 68-84.
- [24] V. B. Chen, W. B. Arendall, 3rd, J. J. Headd, D. A. Keedy, R. M. Immormino, G. J. Kapral, L. W. Murray, J. S. Richardson, D. C. Richardson, *Acta Crystallogr. D. Biol. Crystallogr.* **2010**, *66*, 12-21.
- [25] S. McNicholas, E. Potterton, K. S. Wilson, M. E. Noble, *Acta Crystallogr. D. Biol. Crystallogr.* **2011**, *67*, 386-394.
- [26] J. Hanske, S. Aleksic, M. Ballaschk, M. Jurk, E. Shanina, M. Beerbaum, P. Schmieder, B. G. Keller, C. Rademacher, *J. Am. Chem. Soc.* **2016**, *138*, 12176-12186.

Supporting Information to Tolcapone-like Glycomimetics as Potent LecA Inhibitors

Eike Siebs^{1,2,3}, Sakonwan Kuhaudomlarp^{4,5}, Elena Shanina^{6,7}, Annabelle Varrot⁴, Priscila da Silva Figueiredo Celestino Gomes⁸, Christian Lerner⁹, Uwe Grether⁹, Christoph Rademacher^{6,7,10,11}, Didier Rognan⁸, Anne Imberty⁴, Alexander Titz^{1,2,3}

¹Chemical Biology of Carbohydrates (CBCH) Helmholtz-Institute for Pharmaceutical Research Saarland (HIPS) Helmholtz Centre for Infection Research, 66123 Saarbrücken (Germany);

²Department of Chemistry, Saarland University 66123 Saarbrücken (Germany)

³Deutsches Zentrum für Infektionsforschung (DZIF) Standort Hannover-Braunschweig (Germany);

⁴Université Grenoble Alpes, CNRS, CERMAV 38000 Grenoble (France);

⁵Department of Biochemistry and Centre for Excellence in Protein and Enzyme Technology, Faculty of Science Mahidol University, Bangkok (Thailand);

⁶Department of Biomolecular Systems Max Planck Institute of Colloids and Interfaces, 14424 Potsdam (Germany);

⁷Institute of Chemistry and Biochemistry Department of Biology, Chemistry and Pharmacy Freie Universität Berlin, 14195 Berlin (Germany);

⁸Laboratoire d'Innovation Thérapeutique UMR 7200 CNRS-Université de Strasbourg, Strasbourg, 67400 Illkirch (France);

⁹Roche Innovation Center Basel, F. Hoffman-La Roche Ltd., Grenzacherstrasse 124, Basel 4070, Switzerland;

¹⁰Department of Pharmaceutical Sciences University of Vienna, Althanstrasse 14, 1090 Vienna (Austria);

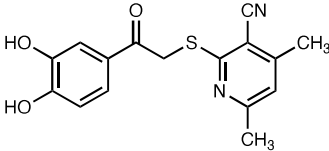
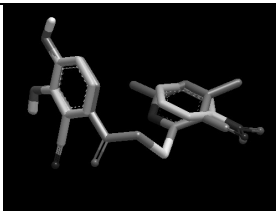
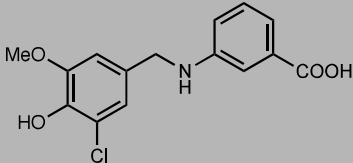
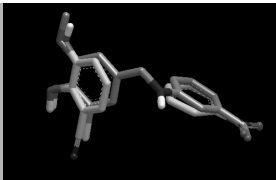
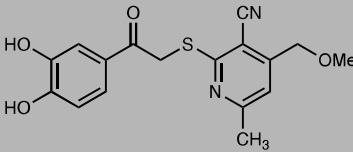
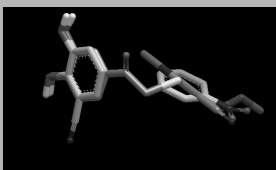
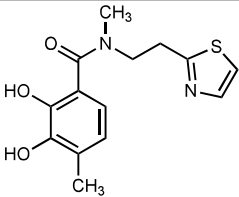
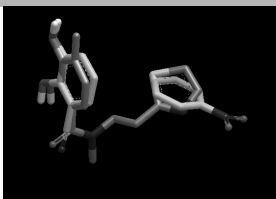
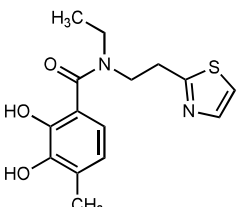
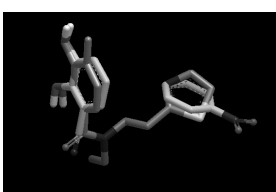
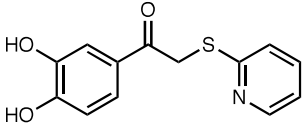
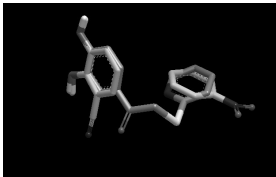
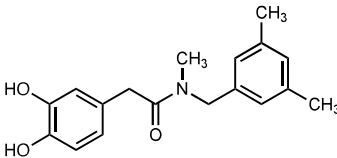
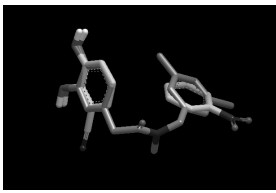
¹¹Department of Microbiology, Immunology and Genetics University of Vienna, Max F. Perutz Labs, Biocenter 5, 1030 Vienna (Austria).

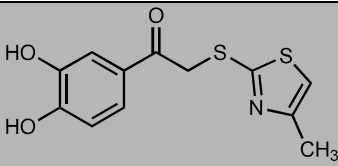
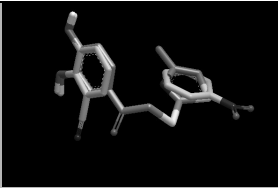
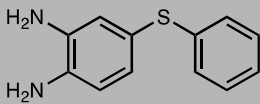
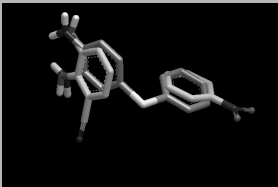
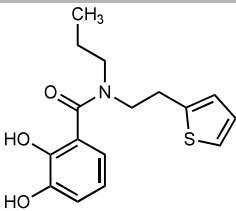
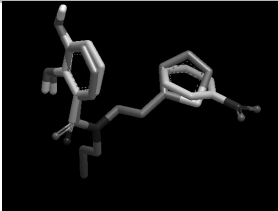
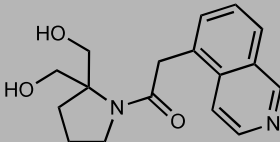
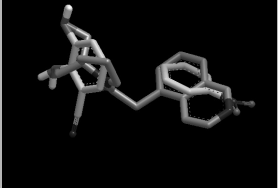
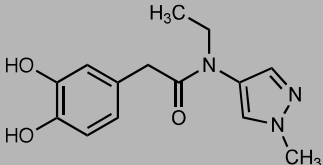
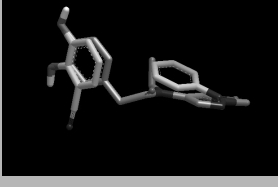
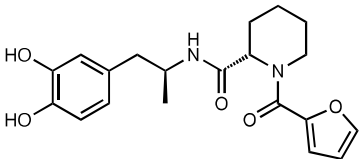
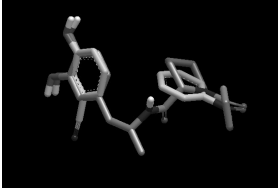
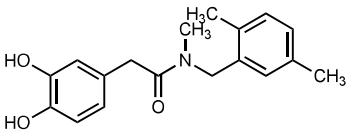
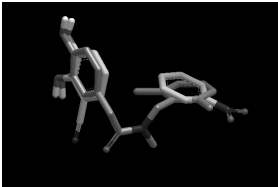
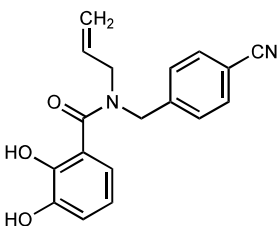
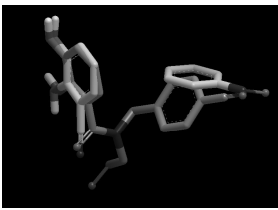
Corresponding e-mail: alexander.titz@helmholtz-hzi.de

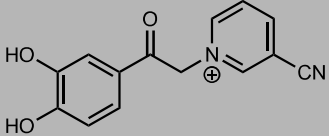
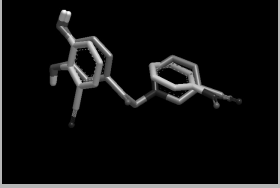
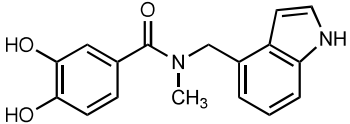
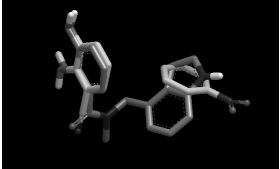
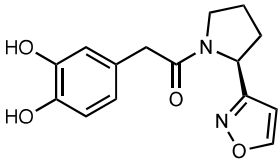
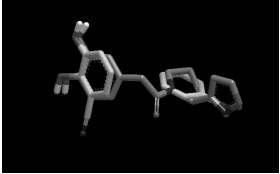
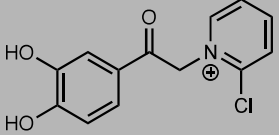
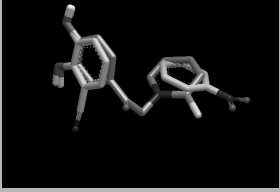
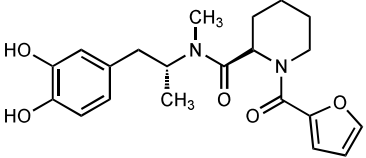
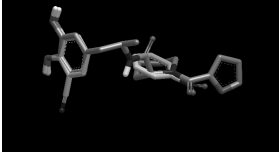
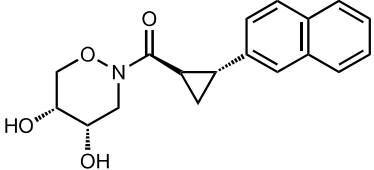
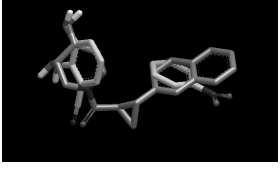
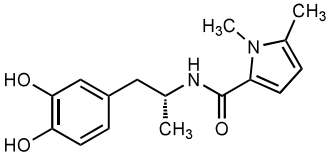
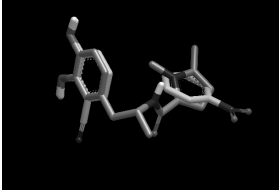
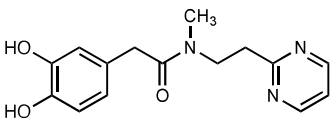
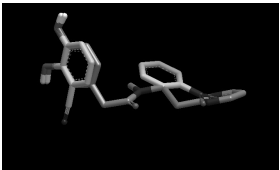
Keywords: glycomimetics, tolcapone, LecA

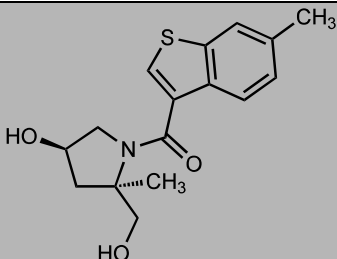
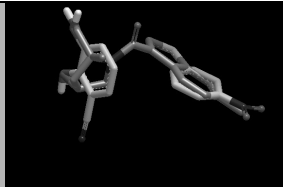
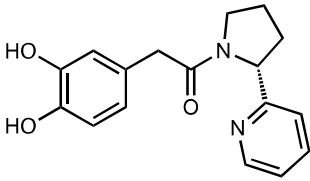
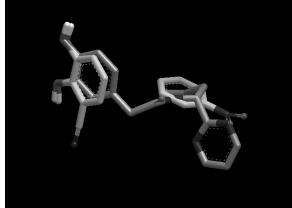
Docking

Table S1: All 25 commercial compounds that align sufficiently with the catechol (pdb: 6yo3) and the aromatic ring of pNPG (pdb: 3zyf) from the two co-crystal structures.

ID number	Supplier	Structure	Alignment
4296-0131	Chemdiv		
B29966	AsisChem		
3294-0014	Chemdiv		
Z1633599741	Enamine		
Z163360185	Enamine		
F3253-0040	Life Chemicals		
IMED335349831 4	Intermed		

ID number	Supplier	Structure	Alignment
Z25108744	Enamine		
DS-17543	Bionet		
Z1932917111	Enamine		
CSC150771429	Chemspcae		
CSC153371590	Chemspace		
Z1662882863	Enamine		
ABB-3369276785	Abamache m		
Z1931724628	Enamine		

ID number	Supplier	Structure	Alignment
AT-051/J236094	Biospecs		
Z1862883625	Enamine		
ABB-3585668028	Abamache m		
AT-051/J235548	Biospecs		
Z1662882863	Enamine		
ABB-3595482018	Abamache m		
Z1727762663	Enamine		
ABB-3371804579	Abamache m		

ID number	Supplier	Structure	Alignment
IMED291489853 1	Chemspace		
BD3359409538	BCH Research		

Green backgrounds: Chosen set of nine structurally divers compounds for biophysical tests.

Biophysical Validation

Stability Control of Tolcapone

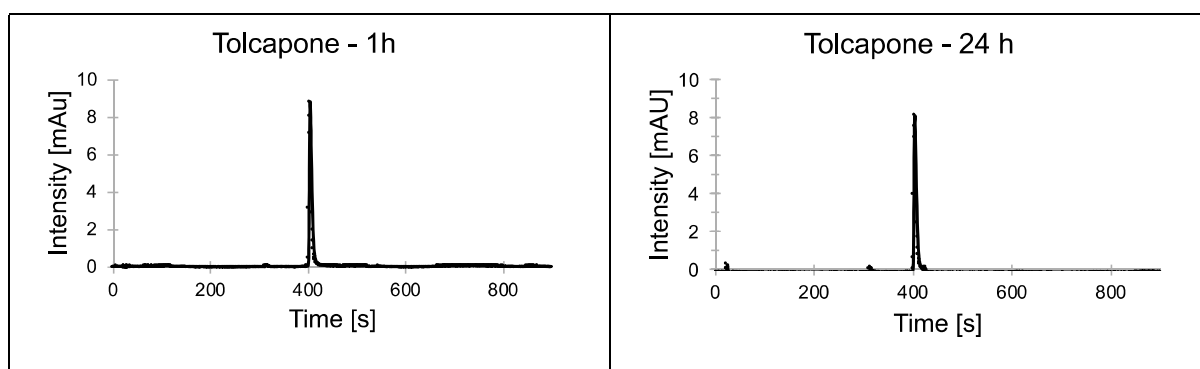


Figure S1: HPLC-UV chromatogram of tolcapone under fluorescence polarization assay conditions (TBS/Ca²⁺-buffer supplemented with 25% DMSO) after 1 and 24 h.

Protein Observed ^{19}F (PrOF) NMR

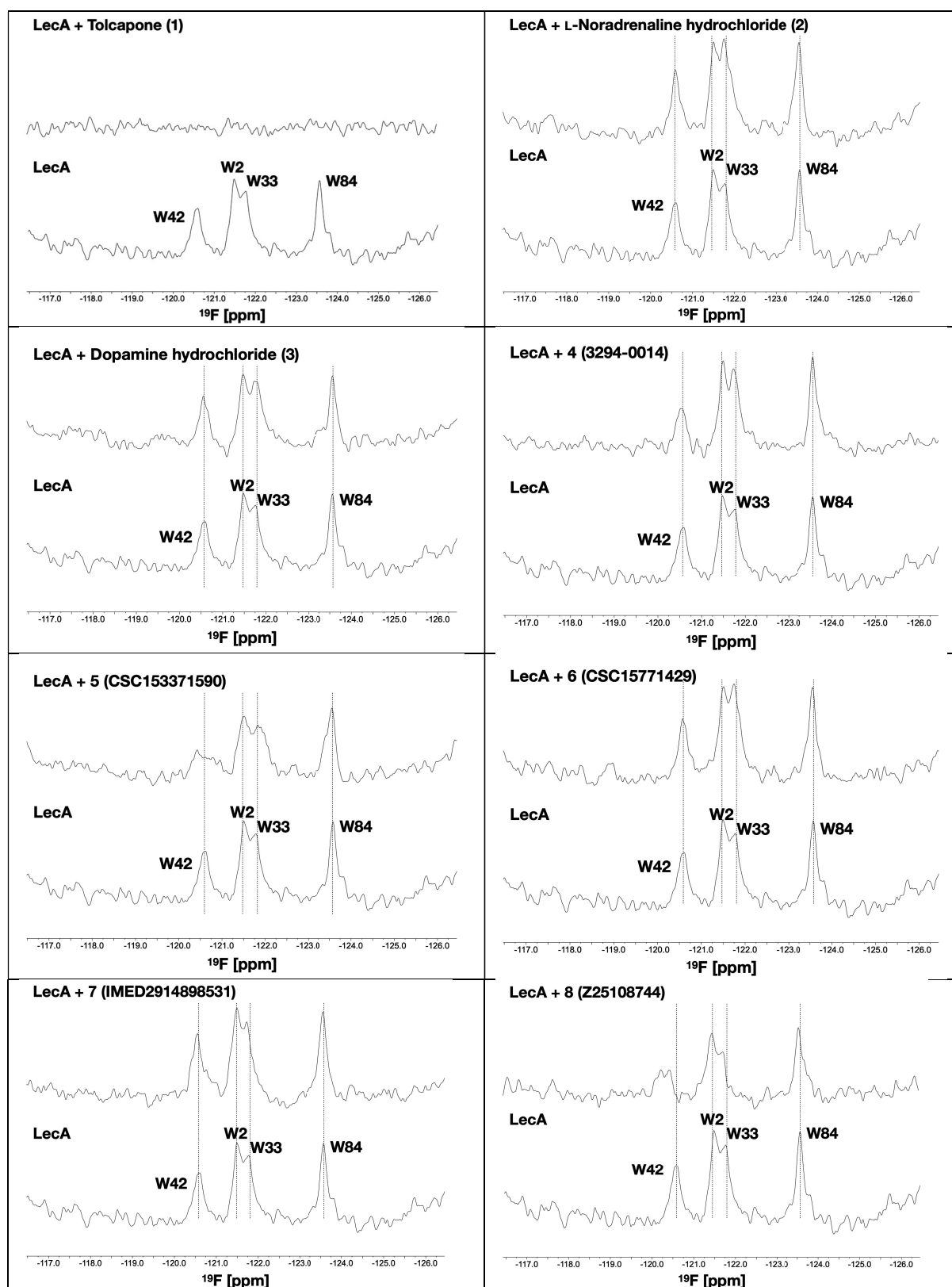
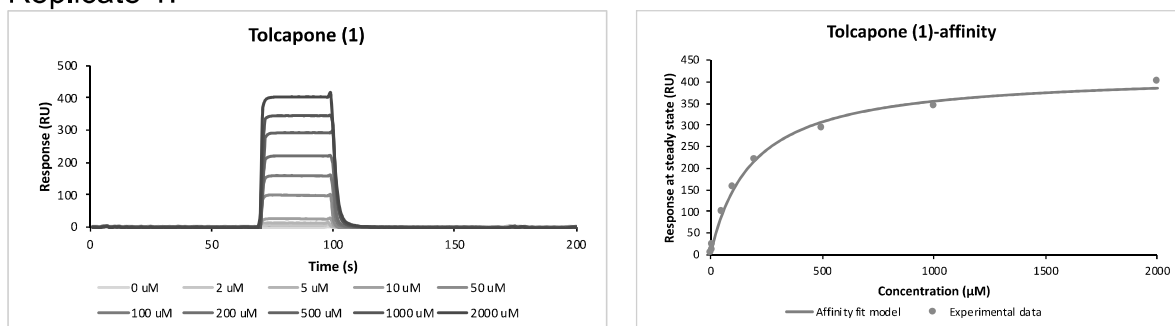


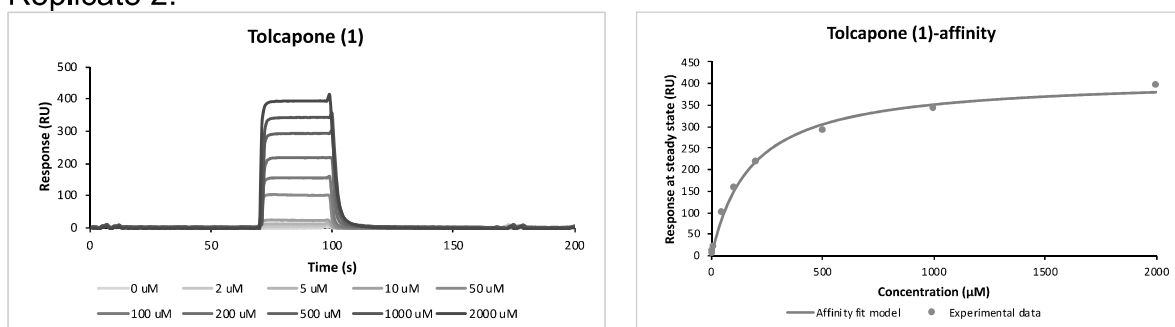
Figure S2: Protein-observed ^{19}F (PrOF) NMR of eight catechol-like compounds with 5FW-LecA.

Surface Plasmon Resonance

Replicate 1:



Replicate 2:



Replicate 3:

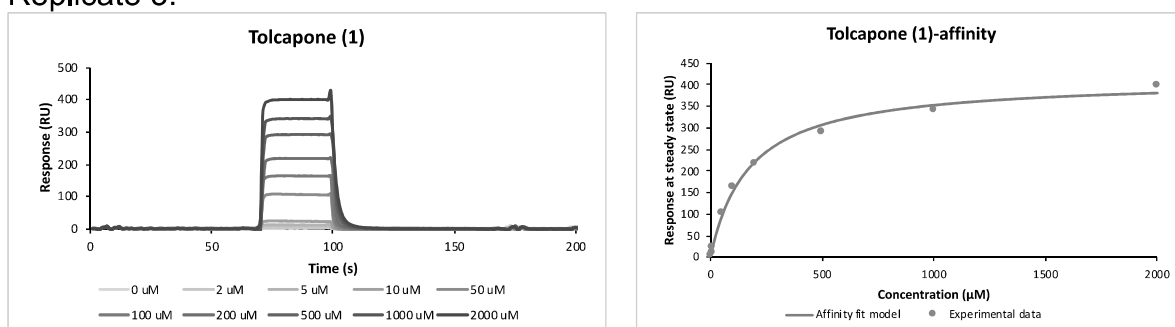
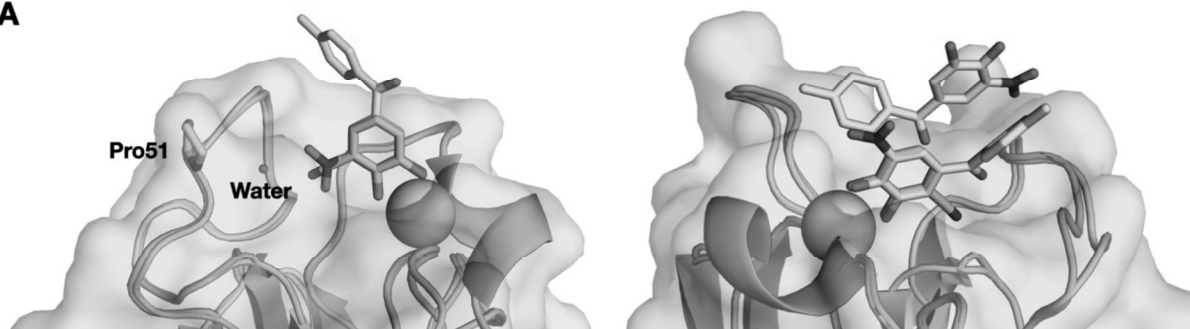


Figure S3: Individual sensorgrams and fits of all replicates for binding of tolcapone to LecA.

X-Ray Crystallography

A



B

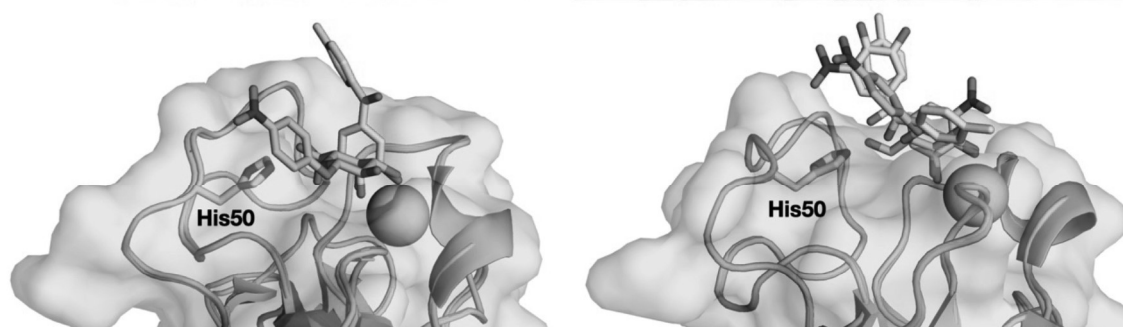


Figure S4: Superimposition of complexes LecA-tolcapone (pdb: 8GUV, cyan, protomer A (left) or B (right), transparent surface grey, calcium ion green sphere) and **A**) LecA-3-cyanocatechol (pdb: 6yo3, light blue, protomer A) or **B**) LecA-pNPG (pdb: 3zyf, light blue, protomer A).

Table S2: Summary of data processing and refinement statistics.

	LecA in complex with tolcapone (PDB: 8GUV)
Data Collection	
Beamline	PROXIMA1 (SOLEIL)
Wavelength (Å)	0.9786
Detector	Pilatus 6M
Resolution (Å) ^a	42.80-1.32 (1.34-1.32)
Space Group	C 2 2 21
a, b, c (Å)	57.65, 164.63, 50.11
α, β, γ (°)	90.0, 90.0, 90.0
Total observations	722,712
Unique reflections	56,290
Multiplicity ^a	12.8 (11.1)
Mean $I/\sigma(I)$ ^a	25.2 (4.1)
Completeness (%) ^a	99.8 (96.4)
R_{merge} ^{a,b}	0.052 (0.451)
$CC_{1/2}$ ^{a,c}	0.999 (0.960)
Refinement	
Reflections: working/free ^d	56272/2793
$R_{\text{work}}/R_{\text{free}}$ ^e	0.198/0.222
Ramachandran plot: allowed/favoured/outliers (%)	2/98/0
R.m.s. bond deviations (Å)	0.0119
R.m.s. angle deviations (°)	1.722
Mean B -factors: protein/ligand ^f / /water (Å ²)	15/23/26

^a Values for the outer resolution shell are given in parentheses.

^b $R_{\text{merge}} = \sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} \sum_i I_i(hkl)$.

^c $CC_{1/2}$ is the correlation coefficient between symmetry-related intensities taken from random halves of the dataset.

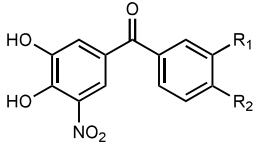
^d The data set was split into "working" and "free" sets consisting of 95 and 5% of the data, respectively. The free set was not used for refinement.

^e The R -factors R_{work} and R_{free} are calculated as follows: $R = \sum(|F_{\text{obs}} - F_{\text{calc}}|) / \sum |F_{\text{obs}}|$, where F_{obs} and F_{calc} are the observed and calculated structure factor amplitudes, respectively

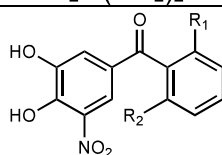
^f refers to ligands bound in the active site and potential surface binding sites

Screening of Catechol-like Compound Library against LecA

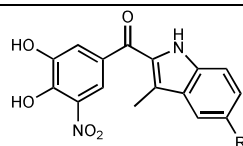
Table S3: %Inhibition of top 48 compounds of the Roche library (3564 compounds) that showed at least 50% inhibition after 1 h compared to pNPG (500 μ M) in the fluorescence polarization assay against LecA. The experiment was conducted in four independent experiments for the concentrations 100 and 25 μ M, in three at 125 μ M and in two at 32 μ M.

				
No.	R ₁	R ₂	%Inhibition at 100 μ M (at 125 μ M)	%Inhibition at 25 μ M (at 32 μ M)
19	-CO ₂ (CH ₂) ₄ CH ₃	-H	92 \pm 10 (82 \pm 3)	49 \pm 5 (42)
S1	-CO ₂ (CH ₂) ₃ CH ₃	-H	87 \pm 10 (74 \pm 3)	56 \pm 2 (38)
S2	-CO ₂ iPr	-H	70 \pm 5	30 \pm 2
S3	-CON((CH ₂) ₂ CH ₃) ₂	-H	67 \pm 4 (62 \pm 14)	30 \pm 5 (28)
S4	-CON(CH ₂ CH ₃) ₂	-H	67 \pm 6	29 \pm 6
S5	-CO ₂ CH ₃	-H	65 \pm 4 (55 \pm 9)	34 \pm 2 (21)
23	-CO ₂ Allyl	-H	63 \pm 4 (59 \pm 3)	26 \pm 5 (31)
S6	-CO ₂ CH ₂ CH ₃	-H	62 \pm 4 (52 \pm 10)	39 \pm 8 (21)
S7	-CO ₂ (CH ₂) ₂ CHOH _(S) CH ₃	-H	62 \pm 1) (55 \pm 9)	27 \pm 7 (27)
S8	-CO ₂ CH ₂ C(CH ₃) ₃	-H	61 \pm 4 (59 \pm 5)	26 \pm 3 (26)
S9	-CONHCy	-H	61 \pm 4	31 \pm 5
21	-CO ₂ (CH ₂) ₂ CHOH _(S) CCF ₃	-H	60 \pm 3	24 \pm 5
S10	-CO ₂ Ph	-H	57 \pm 1	28 \pm 2
S11	-CONHCH ₂ cC ₃ H ₅	-H	57 \pm 6	30 \pm 3
S12	-CONH(CH ₂) ₃ CH ₃ (E)CHCHCO ₂ CH ₃	-H -H	54 \pm 3 (53 \pm 12)	16 \pm 12 (19)
S13	-CH ₂ OH	-H	52 \pm 3	19 \pm 6
24	-F	-H	52 \pm 5	23 \pm 4
S14	-CONH(CH ₂) ₂ CH ₃	-H	51 \pm 3	24 \pm 5
S15	-CONHCH ₂ Ph	-H	51 \pm 4	24 \pm 4
17	-H	-CO ₂ Allyl	103 \pm 3 (99 \pm 2)	17 \pm 5 (39)
18	-H	-CO ₂ CH ₂ Ph	94 \pm 8 (87 \pm 5)	88 \pm 7 (88)
20	-H	-CO ₂ (CH ₂) ₄ CH ₃	89 \pm 1 (81 \pm 1)	28 \pm 6 (12)
S16	-H	-iPr	79 \pm 3 (56 \pm 6)	39 \pm 4 (22)
22	-H	-	71 \pm 4	16 \pm 4
S17	-H	CO ₂ (CH ₂) ₂ CHOH _(S) CF ₃ -tBu	62 \pm 4 (54 \pm 8)	31 \pm 10 (14)

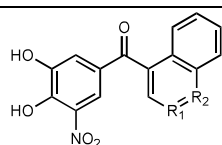
S18	-H	-CO ₂ CH ₃	53 ± 1	22 ± 7
25	-H	-F	55 ± 5	21 ± 5
S19	-H	-CON(CH ₂ CH ₃) ₂	55 ± 4	26 ± 12
S20	-H	-CH ₂ O(CH ₂) ₂ OCH ₃	52 ± 3	21 ± 3



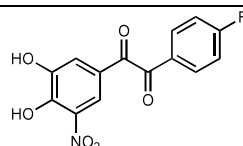
No.	R ₁	R ₂	at 100 μM (at 125 μM)	at 25 μM (at 32 μM)
26	-CF ₃	-H	86 ± 0 (81 ± 5)	52 ± 5 (47)
S21	-F	-F	85 ± 3 (73 ± 6)	57 ± 7 (47)
S22	-Cl	-H	81 ± 7 (73 ± 7)	55 ± 7 (37)
27	-CH ₃	-H	79 ± 3 (63 ± 3)	44 ± 7 (38)
28	-CN	-H	72 ± 2 (53 ± 2)	48 ± 3 (26)



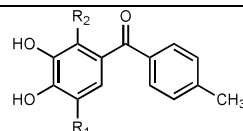
No.	R ₁	R ₂	at 100 μM (at 125 μM)	at 25 μM (at 32 μM)
13	-Br	-H	97 ± 4 (92 ± 3)	43 ± 14 (49)
14	-H	-H	49 ± 5	18 ± 5



No.	R ₁	R ₂	at 100 μM (at 125 μM)	at 25 μM (at 32 μM)
15	-H	-N	78 ± 5 (73 ± 4)	42 ± 2 (37)
16	-N	-H	62 ± 3 (61 ± 10)	26 ± 4 (24)



No.	R ₁	R ₂	at 100 μM (at 125 μM)	at 25 μM (at 32 μM)
30	-CH ₃	-H	69 ± 2 (74 ± 3)	36 ± 3 (29)
29	-H	-H	(66 ± 4)	(31)



No.	R ₁	R ₂	at 100 μM (at 125 μM)	at 25 μM (at 32 μM)
-----	----------------	----------------	--------------------------	------------------------

10	5-(3-methyl-1,2,4-oxadiazole)	-H	58 ± 4	19 ± 4
9	-S(O) ₂ CF ₃	-H	60 ± 7	20 ± 4
S23	-NO ₂	-Cl	(88 ± 3)	(59)

No.	R₁	R₂	at 100 μM (at 125 μM)	at 25 μM (at 32 μM)
11	-H	-C ₄ H ₆ -pCH ₃	58 ± 6 (61 ± 10)	35 ± 11 (21)
12	-Br	-N((CH ₂) ₂ CH ₃) ₂	58 ± 5	24 ± 5

No.	R₁	R₂	at 100 μM (at 125 μM)	at 25 μM (at 32 μM)
31	-F	-H	(35 ± 6)	(4)
32	-H	-CH ₃	20 ± 5	4 ± 6

ChemBioChem

Supporting Information

Targeting the Central Pocket of the *Pseudomonas aeruginosa* Lectin LecA

Eike Siebs, Elena Shanina, Sakonwan Kuhaudomlarp,
Priscila da Silva Figueiredo Celestino Gomes, Cloé Fortin, Peter H. Seeberger, Didier Rognan,
Christoph Rademacher, Anne Imberty, and Alexander Titz*

32	Table of Contents	
33	<i>Chemical Synthesis</i>	3
34	<i>Molecular Dynamics</i>	24
35	<i>Isothermal Titration Calorimetry (ITC)</i>	27
36	<i>¹⁹F ProOF NMR</i>	32
37	<i>Surface Plasmon Resonance (SPR)</i>	33
38	<i>X-Ray Crystallography</i>	38
39	<i>Reference</i>	40
40	<i>Spectra</i>	41
41		
42		

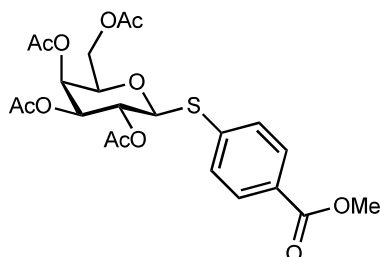
Chemical Synthesis

General information

Chemicals and solvents were obtained from TCI Chemicals, Sigma Aldrich or Carl Roth and deuterated solvents from Eurisotop (Saarbrücken, Germany). Commercial chemicals were used without further purification. Reactions were monitored either by HPLC-MS or by TLC (silica gel 60 Å coated aluminum plates, Macherey-Nagel, Düren, Germany) and visualization with UV light (254 nm) followed by staining with permanganate (1.5 g KMnO_4 , 10 g, K_2CO_3 , 1.25 mL 10% NaOH in 200 mL H_2O) or a molybdate solution (0.02 M $\text{Ce}(\text{NH}_4)_4(\text{SO}_4)_4 \cdot 2\text{H}_2\text{O}$ and $(\text{NH}_4)_6\text{Mo}_7\text{O}_{24} \cdot 4\text{H}_2\text{O}$ in aq. 10% H_2SO_4). MPLC was used for purification (Teledyne Isco CombiFlash Rf 200) using silica pore size 60 Å, particle size 400 mesh, from Fluka. The final test compounds were purified by preparative HPLC (Waters 2545 Binary Gradient Module, Waters 2489 UV/Visible detector, C18 column (250/21 Nucleodur C18 Gravity-SB, 5 μm , Macherey-Nagel, Germany). Purity of all test compounds was > 95% determined by HPLC-UV and/or ^1H -NMR. All synthesized compounds were analyzed by NMR (Bruker Avance III 500 UltraShield spectrometer) at 500 MHz for ^1H -NMR and 126 MHz for ^{13}C -NMR. The spectra were analyzed with MestReNova (Version 12.0.2) and chemical shifts were calibrated on residual solvent peaks as internal standard according to literature^[1] (CDCl_3 = 7.26 ppm and 77.0 ppm, $\text{DMSO}-d_6$ = 2.50 ppm and 39.52 and $\text{MeOH}-d_4$ = 3.31 ppm and 49.00 ppm). Multiplicity is specified as s (singlet), d (doublet), t (triplet), q (quartet) or m (multiplet). Resonance assignment follows the numbering of galactose.

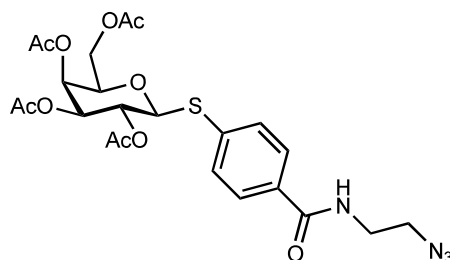
HPLC-MS was performed on a Thermo Dionex UltiMate 3000 HPLC coupled to a Bruker amaZon SL mass spectrometer and UV detection at 254 nm, using a C18 column (100/2 NucleoShell RP 18plus, 2.7 μm Macherey-Nagel, Germany) as stationary phase. High resolution MS was measured on a Bruker maxis 4G hr-QqTOF and analyzed using DataAnalysis (Bruker Daltonics, Bremen, Germany).

4-Methyloxycarbonylphenyl 2,3,4,6-tetra-O-acetyl 1-thio- β -D-galactopyranoside (S3)



The synthesis was performed in analogy to Winssinger *et al.*^[2] Galactose pentaacetate (1.16 g, 2.97 mmol, 1.0 eq.) and $\text{BF}_3 \cdot \text{OEt}_2$ (4.0 eq.) were dissolved in CH_2Cl_2 and methyl-4-sulfanylbzenzoate (3.0 eq.) was added dropwise at 0 °C. The reaction was taken up in CH_2Cl_2 after 18 h, washed with NaHCO_3 (2x – strong gas evolution), water and brine, the org. phase was dried over Na_2SO_4 , filtered and the volatiles removed in vacuo. The residue was purified by flash chromatography (EtOAc/Tol 1:5, R_f = 0.4). Compound **S3** was obtained as a white solid (1.37 g, 2.75 mmol, 95%). The analytical data match the reported spectra in Kuhaudomlarp *et al.*^[3]

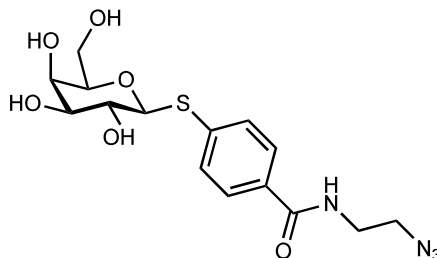
Azidoethyl derivative S5



Azidoethyl derivative **S5** was obtained following the procedure from Kuhaudomlarp *et al.* as a white solid.^[3] (324 mg, 0.59 mmol, 57% o2s from **S3**). The employed 2-azido-ethylamine (**S4**) has an N/C ratio of 2 and is potentially explosive and must be handled with care. ^1H NMR (500 MHz, $\text{DMSO}-d_6$) δ 8.75 (s, 1H, NH), 7.82 (d, J = 8.6 Hz, 2H, 2x ArCH), 7.52 (d, J = 8.5 Hz, 2H, 2x ArCH), 5.38 (d, J = 10.0 Hz, 1H, H-1), 5.35 (dd, J = 3.5, 1.1 Hz, 1H, H-4), 5.28 (dd, J = 9.9, 3.5 Hz, 1H, H-3), 5.08 (t, J = 9.9 Hz, 1H, H-2), 4.38 (td, J = 6.2, 1.1 Hz, 1H, H-5), 4.07 (d, J = 6.2 Hz, 2H, H-6), 3.50 – 3.44 (m, 4H, $\text{NHCH}_2\text{CH}_2\text{N}_3$), 2.13 (s, 3H, CH_3), 2.04 (s, 2H, CH_3), 2.01 (s, 3H, CH_3), 1.93 (s, 3H, CH_3). ^{13}C NMR (126 MHz, $\text{DMSO}-d_6$) δ 169.95 ($\text{C}=\text{O}$), 169.88 ($\text{C}=\text{O}$), 169.48 ($\text{C}=\text{O}$), 169.31 ($\text{C}=\text{O}$), 165.83 ($\text{C}=\text{O}$), 137.26 (ArC), 132.52 (ArC), 128.94 (2x ArCH), 127.75 (2x ArCH), 83.09 (C-1), 73.63 (C-5), 70.97 (C-3), 67.57 (C-4), 66.79 (C-2), 61.81 (C-6),

49.77 (NHCH₂CH₂N₃), 39.02 (NHCH₂CH₂N₃), 20.52 (CH₃), 20.50 (CH₃), 20.42 (CH₃), 20.35 (CH₃).

Deprotected azidoethyl galactoside **S6**

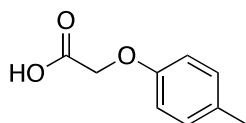


Deprotected azidoethyl galactoside **S6** was obtained as previously reported in good yields (176 mg, 0.45 mmol, 84%). The analytical data match the reported spectra in Kuhaudomlarp *et al.*^[3]

General procedure for phenoxy acetic acids

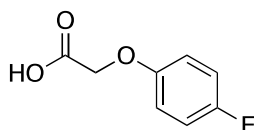
The reactions were performed in analogy to Anbarasa *et al.* and Dewei *et al.*^[4,5] *p*-Cresol or 4-fluorophenol (1 g, 1.0 eq.) and NaOH (849 mg, 21.2 mmol, 2.2 eq.) were dissolved in water (15 mL). Then, a solution of chloroacetic acid (1.04 g, 11.1 mmol, 1.2 eq.) in water (3.5 mL) was added and the mixture was heated to 70 °C. After 16 h, the reaction was stopped by adding 2 M HCl and adjusting the pH to 1–2. The precipitate was collected and dried at 40 °C for 2 h. The products were purified by dissolution under basic conditions (pH = 8–9, K₂CO₃ aq.) and repeated precipitation at pH = 1–2 (HCl aq.).

2-(*p*-tolxyloxy)acetic acid (**SI-1**)



Compound **SI-1** was obtained as a white solid (1.03 g, 6.17 mmol, 67%). ¹H NMR (500 MHz, CDCl₃) δ 7.11 (d, *J* = 8.0 Hz, 2H, 2x ArCH), 6.92 – 6.65 (m, 2H, 2x ArCH), 4.65 (s, 2H, CH₂), 2.30 (s, 3H, CH₃). ¹³C NMR (126 MHz, CDCl₃) δ 172.97 (C=O), 155.33 (ArC), 131.76 (ArC), 130.30 (2x ArCH), 114.67 (2x ArCH), 65.18 (CH₂), 20.65 (CH₃).

2-(4-Fluorophenoxy)acetic acid (SI-2)

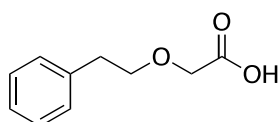


Compound **SI-2** was obtained as a white solid (518 mg, 6.17 mmol, 56%). ¹H NMR (500 MHz, CDCl₃) δ 7.06 – 6.95 (m, 2H, 2x ArCH), 6.95-6.77 (2H, 2x ArCH), 4.66 (s, 2H, CH₂). ¹³C NMR (126 MHz, CDCl₃) δ 172.51 (C=O), 158.04 (d, *J* = 240.4 Hz, ArCF), 153.50 (d, *J* = 2.5 Hz, ArC), 116.28 (d, *J* = 33.4 Hz, 2x ArCH), 116.02 (d, *J* = 18.1 Hz, 2x ArCH), 65.51 (CH₂).

General procedure for phenylalkoxy acetic acids

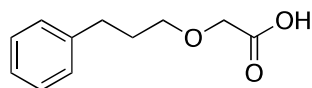
The synthesis was performed in analogy to Machin *et al.* ^[6] The alcohol (1.0 g, 1.0 eq) was dissolved in DMF (final conc. = 0.5 M) and treated with NaH (2.0 eq.). This mixture was stirred at 60 °C until gas evolution stopped (15-30 min). Then, chloroacetic acid (1.0 eq.) dissolved in DMF (2 mL) was added carefully. This reaction was stirred at 60 °C for ca. 60 min. The solvent was removed under reduced pressure and the residue was taken up in H₂O and washed with ether. The aq. phase was acidified with 2 N HCl (pH = 1) and extracted with EtOAc. The org. phase was dried over MgSO₄, filtered and the solvent was removed in vacuo. The crude was purified by flash chromatography (PE/ EtOAc 7:3) to give the desired ethers (24–46%).

2-Phenylethoxyacetic acid (SI-3)



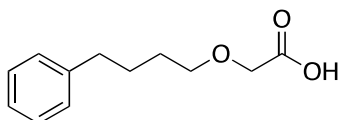
A white solid with fresh flower odor was obtained (684 mg, 3.79 mmol, 46%). ¹H NMR (500 MHz, DMSO-*d*₆) δ 12.58 (s, 1H, COOH), 7.31 – 7.22 (m, 4H, 4x ArCH), 7.22 – 7.16 (m, 1H, ArCH), 4.00 (s, 2H, COCH₂O), 3.67 (t, *J* = 7.0 Hz, 2H, PhCH₂CH₂O), 2.83 (t, *J* = 7.0 Hz, 2H, PhCH₂CH₂O). ¹³C NMR (126 MHz, DMSO-*d*₆) δ 171.66 (COOH), 138.80 (ArC), 128.81 (2x ArCH), 128.21 (2x ArCH), 126.05 (ArCH), 71.26 (PhCH₂CH₂O), 67.38 (COCH₂O), 35.40 (PhCH₂CH₂O). HR-MS (ESI): calcd. for **SI-3** [C₁₀H₁₂O₃ - H]⁺ 179.0714; found 179.0705.

3-Phenylpropoxyacetic acid (SI-4)



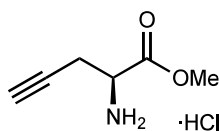
SI-4 was obtained as a white solid (452 mg, 2.23 mmol, 32%). ^1H NMR (500 MHz, $\text{DMSO}-d_6$) δ 12.57 (s, 1H, COOH), 7.27 (t, $J = 7.5$ Hz, 2H, 2x ArCH), 7.23 – 7.14 (m, 3H, 3x ArCH), 3.98 (s, 2H, COCH_2O), 3.45 (t, $J = 6.4$ Hz, 2H, $\text{PhCH}_2\text{CH}_2\text{CH}_2\text{O}$), 2.66 – 2.59 (m, 2H, $\text{PhCH}_2\text{CH}_2\text{CH}_2\text{O}$), 1.84 – 1.76 (m, 2H, $\text{PhCH}_2\text{CH}_2\text{CH}_2\text{O}$). ^{13}C NMR (126 MHz, $\text{DMSO}-d_6$) δ 171.77 (COOH), 141.77 (ArC), 128.31 (2x ArCH), 128.28 (2x ArCH), 125.72 (ArCH), 69.86 ($\text{PhCH}_2\text{CH}_2\text{CH}_2\text{O}$), 67.43 (COCH_2O), 31.66 ($\text{PhCH}_2\text{CH}_2\text{CH}_2\text{O}$), 31.01 ($\text{PhCH}_2\text{CH}_2\text{CH}_2\text{O}$). HR-MS (ESI): calcd. for **SI-4** $[\text{C}_{11}\text{H}_{14}\text{O}_3 - \text{H}]^-$ 193.0870; found 193.0860.

4-Phenylbutoxyacetic acid (SI-6)



Compound **SI-6** was obtained as an oil/white solid (330 mg, 1.58 mmol, 24%). ^1H NMR (500 MHz, $\text{DMSO}-d_6$) δ 12.55 (s, 1H, COOH), 7.29 – 7.24 (m, 2H, 2x ArCH), 7.21 – 7.13 (m, 3H, 3x ArCH), 3.95 (s, 2H, COCH_2O), 3.45 (t, $J = 6.4$ Hz, 2H, $\text{PhCH}_2\text{CH}_2\text{CH}_2\text{CH}_2\text{O}$), 2.58 (t, $J = 7.6$ Hz, 2H, $\text{PhCH}_2\text{CH}_2\text{CH}_2\text{CH}_2\text{O}$), 1.65 – 1.55 (m, 2H, $\text{PhCH}_2\text{CH}_2\text{CH}_2\text{CH}_2\text{O}$), 1.52 (tt, $J = 8.3$, 6.2 Hz, 2H, $\text{PhCH}_2\text{CH}_2\text{CH}_2\text{CH}_2\text{O}$). ^{13}C NMR (126 MHz, $\text{DMSO}-d_6$) δ 171.78 (COOH), 142.18 (ArC), 128.30 (2x ArCH), 128.23 (2x ArCH), 125.64 (ArCH), 70.23 ($\text{PhCH}_2\text{CH}_2\text{CH}_2\text{CH}_2\text{O}$), 67.40 (COCH_2O), 34.83 ($\text{PhCH}_2\text{CH}_2\text{CH}_2\text{CH}_2\text{O}$), 28.70 ($\text{PhCH}_2\text{CH}_2\text{CH}_2\text{CH}_2\text{O}$), 27.61 ($\text{PhCH}_2\text{CH}_2\text{CH}_2\text{CH}_2\text{O}$). HR-MS (ESI): calcd. for **SI-6** $[\text{C}_{12}\text{H}_{16}\text{O}_3 - \text{H}]^-$ 207.1027; found 207.1018.

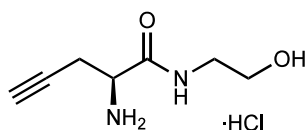
L-H-Pra-OMe *HCl (S8)



The reaction was performed in analogy to Kristensen *et al.*^[7] L-H-Pra-OH*HCl (200 mg, 1.77 mmol, 1.0 eq.) was suspended in MeOH (5 mL, final conc. 0.8 M) and thionyl chloride (150 μL , 1.2 eq.) was added dropwise over 15 min at 0 $^\circ\text{C}$. The reaction was warmed to 25 $^\circ\text{C}$ and

stirred for 18 h. Volatiles were removed under reduced pressure and the residue was washed with Et₂O and dried to obtain compound **S8** as an oil (283 mg, 1.73 mmol, 98%, *R*_f = 0.74 CH₂Cl₂/MeOH 9:1). ¹H NMR (500 MHz, CDCl₃) δ 8.84 (s, 3H, NH₃), 4.45 (s, 1H, H_α), 3.88 (s, 3H, CH₃), 3.49 (s, 1H, C≡CH), 3.36 – 2.79 (m, 2H, CH₂). ¹³C NMR (126 MHz, CDCl₃) δ 168.42 (C=O), 76.35 (C≡CH), 74.72 C≡CH, 54.00 (CH_α), 52.21 (CH₃), 20.95 (CH₃). LR-MS: calcd. **S8** [C₆H₉NO₂ + H]⁺ 128.07; found 128.49.

L-H-Pra-NH(CH₂)₂OH*HCl (**S9**)



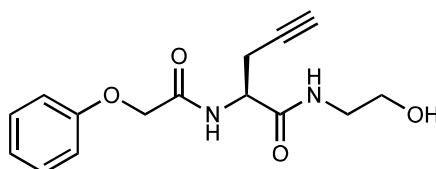
This reaction was performed in analogy to Kristensen *et al.*^[7] L-H-Pra-OMe* HCl (**S8**) (283 mg, 1.73 mmol, 1.0 eq.) was dissolved in pure ethanolamine (1.8 mL, final conc. of **S8** = 1 M) and stirred for 18 h at 25 °C. The mixture was diluted with CH₂Cl₂ and washed with aq. K₂CO₃ (satd.), the aq. phase was extracted with CH₂Cl₂ and the combined organic phases were dried over Na₂SO₄, filtered and the solvents were removed under reduced pressure. The obtained yellow oil was washed with Et₂O and dried to give compound **S9** (233 mg, 1.21 mmol, 70%, *R*_f = 0.56 CH₂Cl₂/MeOH 8:2). ¹H NMR (500 MHz, DMSO-*d*₆) δ 7.94 (br s, 1H, NH), 4.68 (br s, 1H, OH), 3.40 (q, *J* = 5.9 Hz, 2H, NHCH₂CH₂OH), 3.31 – 3.24 (m, 1H, H_α), 3.20 – 3.06 (m, 2H, NHCH₂CH₂OH), 2.83 (t, *J* = 2.6 Hz, 1H, C≡CH), 2.47 – 2.33 (m, 2H, CCH₂CH_α), 1.92 (s, 2H, NH₂). ¹³C NMR (126 MHz, DMSO-*d*₆) δ 173.26 (C=O), 81.93 (C≡CH), 73.43 (C≡CH), 60.29 (NHCH₂CH₂OH), 53.97 (CH_α), 41.77 (NHCH₂CH₂OH), 25.25 (CCH₂CH_α). HR-MS (ESI): calcd. for **S9** [C₇H₁₂N₂O₂ + H]⁺ 157.0972; found 157.0972.

General procedure for the (*S*)-*N*-(2-hydroxyethyl)-2-(acetamido)pent-4-ynamides

The peptide couplings were performed in analogy as follows: L-H-Pra-NH(CH₂)₂OH (**S9**) (50 mg, 0.32 mmol) was dissolved in DMF (final conc. of **S9** = 0.1 M) and EDC*HCl (1.2 eq.), HOBt (1.2 eq.), the individual acids (2.0 eq.) and DIPEA (2.5 eq.) were added. The mixture was stirred at 25 °C for 24 h. The solvent was removed under reduced pressure and the crude was taken up in a mixture of EtOAc and 1 N HCl (1:1) and the phases were separated. The aq. phase was extracted

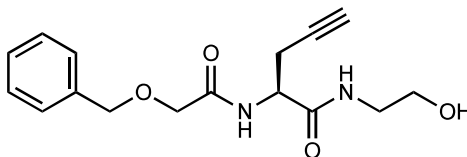
with EtOAc (2x). The combined org. phases were washed with NaHCO₃ (aq.), brine, dried over Na₂SO₄, filtered and concentrated under reduced pressure. The residue was purified by flash chromatography (CH₂Cl₂/MeOH) to give the products as white solids (36–71%).

(S)-N-(2-hydroxyethyl)-2-(2-phenoxyacetamido)pent-4-ynamide (S10-a)



125 mg of **S9** (0.8 mmol) were used for the reaction and 1 eq. of phenoxyacetic acid. Different to the general procedure, the white precipitate containing the product was filtered and washed with Et₂O to give compound **S10-1** (155 mg, 0.53 mmol, 67%, *R*_f = 0.60 CH₂Cl₂/MeOH 9:1). ¹H NMR (500 MHz, CDCl₃) δ 7.39 (d, *J* = 7.6 Hz, 1H, CHNHCO), 7.33 (dd, *J* = 8.7, 7.3 Hz, 2H, 2x ArCH), 7.04 (td, *J* = 7.4, 1.1 Hz, 1H, ArCH), 6.98 – 6.90 (m, 2H, 2x ArCH), 6.57 (s, 1H, CONHCH₂), 4.62 (td, *J* = 7.5, 5.6 Hz, 1H, CH₂CHNHCO), 4.56 (s, 2H, OCH₂CO), 3.73 (t, *J* = 5.0 Hz, 2H, NHCH₂CH₂OH), 3.45 (t, *J* = 5.3 Hz, 2H, NHCH₂CH₂OH), 2.91 – 2.80 (m, 1H, CHCH₂C≡CH), 2.69 – 2.60 (m, 1H, CHCH₂C≡CH), 2.10 (t, *J* = 2.6 Hz, 1H, C≡CH). ¹³C NMR (126 MHz, CDCl₃) δ 170.16 (C=O), 168.94 (C=O), 157.13 (ArC), 130.02 (2x ArCH), 122.51 (ArCH), 114.87 (2x ArCH), 79.16 (CH₂CCH), 72.18 (C≡CH), 67.34 (COCH₂O), 61.99 (NHCH₂CH₂OH), 51.49 (CH₂CHNHCO), 42.64 (NHCH₂CH₂OH), 22.24 (CHCH₂CCH). HR-MS (ESI): calcd. for **S10-1** [C₁₅H₁₉N₂O₄ + H]⁺ 291.1339; found 291.1337.

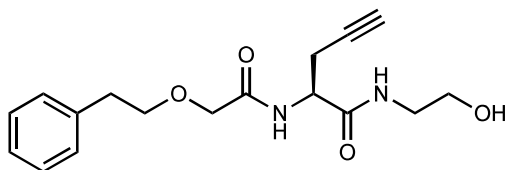
(S)-2-(2-(benzyloxy)acetamido)-N-(2-hydroxyethyl)pent-4-ynamide (S10-b)



The crude product was purified by flash chromatography to give a yellow oily solid (48.6 mg, 0.16 mmol, 50%, *R*_f = 0.39; CH₂Cl₂/MeOH 19:1). ¹H NMR (500 MHz, DMSO-*d*₆) δ 8.09 (t, *J* = 5.7 Hz, 1H, CONHCH₂), 7.84 (d, *J* = 8.3 Hz, 1H, CHNHCO), 7.38 (d, *J* = 4.4 Hz, 4H, 4x ArCH), 7.35 – 7.29 (m, 1H, ArCH), 4.68 (t, *J* = 5.5 Hz, 1H, NHCH₂CH₂OH), 4.57 (d, *J* = 1.3 Hz, 2H,

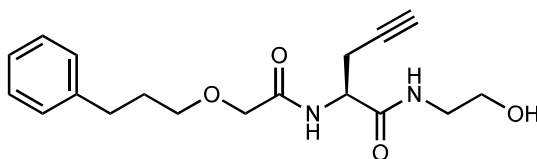
COCH₂O), 4.49 – 4.43 (m, 1H, CH₂CHNHCO), 4.04 – 3.91 (m, 2H, OCH₂Ar), 3.42 – 3.37 (m, 2H, NHCH₂CH₂OH), 3.20 – 3.09 (m, 2H, NHCH₂CH₂OH), 2.86 (t, *J* = 2.6 Hz, 1H, HC≡C), 2.65 – 2.53 (m, 2H, HCCCH₂CH). ¹³C NMR (126 MHz, DMSO-*d*₆) δ 169.42 (C=O), 168.75 (C=O), 137.56 (ArC), 128.37 (2x ArCH), 127.85 (2x ArCH), 127.76 (ArCH), 80.35 (HC≡CCH₂), 73.15 (HC≡C), 72.37 (COCH₂O), 69.01 (OCH₂Ar), 59.64 (NHCH₂CH₂OH), 50.49 (CH₂CHNHCO), 41.61 (NHCH₂CH₂OH), 22.16 (HC≡CCH₂CH). HR-MS (ESI): calcd. for **S10-2** [C₁₆H₂₀N₂O₄ + H]⁺ 305.1496; found 305.1491.

(S)-N-(2-hydroxyethyl)-2-(2-phenethoxyacetamido)pent-4-ynamide (S10-c)



The product was obtained as a white powder (37 mg, 0.12 mmol, 36%, *R*_f = 0.39; CH₂Cl₂/MeOH 9:1). ¹H NMR (500 MHz, MeOH-*d*₄) δ 7.31 – 7.25 (m, 4H, 4x ArCH), 7.23 – 7.16 (m 1H, ArCH), 4.53 (td, *J* = 7.0, 5.8 Hz, 1H, CH₂CHCONH), 4.06 – 3.92 (m, 2H, COCH₂O), 3.77 (t, *J* = 6.8 Hz, 2H, OCH₂CH₂Ph), 3.60 (td, *J* = 5.9, 1.5 Hz, 2H, NHCH₂CH₂OH), 3.33 (d, *J* = 5.9 Hz, 2H, NHCH₂CH₂OH), 2.94 (t, *J* = 6.8 Hz, 1H, OCH₂CH₂Ph), 2.74 – 2.54 (m, 2H, CHCH₂C≡CH), 2.39 (t, *J* = 2.6 Hz, 1H, C≡CH). ¹³C NMR (126 MHz, MeOH-*d*₄) δ 172.54 (C=O), 172.07 (C=O), 140.05 (ArC), 130.07 (2x ArCH), 129.50 (2x ArCH), 127.35 (ArCH), 79.93 (CH₂C≡CH), 73.78 (OCH₂CH₂Ph), 72.60 (C≡CH), 70.84 (COCH₂O), 61.40 (NHCH₂CH₂OH), 52.61 (NHCH₂CH₂OH), 43.05 (NHCH₂CH₂OH), 37.10 (OCH₂CH₂Ph), 23.03 (CHCH₂C≡CH). HR-MS (ESI): calcd. for **S10-3** [C₁₇H₂₂N₂O₄ + H]⁺ 319.165; found 319.165.

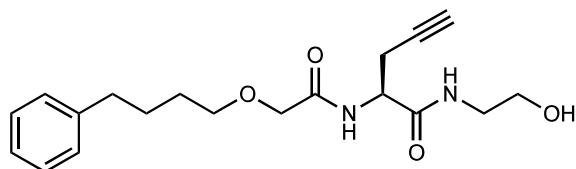
(S)-N-(2-hydroxyethyl)-2-(2-(3-phenylpropoxy)acetamido)pent-4-ynamide (S10-d)



The product was obtained as a white powder (55 mg, 0.17 mmol, 52%, *R*_f = 0.42; CH₂Cl₂/MeOH 9:0.25). ¹H NMR (500 MHz, DMSO-*d*₆) δ 8.08 (t, *J* = 5.7 Hz, 1H, NHCH₂CH₂OH), 7.74 (d, *J* =

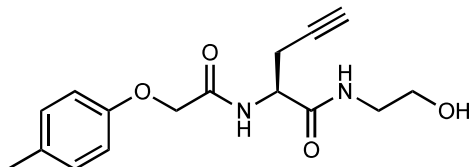
8.3 Hz, 1H, CONHCH), 7.28 (t, $J = 7.5$ Hz, 2H, 2x ArCH), 7.25 – 7.14 (m, 3H, 3x ArCH), 4.68 (t, $J = 5.5$ Hz, 1H, NHCH₂CH₂OH), 4.48 – 4.41 (m, 1H, NHCHCH₂CO), 3.95 – 3.85 (m, 2H, COCH₂O), 3.47 (tt, $J = 6.2, 3.1$ Hz, 2H, OCH₂CH₂CH₂Ph), 3.43 – 3.35 (m, 2H, NHCH₂CH₂OH), 3.19 – 3.09 (m, 2H, NHCH₂CH₂OH), 2.83 (t, $J = 2.6$ Hz, 1H, C≡CH), 2.70 – 2.62 (m, 2H, OCH₂CH₂CH₂Ph), 2.62 – 2.56 (m, 2H, CHCH₂C≡CH), 1.90 – 1.81 (m, 2H, OCH₂CH₂CH₂Ph). ¹³C NMR (126 MHz, DMSO-*d*₆) δ 169.86 (C=O), 169.45 (C=O), 142.11 (ArC), 128.78 (4x ArCH), 126.22 (ArCH), 80.71 (CH₂C≡CH), 73.55 (C≡CH), 70.61 (OCH₂CH₂CH₂Ph), 70.03 (COCH₂O), 60.10 (NHCH₂CH₂OH), 50.85 (NHCHCH₂CO), 42.07 (NHCH₂CH₂OH), 32.04 (OCH₂CH₂CH₂Ph), 31.26 (OCH₂CH₂CH₂Ph), 22.65 (CHCH₂C≡CH). HR-MS (ESI): calcd. for **S10-4** [C₁₈H₂₂N₂O₄ + H]⁺ 333.1809; found 333.1804.

(S)-N-(2-hydroxyethyl)-2-(2-(4-phenylbutoxy)acetamido)pent-4-ynamide (S10-e)



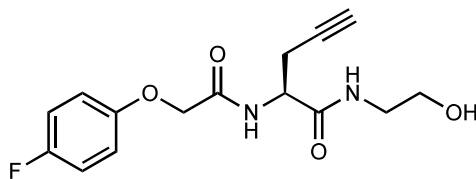
The product was obtained as a white powder (44.3, 0.13 mmol, 40%, $R_f = 0.59$; CH₂Cl₂/MeOH 19:1). ¹H NMR (500 MHz, MeOH-*d*₄) δ 7.23 (t, $J = 7.6$ Hz, 2H, 2x ArCH), 7.20 – 7.15 (m, 2H, 2x ArCH), 7.15 – 7.09 (m, 1H, ArCH), 4.54 (t, $J = 6.3$ Hz, 1H, CH₂CHCONH), 4.01 – 3.90 (m, 2H, COCH₂O), 3.60 – 3.57 (m, 2H, NHCH₂CH₂OH), 3.55 (t, $J = 6.3$ Hz, 2H, OCH₂CH₂CH₂CH₂Ph), 3.36 – 3.27 (m, 2H, NHCH₂CH₂OH), 2.71 – 2.66 (m, 2H, CHCH₂C≡CH), 2.63 (t, $J = 7.4$ Hz, 2H, OCH₂CH₂CH₂CH₂Ph), 2.31 (t, $J = 2.7$ Hz, 1H, C≡CH), 1.77 – 1.68 (m, 2H, OCH₂CH₂CH₂CH₂Ph), 1.68 – 1.60 (m, 2H, OCH₂CH₂CH₂CH₂Ph). ¹³C NMR (126 MHz, MeOH-*d*₄) δ 172.62 (C=O), 172.06 (C=O), 143.59 (ArC), 129.46 (2x ArCH), 129.30 (2x ArCH), 126.75 (ArCH), 79.86 (CH₂C≡CH), 72.73 (C≡CH), 72.61 (OCH₂CH₂CH₂CH₂Ph), 70.82 (COCH₂O), 61.39 (NHCH₂CH₂OH), 52.52 (CH₂CHCONH), 43.05 (NHCH₂CH₂OH), 36.59 (OCH₂CH₂CH₂CH₂Ph), 30.13 (OCH₂CH₂CH₂CH₂Ph), 29.05 (OCH₂CH₂CH₂CH₂Ph), 23.09 (CHCH₂CCH). HR-MS (ESI): calcd. for **S10-e** [C₁₉H₂₆N₂O₄ + H]⁺ 347.1965; found 347.1960.

(S)-N-(2-hydroxyethyl)-2-(2-(*p*-toloxy)acetamido)pent-4-ynamide (S10-f)



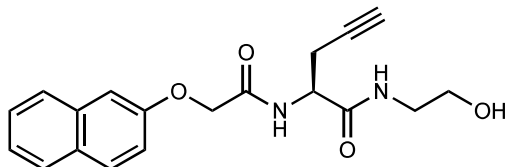
The product was obtained as a white solid (64.1 mg, 0.21 mmol, 71%, $R_f = 0.28$; $\text{CH}_2\text{Cl}_2/\text{MeOH}$ 19:1). ^1H NMR (500 MHz, CDCl_3) δ 7.39 (d, $J = 7.6$ Hz, 1H, CHNHCO), 7.12 (d, $J = 8.5$ Hz, 2H, 2x ArCH), 6.84 (d, $J = 8.7$ Hz, 2H, 2x ArCH), 6.58 (s, 1H, CONHCH_2), 4.61 (td, $J = 7.6, 5.6$ Hz, 1H, CH_2CHNHCO), 4.52 (s, 2H, COCH_2OAr), 3.72 (t, $J = 4.9$ Hz, 2H, $\text{NHCH}_2\text{CH}_2\text{OH}$), 3.44 (t, $J = 5.2$ Hz, 1H, $\text{NHCH}_2\text{CH}_2\text{OH}$), 2.91 – 2.59 (m, 2H, $\text{HC}\equiv\text{CCH}_2\text{CH}$), 2.30 (s, 3H, CH_3), 2.10 (d, $J = 2.6$ Hz, 1H, $\text{C}\equiv\text{CH}$). ^{13}C NMR (126 MHz, CDCl_3) δ 170.22 ($\text{C}=\text{O}$), 169.19 ($\text{C}=\text{O}$), 155.05 (ArCOR), 131.90 (ArCCH_3), 130.41 (2x ArCH), 114.69 (2x ArCH), 79.14 ($\text{HC}\equiv\text{CCH}_2$), 72.17 ($\text{HC}\equiv\text{C}$), 67.47 (COCH_2OAr), 61.95 ($\text{NHCH}_2\text{CH}_2\text{OH}$), 51.44 (CH_2CHNHCO), 42.63 ($\text{NHCH}_2\text{CH}_2\text{OH}$), 22.17 ($\text{HC}\equiv\text{CCH}_2$), 20.64 (CH_3). HR-MS (ESI): calcd. for **S10-6** [$\text{C}_{16}\text{H}_{20}\text{N}_2\text{O}_4 + \text{H}$] $^+$ 305.1496; found 305.1491.

L-*p*-Fluorophenoxy-Pra-N(CH_2) $_2$ OH (S10-g)



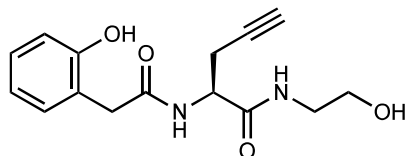
The compounds was obtained as a yellowish powder (64.1 mg, 0.21 mmol, 65%, $R_f = 0.31$; $\text{CH}_2\text{Cl}_2/\text{MeOH}$ 19:1). ^1H NMR (500 MHz, $\text{MeOH}-d_4$) δ 7.09 – 7.01 (m, 4H, 4x ArCH), 4.67 – 4.61 (m, 1H, CH_2CHNHCO), 4.59 (d, $J = 3.8$ Hz, 2H, COCH_2OAr), 3.62 (td, $J = 5.9, 1.4$ Hz, 2H, $\text{NHCH}_2\text{CH}_2\text{OH}$), 3.36 – 3.34 (m, 2H, $\text{NHCH}_2\text{CH}_2\text{OH}$), 2.81 – 2.66 (m, 2H, $\text{HC}\equiv\text{CCH}_2\text{CH}$), 2.40 (t, $J = 2.7$ Hz, 1H, $\text{HC}\equiv\text{C}$). ^{13}C NMR (126 MHz, $\text{MeOH}-d_4$) δ 170.73 ($\text{C}=\text{O}$), 169.59 ($\text{C}=\text{O}$), 156.96 (ArC), 154.00 (ArC), 115.96 (ArCH), 115.90 (ArCH), 115.60 (ArCH), 115.41 (ArCH), 78.63 ($\text{HC}\equiv\text{CCH}_2$), 71.08 ($\text{HC}\equiv\text{C}$), 67.39 (COCH_2OAr), 59.98 ($\text{NHCH}_2\text{CH}_2\text{OH}$), 51.58 (CH_2CHNHCO), 41.65 ($\text{NHCH}_2\text{CH}_2\text{OH}$), 21.44 ($\text{HC}\equiv\text{CCH}_2\text{CH}$). HR-MS (ESI): calcd. for **S10-7** [$\text{C}_{15}\text{H}_{17}\text{FN}_2\text{O}_4 - \text{H}$] $^-$ 307.1100; found 307.1089.

(S)-N-(2-hydroxyethyl)-2-(2-(naphthalen-2-yloxy)acetamido)pent-4-ynamide (S10-h)



The product was obtained as a white powder (51.0 mg, 0.15 mmol, 47%, R_f = 0.19; $\text{CH}_2\text{Cl}_2/\text{MeOH}$ 19:1). ^1H NMR (500 MHz, $\text{MeOH}-d_4$) δ 7.86 – 7.77 (m, 3H, 3x naphthyl-H), 7.46 (ddd, J = 8.2, 6.8, 1.3 Hz, 1H, naphthyl-H), 7.37 (ddd, J = 8.1, 6.8, 1.3 Hz, 1H, naphthyl-H), 7.32 – 7.28 (m, 2H, 2x naphthyl-H), 4.75 (d, J = 3.4 Hz, 2H, COCH_2OR), 4.70 – 4.62 (m, 1H, CH_2CHNHCO), 3.60 (td, J = 5.9, 1.5 Hz, 2H, $\text{NHCH}_2\text{CH}_2\text{OH}$), 3.38 – 3.34 (m, 2H, $\text{NHCH}_2\text{CH}_2\text{OH}$), 2.88 – 2.62 (m, 2H, $\text{HC}\equiv\text{CCH}_2\text{CH}$), 2.36 (s, 1H, $\text{HC}\equiv\text{C}$). ^{13}C NMR (126 MHz, $\text{MeOH}-d_4$) δ 170.76 ($\text{C}=\text{O}$), 169.66 ($\text{C}=\text{O}$), 155.51 (naphthyl-C), 134.49 (naphthyl-C), 129.58 (naphthyl-C), 129.32 (naphthyl-CH), 127.23 (naphthyl-CH), 126.64 (naphthyl-CH), 126.17 (naphthyl-CH), 123.79 (naphthyl-CH), 118.12 (naphthyl-CH), 107.18 (naphthyl-CH), 78.63 $\text{HC}\equiv\text{CCH}_2$, 71.06 $\text{HC}\equiv\text{C}$, 66.74 (COCH_2Ar), 59.97 ($\text{NHCH}_2\text{CH}_2\text{OH}$), 51.63 (CH_2CHNHCO), 41.65 ($\text{NHCH}_2\text{CH}_2\text{OH}$), 21.45 ($\text{HC}\equiv\text{CCH}_2\text{CH}$). HR-MS (ESI): calcd. for **S10-8** [$\text{C}_{19}\text{H}_{20}\text{N}_2\text{O}_4 + \text{H}$] $^+$ 341.1496; found 341.1491.

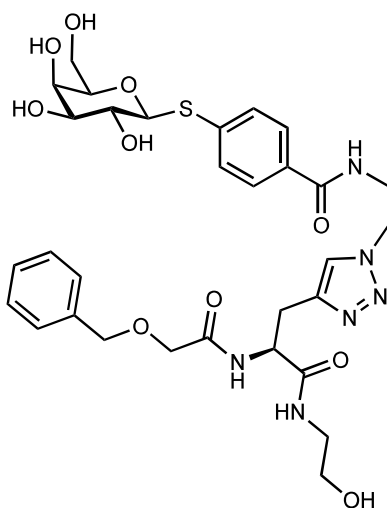
(S)-N-(2-hydroxyethyl)-2-(2-(2-hydroxyphenyl)acetamido)pent-4-ynamide (S10-i)



The product was obtained as a white powder (38.7 mg, 0.13 mmol, 42%, R_f = 0.17; $\text{CH}_2\text{Cl}_2/\text{MeOH}$ 19:1). ^1H NMR (500 MHz, $\text{DMSO}-d_6$) δ 9.64 (s, 1H, ArOH), 8.17 (d, J = 8.1 Hz, 1H, CHNHCO), 8.00 (t, J = 5.8 Hz, 1H, CONHCH_2), 7.11 – 7.05 (m, 1H, ArCH), 7.05 – 7.01 (m, 1H, ArCH), 6.81 – 6.76 (m, 1H, ArCH), 6.76 – 6.70 (m, 1H, ArCH), 4.68 (t, J = 5.3 Hz, 1H, CH_2OH), 4.39 (td, J = 7.6, 5.9 Hz, 1H, CH_2CHNHCO), 3.51 – 3.41 (m, 2H, COCH_2Ar), 3.41 – 3.36 (m, 2H, $\text{NHCH}_2\text{CH}_2\text{OH}$), 3.19 – 3.06 (m, 2H, $\text{NHCH}_2\text{CH}_2\text{OH}$), 2.81 (t, J = 2.6 Hz, 1H, $\text{HC}\equiv\text{C}$), 2.57 – 2.41 (m, 2H, $\text{HC}\equiv\text{CCH}_2\text{CH}$). ^{13}C NMR (126 MHz, $\text{DMSO}-d_6$) δ 171.29 ($\text{C}=\text{O}$), 170.16 ($\text{C}=\text{O}$), 155.86 (ArC), 131.12 (ArCH), 128.15 (ArCH), 122.95 (ArC), 119.30 (ArCH), 115.58 (ArCH), 80.98 $\text{HC}\equiv\text{CCH}_2$, 73.41 ($\text{HC}\equiv\text{C}$), 60.11 ($\text{NHCH}_2\text{CH}_2\text{OH}$), 51.95 (CH_2CHNHCO), 42.00

= 6.2 Hz, 2H, NHCH₂CH₂triazol), 3.72 (t, *J* = 3.8 Hz, 1H, H-4?), 3.69 – 3.55 (m, 2H, NHCH₂CH₂triazol), 3.57 – 3.42 (m, 4H, H-2, H-3?, NHCH₂CH₂OH), 3.41 – 3.34 (m, 3H, H-5?, H-6), 3.18 – 3.10 (m, 1H, 1x triazoleCH₂CH), 3.10 – 3.02 (m, 1H, NHCH₂CH₂OH), 2.98 (dd, *J* = 14.7, 8.1 Hz, 1H, 1x triazoleCH₂CH). ¹³C NMR (126 MHz, DMSO-*d*₆) δ 170.22 (C=O), 167.55 (C=O), 166.08 (C=O), 157.63 OArC, 142.65 (SArC), 140.15 (SArC), 131.13 (triazole-C), 129.49 (2x OArCH), 127.73 (2x SArCH), 127.57 (2x SArCH), 123.14 (triazole-CH), 121.17 (OArCH), 114.64 (2x OArCH), 86.67 (C-1), 79.28 (C-5), 74.66 (C-3)), 69.12 (C-2), 68.37 (C-4), 66.60 (COCH₂O), 60.59 (NHCH₂CH₂OH), 59.60 (C-6), 52.15 (CH₂CHNHCO), 48.51 (NHCH₂CH₂triazole), 41.56 (NHCH₂CH₂OH), 39.35 (NHCH₂CH₂triazole), 28.31 (triazoleCH₂CH). HRMS (ESI) calcd. for **1** [C₃₀H₃₈N₆O₁₀S + H]⁺ 675.2443; found, 675.2437.

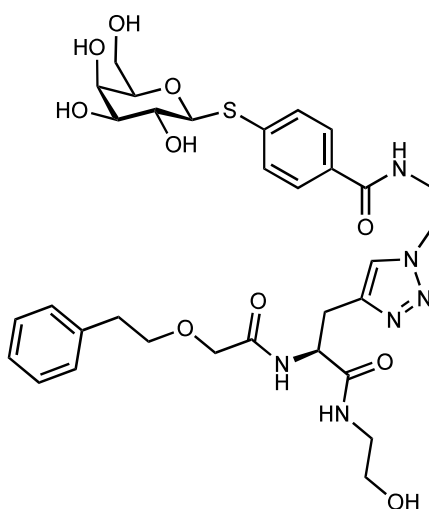
Benzyloxy-modified galactoside **2**



The product was obtained as a white powder (21.3 mg, 0.031 mmol, 80%, *R*_f = 0.23; CH₂Cl₂/MeOH 19:1). ¹H NMR (500 MHz, DMSO-*d*₆) δ 8.56 (t, *J* = 5.7 Hz, 1H, NHCH₂CH₂triazole), 8.03 (t, *J* = 5.7 Hz, 1H, NHCH₂CH₂OH), 7.83 (d, *J* = 8.2 Hz, 1H, CHNHCO), 7.80 (s, 1H, triazole-H), 7.73 – 7.69 (m, 2H, 2x SArCH), 7.50 – 7.46 (m, 2H, 2x SArCH), 7.38 – 7.27 (m, 5H, 5x ArCH), 5.22 (d, *J* = 6.1 Hz, 1H, OH-2), 4.93 (d, *J* = 5.6 Hz, 1H, OH-3), 4.76 – 4.63 (m, 3H, H-1, OH-4, NHCH₂CH₂OH), 4.56 – 4.51 (m, 2H, OH-6?, CH₂CHCO), 4.49 (d, *J* = 0.9 Hz, 2H, COCH₂O), 4.47 (t, *J* = 6.3 Hz, 2H, NHCH₂CH₂triazole), 3.94 – 3.84 (m, 2H, OCH₂Ar), 3.72 (t, *J* = 4.0 Hz, 1H, H-4), 3.69 – 3.56 (m, 2H, NHCH₂CH₂triazole), 3.56 – 3.43 (m, 4H, H-2, H-3, NHCH₂CH₂OH), 3.40 – 3.36 (m, 3H, H-5?, H-6?), 3.19 – 3.11 (m, 1H, 1x triazoleCH₂CH),

3.11 – 3.03 (m, 2H, NHCH₂CH₂OH), 2.99 (dd, *J* = 14.7, 7.7 Hz, 1H, 1x triazoleCH₂CH). ¹³C NMR (126 MHz, DMSO-*d*₆) δ 170.25 (C=O), 168.78 NHCOCH₂O, 166.12 (C=O), 142.61 (SArC), 140.20 (SArC), 137.62 (ArC), 131.14 (triazole-C), 128.34 (2x ArCH), 127.78 (2x ArCH), 127.74 (ArCH), 127.71 (2x SArCH), 127.61 (2x SArCH), 123.24 (triazole-CH), 86.68 (C-1), 79.30 (C-5), 74.68 (C-3), 72.32 (OCH₂Ar), 69.14 (C-2), 69.05 (COCH₂O), 68.41 (C-4), 60.63 (NHCH₂CH₂OH), 59.63 (C-6), 51.90 (CH₂CHNHCO), 48.51 (NHCH₂CH₂triazole), 41.58 (NHCH₂CH₂OH), 39.78 (NH₂CH₂CH₂triazole) 28.43 (triazoleCH₂CH). HRMS (ESI) calcd. for **2** [C₃₁H₄₀N₆O₁₀S + H]⁺ 689.2599; found, 689.2592.

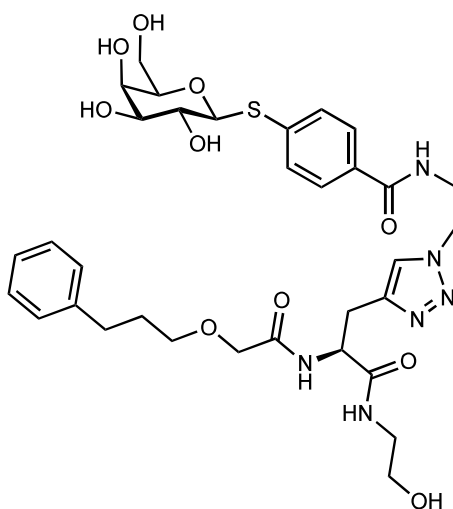
Phenylethoxy-modified galactoside **3**



The product was obtained as a white powder (24.5 mg, 0.035 mmol, 90%, *R*_f = 0.33; CH₂Cl₂/MeOH 19:1). ¹H NMR (500 MHz, DMSO-*d*₆) δ 8.55 (t, *J* = 5.6 Hz, 1H, NHCH₂CH₂triazole), 8.00 (t, *J* = 5.7 Hz, 1H, NHCH₂CH₂OH), 7.77 (s, 1H, triazole-H), 7.75 – 7.69 (m, 2H, 2x SArCH), 7.68 (d, *J* = 8.3 Hz, 1H, CONHCH), 7.51 – 7.46 (m, 2H, 2x SArCH), 7.30 – 7.21 (m, 4H, 4x ArCH), 7.21 – 7.16 (m, 1H, ArCH), 4.70 (d, *J* = 9.6 Hz, 1H, H-1), 4.51 (td, *J* = 7.9, 5.4 Hz, 1H, NHCHCH₂CO), 4.45 (t, *J* = 6.4 Hz, 2H, NHCH₂CH₂triazole), 3.90 – 3.79 (m, 2H, COCH₂O), 3.72 (d, *J* = 3.2 Hz, 1H, H-4), 3.69 – 3.55 (m, 2H, NHCH₂CH₂triazole), 3.60 (t, *J* = 7.0 Hz, 2H, PhCH₂CH₂O), 3.56 – 3.44 (m, 4H, H-2, H-5, H-6), 3.44 – 3.33 (m, 4H, NHCH₂CH₂OH, H-3), 3.18 – 2.91 (m, 4H, NHCH₂CH₂OH, CHCH₂triazole), 2.83 (t, *J* = 7.0 Hz, 2H, PhCH₂CH₂O). ¹³C NMR (126 MHz, DMSO-*d*₆) δ 170.20 (C=O), 168.94 (C=O), 166.08 (C=O), 142.56 (ArC), 140.17 (SArC), 138.66 (triazole-C), 131.13 (SArC), 128.89 (2x ArCH),

128.26 (2x ArCH), 127.73 (2x SArCH), 127.58 (2x SArCH), 126.10 (ArCH), 123.17 (triazole-CH), 86.67 (C-1), 79.28 (C-5), 74.66 (C-3), 71.66 (PhCH₂CH₂O), 69.51 (COCH₂O), 69.12 (C-2), 68.38 (C-4), 60.60 (C-6), 59.61 (NHCH₂CH₂OH), 51.81 (NHCHCH₂CO), 48.47 (NHCH₂CH₂triazole), 41.54 (NHCH₂CH₂OH), 39.61 (NH₂CH₂CH₂triazole), 35.32 (PhCH₂CH₂OH), 28.42 (CHCH₂triazole). HR-MS (ESI): calcd. for **3** [C₃₂H₄₂N₆O₁₀S + H]⁺ 703.2756; found 703.2751.

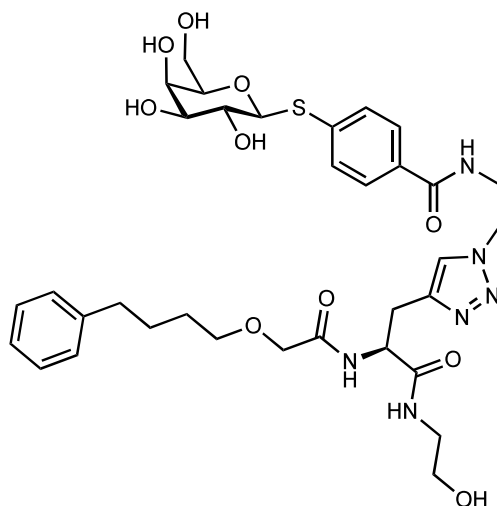
Phenylpropoxy-modified galactoside **4**



The product was obtained as a white powder (25.6 mg, 0.036 mmol, 92%, *R_f* = 0.41; CH₂Cl₂/MeOH 19:1). ¹H NMR (500 MHz, DMSO-*d*₆) δ 8.55 (t, *J* = 5.7 Hz, 1H, NHCH₂CHtriazole), 8.00 (t, *J* = 5.7 Hz, 1H, NHCH₂CH₂OH), 7.80 (s, 1H, triazole-H), 7.73 (d, *J* = 10.7 Hz, 1H, NHCHCH₂CO), 7.71 (m, 2H, 2x SArCH), 7.51 – 7.46 (m, 2H, 2x SArCH), 7.27 (t, *J* = 7.5 Hz, 2H, 2x ArCH), 7.23 – 7.13 (m, 3H, 3x ArCH), 4.70 (d, *J* = 9.6 Hz, 1H, H-1), 4.52 (td, *J* = 7.8, 5.4 Hz, 1H, NHCHCH₂CO), 4.46 (t, *J* = 6.4 Hz, 2H, NHCH₂CH₂triazole), 3.83 (q, *J* = 15.1 Hz, 2H, COCH₂O), 3.72 (d, *J* = 3.2 Hz, 1H, H-4), 3.63 (tdd, *J* = 13.7, 7.6, 5.9 Hz, 2H, NHCH₂CH₂triazole), 3.55 – 3.44 (m, 5H, H-2, H-5, H-6), 3.43 – 3.32 (m, 5H, H-3, NHCH₂CH₂OH, PhCH₂CH₂CH₂O), 3.18 – 2.95 (m, 4H, CHCH₂triazole, NHCH₂CH₂OH), 2.64 – 2.58 (m, 2H, PhCH₂CH₂CH₂O), 1.84 – 1.77 (m, 2H, PhCH₂CH₂CH₂O). ¹³C NMR (126 MHz, DMSO-*d*₆) δ 170.20 (C=O), 169.00 (C=O), 166.07 (C=O), 142.54 (ArC), 141.69 (SArC), 140.16 (triazole-C), 131.12 (SArC), 128.31 (2x ArCH), 128.29 (2x ArCH), 127.72 (2x SArCH), 127.57 (2x SArCH), 125.72 (ArCH), 123.19 (triazole-CH), 86.67 (C-1), 79.27 (C-5), 74.66 (C-3), 70.13

(PhCH₂CH₂CH₂O), 69.57 (COCH₂O), 69.12 (C-2), 68.38 (C-4), 60.59 (C-6), 59.61 (NHCH₂CH₂OH), 51.81 (NHCHCH₂CO), 48.48 (NHCH₂CH₂triazole), 41.54 (NHCH₂CH₂OH), 39.78 (NHCH₂CH₂triazole), 31.54 (PhCH₂CH₂CH₂O), 30.80 (PhCH₂CH₂CH₂O), 28.41 (CHCH₂triazole). HR-MS (ESI): calcd. for **4** [C₃₃H₄₄N₆O₁₀S + H]⁺ 717.2912; found 717.2902.

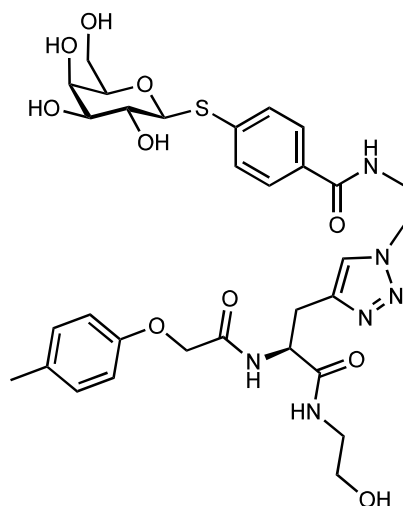
Phenylbutoxy-modified galactoside **5**



The product was obtained as a white powder (24.2 mg, 0.033, 85%, *R*_f = 0.44; CH₂Cl₂/MeOH 19:1). ¹H NMR (500 MHz, DMSO-*d*₆) δ 8.55 (t, *J* = 5.6 Hz, 1H, NHCH₂CH₂triazole), 8.01 (t, *J* = 5.7 Hz, 1H, NHCH₂CH₂OH), 7.79 (s, 1H, triazole-H), 7.71 (d, *J* = 8.5 Hz, 2H, 2x SArCH), 7.68 (d, *J* = 8.2 Hz, 1H, CONHCH), 7.48 (d, *J* = 8.6 Hz, 2H, 2x SArCH), 7.26 (t, *J* = 7.5 Hz, 2H, 2x ArCH), 7.23 – 7.12 (m, 3H, 3x ArCH), 5.20 (s, 1H, OH-2), 4.90 (s, 1H, OH-3), 4.70 (d, *J* = 9.6 Hz, 1H, H-1), 4.66 (s, 1H, OH-4), 4.51 (td, *J* = 7.8, 5.4 Hz, 1H, NHCHCH₂OH), 4.45 (t, *J* = 6.3 Hz, 2H, NHCH₂CH₂triazole), 3.88 – 3.75 (m, 2H, COCH₂O), 3.72 (d, *J* = 3.2 Hz, 1H, H-4), 3.69 – 3.56 (m, 2H, NHCH₂CH₂triazole), 3.55 – 3.44 (m, 5H, H-2, H-5, H-6), 3.42 – 3.27 (m, 6H, NHCH₂CH₂OH, PhCH₂CH₂CH₂CH₂O, H-3), 3.18 – 2.94 (m, 4H, NHCH₂CH₂OH, CHCH₂triazole), 2.57 (t, *J* = 7.5 Hz, 2H, PhCH₂CH₂CH₂CH₂O), 1.65 – 1.56 (m, 2H, PhCH₂CH₂CH₂CH₂O), 1.56 – 1.47 (m, 2H, PhCH₂CH₂CH₂CH₂O). ¹³C NMR (126 MHz, DMSO-*d*₆) δ 170.19 (C=O), 169.02 (C=O), 166.07 (C=O), 142.52 (ArC), 142.09 (SArC), 140.17 (triazole-C), 131.13 (SArC), 128.30 (2x ArCH), 128.23 (2x ArCH), 127.73 (2x SArCH), 127.58 (2x SArCH), 125.64 (ArCH), 123.16 (triazole-CH), 86.68 (C-1), 79.28 (C-5), 74.66 (C-3), 70.62 (PhCH₂CH₂CH₂CH₂O), 69.58 (COCH₂O), 69.12 (C-2), 68.38 (C-4), 60.60 (C-6), 59.60

(NHCH₂CH₂OH), 51.79 (NHCHCH₂CO), 48.47 (NHCH₂CH₂triazole), 41.55 (NHCH₂CH₂OH), 39.78 (NHCH₂CH₂triazole), 34.84 (PhCH₂CH₂CH₂CH₂O), 28.59 (PHCH₂CH₂CH₂CH₂O), 28.45 (CHCH₂triazole), 27.36 (PhCH₂CH₂CH₂CH₂O). HR-MS (ESI): calcd. for **5** [C₃₄H₄₆N₆O₁₀S + H]⁺ 731.3069; found 731.3061.

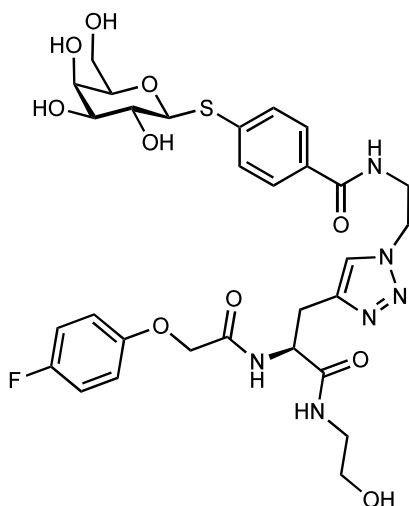
p-Tolyloxy-modified galactoside **6**



The product was obtained as a white powder (22.9 mg, 0.033 mmol, 85%, *R_f* = 0.28; CH₂Cl₂/MeOH 19:1). ¹H NMR (500 MHz, DMSO-*d*₆) δ 8.56 (t, *J* = 5.7 Hz, 1H, NHCH₂CH₂triazole), 8.13 (d, *J* = 8.2 Hz, 1H, CHNHCO), 8.00 (t, *J* = 5.7 Hz, 1H, NHCH₂CH₂OH), 7.74 (s, 1H, triazole-H), 7.71 (d, *J* = 8.6 Hz, 2H, 2x SArCH), 7.48 (d, *J* = 8.6 Hz, 2H, 2x SArCH), 7.10 – 7.05 (m, 2H, 2x OArCH), 6.79 (d, *J* = 8.6 Hz, 2H, 2x OArCH), 5.22 (d, *J* = 6.0 Hz, 1H, OH-2), 4.92 (d, *J* = 5.6 Hz, 1H, OH-3), 4.70 (d, *J* = 9.2 Hz, 1H, H-1), 4.66 (t, *J* = 5.8 Hz, 1H, OH-6), 4.55 – 4.49 (m, 2H, CH₂CHNH, OH-4), 4.48 – 4.39 (m, 4H, COCH₂O, NHCH₂CH₂triazole), 3.72 (t, *J* = 4.0 Hz, 1H, H-4), 3.69 – 3.55 (m, 2H, NHCH₂CH₂triazole), 3.56 – 3.44 (m, 4H, H-2, H-3, NHCH₂CH₂OH), 3.40 – 3.36 (m, 1H, H-5?), 3.32 (s, 2H, H-6), 3.17 – 3.09 (m, 1H, 1x triazoleCH₂CH), 3.09 – 3.03 (m, 2H, NHCH₂CH₂OH), 2.97 (dd, *J* = 14.7, 8.1 Hz, 1H, 1x triazoleCH₂CH), 2.22 (s, 3H, CH₃). ¹³C NMR (126 MHz, DMSO-*d*₆) δ 170.21 (C=O), 167.67 (C=O), 166.08 (C=O), 155.56 (OArC), 142.64 (SArC), 140.15 (SArC), 131.13 (triazole-C), 129.90 (OArCCH₃), 129.82 (2x OArCH), 127.73 (2x SArCH), 127.57 (2x SArCH), 123.16 (triazole-CH), 114.50 (2x OArCH), 86.67 (C-1), 79.28 (C-5), 74.66 (C-3), 69.12 (C-2), 68.39

(C-4), 66.81 (COCH₂O, 60.60 (C-6), 59.60 (NHCH₂CH₂OH, 52.12 (NHCHCH₂CO), 48.50 (NHCH₂CH₂triazole), 41.55 (NHCH₂CH₂OH), 39.61 (NHCH₂CH₂triazole), 28.30 (CHCH₂triazole), 20.08 (OArCCH₃). HRMS (ESI) calcd. For **6** [C₃₁H₄₀N₆O₁₀S + H]⁺ 689.2599; found, 689.2593.

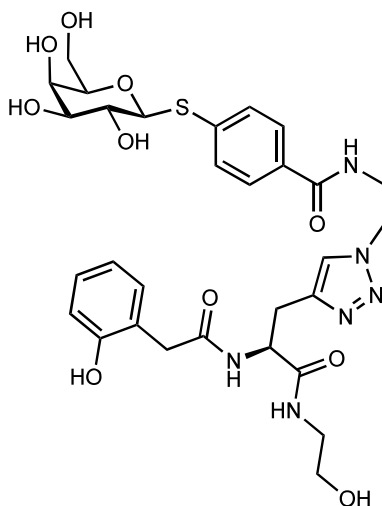
4-Fluorophenylxy-modified galactoside **7**



The product was obtained as a white powder (25.1 mg, 0.036 mmol, 92%, *R*_f = 0.33; CH₂Cl₂/MeOH 19:1). ¹H NMR (500 MHz, DMSO-*d*₆) δ 8.58 (t, *J* = 5.7 Hz, 1H, NHCH₂CH₂triazole), 8.20 (d, *J* = 8.1 Hz, 1H, CHNHCO), 8.02 (t, *J* = 5.7 Hz, 1H, NHCH₂CH₂OH), 7.79 (s, 1H, triazole-H), 7.72 (d, *J* = 8.6 Hz, 1H, 2x SArCH), 7.49 (d, *J* = 8.6 Hz, 1H, 2x SArCH), 7.17 – 7.10 (m, 2H, 2x OArCH), 6.96 – 6.90 (m, 2H, 2x OArCH), 5.23 (d, *J* = 6.1 Hz, 1H, OH-2), 4.94 (d, *J* = 5.7 Hz, 1H, OH-3), 4.74 – 4.67 (m, 3H, H-1, OH-4?, OH-6), 4.57 – 4.51 (m, 2H, CH₂CHNHCO, OH-5?), 4.50 – 4.44 (m, 4H, COCH₂O, NHCH₂CH₂triazole), 3.73 (t, *J* = 4.0 Hz, 1H, H-4), 3.71 – 3.57 (m, 2H, NHCH₂CH₂triazole), 3.57 – 3.44 (m, 4H, NHCH₂CH₂OH, H-2, H-3), 3.41 – 3.38 (m, 1H, H-5?, H-6), 3.19 – 3.12 (m, 1H, 1x triazoleCH₂CH), 3.12 – 3.05 (m, 2H, NHCH₂CH₂OH), 3.02 – 2.94 (m, 1H, 1x triazoleCH₂CH). ¹³C NMR (126 MHz, DMSO-*d*₆) δ 170.25 (C=O), 167.48 (C=O), 166.12 (C=O), 156.82 (d, *J* = 236.44 Hz, OArC), 154.02 (d, *J* = 1.90 Hz, OArCF), 142.70 (SArC), 140.19 (SArC), 131.14 (triazole-C), 127.73 (2x SArCH), 127.60 (2x SArCH), 123.21 (triazole-CH), 116.01 (d, *J* = 14.34 Hz, 2x OArCH), 115.88 (d, *J* = 29.25 Hz, 2x OArCH), 86.68 (C-1), 79.30 (C-5), 74.68 (C-3), 69.13 (C-2), 68.41 (C-4), 67.23 (COCH₂Ar), 60.63 (NHCH₂CH₂OH), 59.63 (C-6), 52.20 (NHCHCO), 48.54 (NHCH₂CH₂triazole), 41.58

(NHCH₂CH₂triazole), 41.59 (NHCH₂CH₂OH), 39.78 (NHCH₂CH₂triazole) 28.32
 (triazoleCH₂CH). HRMS (ESI) calcd. for **8** [C₃₄H₄₀N₆O₁₀S + H]⁺ 725.2599; found, 725.2591.

o*-Hydroxyphenyl-modified galactoside **9*



The product was obtained as a white powder (19.1 mg, 0.028 mmol, 73%, *R*_f = 0.21; CH₂Cl₂/MeOH 19:1). ¹H NMR (500 MHz, DMSO-*d*₆) δ 9.63 (s, 1H, ArOH), 8.58 (t, *J* = 5.7 Hz, 1H, NHCH₂CH₂triazole), 8.11 (d, *J* = 8.2 Hz, 1H, CHNHCO), 7.94 (t, *J* = 5.7 Hz, 1H, NHCH₂CH₂OH), 7.72 (d, *J* = 8.5 Hz, 1H, SArCH), 7.65 (s, 1H, triazole-H), 7.48 (d, *J* = 8.5 Hz, 1H, SArCH), 7.04 (td, *J* = 7.7, 1.7 Hz, 1H, ArCH), 6.99 (dd, *J* = 7.6, 1.7 Hz, 1H, ArCH), 6.78 (dd, *J* = 8.1, 1.3 Hz, 1H, ArCH), 6.71 (td, *J* = 7.4, 1.3 Hz, 1H, ArCH), 5.22 (d, *J* = 6.1 Hz, 1H, OH-2), 4.93 (d, *J* = 5.7 Hz, 1H, OH-3), 4.70 (d, *J* = 9.7 Hz, 1H, H-1), 4.68 (s, 1H, OH-6), 4.54 (d, *J* = 4.4 Hz, 1H, OH-4), 4.49 – 4.41 (m, 3H, NHCH₂CH₂triazole, CH₂CHNHCO), 3.72 (t, *J* = 3.8 Hz, 1H, H-4), 3.70 – 3.55 (m, 1H, NHCH₂CH₂triazole), 3.55 – 3.44 (m, 4H, H-2, H-3?, NHCH₂CH₂OH), 3.41 (s, 2H, COCH₂Ar), 3.39 – 3.37 (m, 3H, H-5, H-6), 3.19 – 3.10 (m, 1H, 1x triazoleCH₂CH), 3.10 – 2.98 (m, 2H, NHCH₂CH₂OH), 2.93 – 2.84 (m, 1H, 1x triazoleCH₂CH). ¹³C NMR (126 MHz, DMSO-*d*₆) δ 170.74 (C=O), 170.52 (C=O), 166.15 (C=O), 155.27 (ArCOH), 142.73 (SArC), 140.22 (SArC), 131.14 (triazole-C), 130.70 (ArCH), 127.78 (ArCH), 127.76 (2x SArCH), 127.63 (2x SArCH), 123.11 (ArC), 122.56 (triazole-CH), 118.99 (ArCH), 115.16 (ArCH), 86.68 (C-1), 79.31 (C-5), 74.68 (C-3), 69.15 (C-2), 68.43 (C-4), 60.65 (NHCH₂CH₂OH), 59.63 (C-6), 52.51 (CH₂CHNHCO), 48.52 (NHCH₂CH₂triazole), 41.52 (NHCH₂CH₂OH), 39.78

519 (NHCH₂CH₂triazole), 37.17 (COCH₂Ar), 28.55 (triazoleCH₂CH). HRMS (ESI) calcd. for **9**
520 [C₃₀H₃₈N₆O₁₀S + H]⁺ 675.2443; found, 675.2433.
521

522 **Molecular Dynamics**523 **Table S1:** Analyzed LecA PDB structures and corresponding co-crystallized ligands, when present.

Ligand	PDB ID	Resolution (Å)
Apo	1l7l	1.5
Apo	1uoj	2.4
β -D-galactose / α -D-galactose	1oko	1.6
Gal α (1 \rightarrow 3) Gal β (1 \rightarrow 4) Glc	2vxj	1.9
Gal- α (1 \rightarrow 2) Gal β -O-Me	2wyf	2.4
4-Hydroxybenzoic acid 4-O-galactoside	3zyb	2.29
4-Nitrophenyl β -D-galactopyranoside	3zyf	1.94
3-(β -D-galactopyranosylthio)propanoic acid	3zyh	1.5
Naphthalen-2-YL-Thio- β -D-galactopyranoside	4a6s	2.15
Melibiose (Gal α (1 \rightarrow 6) Glc)	4al9	1.75
(4S)-N-Ethyl-4-{[N-methyl-3-(1-{2-[(4-sulfanylbzoyl)-aminoethyl]-1H-1,2,3-triazol-4-yl)-L-alanyl]amino}-L-prolinamide- β -D-galactoside	4cp9	1.65
(4S)-N-Ethyl-4-{[N-methyl-3-(1-{2-[(4-sulfanylbzoyl)-aminoethyl]-1H-1,2,3-triazol-4-yl)-L-alanyl]amino}-L-prolinamide- β -D-galactoside	4cpb	1.57
N-Methyl-3-indolyl β -D-galactopyranoside	4ljh	1.45
Chlorophenol Red- β -D-galactopyranoside	4lk6	2.86
Resorufin- β -D-galactopyranoside	4lk7	1.76
4-Hydroxybenzoic acid 4-O-galactoside	4lkd	2.31
4-Hydroxybenzoic acid 4-O-galactoside	4lke	1.65
4-Hydroxybenzoic acid 4-O-galactoside	4lkf	1.64
N-[(2S)-6-amino-1-oxo-1-(pyrrolidin-1-yl)hexan-2-yl]-4-(beta-D- galactopyranosyloxy)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)-3,4- dihydroxy-6-(hydroxymethyl)tetrahydro-2H-pyran-2,5-diyl]bis{1H- 1,2,3-triazole-1,4-diyl[(2S,3R,4S,5S,6S)-3,4-	4yw7	1.82

dihydroxy-6- (hydroxymethyl)tetrahydro-2H-pyran-2,5-diyl]- 1H-1,2,3-triazole-1,4- diylpropane-3,1-diylloxy})bis[6- (hydroxymethyl)tetrahydro-2H-pyran- 3,4,5-triol] Phenyl 6,7-dideoxy-6,7-epoxy-beta-D-galacto- heptopyranoside(6D)	5mih	1.8
--	------	-----

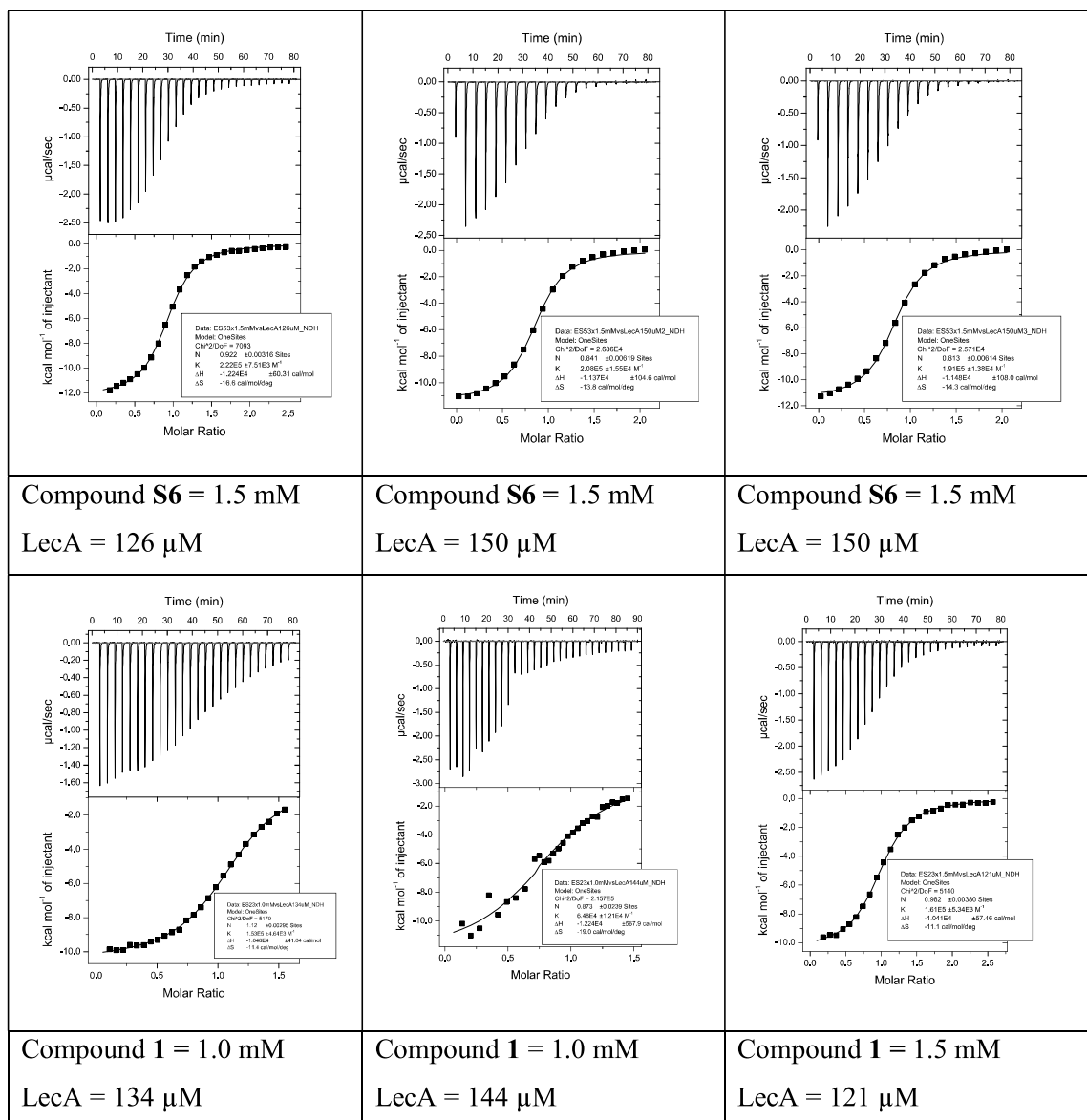
Table S2: Analyzed dimer LecA structures of different PDB files to determine the volume of the cavity.

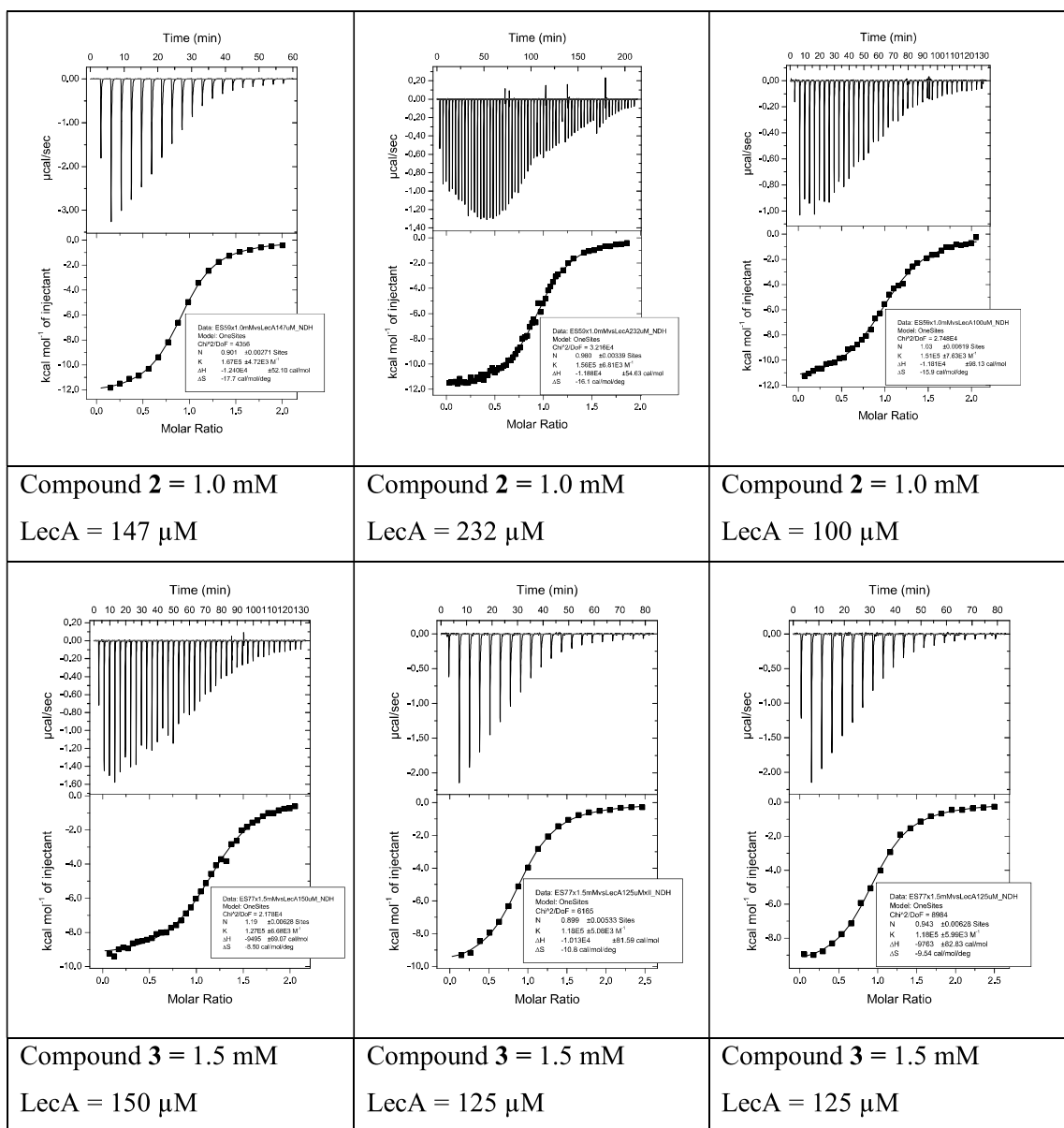
PDB ID	Number of cavities	Nature	Volume (Å ³)
117l	0		
1oko	1	Central pocket	236
1uoj	0		
2vxj	1	Central pocket	297
2wyf	1	Central pocket	307
2wyf	2	Central pocket	297
3zyf	0		-
3zyh	1	Additional cavity	263
3zyh	2	Central pocket	243
4a6s	0		-
4al9	1	Central pocket	280
4cp9	1	Central pocket	303
4cp9	2	Central pocket	286
4cpb	1	Central pocket	253
4ljh	1	Central pocket	314
4ljh	2	Central pocket	273
4lk6	1	Central pocket	344
4lk6	2	Central pocket	297
4lk7	1	Central pocket	294
4lk7	2	Central pocket	260
4lkd	1	Central pocket	284
4lke	1	Central pocket	341

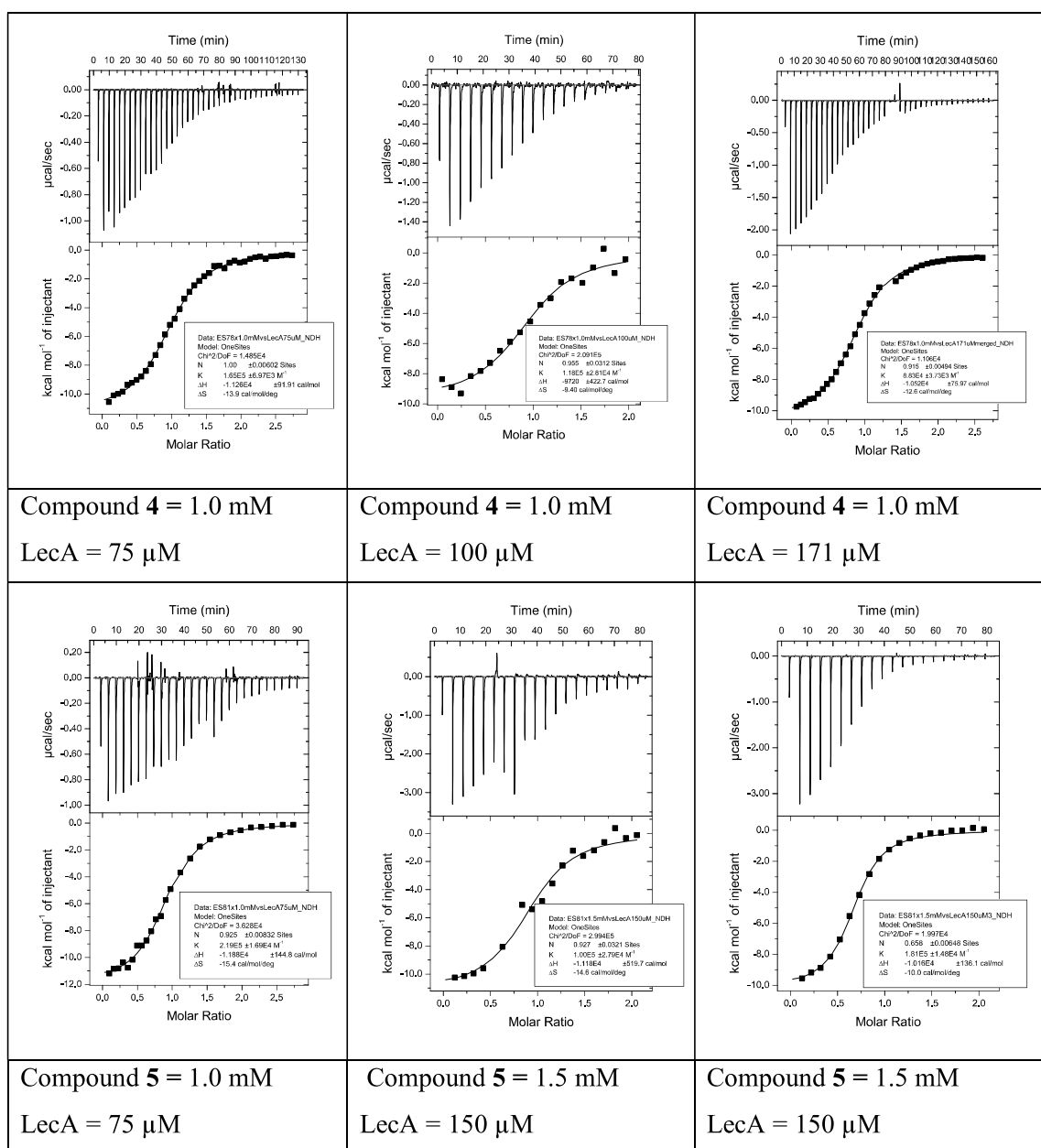
4lke	2	Central pocket	314
4lkf	1	Central pocket	297
4lkf	2	Central pocket	287
4yw6	0		
4yw7	1	Central pocket	300
4ywa	1	Central pocket	182
5d21	0		
5mih	1	Central pocket	277
5mih	2	Central pocket	263

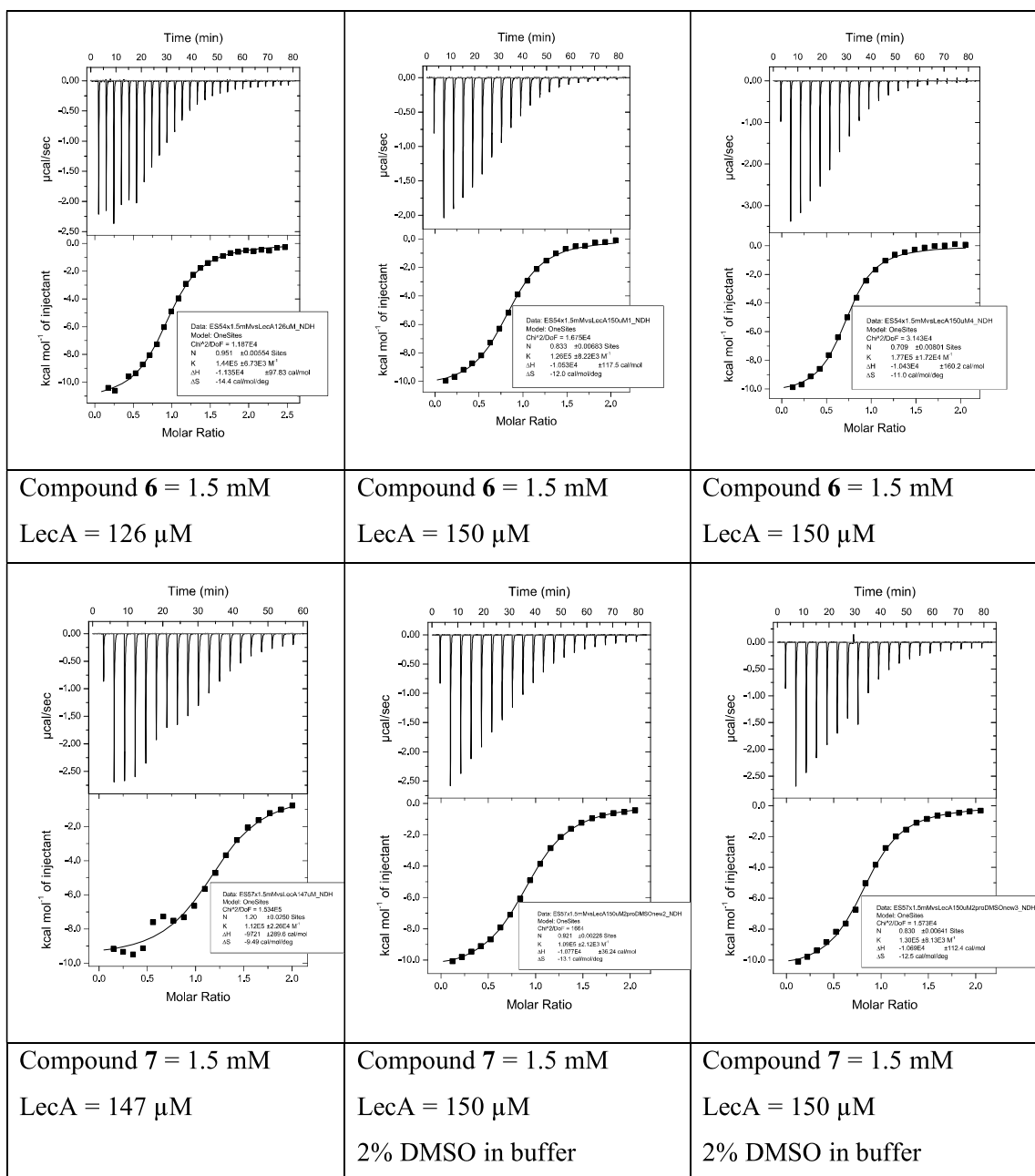
527

528 Isothermal Titration Calorimetry (ITC)









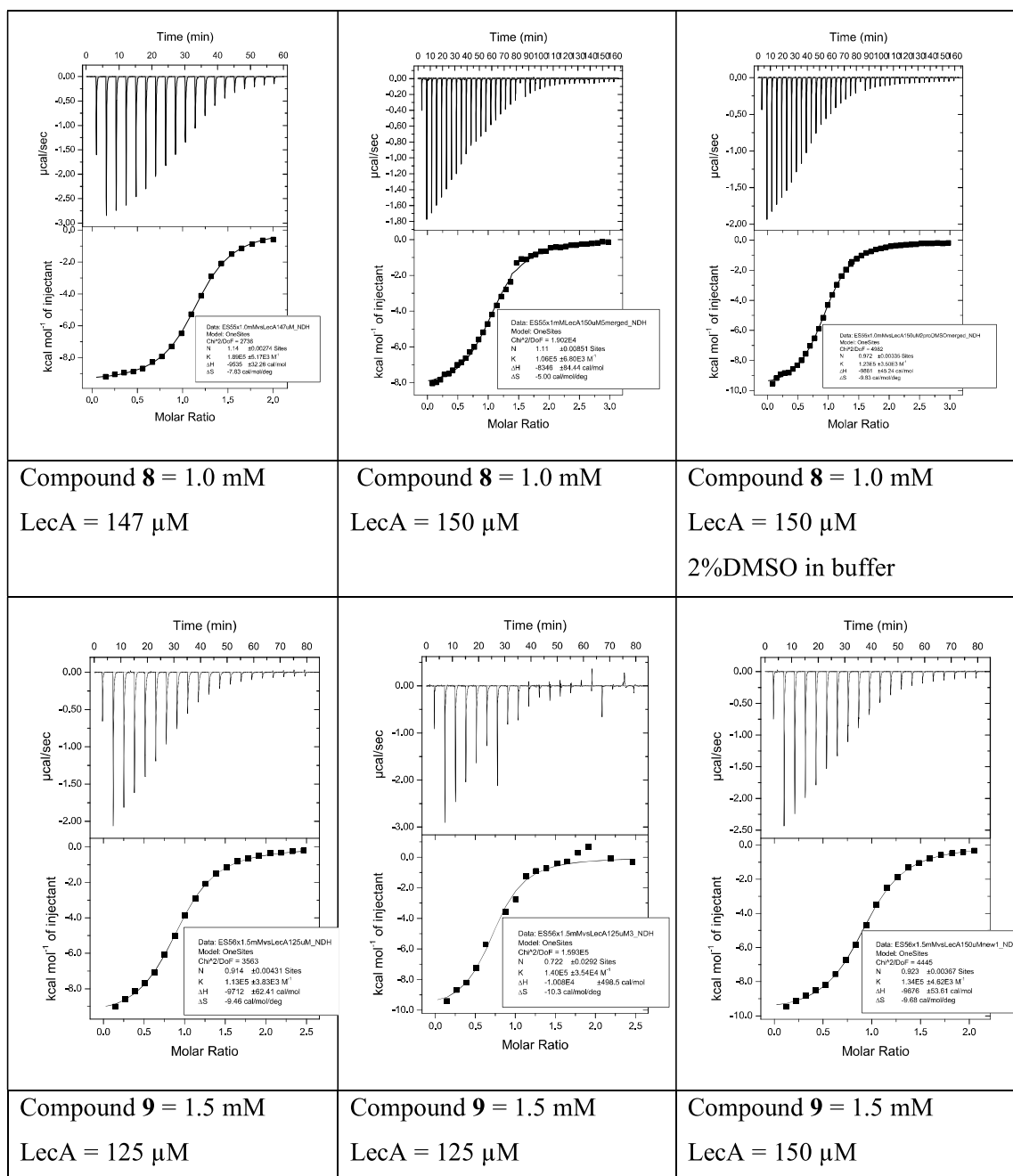
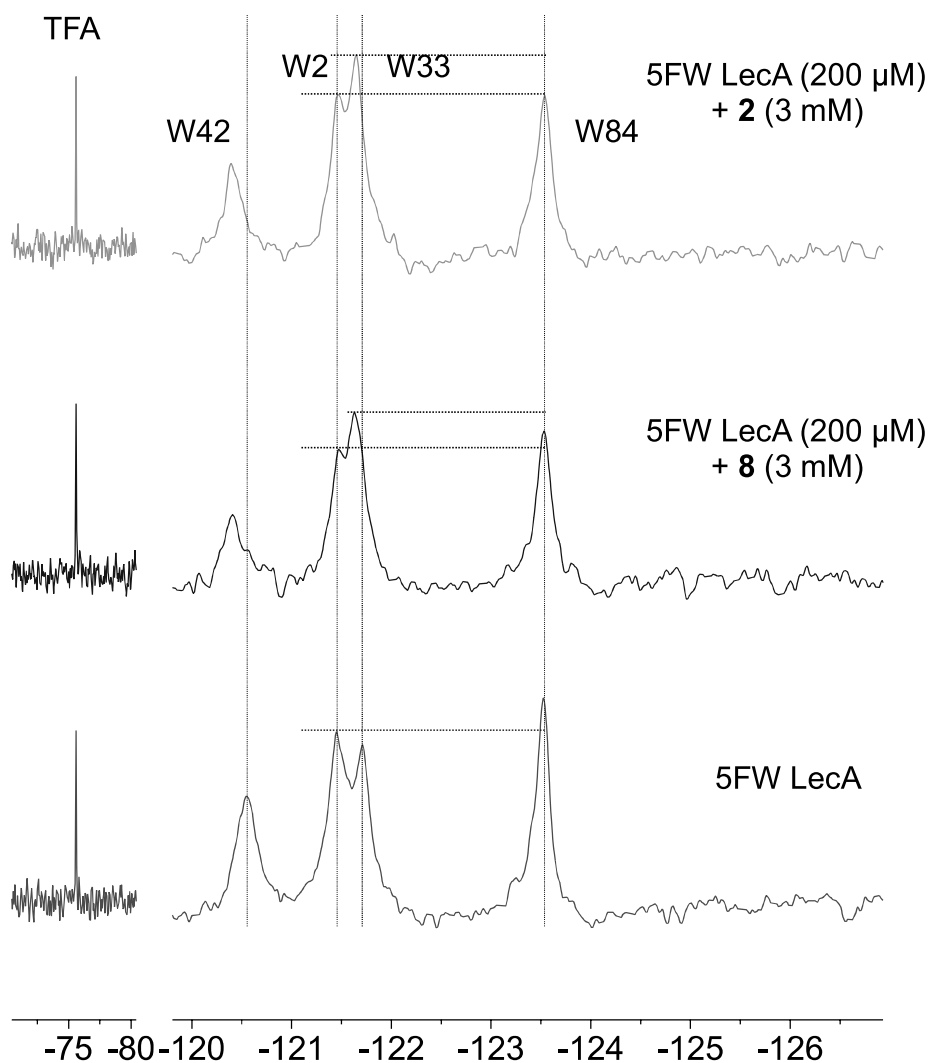


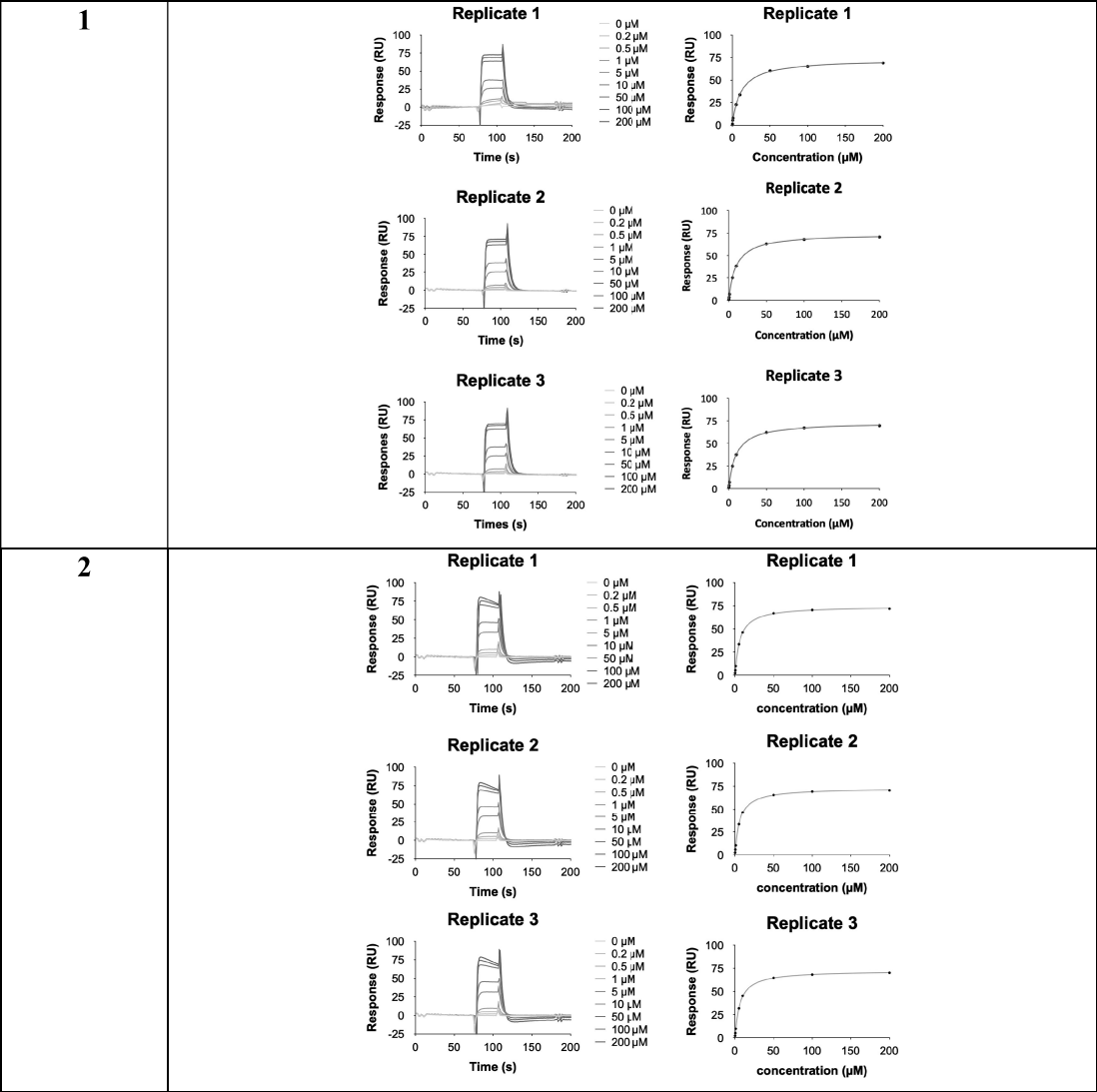
Figure S1: ITC titration experiments with LecA.

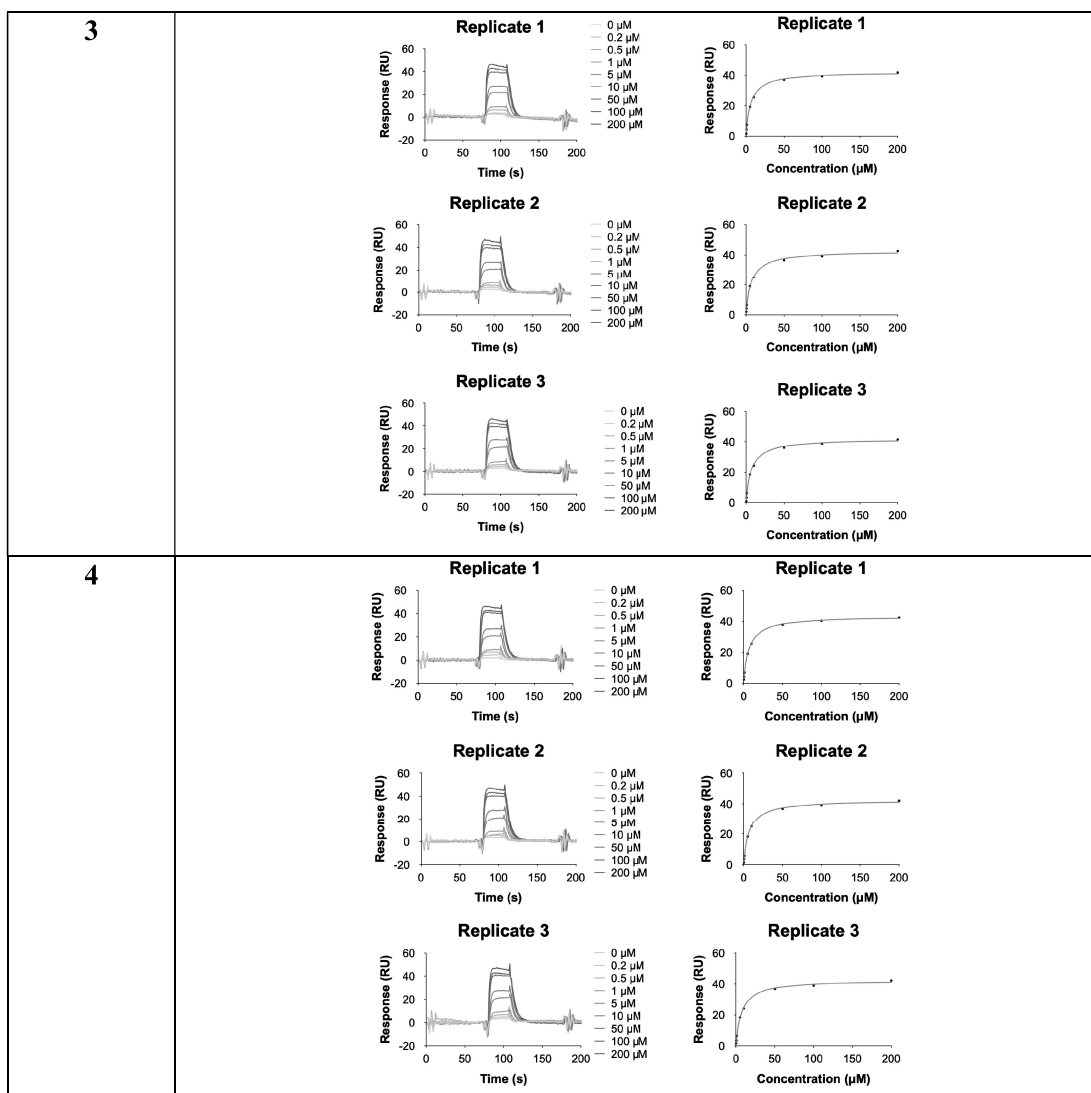


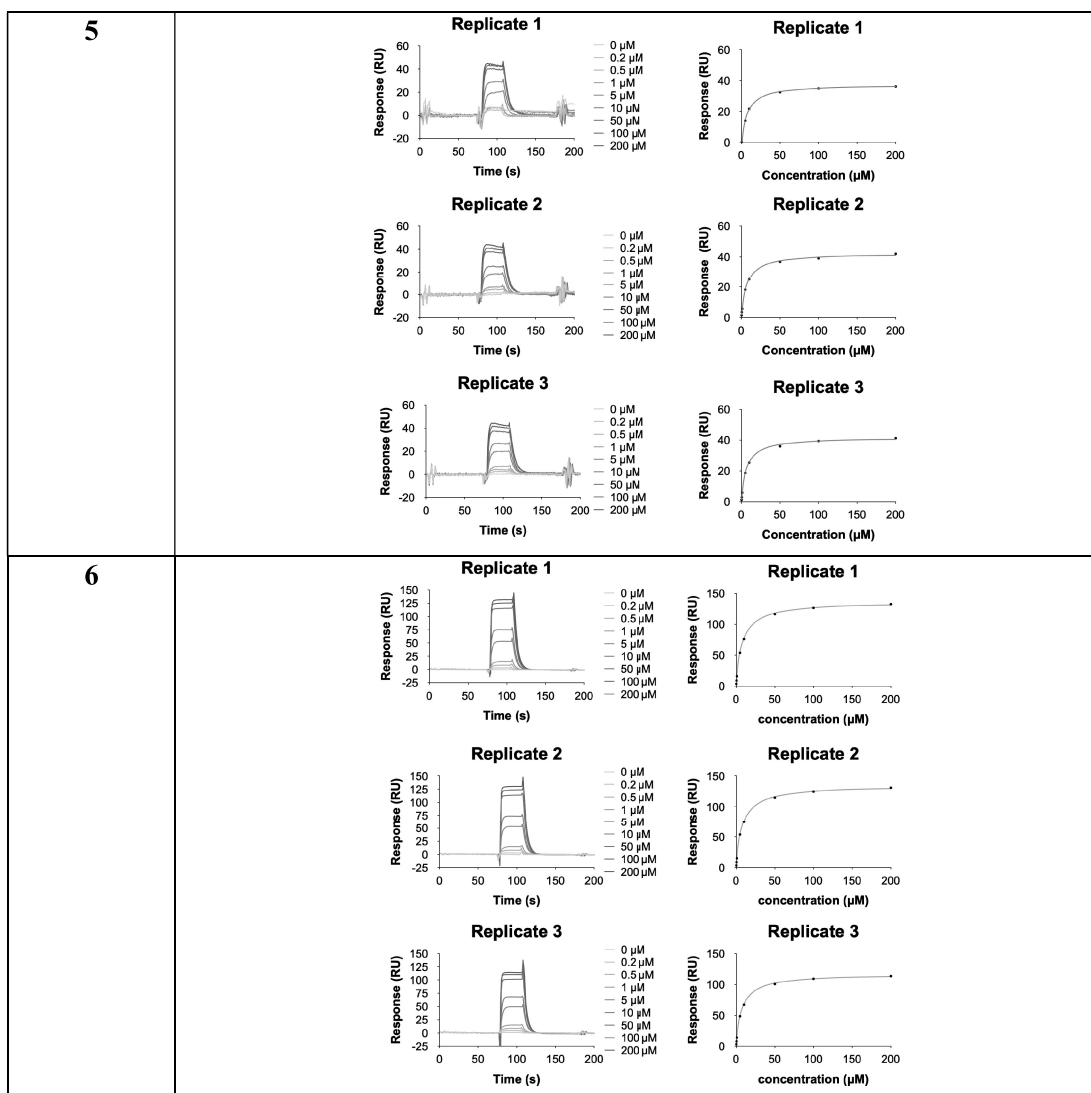
532
533 **Figure S2:** ¹⁹F PrOF NMR of 5FW LecA: without ligand (gray), with the benzyl ether **2** (orange)
534 and the naphthyl compound **8** (blue). A chemical shift perturbation of Trp42 and Trp33 has been
535 observed in presence of both ligands indicating that ligands target the carbohydrate and the central
536 pockets, respectively.

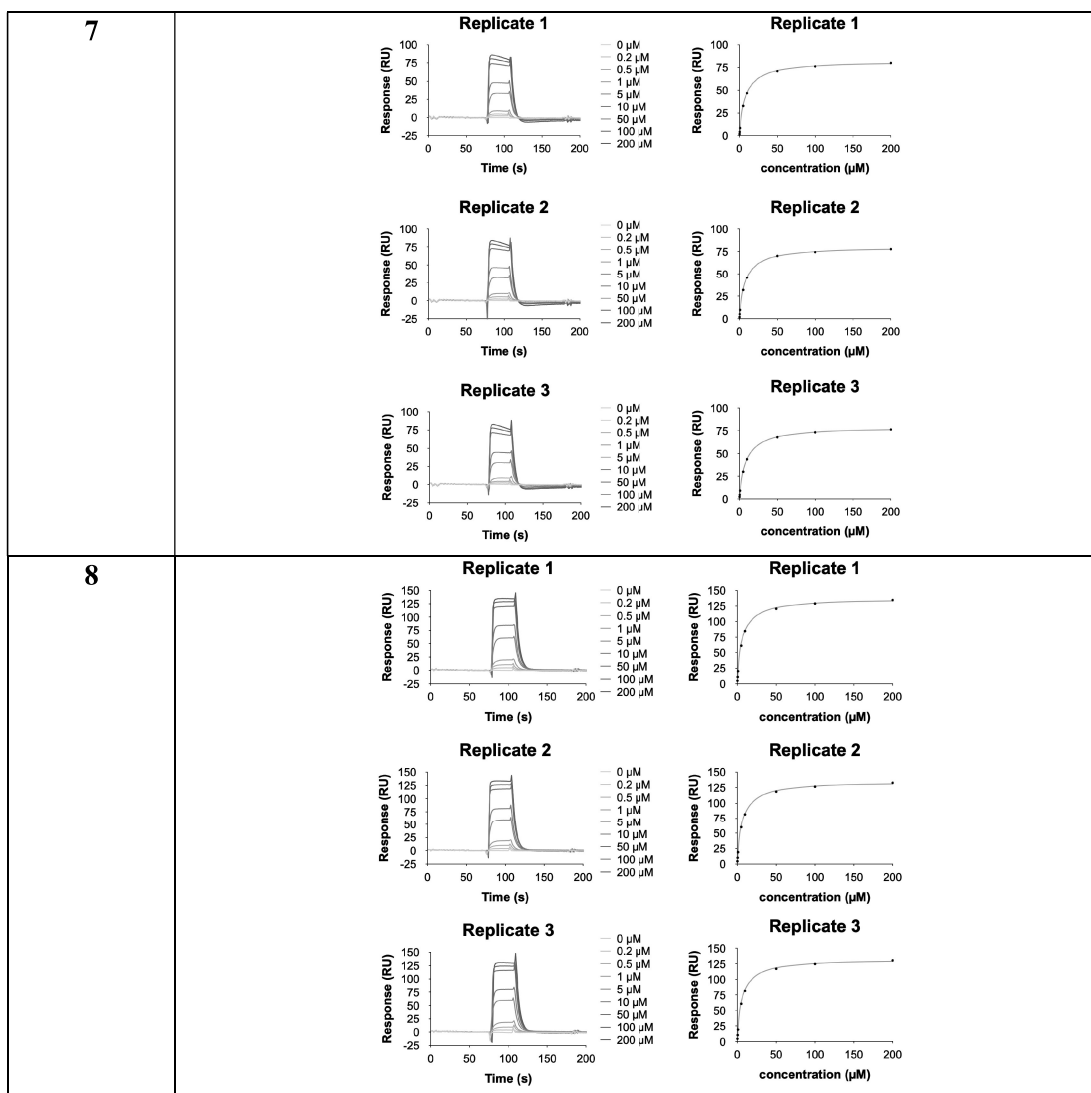
537 Surface Plasmon Resonance (SPR)

538









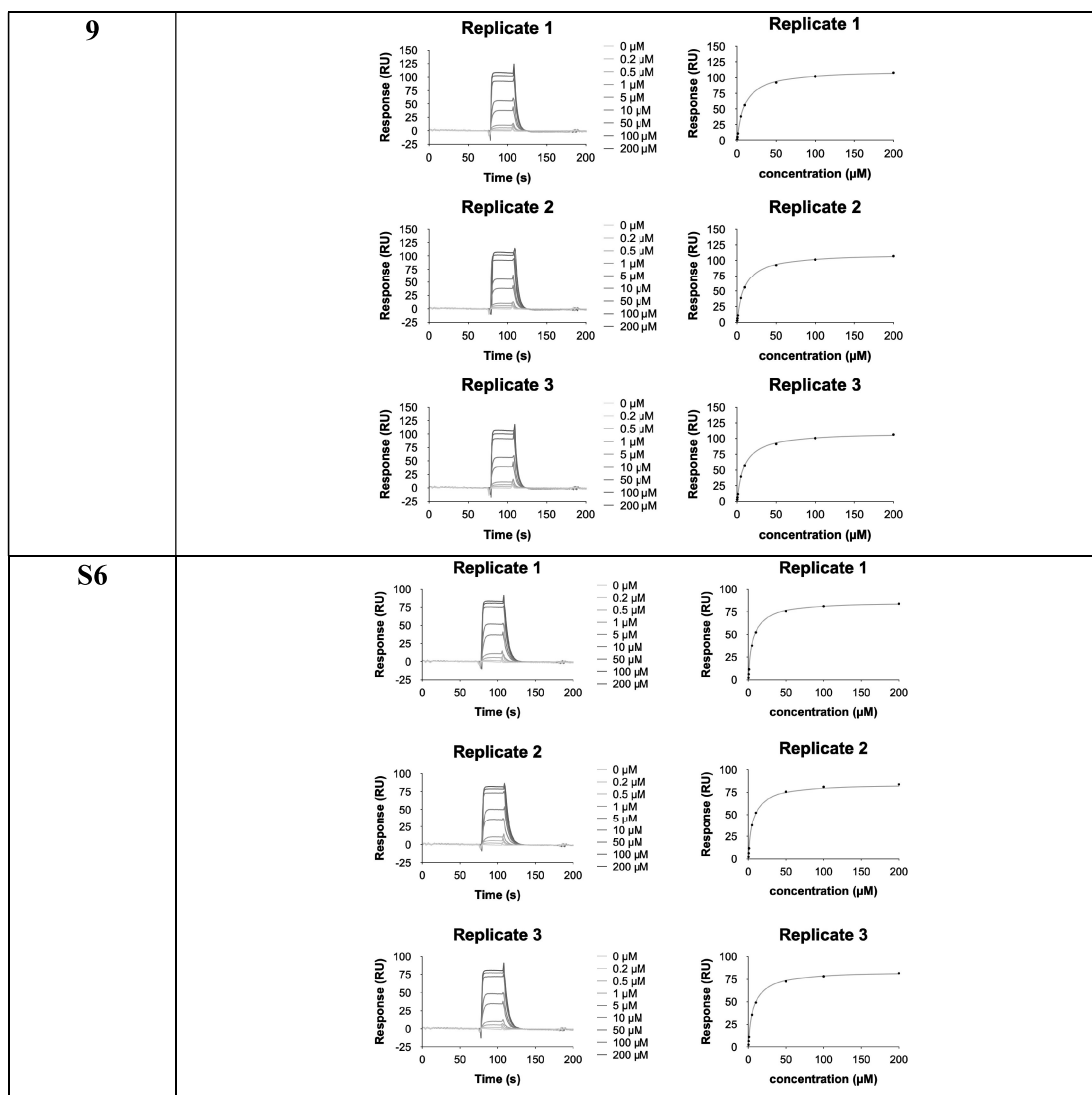


Figure S3: Individual sensorgrams and fits of of all replicates for binding to LecA.

X-Ray Crystallography

Table S3. X-ray data collection and processing of LecA in complex with **1**

PDB code	7FIO
Data Collection	
Beamline	PROXIMA1 (SOLEIL)
Wavelength (Å)	0.9786
Detector	EigerX-16M
Resolution (Å)^a	47.29-1.50 (1.53-1.50)
Space Group	P2 ₁ 2 ₁ 2 ₁
a, b, c (Å)	49.60 52.68 156.97
α, β, γ (°)	90.0,90.0 ,90.0
Total observations	690,166
Unique reflections	66,700
Multiplicity^a	10.3 (9.3)
Mean I/σ(I)^a	13.5 (1.7)
Completeness (%)^a	99.8 (99.1)
R_{merge}^{a,b}	0.089 (0.999)
CC_{1/2}^{a,c}	1.0 (0.6)
Refinement	
Reflections: working/free^d	66,623/3,419
R_{work}/ R_{free}^e	0.177/0.212
Ramachandran plot: allowed/favoured/outliers (%)	3/97/0
R.m.s. bond deviations (Å)	0.0162
R.m.s. angle deviations (°)	1.896
Mean B-factors: protein/ligand^f/ /water (Å²)	20/41/29

^a Values for the outer resolution shell are given in parentheses.

^b $R_{\text{merge}} = \sum_{\text{hkl}} \sum_i |I_i(\text{hkl}) - \langle I(\text{hkl}) \rangle| / \sum_{\text{hkl}} \sum_i I_i(\text{hkl})$.

^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: $R = \sum (|F_{\text{obs}} - F_{\text{calc}}|) / \sum |F_{\text{obs}}|$, 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

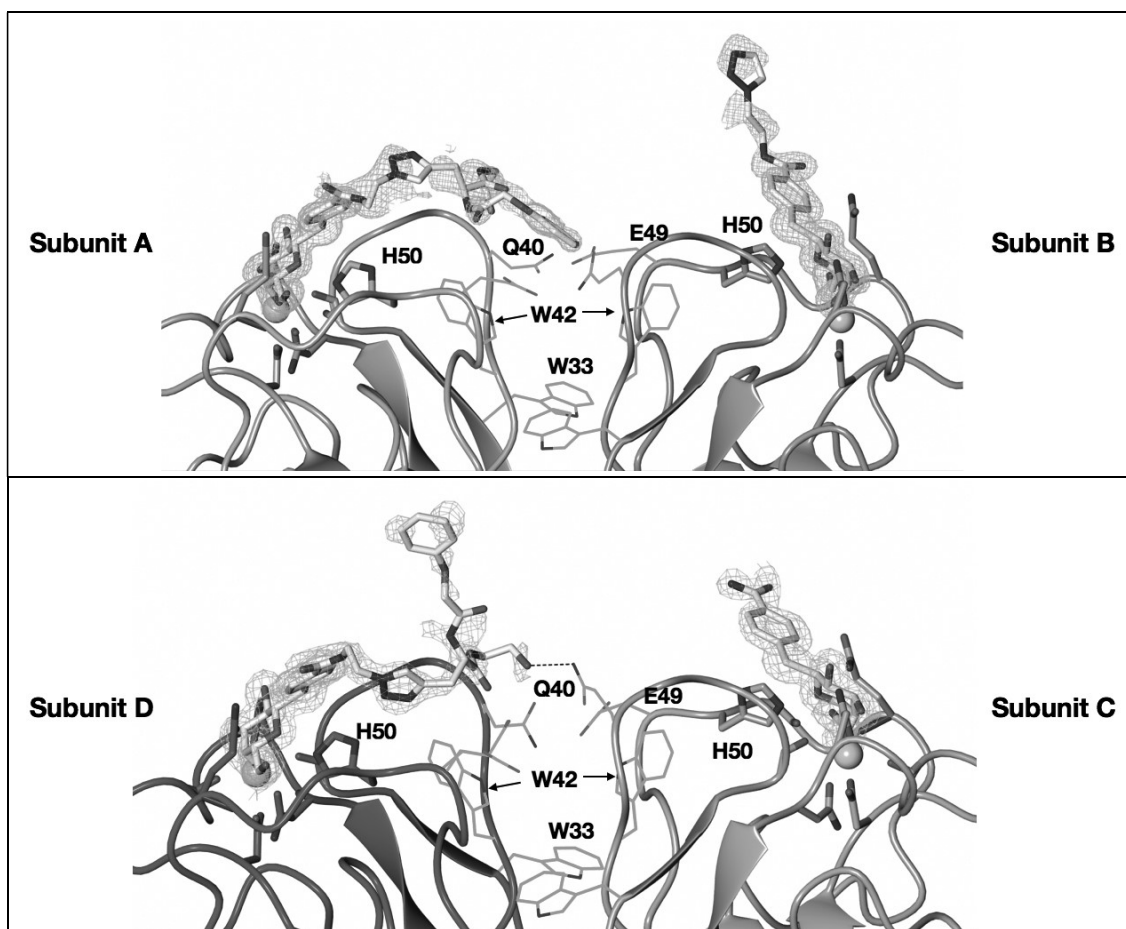
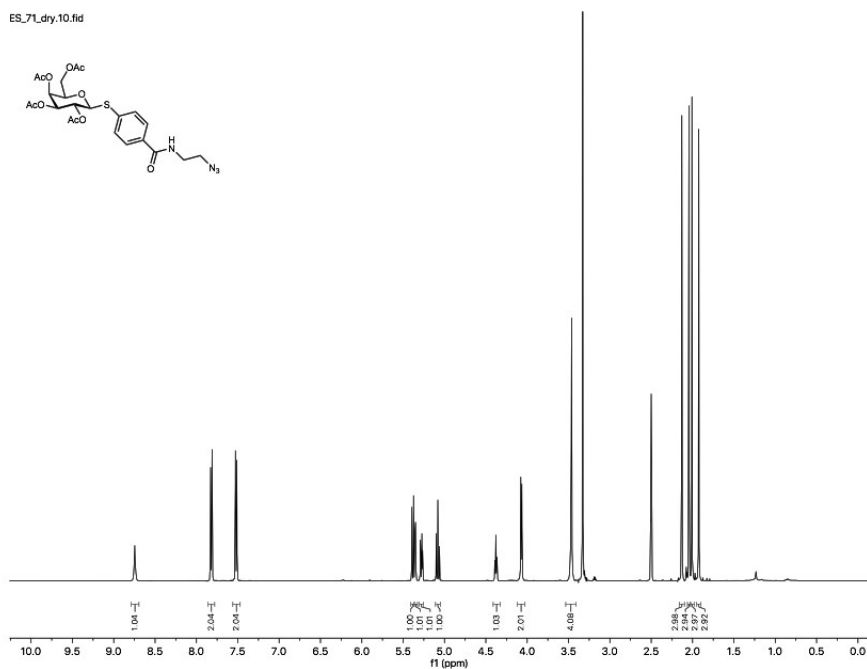


Figure S4: Co-crystal structure of compound **1** and tetrameric LecA. All carbohydrate binding sites of the four monomers are occupied by one ligand by coordination to the calcium ion with the galactose moiety. Subunit A shows the ligand with its side chain at the entrance of the central pocket. Subunit B and C are not completely resolved. Subunit D shows a hydrogen bonding between the ethanol amine function of **1** with Q40 of the neighboring subunit C.

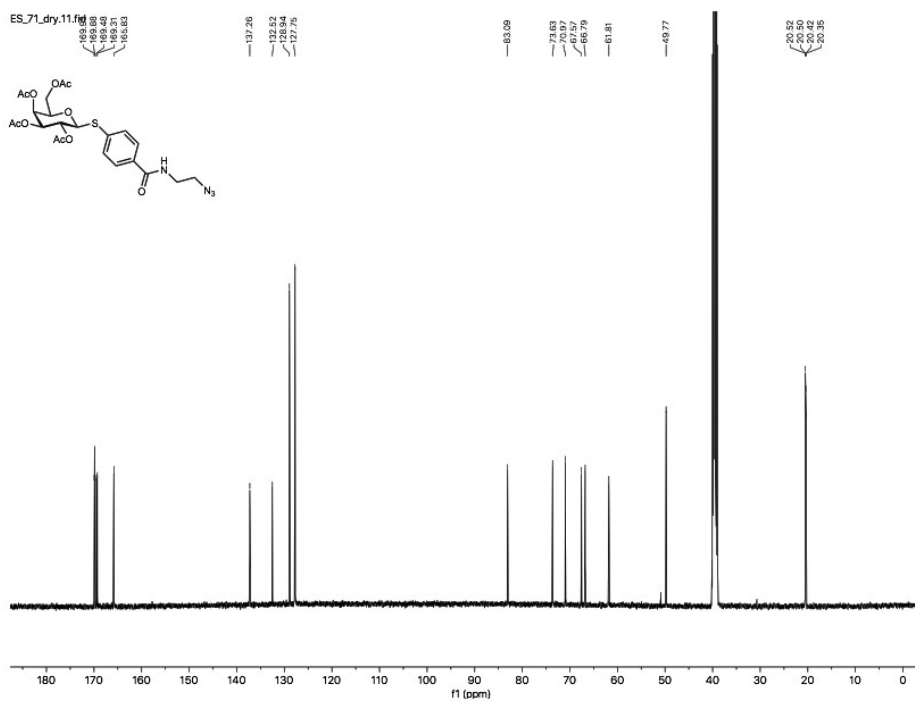
References

- [1] G. R. Fulmer, A. J. M. Miller, N. H. Sherden, H. E. Gottlieb, A. Nudelman, B. M. Stoltz, J. E. Bercaw, K. I. Goldberg, *Organometallics* **2010**, *29*, 2176–2179.
- [2] A. Novoa, T. Eierhoff, J. Topin, A. Varrot, S. Barluenga, A. Imberty, W. Römer, N. Winssinger, *Angew. Chemie Int. Ed.* **2014**, *53*, 8885–8889.
- [3] S. Kuhaudomlarp, E. Siebs, E. Shanina, J. Topin, I. Joachim, P. da Silva Figueiredo Celestino Gomes, A. Varrot, D. Rognan, C. Rademacher, A. Imberty, A. Titz, *Angew. Chemie Int. Ed.* **2021**, *60*, 8104–8114.
- [4] K. Anbarasu, K. K. Ilavenil, *Asian J. Chem.* **2018**, *30*, 2238–2240.
- [5] D. Li, S. Xiong, T. Guo, D. Shu, H. Xiao, G. Li, D. Guo, *Dye. Pigment.* **2018**, *158*, 28–35.
- [6] P. J. Machín, D. N. Hurst, R. M. Bradshaw, L. C. Blaber, D. T. Burden, A. D. Fryer, R. A. Melarange, C. Shivdasani, *J. Med. Chem.* **1983**, *26*, 1570–1576.
- [7] T. E. Kristensen, K. Vestli, M. G. Jakobsen, F. K. Hansen, T. Hansen, *J. Org. Chem.* **2010**, *75*, 1620–1629.
- [8] J. C. Sheehan, G. L. Boshart, P. A. Cruickshank, *J. Org. Chem.* **1961**, *26*, 2525–2528.

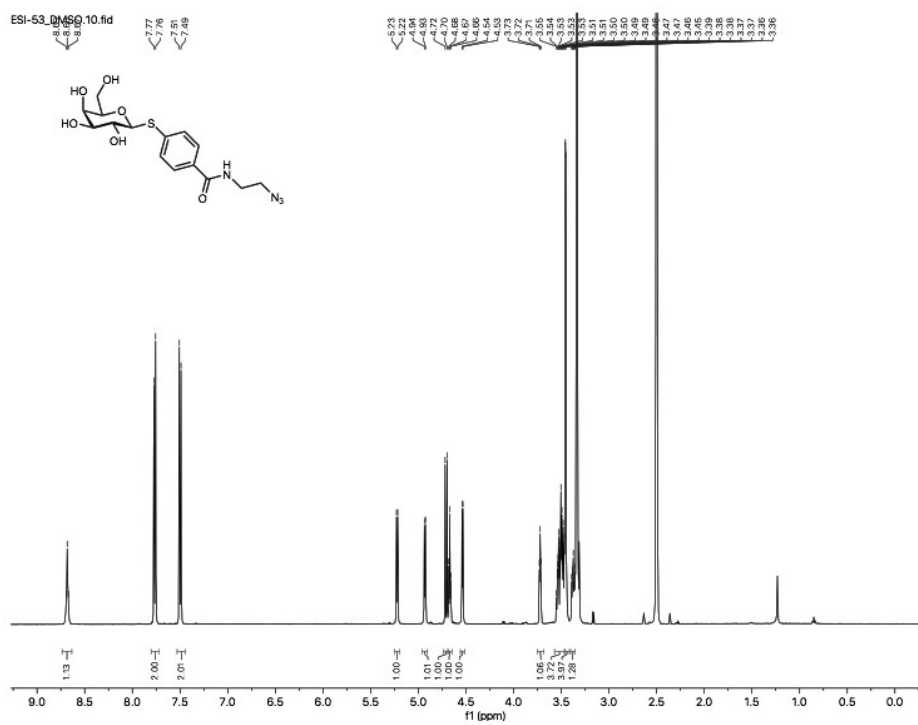
578 Spectra



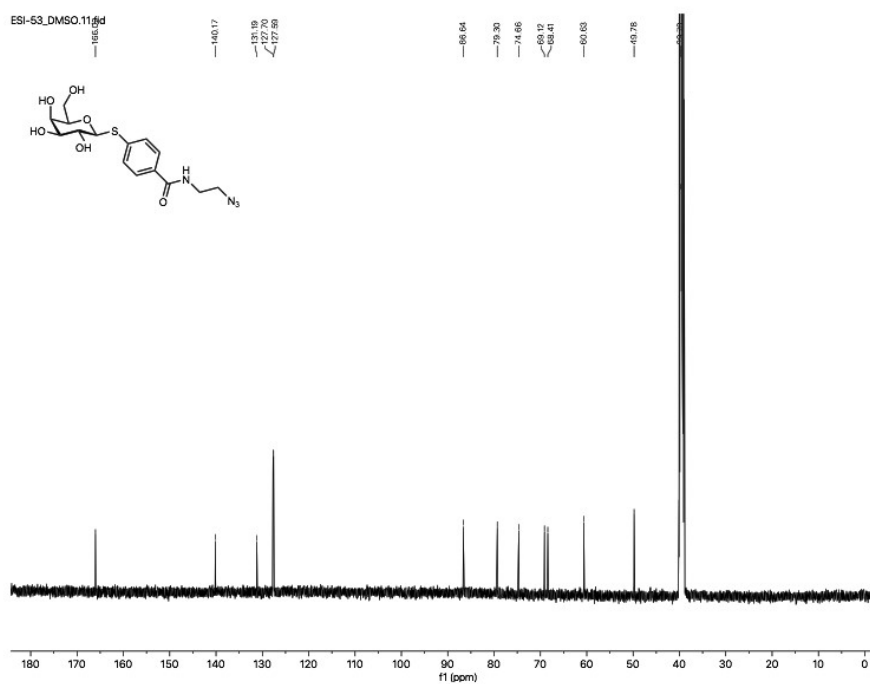
579



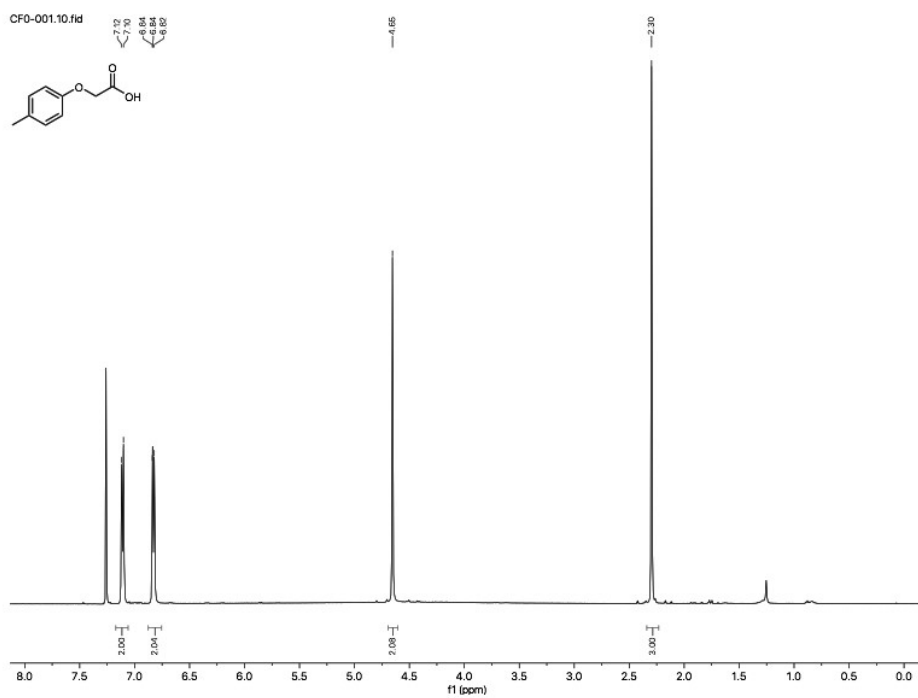
580



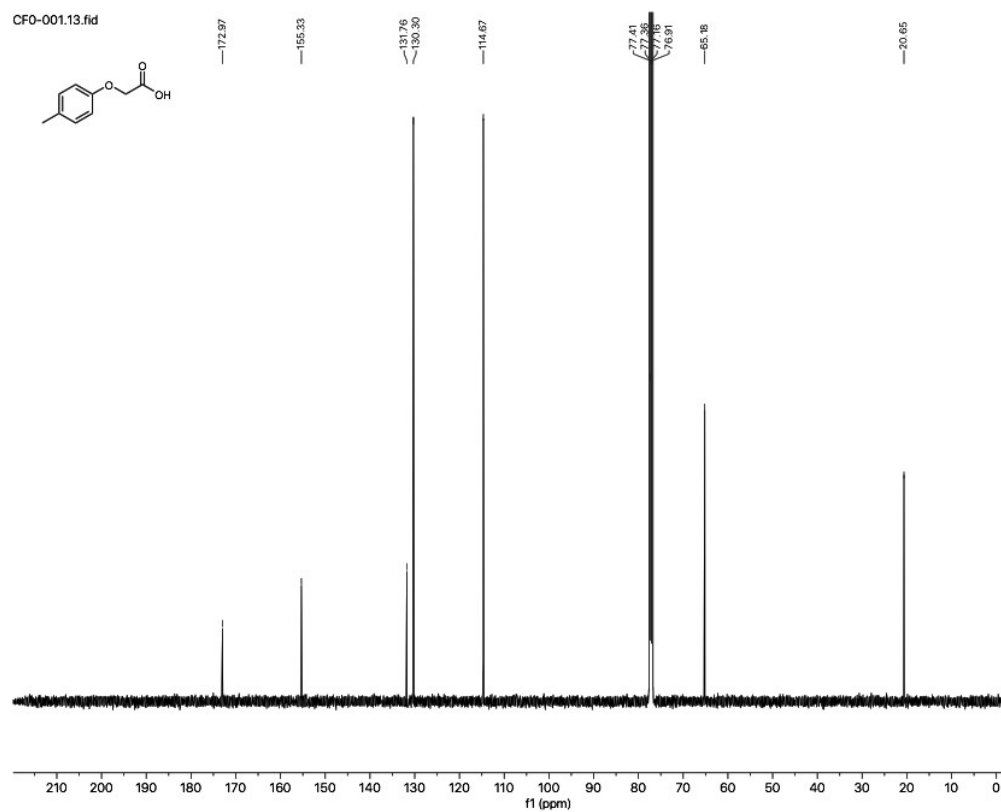
581



582

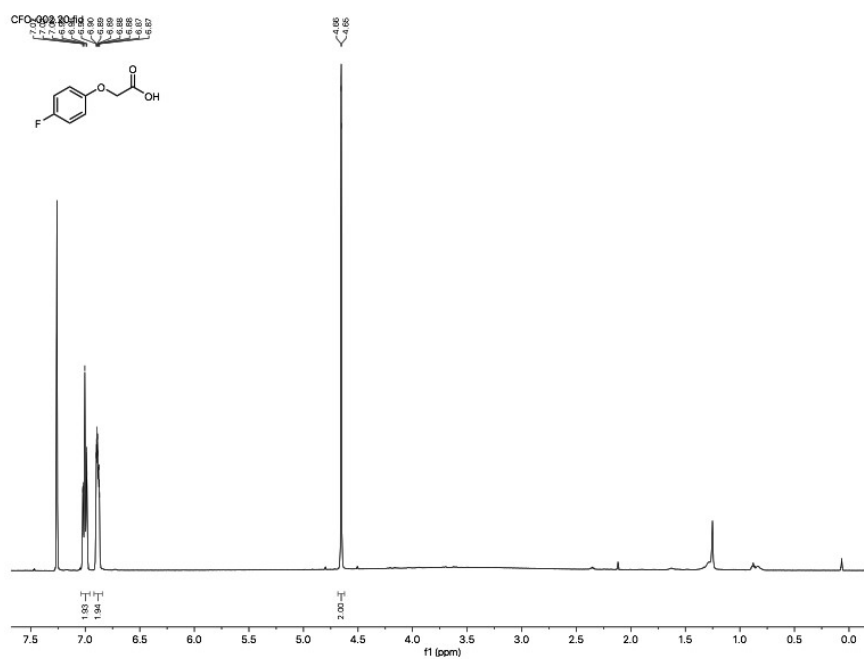


583

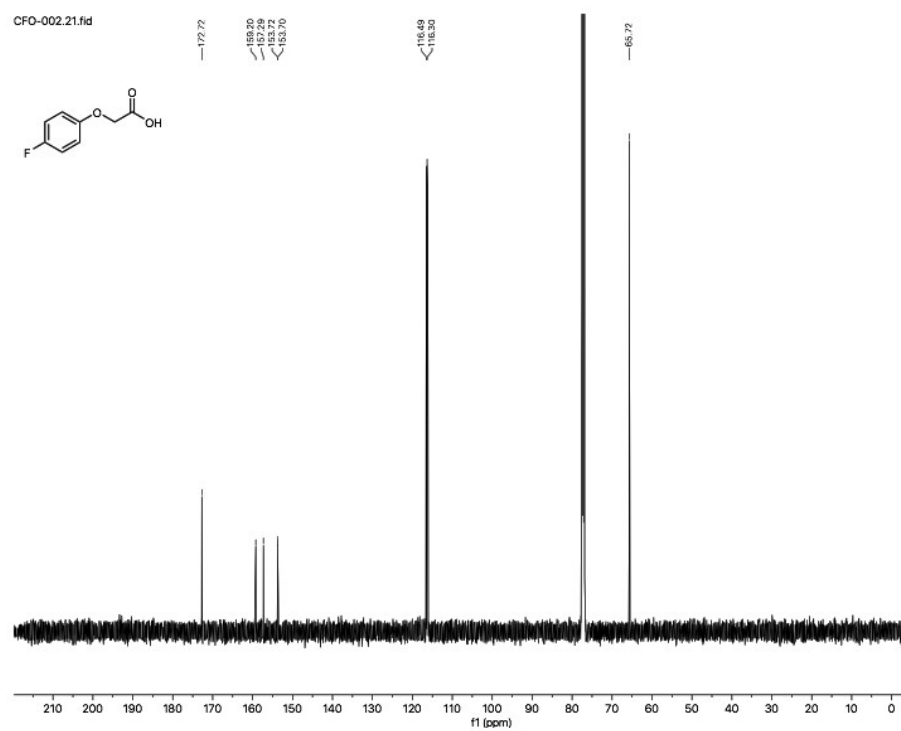


584

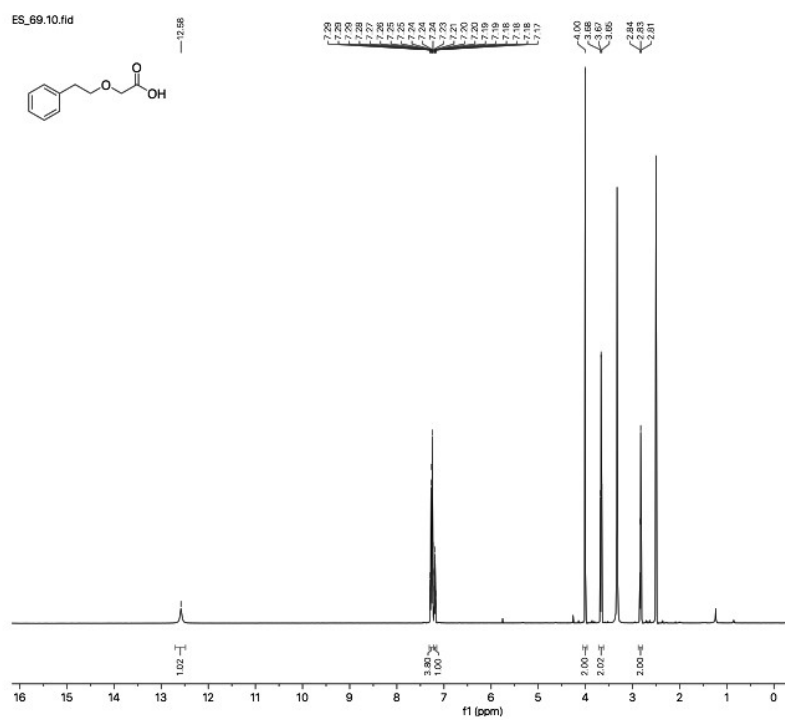
43



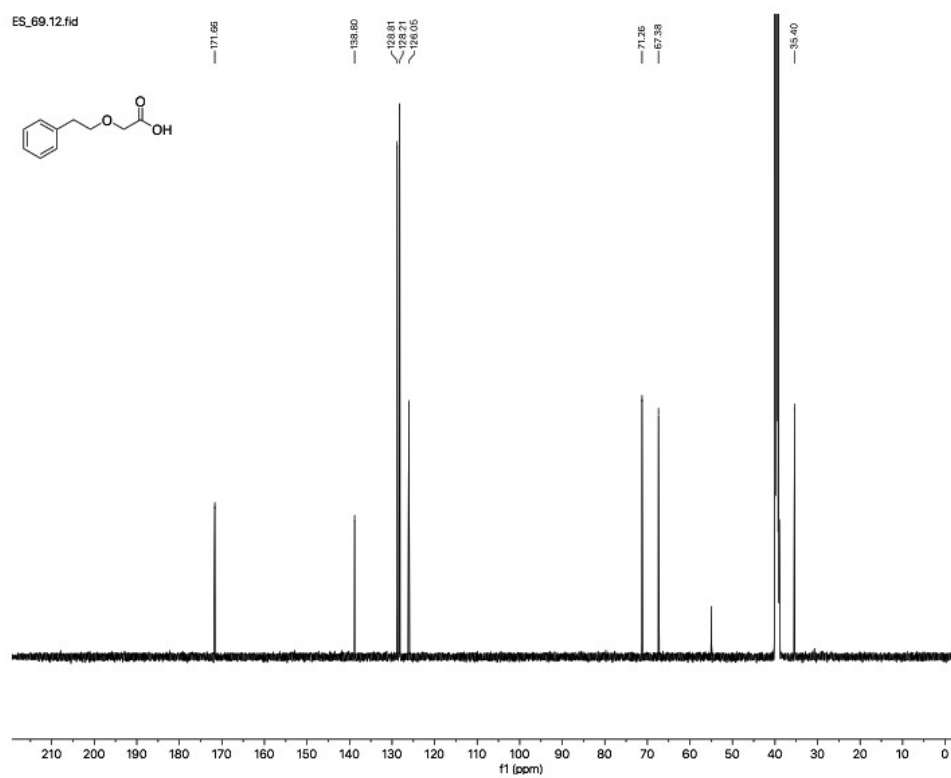
585



586

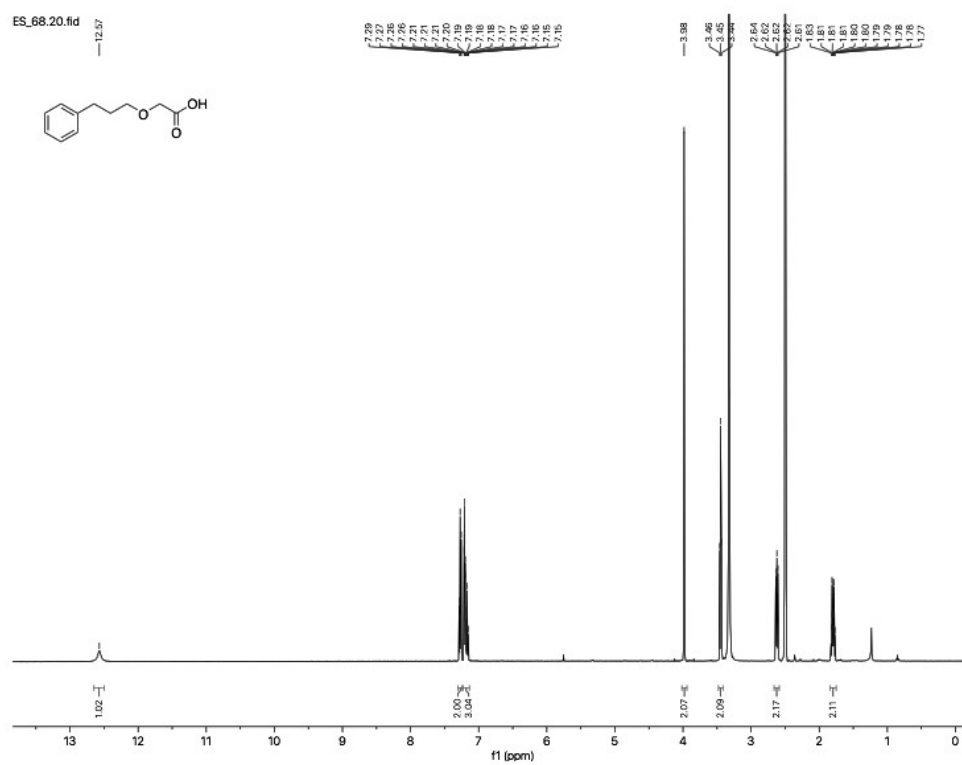


587

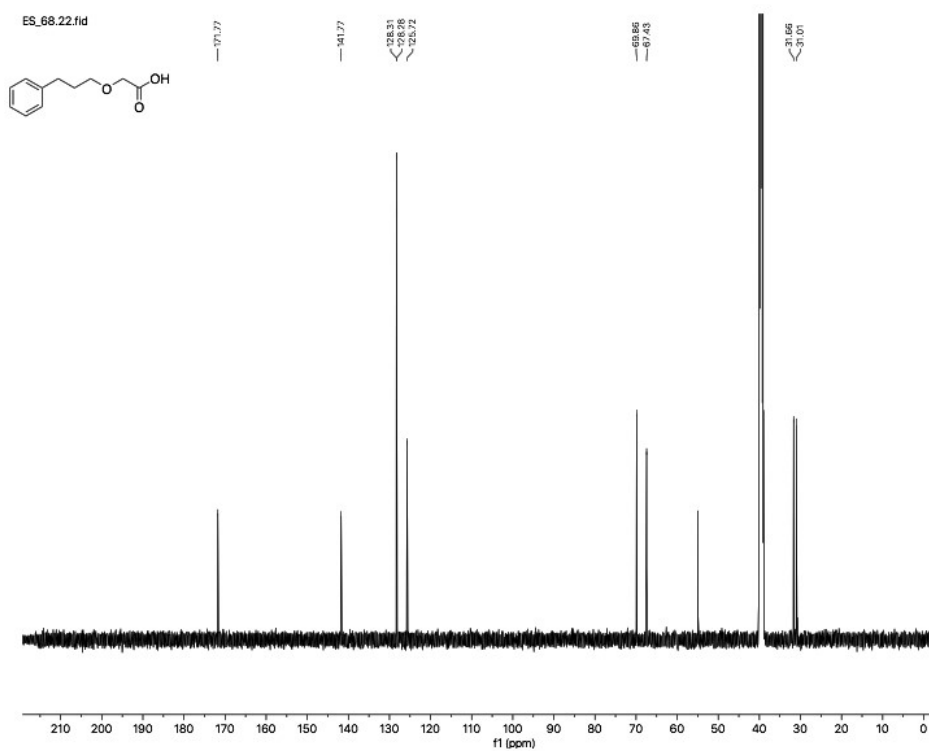


588

45

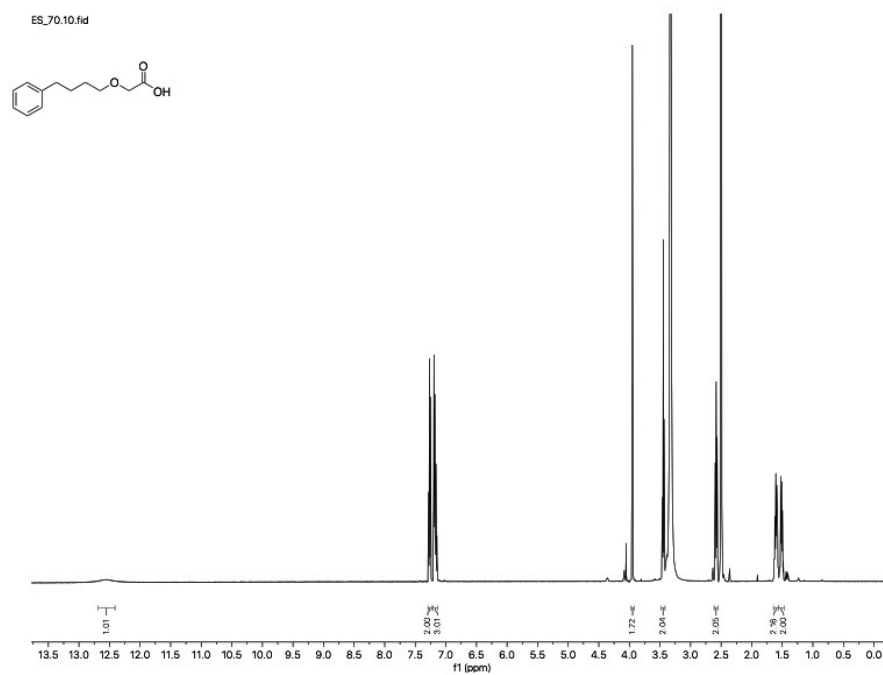


589

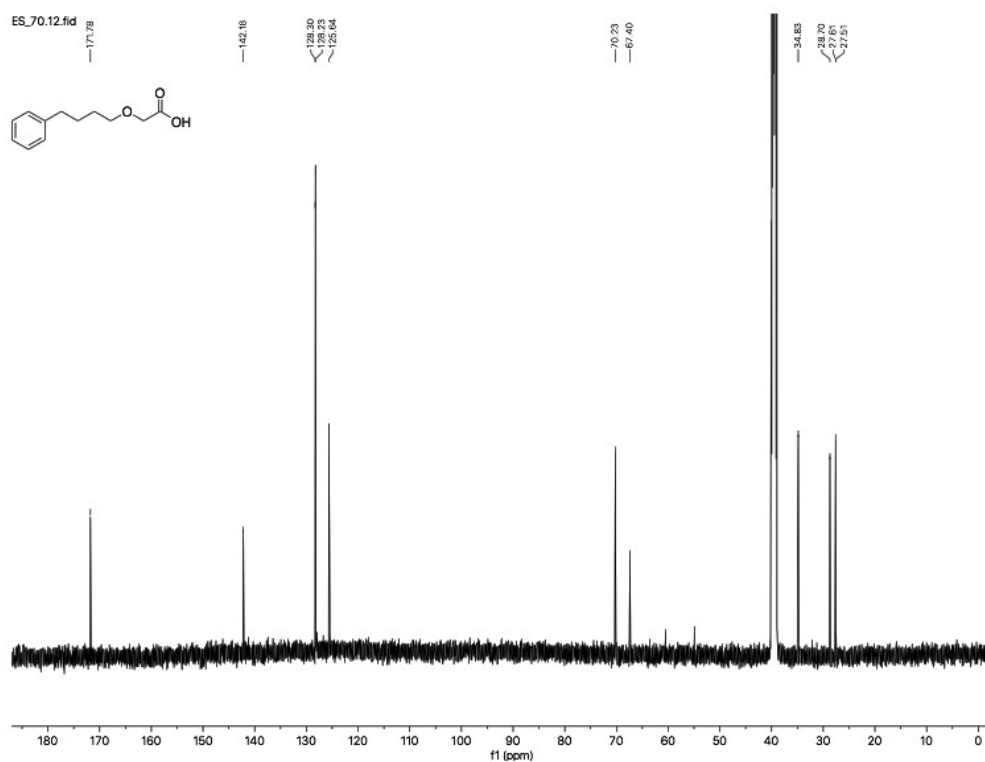


590

46



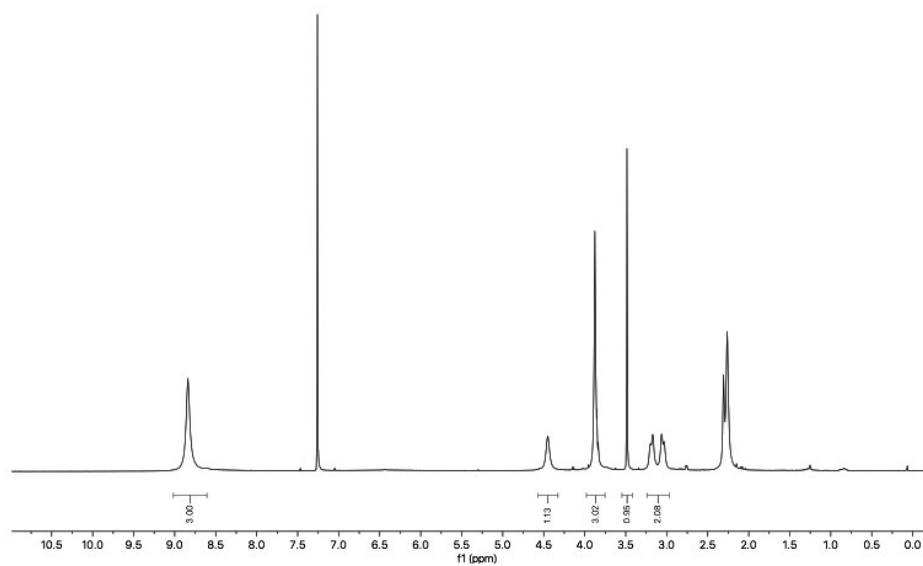
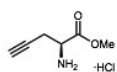
591



592

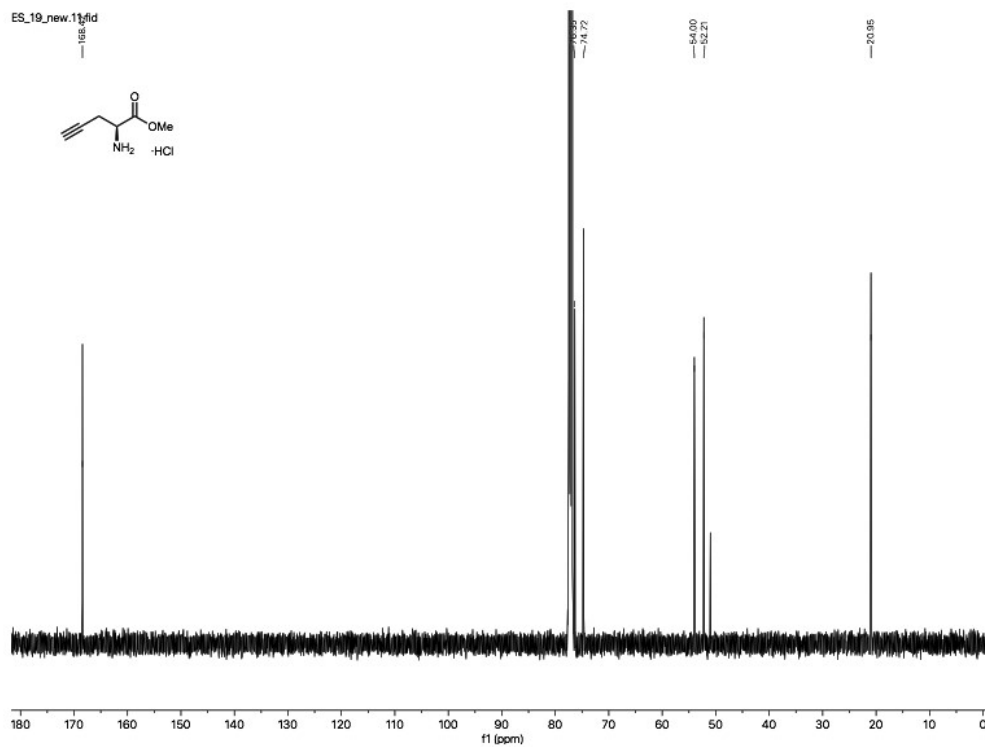
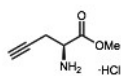
47

ES_19_new.10.fid



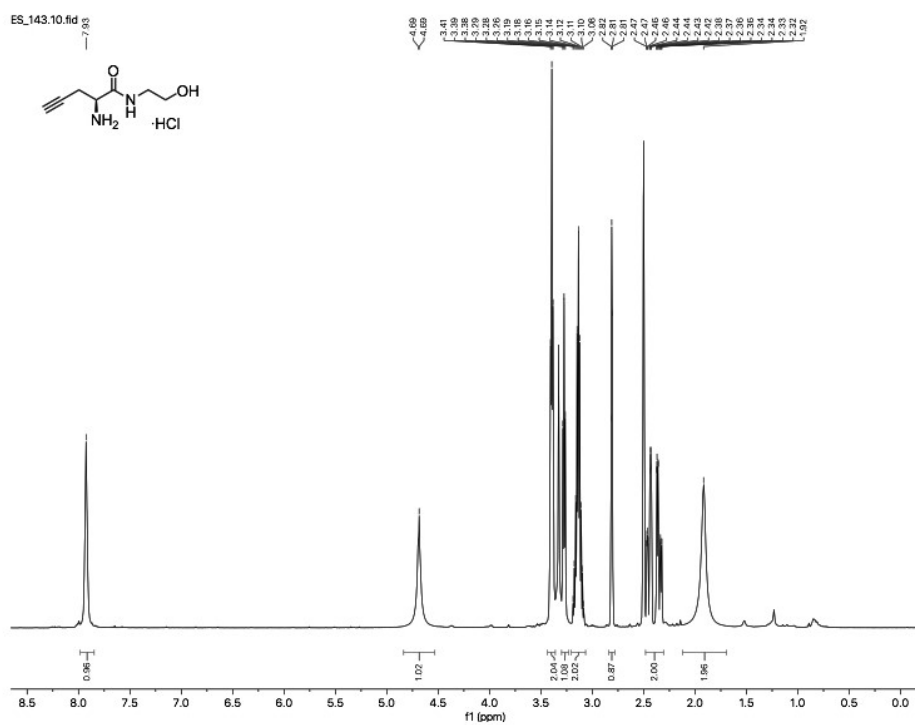
593

ES_19_new.11.fid

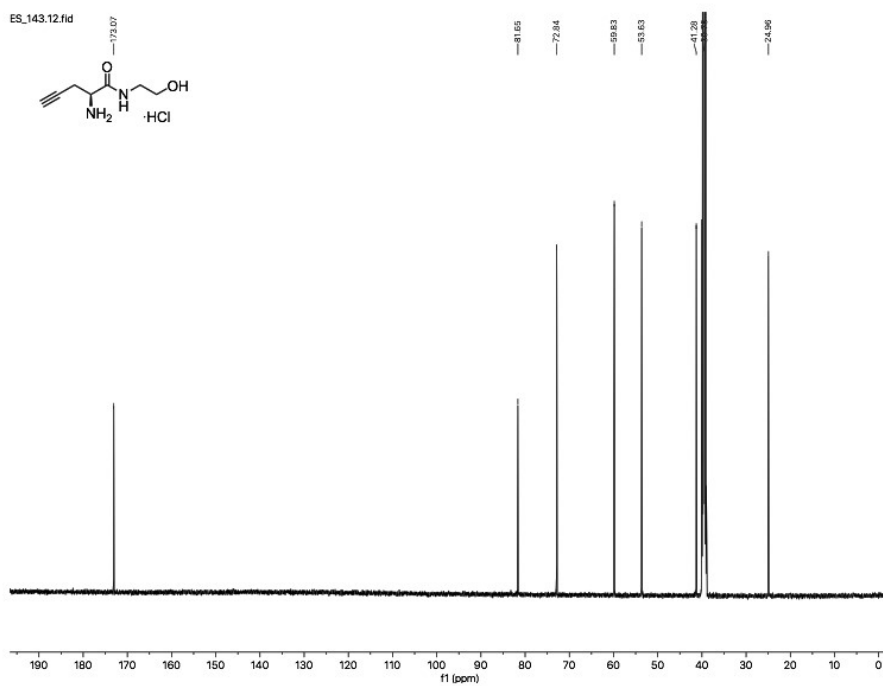


594

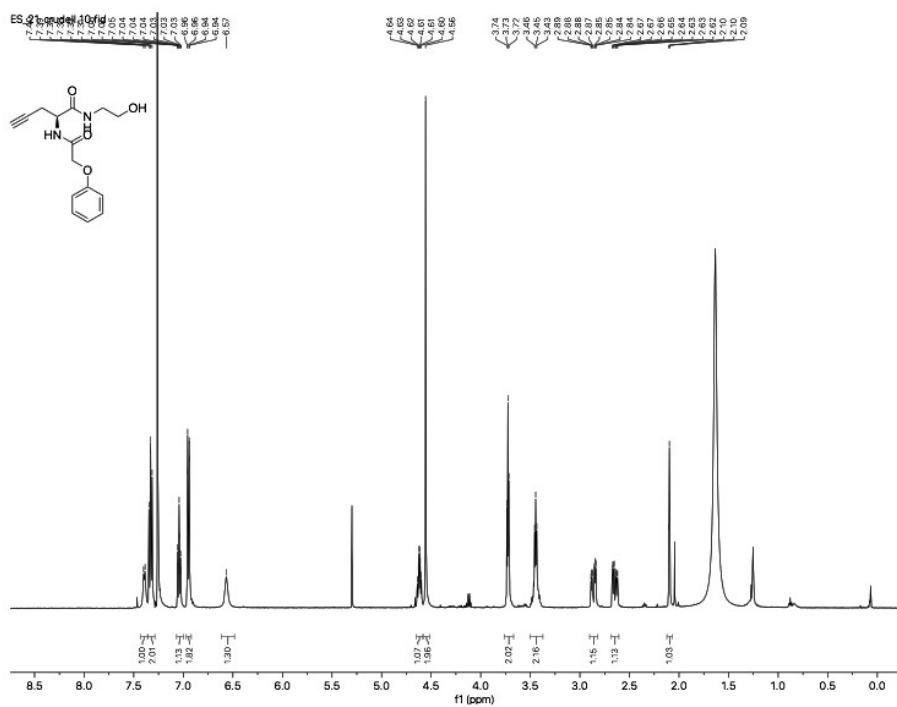
48



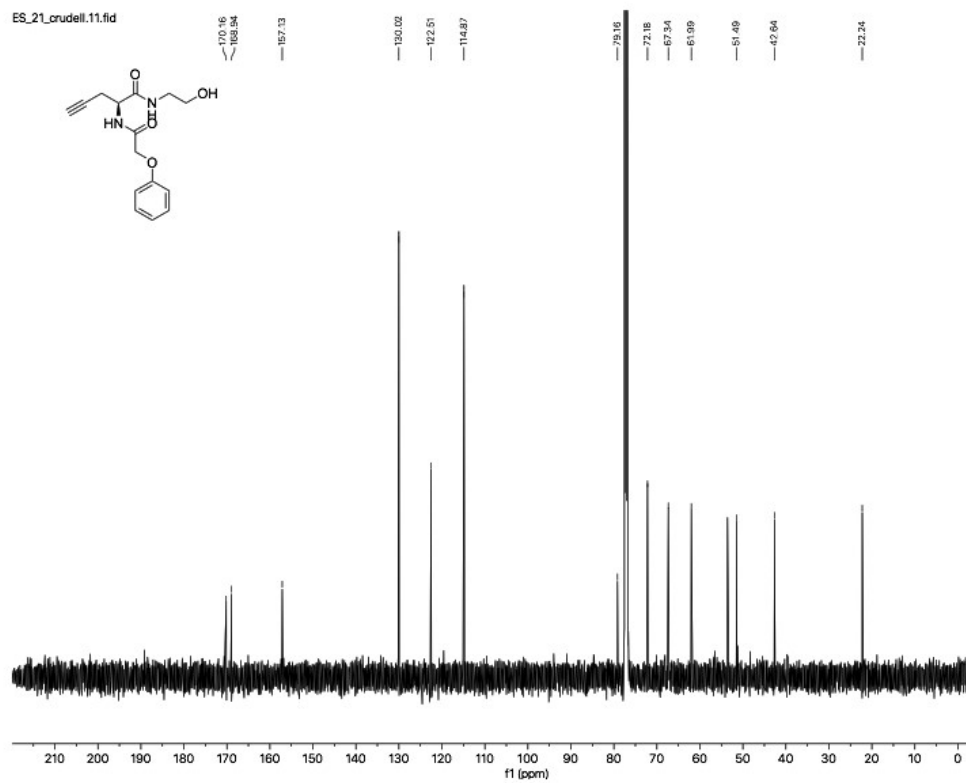
595



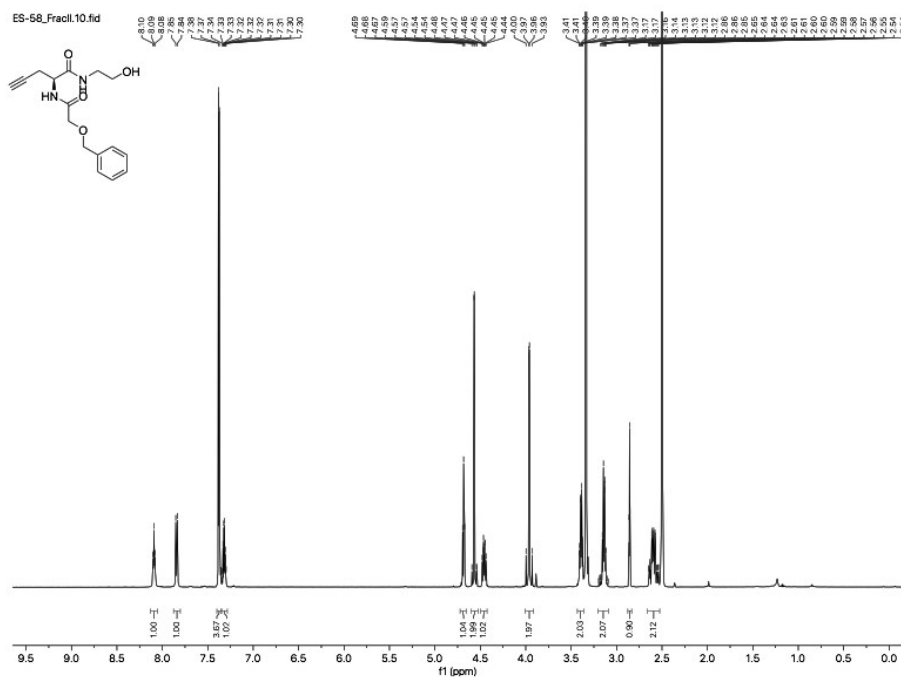
596



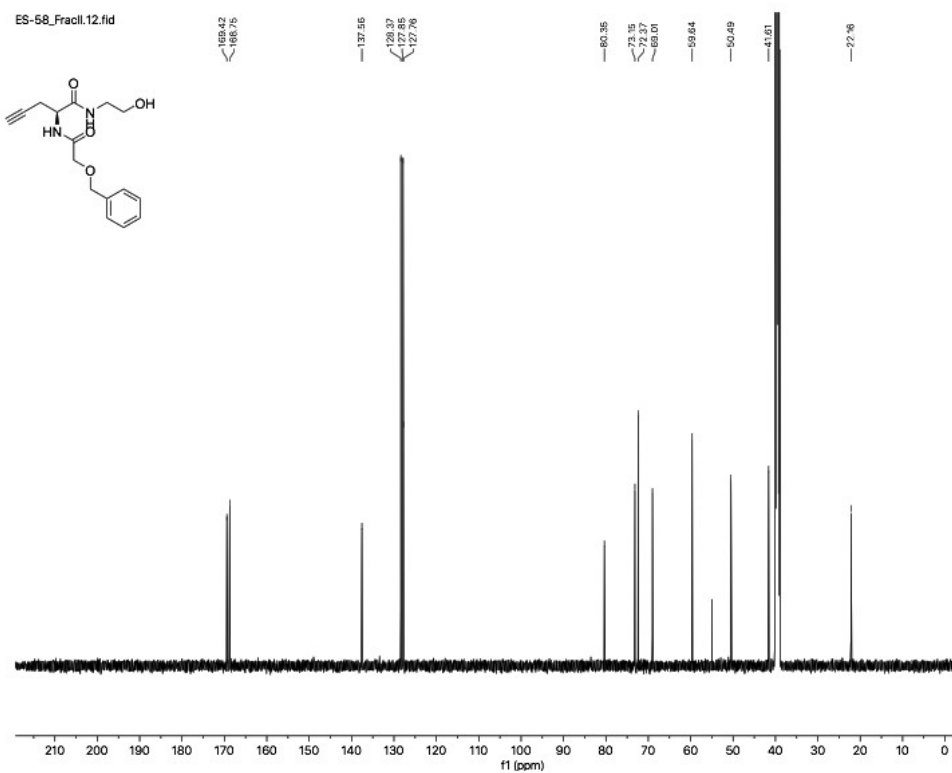
597



50

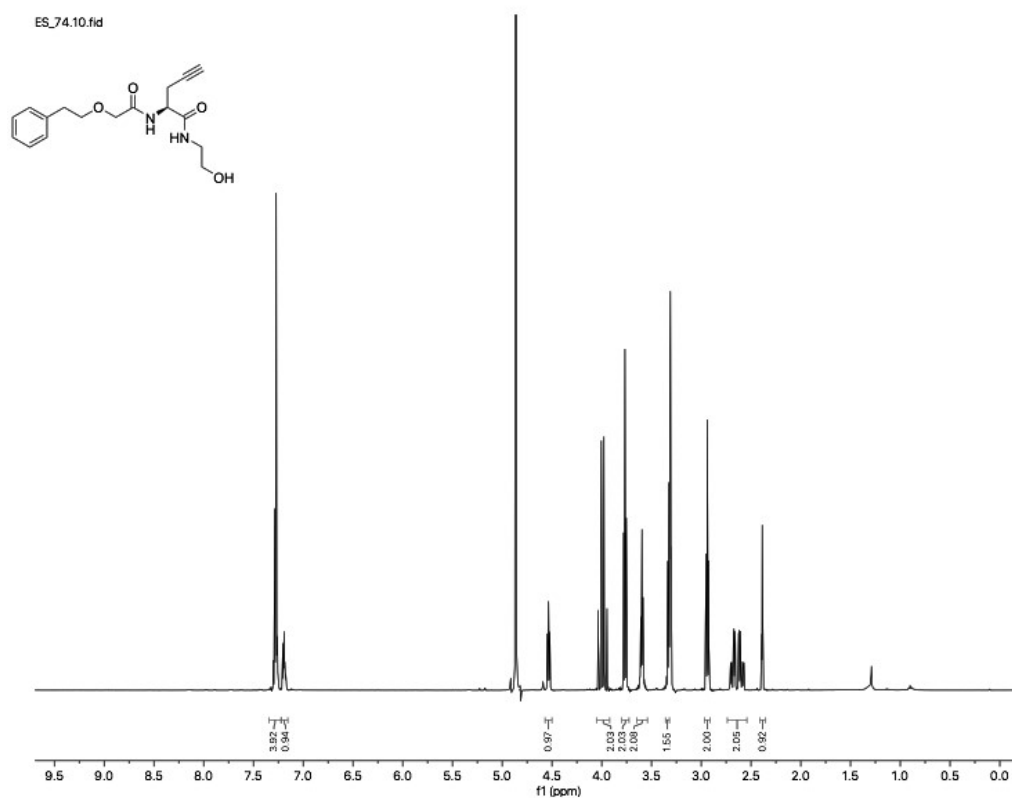


599

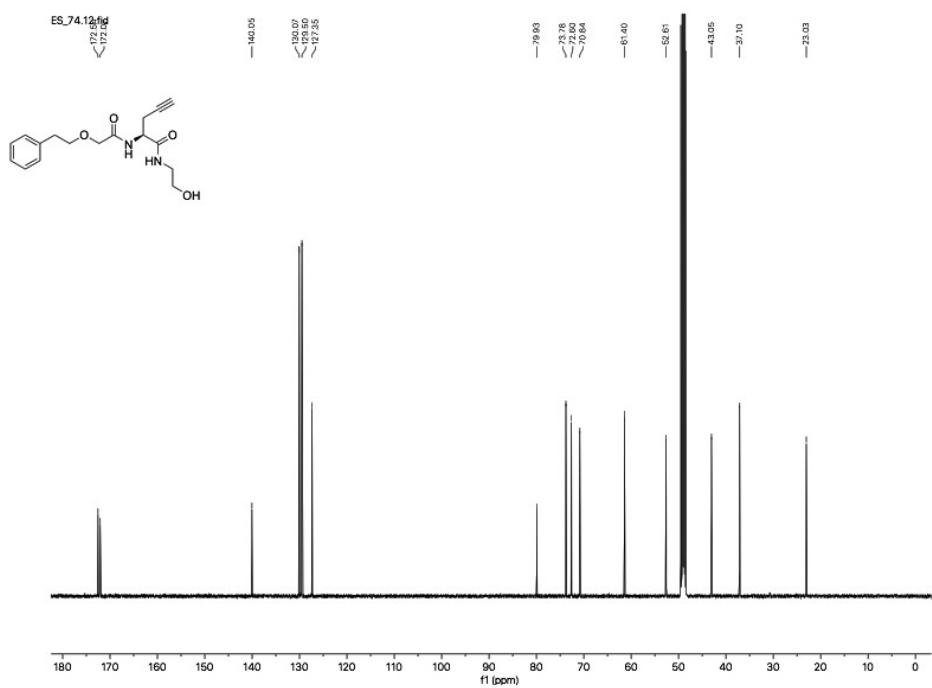


600

51

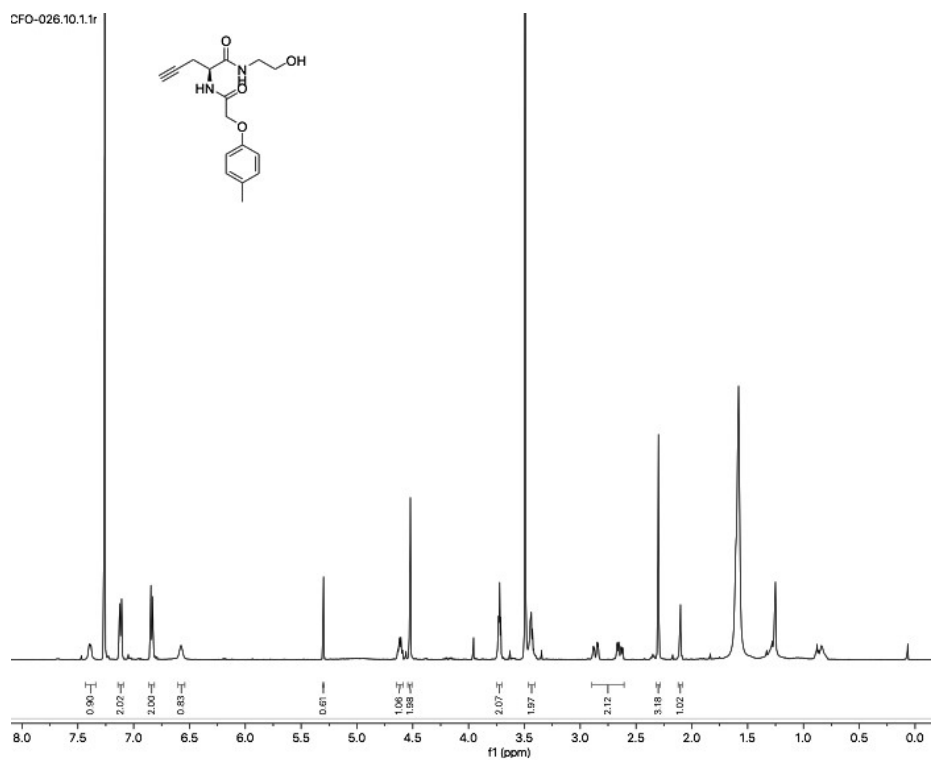


601

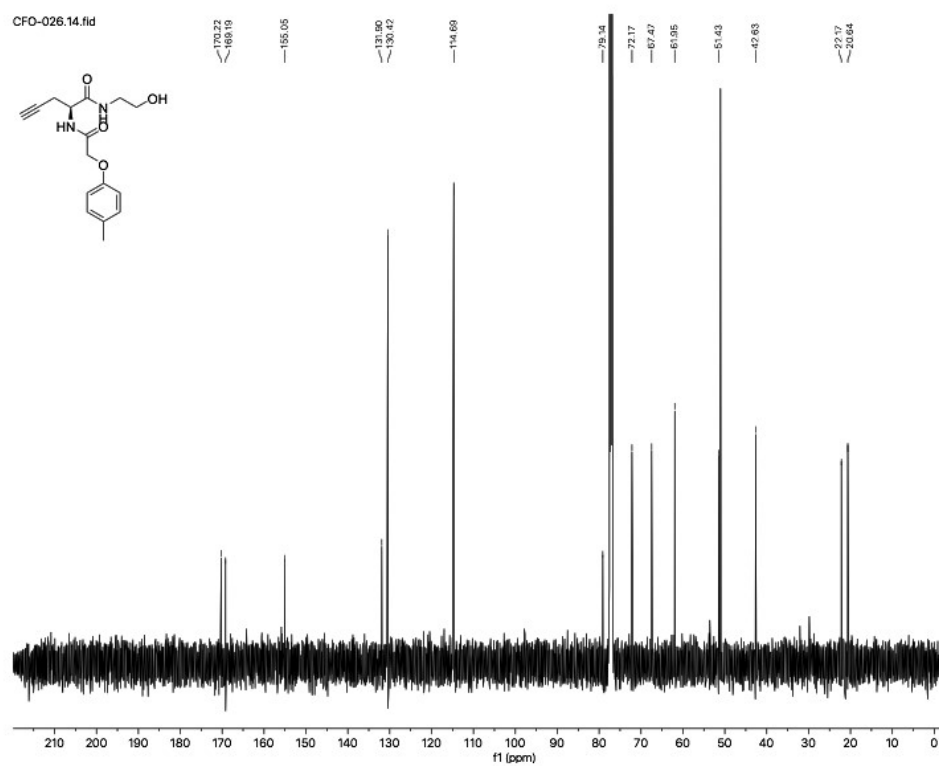


602

52

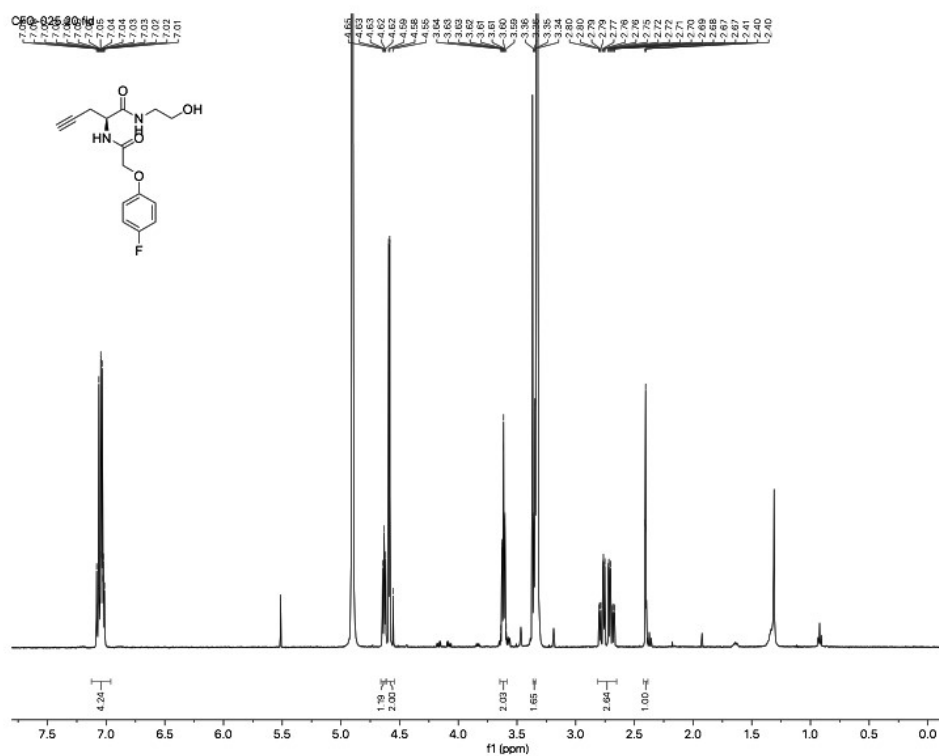


607

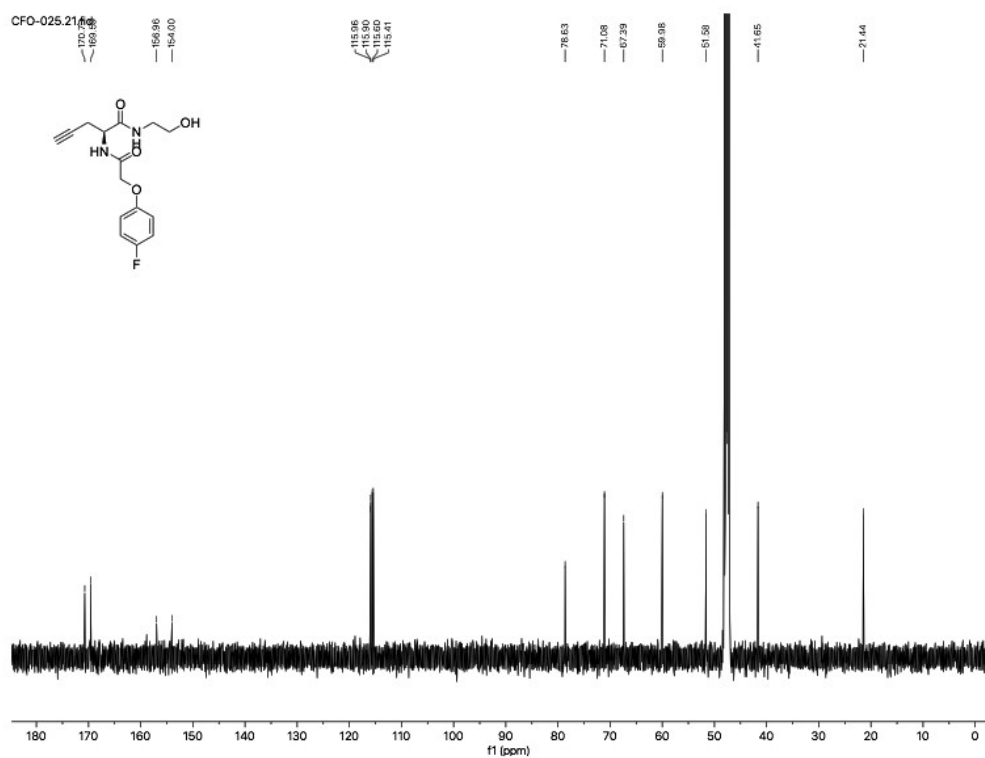


608

55



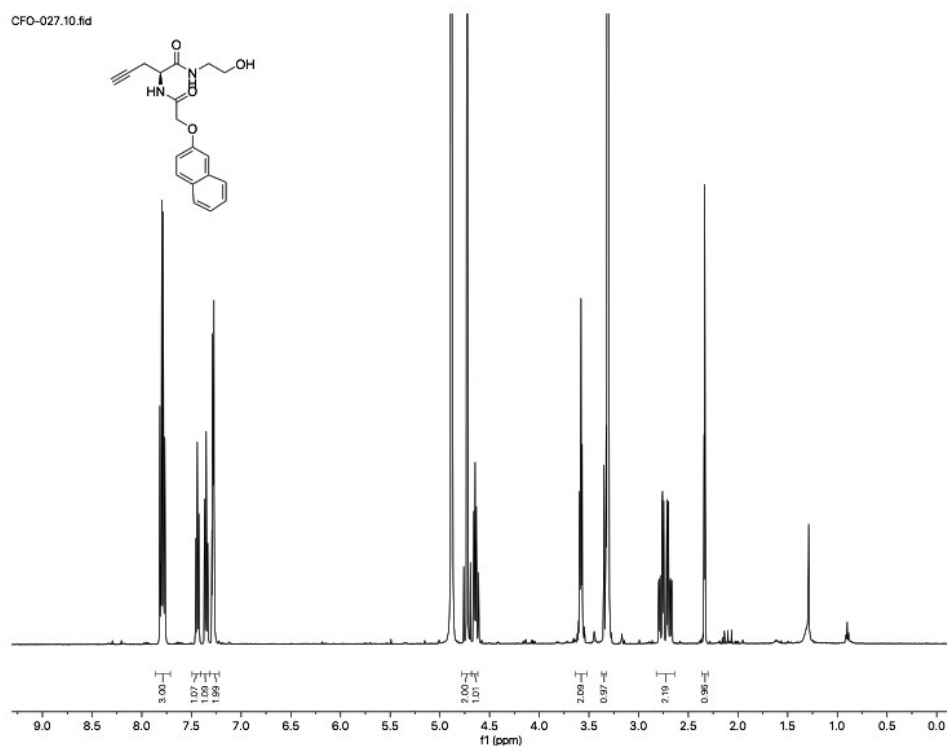
609



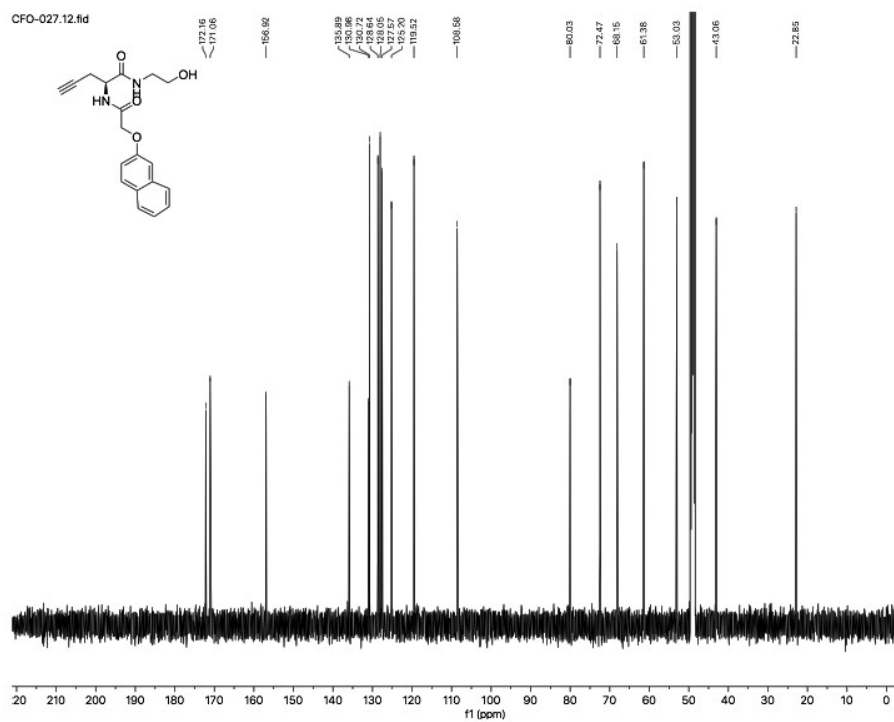
610

56

611



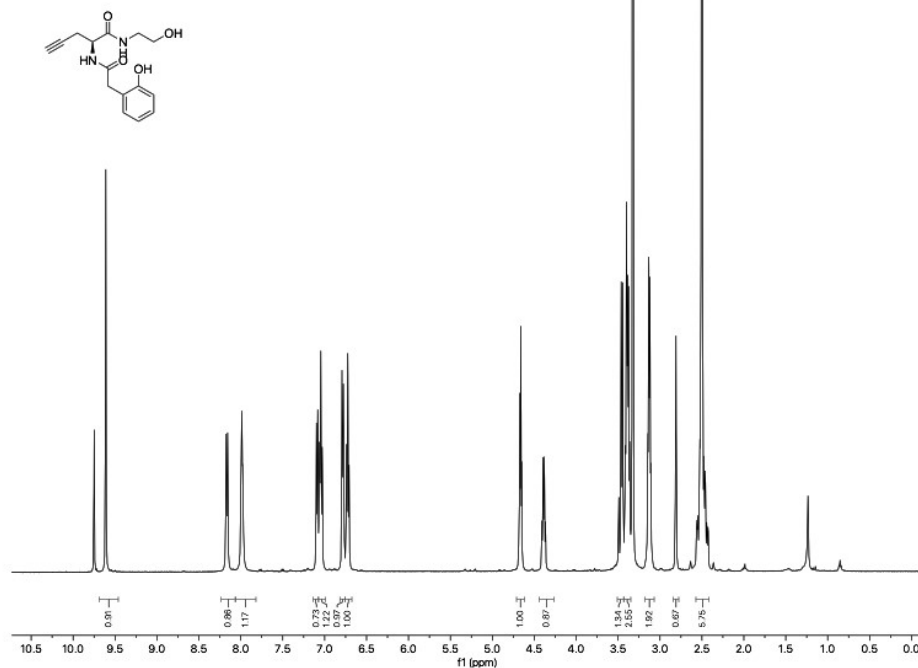
612



613

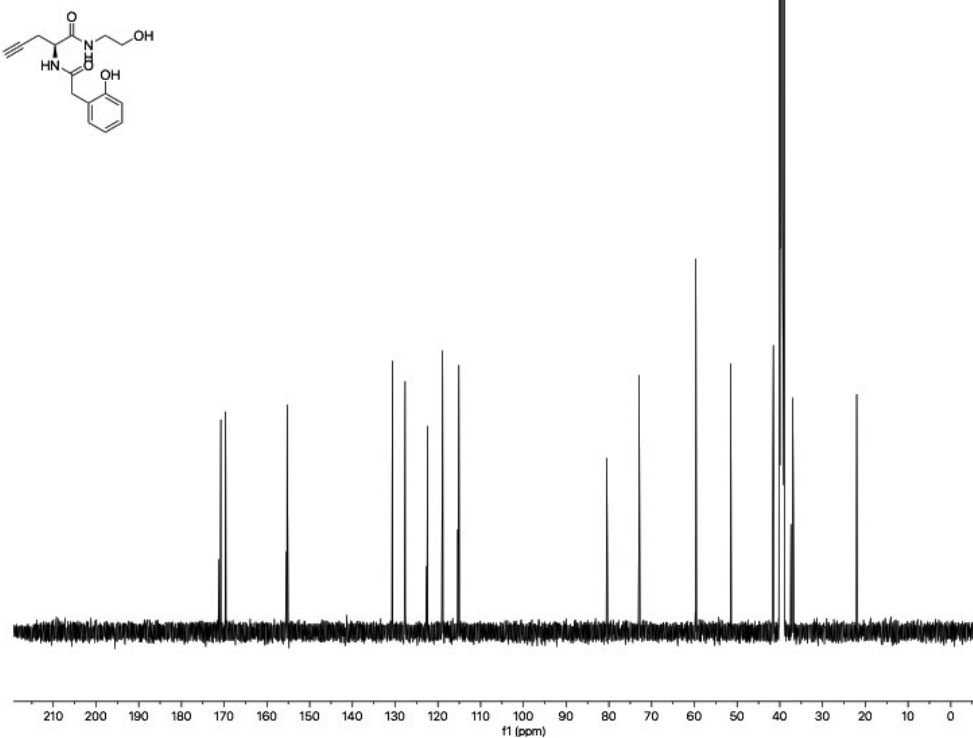
57

ES_CFO_28.10.fid



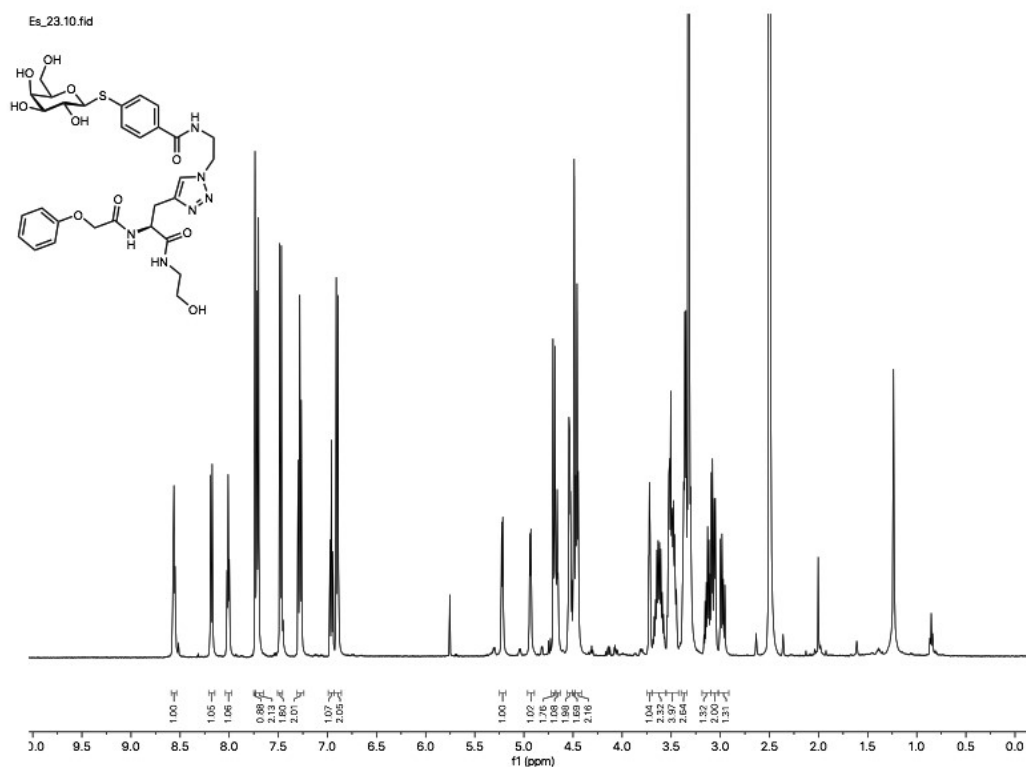
614

ES_CFO_28.11.fid

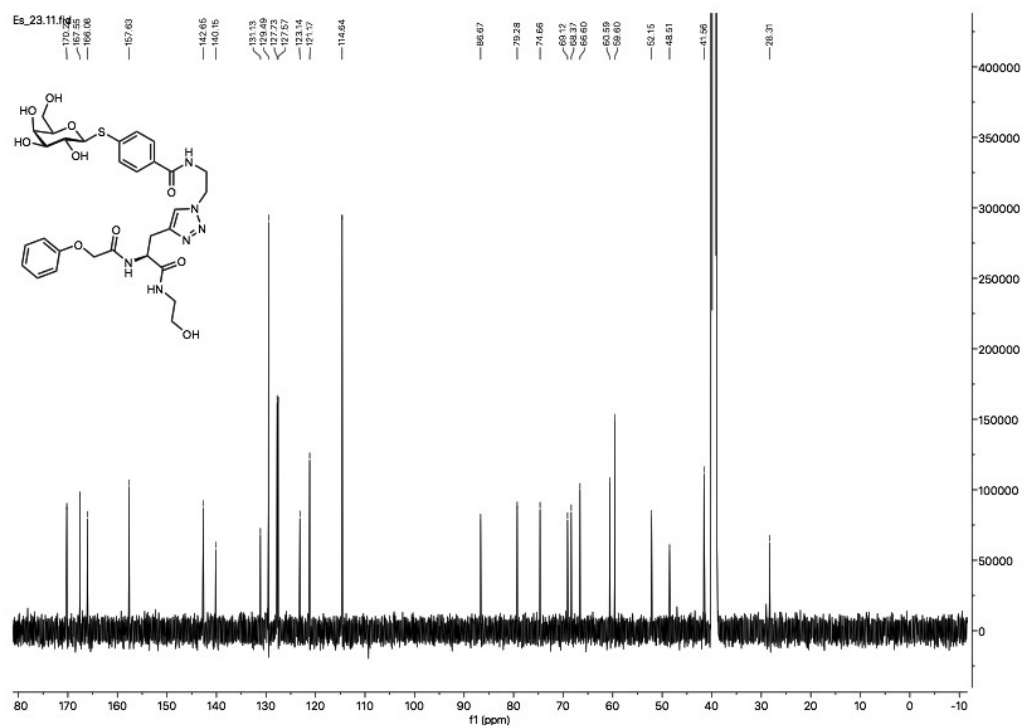


615

58



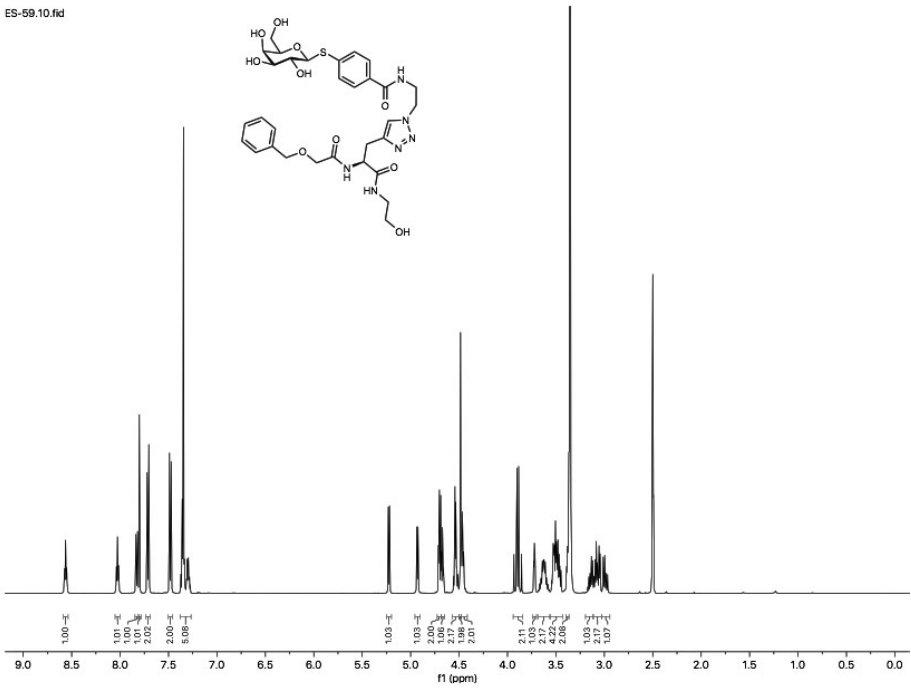
616



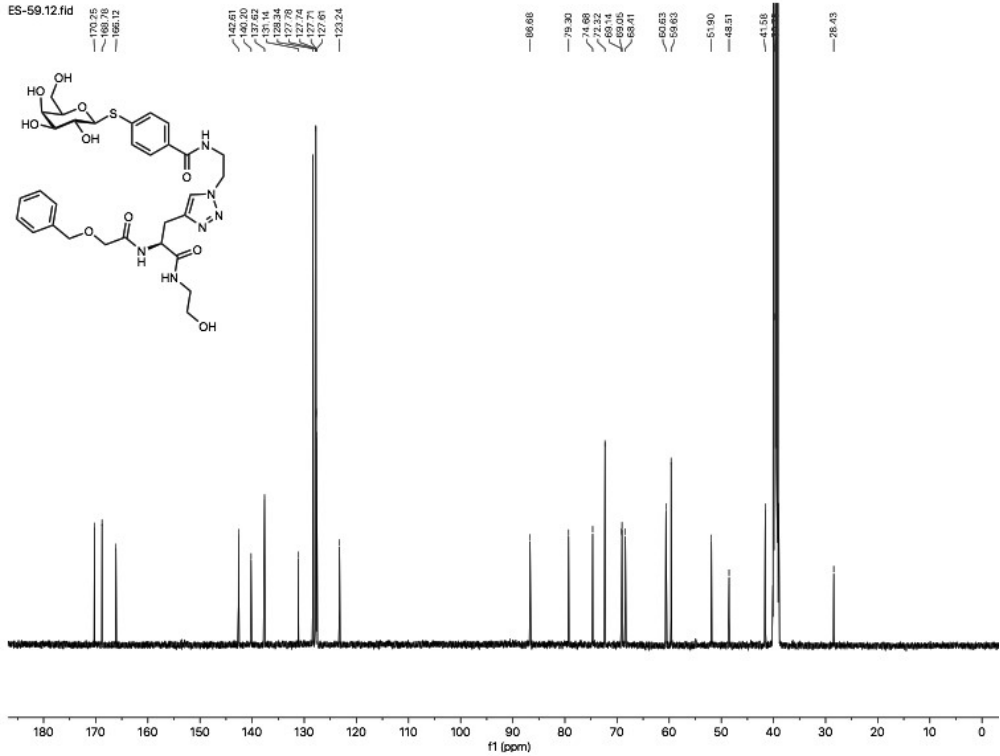
617

59

618



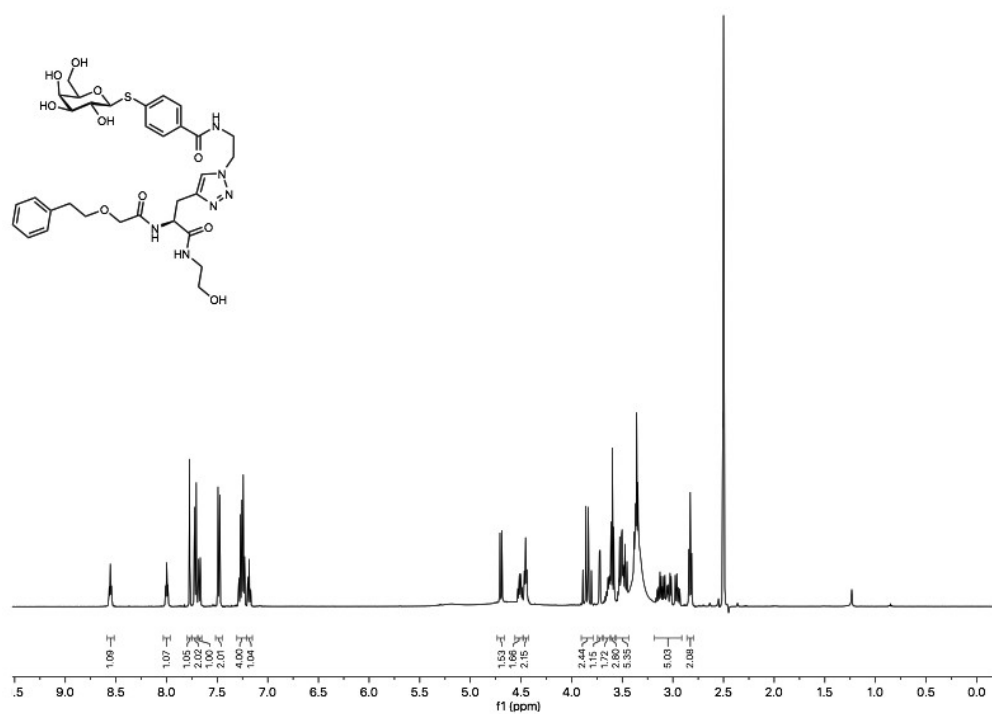
619



620

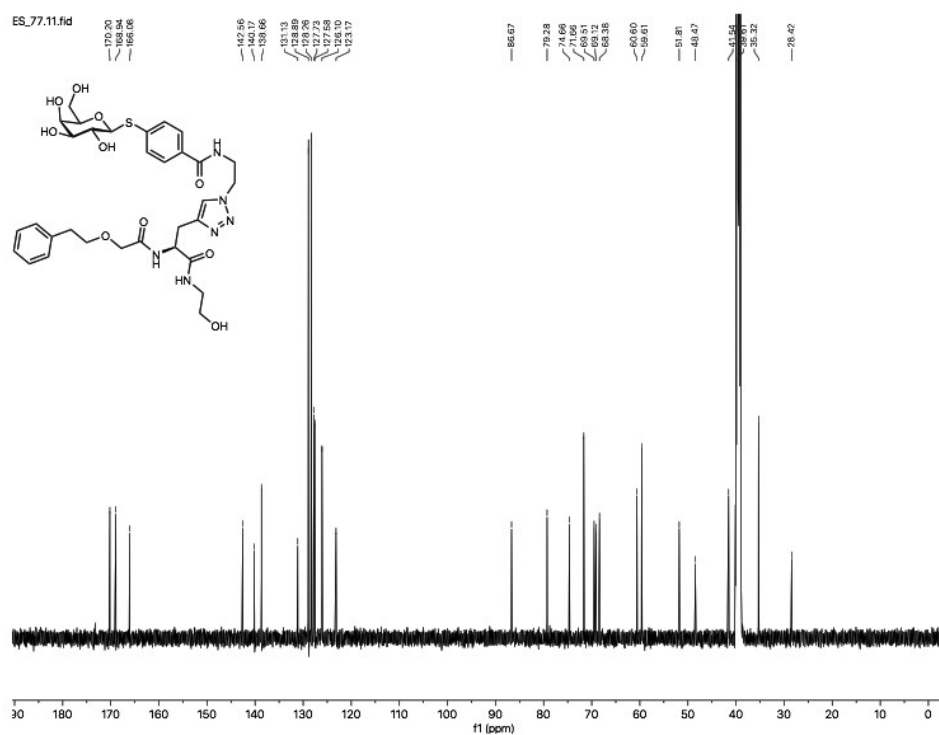
60

ES_77.10.fid



621

ES_77.11.fid



622

61

ES_78.10.fid

OCC(=O)N[C@@H](Cc1c[n+]([n-])cn1)C(=O)NCCOC(=O)CCc2ccccc2

Integration values (from left to right): 1.13, 1.11, 3.00, 2.03, 2.00, 3.09, 1.26, 1.32, 2.38, 2.63, 1.15, 2.80, 8.79, 5.23, 2.07, 2.19.

ES_78.12.fid

170.20
169.00
168.87

142.54
141.69
140.16

131.12
128.31
127.72
127.57
127.22
125.10

86.67

79.27

74.68
70.13
69.57
68.72
66.38

60.59
59.61

51.81
49.48

41.54
39.89

31.54
30.80
28.41

OCCNC(=O)C1C(=O)N(C1)C#N2C=NC=CC2C(=O)NCC(=O)c3ccc(cc3)SC4C(C(C(C(C4O)O)O)O)O

215

ES_8110.fid

Chemical structure of compound 10b (top):

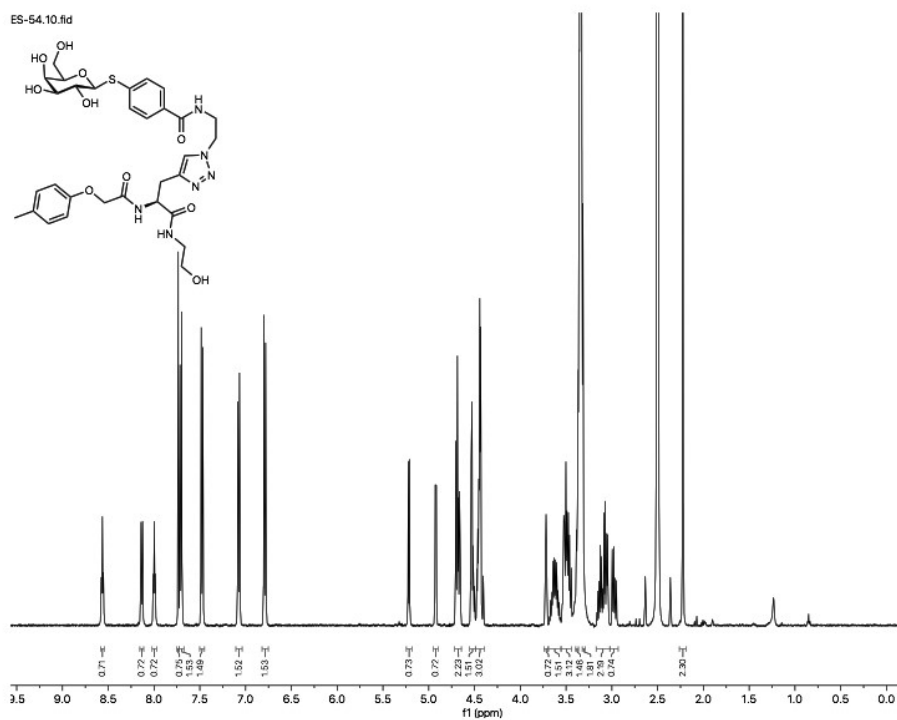
OCCNC(=O)C1C(=O)NC(=O)C1C(=O)N2C=CN=C2C(=O)NCC(=O)c3ccc(cc3)SC[C@H]4O[C@@H](CO)[C@H](O)[C@@H](O)[C@H]4O

¹H NMR spectrum (bottom) showing chemical shifts (ppm) and integrations:

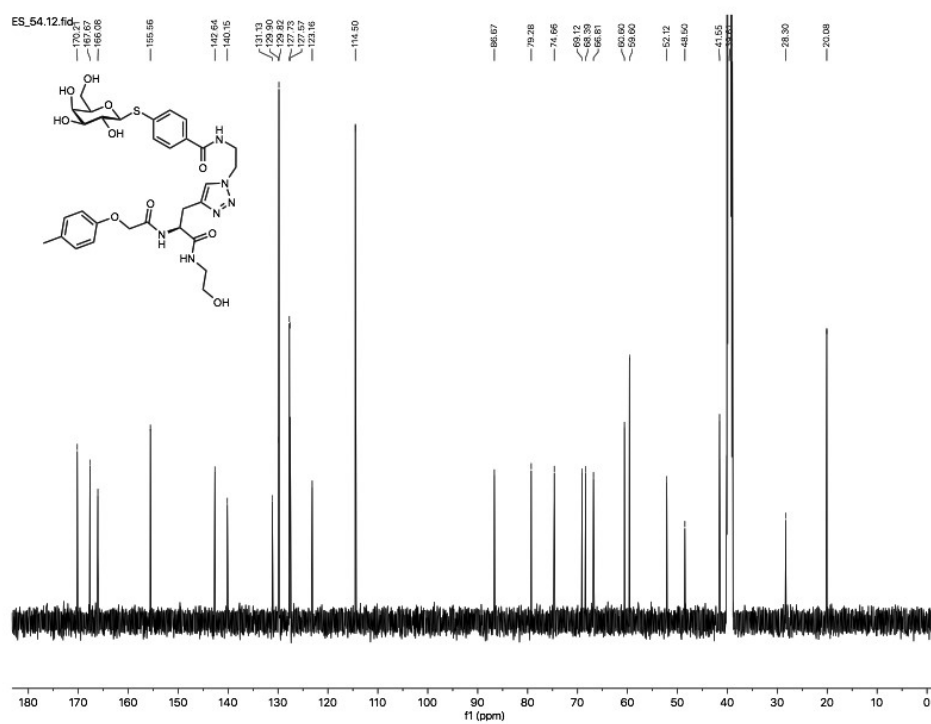
Chemical Shift (ppm)	Integration
~8.5	1.14
~7.8	1.11
~7.6	1.05
~7.4	1.03
~7.2	2.03
~7.0	2.00
~6.8	3.00
~4.8	3.02
~4.6	1.88
~4.4	2.21
~3.8	2.23
~3.6	1.96
~3.4	2.88
~3.2	4.51
~3.0	16.58
~2.8	2.30
~2.6	2.17
~2.5	2.05
~1.6	2.13
~1.4	2.10

ES_81.12.fid

216



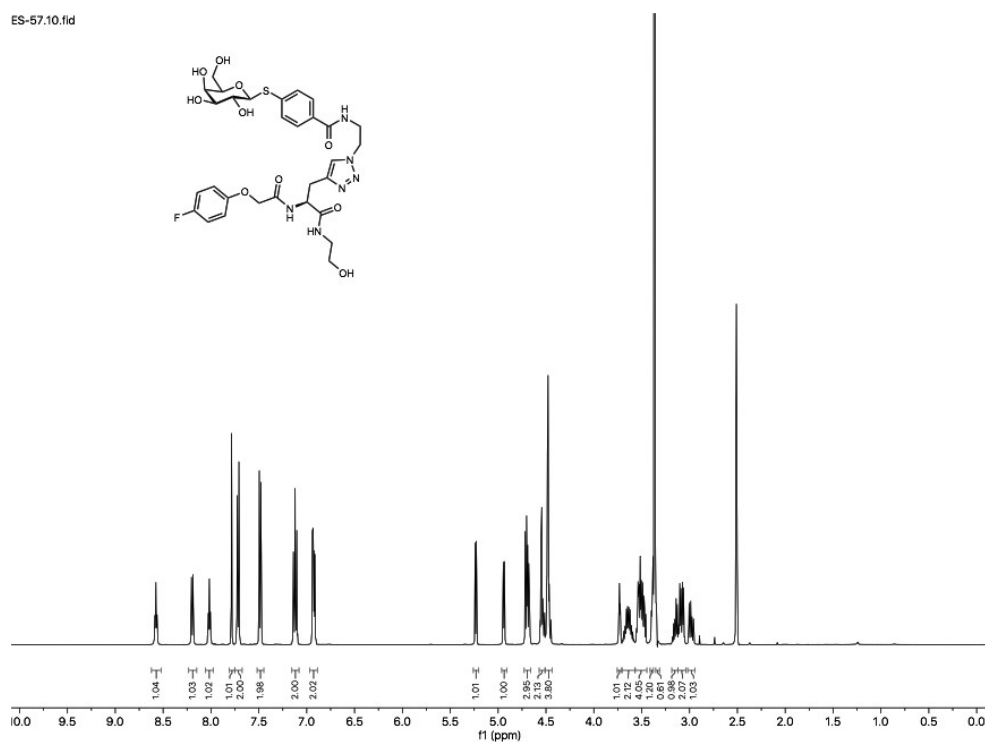
627



628

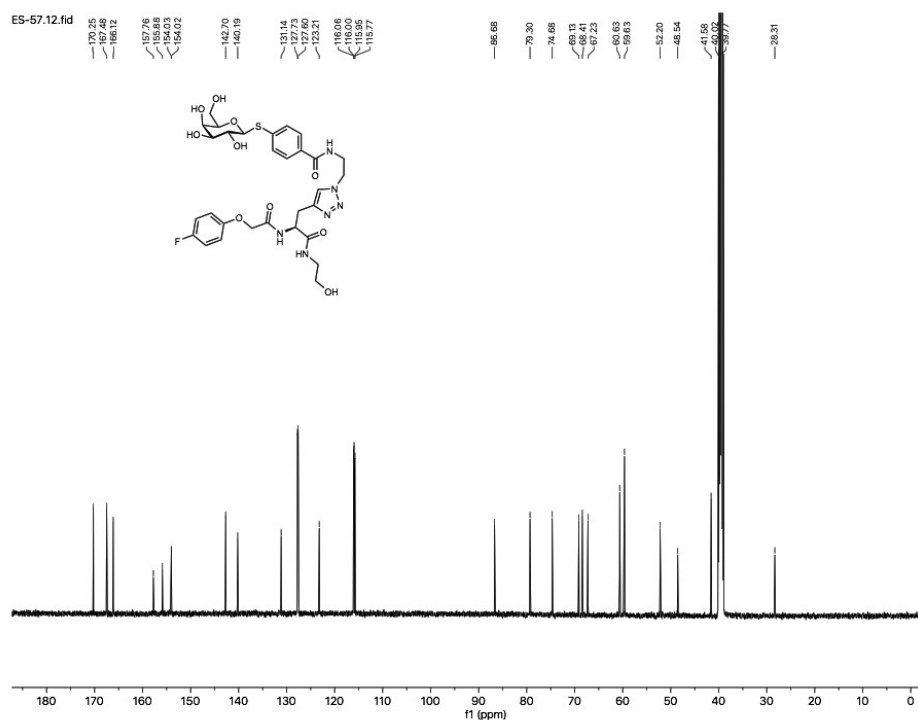
64

ES-57.10.fid



629

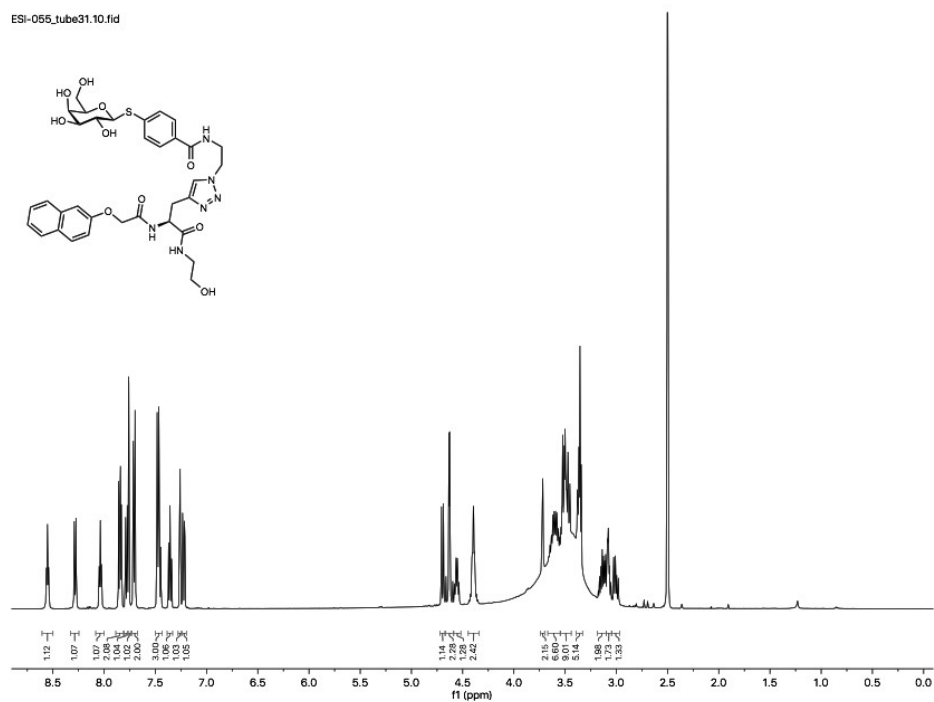
ES-57.12.fid



630

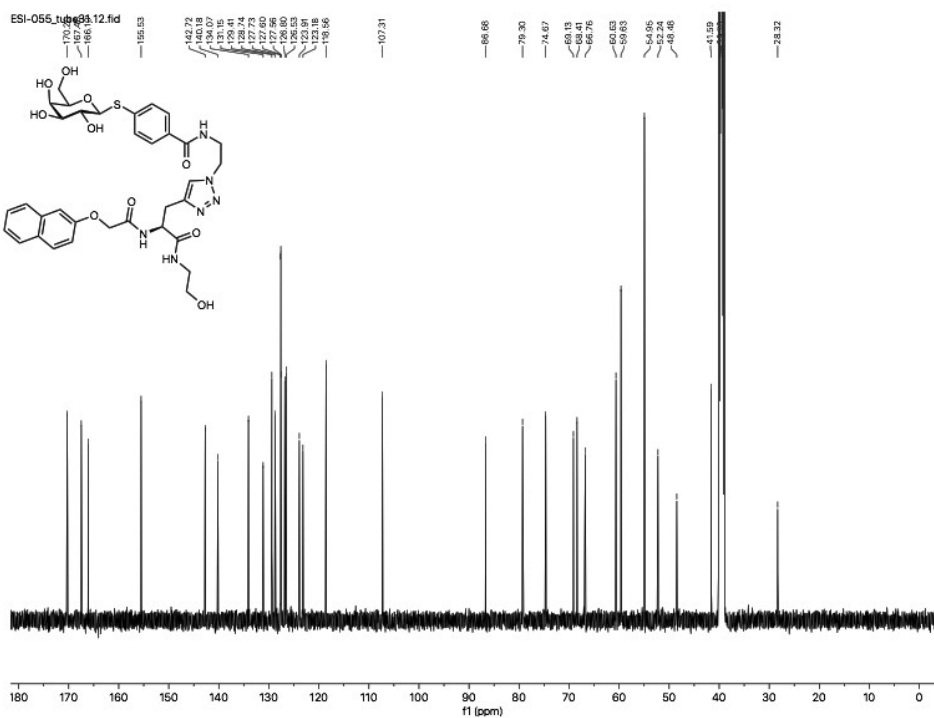
65

ESI-055_tube31.10.fid



631

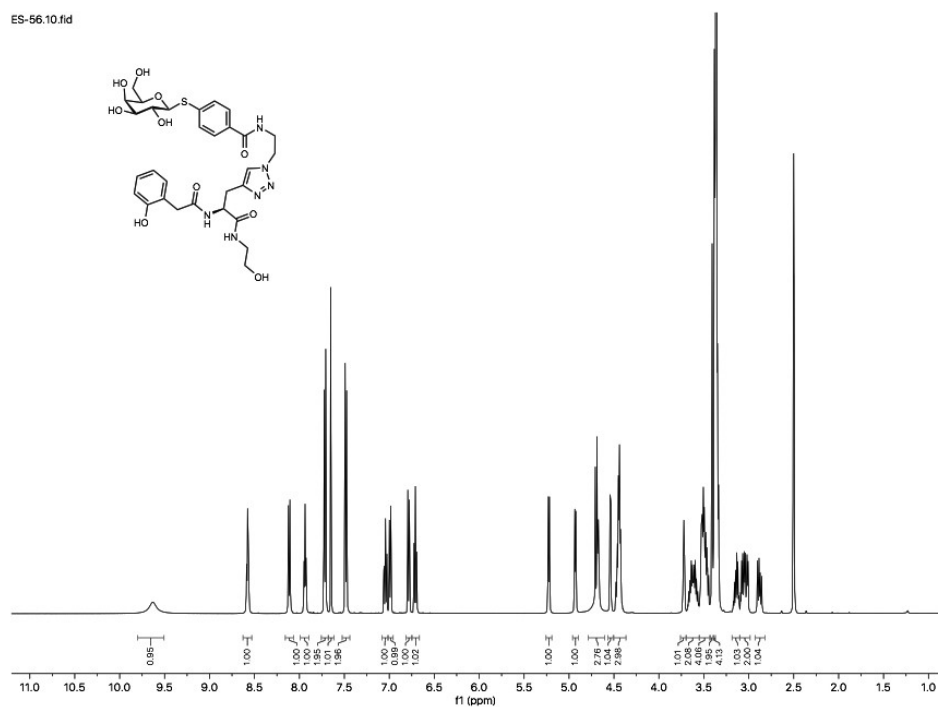
ESI-055_tube31.12.fid



632

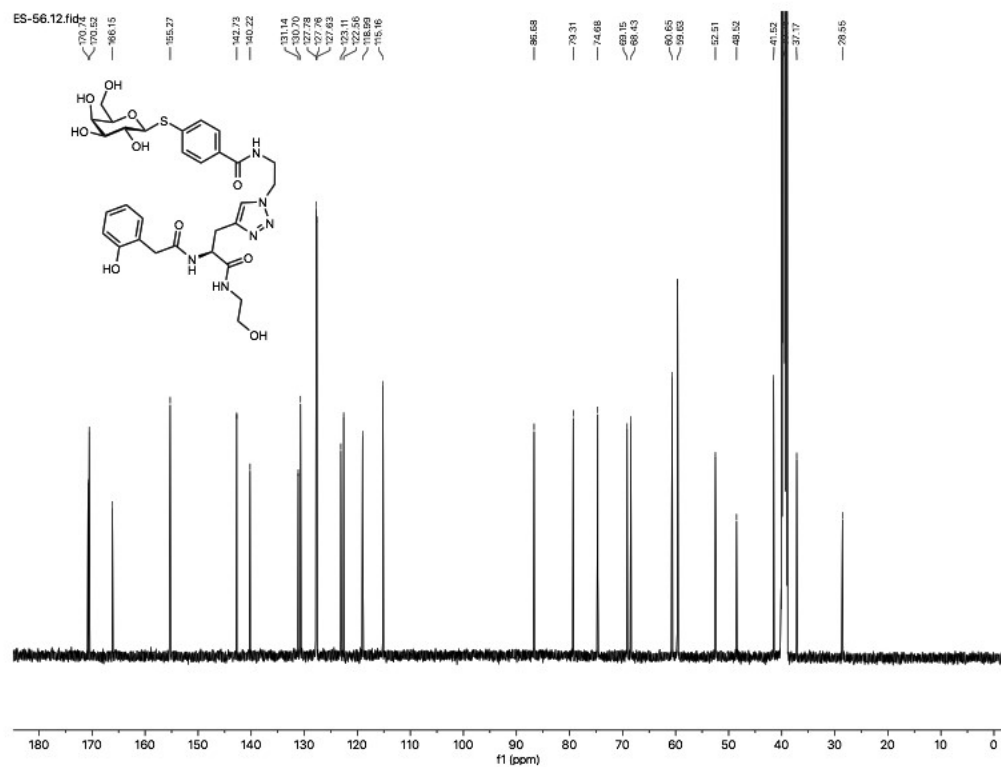
66

ES-56.10.fid



633

ES-56.12.fid



634

67

Supporting Information to Discovery of *N*- β -Fucosyl Amides as High-Affinity Ligands for the *Pseudomonas aeruginosa* Lectin LecB

Patrycja Mała^{#1,2}, Eike Siebs^{#1,3,4}, Joscha Meiers^{1,3,4}, Katharina Rox^{4,5}, Annabelle Varrot⁶, Anne Imberty⁶, Alexander Titz^{1,3,4*}

¹Chemical Biology of Carbohydrates (CBCH), Helmholtz-Institute for Pharmaceutical Research Saarland (HIPS), Helmholtz Centre for Infection Research, 66123 Saarbrücken, Germany;

²Faculty of Chemistry, Adam Mickiewicz University, 61-614 Poznań, Poland;

³Department of Chemistry, Saarland University, 66123 Saarbrücken, Germany;

⁴Deutsches Zentrum für Infektionsforschung (DZIF), Standort 38124 Braunschweig, Germany;

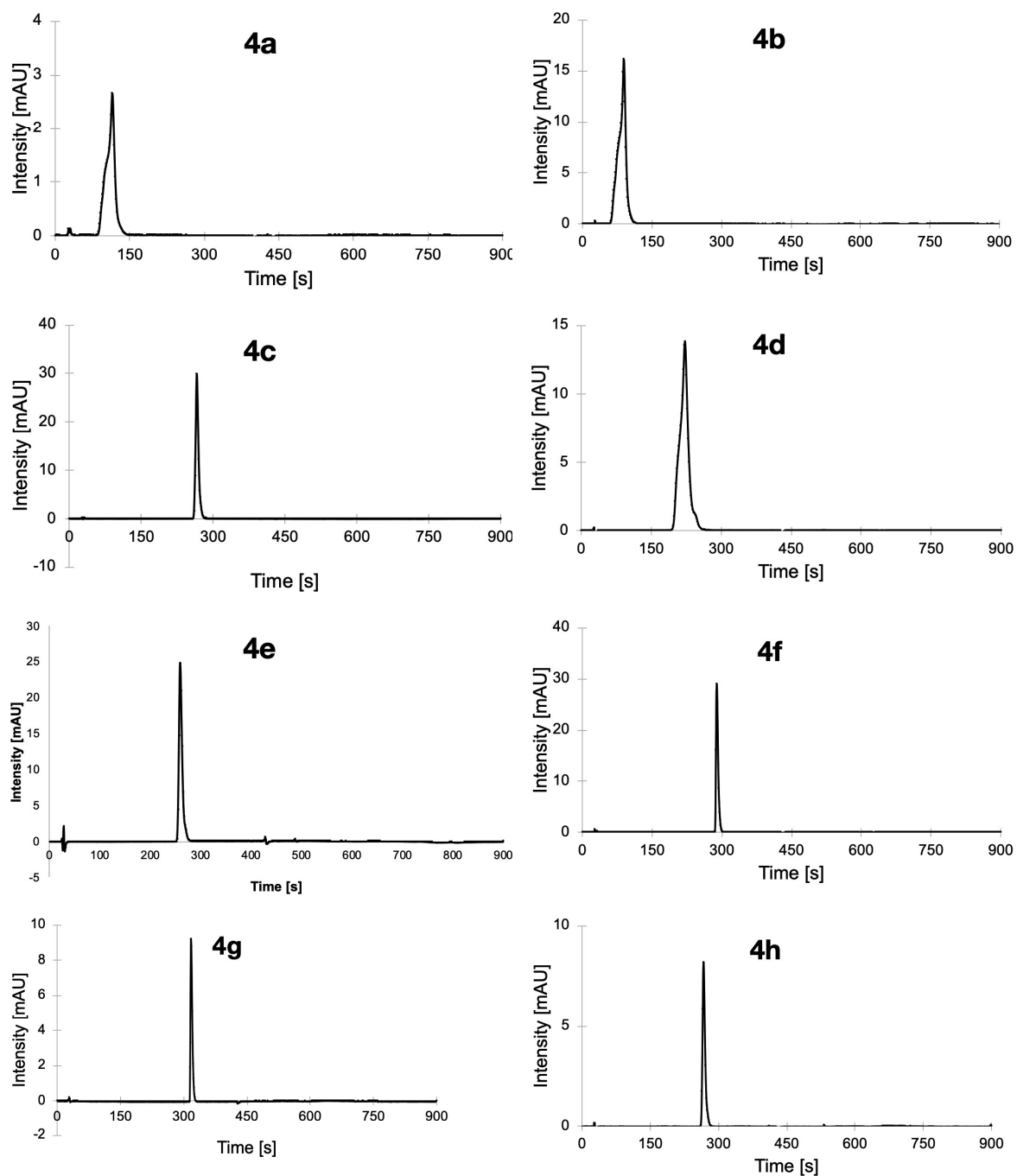
⁵Chemical Biology (CBIO), Helmholtz Centre for Infection Research, 38124 Braunschweig, Germany;

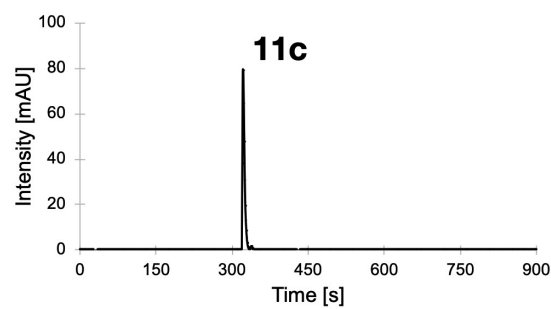
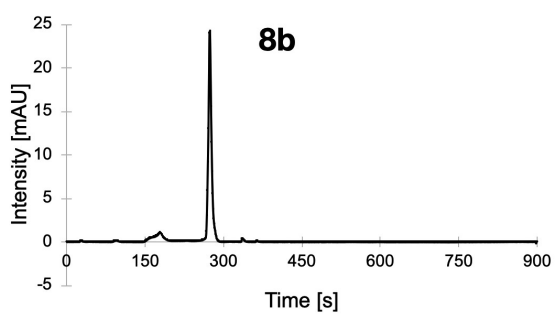
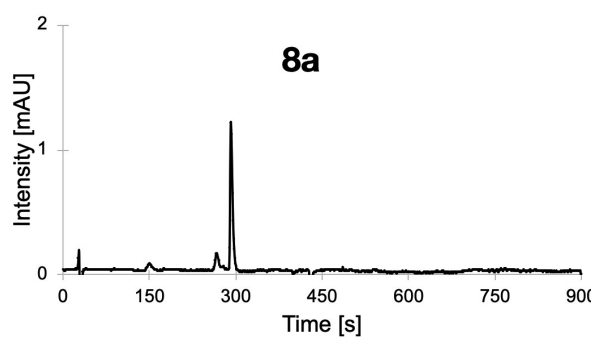
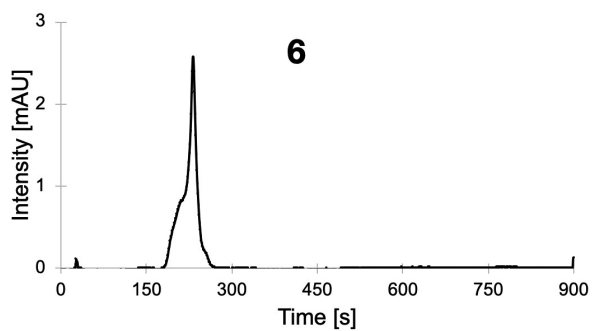
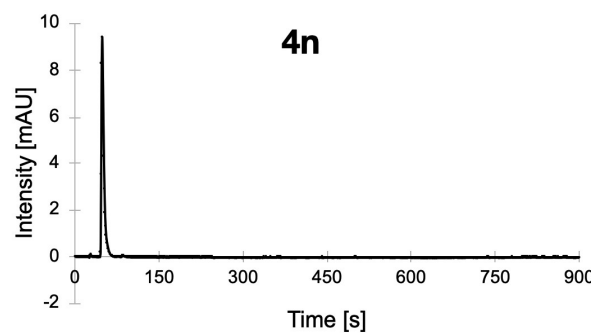
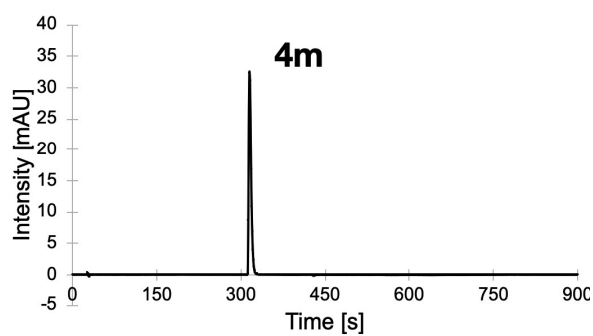
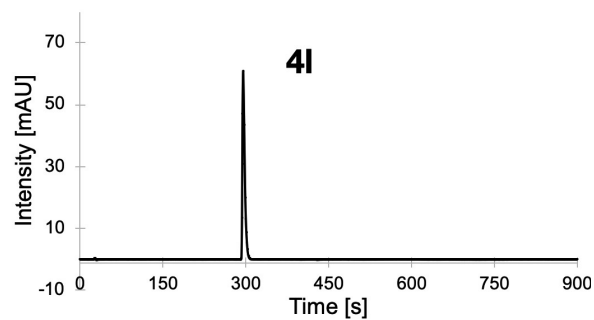
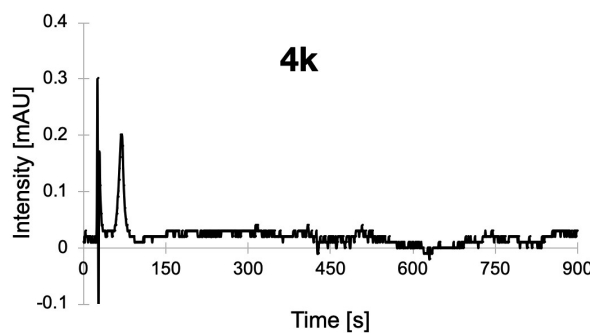
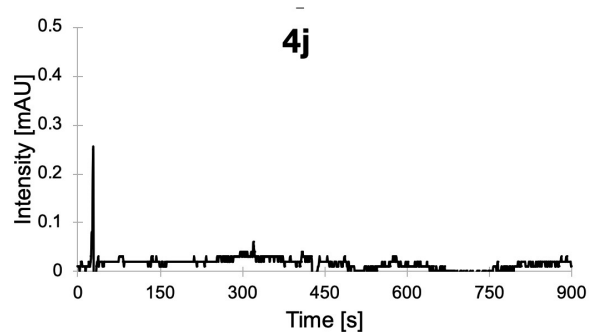
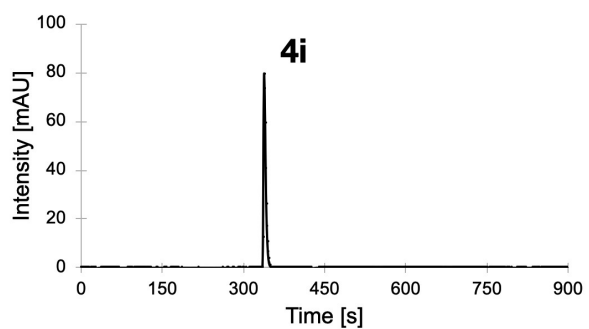
⁶Univ. Grenoble Alpes, CNRS, CERMAV, 38000 Grenoble, France.

*corresponding author e-mail: alexander.titz@helmholtz-hzi.de

Purity Control of Key Compounds

Table S4: Purity of studied compounds by HPLC-UV.





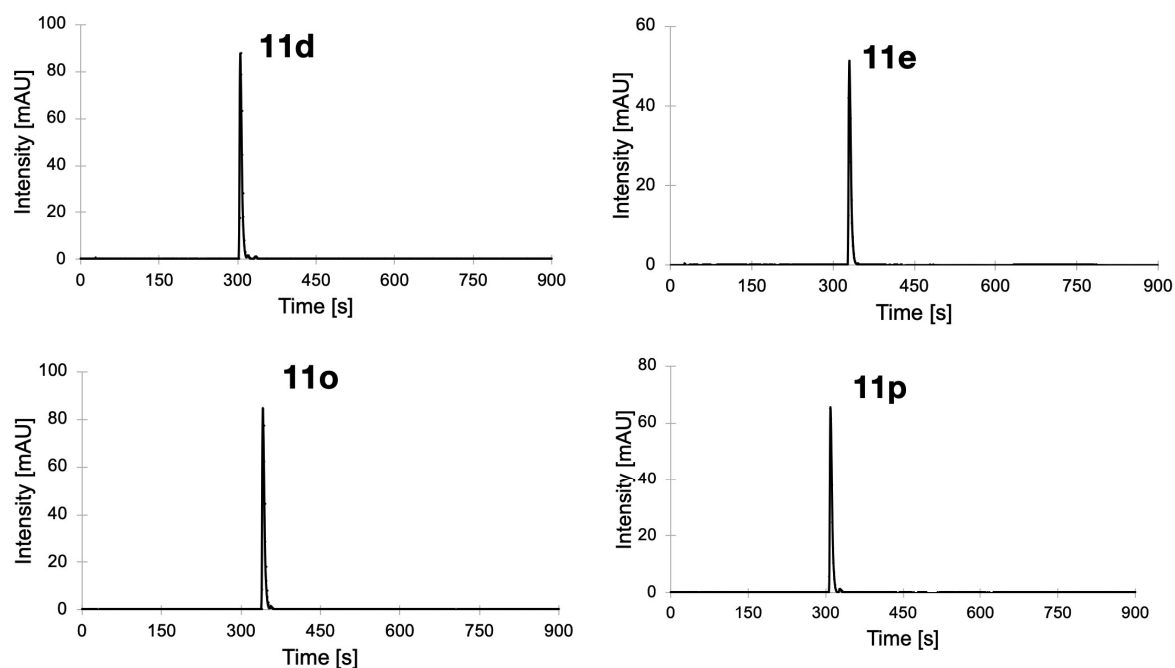


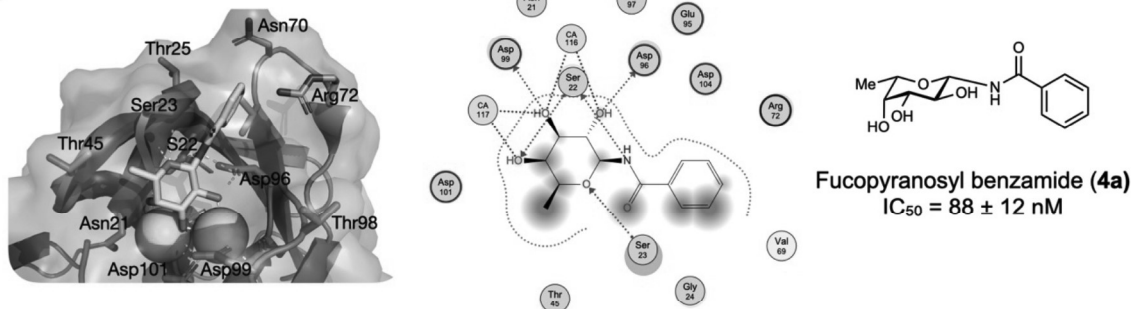
Table S5: Compound purity determination by integration of individual peaks of HPLC-UV chromatograms.

Compound	Purity [%]	Compound	Purity
4a	>99%	4m	>99%
4b	>99%	4n	>99%
4c	>99%	4o	>99%
4d	>99%	6	>99%
4e	>99%	8a	87%*, $\beta = 62\%$, $\alpha = 25\%$
4f	>99%	8b	>99%, $\beta = 92\%$, $\alpha = 8\%$
4g	>99%	11c	>99%
4h	>99%	11d	>99%
4i	>99%	11e	>99%
4j	>99%	11o	>99%
4k	>99%	11p	>99%
4l	>99%		

*pure by NMR (>95%).

Docking

A



B

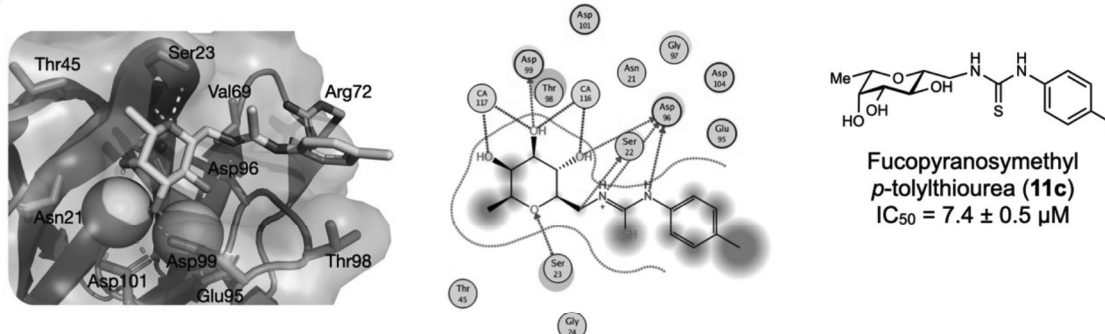


Figure S5: A) Left: β -fucosyl benzamide (**4a**) docked to the carbohydrate binding site of LecB (pdb: 1OXC), center: 2D protein-ligand interaction map of **4a** in complex with LecB in MOE (purple = polar, purple red = acidic, purple blue = basic, green = hydrophobic, blue arrow = backbone acceptor/ donor, green arrow = sidechain acceptor/ donor, blue = ligand exposure, light blue = receptor exposure, grey line = solvent contact), right: structure of β -fucosyl benzamide (**4a**). **B)** Left: β -fucosylmethyl *p*-toloylthiourea (**11c**) docked to the carbohydrate binding site of LecB (pdb: 1OXC), center: 2D protein-ligand interaction map of **11c** in complex with LecB in MOE, right: structure of β -fucosylmethyl thiourea (**11c**).

Isothermal Titration Calorimetry

Table S6: ITC sensorgrams (top) obtained by titration of **4a** or **6** against LecB_{PAO1} with integration of peaks and fits (bottom).

LecB (100 μ M)	LecB (100 μ M)	LecB (100 μ M)
Compound 4a (1500 μ M)	Compound 4a (500 μ M)	Compound 4a (1000 μ M)
LecB (100 μ M)	LecB (100 μ M)	LecB (100 μ M)
Compound 6 (1500 μ M)	Compound 6 (1000 μ M)	Compound 6 (1000 μ M)

X-ray Crystallography

Data Collection and Refinement Statistics

Table S7: X-ray data collection and refinement statistics of LecB complexes.

Complex	LecB _{PAO1-4a}	LecB _{PAO1-4i}	LecB _{PAO1-6}
Crystallization			
Crystallizing solution	30% PEG 8K, 200 mM (NH ₄) ₂ SO ₄ , 100 mM tris pH = 8.5	28% PEG 8K, 200 mM (NH ₄) ₂ SO ₄ , 100 mM tris pH = 8.5	24% PEG 8K, 1M LiCl and 100 mM sodium acetate pH = 4.4
Crystal form:	hexagons	hexagons	plates
Data collection*			
Date	08.04.2021	08.04.2021	09-06-2021
Beamline	Proxima 1 Soleil	Proxima 1 Soleil	Proxima 2 Soleil
Wavelength [Å]	0.97856	0.97856	0.9801
Resolution range [Å]	44.1-2.5 (2.6-2.5)	46.08–1.55 (1.58–1.55)	40.69–1.5 (1.53–1.50)
Space group	P6 ₂ 22	C222 ₁	P1
Unit cell a, b, c (Å)	118.90 118.90 256.57	71.9 76.49 184.31	45.38 51.59 52.60
α, β, γ (°)	90.00 90.00 120.00	90.00 90.00 90.00	101.81 99.47 115.81
Total number of observations	759188 (82753)	711298 (36020)	211861 (9941)
Total number unique	38011 (4202)	73973 (3576)	61588 (2980)
Multiplicity	20 (19.7)	2.0 (2.0)	3.4 (3.3)
Completeness (%)	100 (100)	99.97 (100.00)	95.50 (92.49)
Mean I/sigma(I)	20.8 (4)	13.57 (3.43)	12.9 (3.4)
Wilson B-factor [Å ²]...	44.1	18.9	13.1
R-merge (within I+/I-)	0.114 (0.819)	0.082 (0.557)	0.051 (1.53)
R-meas (within I+/I-)	0.120 (0.862)	0.092 (0.621)	0.072 (0.45)
R-pim (within I+/I-)	0.036 (0.267)	0.041 (0.272)	0.051 (0.318)
CC1/2	0.999 (0.904)	0.998 (0.902)	0.997 (0.881)
Refinement*			
Nb of Reflections	Not refined	70251 (7294)	54017 (5971)
Nb of R-free		3665 (360)	7571 (718)
R-work		0.155	0.1497
R-free		0.175	0.1741
Rms Bond Length [Å]		0.0146	0.0153
Rms Bond Angle [°]		1.7103	1.8239
Rms Chiral Volume		0.0975	0.0949
Ramachandran [%]			
-Favored		97.77	97.32
-Allowed		100	100
-Outliers		0	0
Rotamer outliers [%]		0.27	0.26
Clashscore		3.14	4.21
Nb of atoms / Average			
B-factor [Å ²]...		4094 / 21.6	4032 / 15.5
...-Protein		3447 / 19.6	3427 / 13.9
...-Ligand		118 / 19.95	100 / 20.1
-Calcium		8 / 18.0	8 /
...-Waters		637 / 32.0	505 / 26.0
PDB code		8AIY	8AIJ

*Statistics for the highest-resolution shell are shown in parentheses.

LecB_{PAO1} Protomers in Complex with α -Fucosyl Benzamide (6) or β -Fucosyl Biphenylamide (4i)

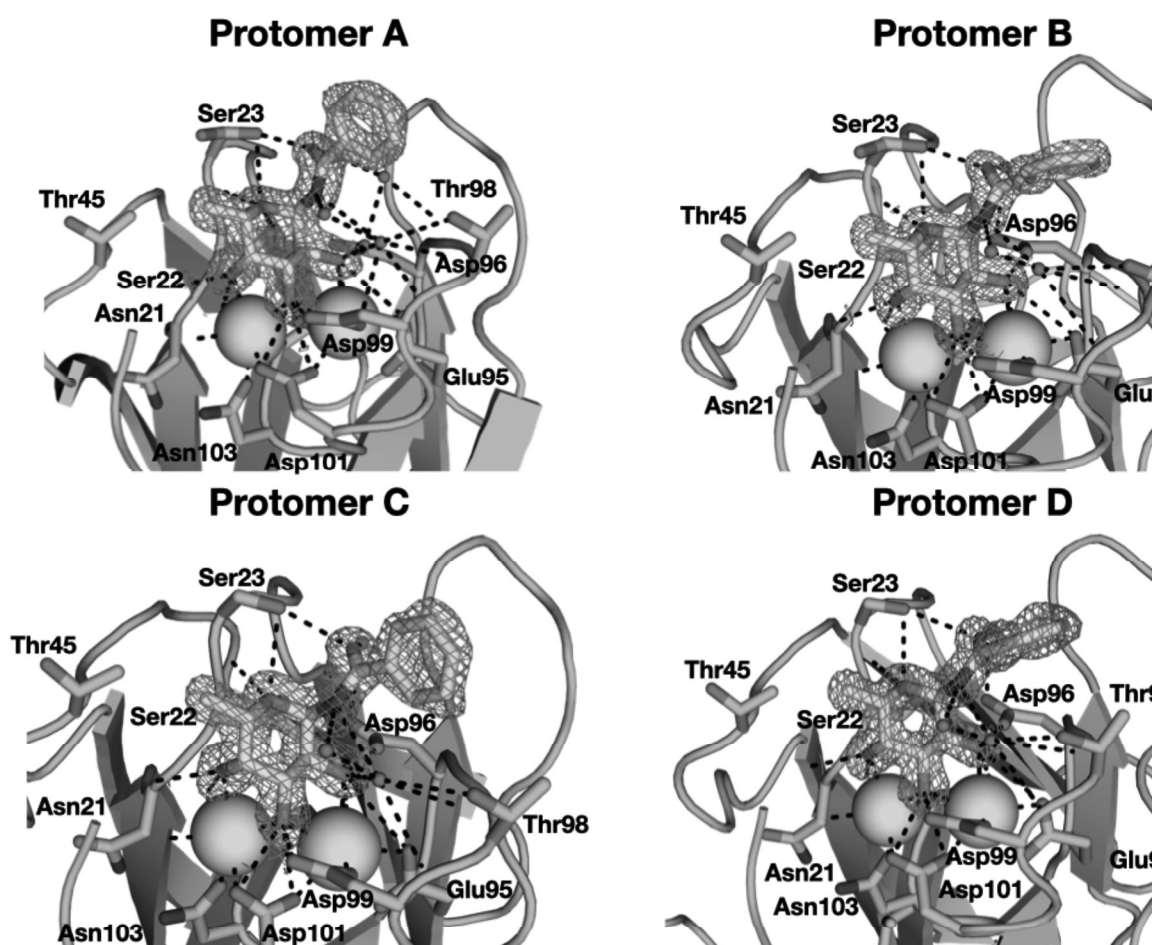


Figure S6: Co-crystal structure of tetrameric LecB_{PAO1} in complex with α -fucosyl benzamide (**6**). All four carbohydrate binding sites of LecB_{PAO1} are occupied by one ligand *via* co-ordination to the calcium ions with its fucose moiety. The aglycon of **6** points away from the surface resulting in different orientations of the phenyl ring (protomers A–D). The ligands are visualized in sticks with the carbons, nitrogens and oxygens atoms colored in cyan, blue and red, respectively and the calcium ions in green spheres. The electron density is displayed as grey mesh at 1σ .

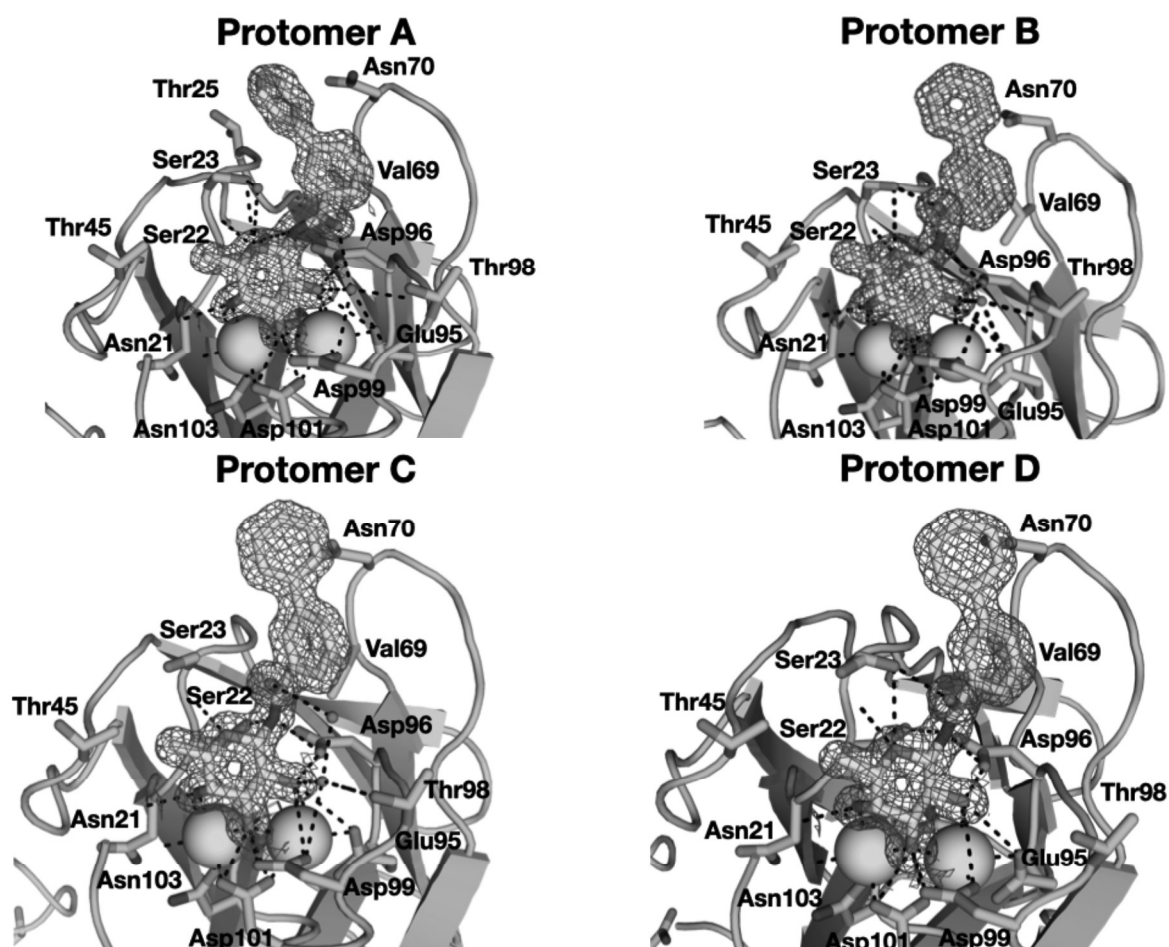


Figure S7: Co-crystal structure of tetrameric LecB_{PAO1} in complex with β -fucosyl biphenylamide (**4i**). All four carbohydrate binding sites of LecB_{PAO1} are occupied by one ligand *via* co-ordination to the calcium ions with the fucose moiety. The nitrogen of the amide function forms a hydrogen bond with Ser22 or an ion-dipole interaction with Asp96 from the cleft, the aromatic ring in proximity to the fucose moiety forms lipophilic contacts with Gly24 and Val69 and forms further hydrophobic interactions with the surface of LecB. The distant ring points from the surface and binds loosely to the surface resulting in different orientations for the second ring (protomers A–D). The ligands are visualized in sticks with the carbons, nitrogens and oxygens atoms colored in cyan, blue and red, respectively and the calcium ions in green spheres. The electron density is displayed as grey mesh at 1σ .

Biphenyl **4i** Superimposed with Fucose-Mannose Hybrid **II** or *Manno*-Cinnamide **S1**



Figure S8: Superposition of complexes LecB_{PAO1}-**4i** (pdb 8AIY, lightblue, protomer D, transparent surface grey) and LecB_{PA14}-**S1** (pdb 5A3O, cyan, protomer B). Amide NH of β -fucosylamide **4i** points towards Ser22 and Asp96, while the introduction of an additional CH₂-spacer in **S1**, reduces these interactions.

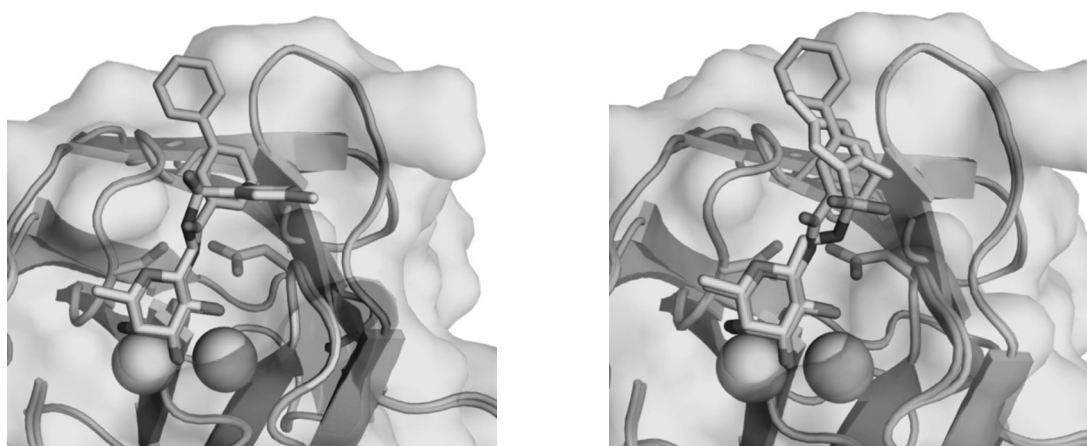


Figure S9: Superposition of complexes LecB_{PAO1}-**4i** (8AIY, lightblue, protomer D, transparent surface grey) and LecB_{PA14}-**II** (pdb 5MAZ, cyan, left: protomer B, right: protomer D). The absence of a CH₂-spacer in fucoside **4i** allows a beneficial orientation of the amide NH towards Ser22 and Asp96. The biphenyl residue of **4i** points towards an opposing direction than the binding pose of the dimethylthiophene in **II** without crystal contacts (left) while the binding pose of **II** with crystal contacts points in the same direction (right).

ADME

Table S8: Mass transitions of the internal standard, caffeine, and the compounds.

	Q1 mass	Q3 mass	DP [volt]	CE [volts]	CXP [volts]
Caffeine	195.024	138.0	130.0	25.0	14.0
Caffeine	195.024	110.0	130.0	31.0	18.0
4a	265.209	96.900	-65.000	-32.000	-11.000
4a	265.209	79.900	-65.000	-90.000	-19.000
4a	265.209	205.100	-65.000	-10.000	-11.000
4b	272.123	167.900	-50.000	-12.000	-9.000
4b	272.123	83.900	-50.000	-28.000	-9.000
4c	280.130	175.900	-50.000	-12.000	-19.000
4c	280.130	157.900	-50.000	-32.000	-17.000
4d	297.132	182.900	-85.000	-42.000	-19.000
4d	297.132	183.900	-85.000	-38.000	-19.000
4d	297.132	118.900	-85.000	-60.000	-13.000
4e	311.180	182.900	-70.000	-44.000	-19.000
4e	311.180	206.900	-70.000	-16.000	-23.000
4f	301.314	255.200	-5.000	-8.000	-13.000
4f	301.314	45.000	-5.000	-26.000	-11.000
4f	301.314	196.800	-5.000	-16.000	-19.000
4g	334.146	229.900	-70.000	-18.000	-11.000
4g	334.146	211.900	-70.000	-32.000	-23.000
4h	280.138	175.900	-30.000	-14.000	-17.000
4h	280.138	157.900	-30.000	-30.000	-17.000

4i	342.145	238.000	-40.000	-16.000	-11.000
4i	342.145	220.000	-40.000	-28.000	-25.000
4j	205.181	189.100	-85.000	-34.000	-19.000
4j	205.181	189.900	-85.000	-26.000	-9.000
4j	205.181	173.200	-85.000	-54.000	-17.000
4k	232.144	128.000	-65.000	-14.000	-15.000
4k	232.144	188.000	-65.000	-6.000	-23.000
4l	292.174	187.900	-45.000	-12.000	-9.000
4l	292.174	160.000	-45.000	-26.000	-17.000
4m	316.185	211.900	-55.000	-16.000	-11.000
4m	316.185	193.900	-55.000	-30.000	-9.000
4n	255.099	211.000	-80.000	-16.000	-11.000
4n	255.099	191.000	-80.000	-32.000	-19.000
6	266.167	161.900	-35.000	-14.000	-17.000
6	266.167	144.000	-35.000	-32.000	-15.000
8a	301.226	255.100	-5.000	-10.000	-13.000
8a	301.226	45.000	-5.000	-28.000	-19.000
8b	308.092	203.900	-20.000	-12.000	-11.000
8b	308.092	146.800	-20.000	-20.000	-15.000

Supporting Information to Ligand-Based Design for Targeting *Burkholderia ambifaria* Lectin BamBL

Synthesis

General Information

Commercial chemicals and solvents were used without further purification. Deuterated solvents were purchased from Eurisotop (Saarbrücken, Germany).

Thin layer chromatography (TLC) was performed on Silica Gel 60 coated aluminum sheets containing a fluorescence indicator (Merck KGaA, Darmstadt, Germany) and developed under UV light (254 nm) and using a molybdate solution (0.02 M solution of $(\text{NH}_4)_4\text{Ce}(\text{SO}_4)_4 \cdot 2\text{H}_2\text{O}$ and $(\text{NH}_4)_6\text{Mo}_7\text{O}_{24} \cdot 4\text{H}_2\text{O}$ in aqueous 10% H_2SO_4) or a potassium permanganate solution (3 g of KMnO_4 , 20 g of K_2CO_3 in 5 mL of 5% NaOH and 300 mL of H_2O) followed by heating.

Teledyne Isco Combiflash Rf200 system using self-packed silica gel columns (60 Å, 400 mesh particle size, Fluka, for normal-phase liquid chromatography) or pre-packed Chromabond Flash RS15 C18 ec columns (60 Å, 15–40 µm particle size, Macherey-Nagel, Germany, for reversed-phase liquid chromatography) was used for preparative medium pressure liquid chromatography (MPLC).

Preparative high-pressure liquid chromatography (HPLC) was performed on Waters 2545 Binary Gradient Module equipped with a Waters 2489 UV/Vis detector using a C18 column (EC HPLC column, 250/21 Nucleodur C18 Gravity SB, 5 µm particle size, Macherey-Nagel, Germany).

Thermo Dionex Ultimate 3000 HPLC (Thermo Scientific, Germany) coupled to a Bruker amaZon SL mass spectrometer, equipped with a UV detector (254 nm) was used for analytical HPLC-MS. HPLC was operated with a C18 column (EC HPLC column, 100/2 Nucleoshell RP18plus, 2.7 µm particle size, Macherey-Nagel, Germany) as a stationary phase. LCMS-grade distilled MeCN and double distilled H_2O were used as mobile phases containing formic acid (0.1%, v/v).

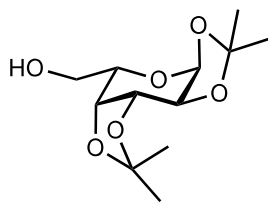
Nuclear magnetic resonance (NMR) spectroscopy was performed on a Bruker Avance III 500 UltraShield spectrometer at 500 MHz (^1H) or 126 MHz (^{13}C). Chemical shifts (δ) are given in parts per million (ppm) and the recorded spectra were referenced to

the respective solvent peak as internal standard^[163]: DMSO-*d*₆ (¹H-NMR δ = 2.50 ppm, ¹³C-NMR δ = 39.52 ppm). Multiplicities were specified as: s (singlet), d (doublet), t (triplet), q (quartet), m (multiplet), br (broad signal). The signals were assigned with the help of ¹H, ¹H-COSY and ¹H, ¹³C-HSQC experiments. Assignment numbering of the fucoside atoms and groups corresponds to the numbering in fucose.

High resolution mass spectra (HRMS) were recorded using an Ultimate 3000 UPLC system coupled to a Q Exactive Focus Orbitrap spectrometer with HESI source (Thermo Fisher, Dreieich, Germany). The UPLC was operated with a C18 column (EC 150/2 Nucleodur C18 Pyramid, 3 μm particle size, Macherey-Nagel, Germany).

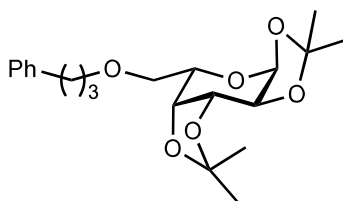
Procedures and Characterizations

1,2:3,4-bis-*O*-isopropylidene-*L*-galactopyranoside (**14**)



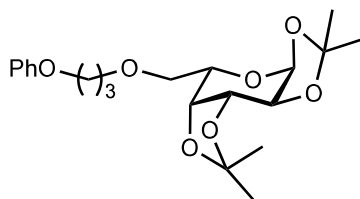
The reaction was performed in analogy to Koch *et al.*^[156] *L*-galactose (503 mg, 2.79 mmol, 1 eq) was suspended in a solution of CuSO₄ (1.18 g, 7.4 mmol, 2.5 eq) in acetone (*c* = 0.2 M) and H₂SO₄ (conc., 1 drop) was added. The mixture was stirred overnight, filtered and the residue was washed with acetone. The filtrate was neutralized with KOH, filtered, and the solvent was evaporated under reduce pressure. The crude yellow oil of **14** was directly used without further purifications (712 mg, 2.74 mmol, 98%). ¹H NMR (500 MHz, DMSO-*d*₆) δ 5.43 (d, *J* = 5.0 Hz, 1H, H-1), 4.68 (dd, *J* = 6.5, 5.1 Hz, 1H, OH-6), 4.56 (dd, *J* = 8.0, 2.3 Hz, 1H, H-3), 4.31 (dd, *J* = 5.1, 2.3 Hz, 1H, H-2), 4.23 (dd, *J* = 8.0, 1.8 Hz, 1H, H-4), 3.68 (td, *J* = 6.5, 1.8 Hz, 1H, H-5) 3.52 – 3.37 (m, 2H, H-6), 1.44 (s, 3H, CH₃), 1.33 (s, 3H, CH₃), 1.27 (d, *J* = 1.7 Hz, 6H, 2x CH₃). ¹³C NMR (126 MHz, DMSO) δ 108.03 (CCH₃), 107.63 (CCH₃), 95.63 (C-1), 70.12 (C-4), 69.97 (C-2, C-3), 68.21 (C-5), 59.94 (C-6), 25.96 (CH₃), 25.86 (CH₃), 24.89 (CH₃), 24.22 (CH₃). LR-MS [C₁₂H₂₀O₆+Na]⁺: calcd. 283.12, found 283.17.

1,2:3,4-Bis-O-Isopropylidene-6-O-(3-Phenylpropyl)- α -L-Galactopyranoside (15a)



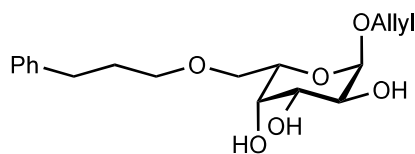
The reaction was performed in analogy to Koch *et al.*^[156] 1,2:3,4-bis-O-isopropylidene-L-galactopyranoside (**14**) (138 mg, 0.53 mmol, 1 eq.) was dissolved in a 1:1 mixture of toluene (1 mL) and 1-bromo-3-phenylpropane (1 mL, 6.58 mmol, 12.4 eq.) and powdered KOH was added (440 mg). The reaction mixture was stirred at 100 °C overnight, the mixture was cooled to room temperature and 10 mL of ice-water was added until the salts dissolved. The solution was diluted in EtOAc (100 mL) and the resulting aqueous layer was extracted with EtOAc (50 mL, 3x). The combined organic phase was dried over anhydrous Na₂SO₄, filtrated, and concentrated under reduced pressure. The crude product was purified by flash-chromatography (petroleum ether/EtOAc 9:1) and the product **15a** was obtained as a clear solution (107 mg, 0.28 mmol, 53%). ¹H NMR (500 MHz, DMSO-*d*₆) δ 7.27 (m, 2H, 2x ArCH), 7.22 – 7.14 (m, 3H, ArCH), 5.45 (d, *J* = 5.0 Hz, 1H, H-1), 4.58 (dd, *J* = 7.9, 2.3 Hz, 1H, H-3), 4.33 (dd, *J* = 5.0, 2.3 Hz, 1H, H-2), 4.22 (dd, *J* = 7.9, 1.9 Hz, 1H, H-4), 3.83 (ddd, *J* = 7.0, 5.2, 1.9 Hz, 1H, H-5), 3.52 (dd, *J* = 10.3, 5.3 Hz, 1H, 1x H-6), 3.45 – 3.34 (m, 3H, 1x H-6, PhCH₂CH₂CH₂), 2.61 (dd, *J* = 8.6, 6.7 Hz, 2H, PhCH₂CH₂CH₂), 1.78 (dq, *J* = 8.4, 6.5 Hz, 2H, PhCH₂CH₂CH₂), 1.42 (s, 3H CH₃), 1.34 (s, 3H, CH₃), 1.28 (s, 6H, 2x CH₃). ¹³C NMR (126 MHz, DMSO) δ 141.74 (ArC), 128.34 (2x ArCH), 128.25 (2x ArCH), 125.69 (ArCH), 108.23 (C(CH₃)₂), 107.70 (C(CH₃)₂), 95.65 (C-1), 70.50 (C-4), 69.99 (C-3), 69.79 (C-2), 69.48 (C-6), 69.04 (PhCH₂CH₂CH₂), 66.22 (C-5), 31.57 (PhCH₂CH₂CH₂), 30.88 (PhCH₂CH₂CH₂), 25.93 (CH₃), 25.86 (CH₃), 24.83 (CH₃), 24.26 (CH₃). HR-MS [C₂₁H₃₀O₆ + NH₄]⁺ calcd. 396.2381, found: 396.2377.

6-O-(3-Phenoxypropyl)- α -L-Galactopyranoside (15b)



The reaction was performed in analogy to Koch *et al.*^[156] 1,2:3,4-bis-O-isopropylidene-L-galactopyranoside (**14**) (159 mg, 0.61 mmol, 1 eq.) was dissolved in a 1:1 mixture of toluene (1 mL) and 3-phenoxypropyl bromide (1 mL, 6.58 mmol, 12.4 eq.) and powdered KOH was added (188 mg, 5.5 eq.). The reaction mixture was stirred at 100 °C overnight, the mixture was cooled to room temperature and 10 mL of ice-water was added until the salts dissolved. The solution was diluted in EtOAc (100 mL) and the resulting aqueous layer was extracted with EtOAc (50 mL, 3x). The combined organic phase was dried over anhydrous Na₂SO₄, filtrated, and concentrated under reduced pressure. The crude product was purified by flash-chromatography (petroleum ether/EtOAc 9:1) and the product **15a** was obtained as a brown oil (146 mg, 0.37 mmol, 61%). ¹H NMR (500 MHz, DMSO-*d*₆) δ 7.30 – 7.23 (m, 2H, 2x ArCH), 6.94 – 6.88 (m, 3H, 2x ArCH), 5.44 (d, *J* = 5.0 Hz, 1H, H-1), 4.56 (dd, *J* = 7.9, 2.3 Hz, 1H, H-3), 4.32 (dd, *J* = 5.1, 2.3 Hz, 1H, H-2), 4.20 (dd, *J* = 7.9, 1.9 Hz, 1H, H-4), 4.01 (t, *J* = 6.4 Hz, 2H, PhOCH₂CH₂CH₂), 3.83 (ddd, *J* = 6.9, 5.0, 1.9 Hz, 1H, H-5), 3.62 – 3.50 (m, 3H, 1x H-6, PhOCH₂CH₂CH₂), 3.41 (dd, *J* = 10.4, 7.0 Hz, 1H, 1x H-6), 1.93 (p, *J* = 6.4 Hz, 2H, PhOCH₂CH₂CH₂), 1.41 (s, 3H, CH₃), 1.32 (s, 3H, 2x CH₃), 1.27 (s, 3H, 2x CH₃), 1.25 (s, 3H, CH₃). ¹³C NMR (126 MHz, DMSO) δ 158.57 (ArC), 129.46 (2x ArCH), 120.44 (ArCH), 114.40 (2x ArCH), 108.25 (C(CH₃)₂), 107.71 (C(CH₃)₂), 95.66 (C-1), 70.50 (C-4), 70.00 (C-3), 69.77 (C-2), 69.18 (C-6), 67.00 (PhOCH₂CH₂CH₂), 66.28 (C-5), 64.36 (PhOCH₂CH₂CH₂), 29.05 (PhOCH₂CH₂CH₂), 25.92 (CH₃), 25.86 (CH₃), 24.83 (CH₃), 24.22 (CH₃). HR-MS [C₂₁H₃₀O₇ + NH₄]⁺ calcd. 412.2330, found: 412.2326.

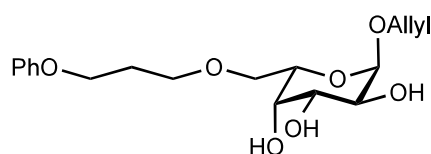
6-O-Phenylpropyl-α-Allyl-L-Galactopyranoside (**16a**)



The reaction was performed in analogy to Koch *et al.*^[156] The isopropylidene protected L-galactopyranoside **15a** (102 mg, 0.27 mmol, 1 eq.) was dissolved in allyl alcohol and Dowex® 50WX4-50 (126 mg) was added. The mixture was refluxed overnight, cooled to room temperature, filtered, and the pure product precipitated from the filtrate. The supernatant was concentrated under reduced pressure and the remaining residue was purified by flash-chromatography (CH₂Cl₂/MeOH 95:5). The anomeric mixture was obtained as a light-brown solid (24.4 mg, 0.08 mmol, 30%) that was further purified by

preparative HPLC (CH₃CN/H₂O gradient: 5–75% CH₃CN) to yield the final α -product **16a** (5.6 mg, 0.019, 7%). ¹H NMR (500 MHz, DMSO-*d*₆) δ 7.30 – 7.24 (m, 2H, 2x ArCH), 7.22 – 7.13 (m, 3H, 3x ArCH), 5.98 – 5.87 (m, 1H, OCH₂CHCH₂), 5.30 (dd, *J* = 17.2, 1.9 Hz, 1H, 1x OCH₂CHCH₂), 5.12 (dd, *J* = 10.5, 1.8 Hz, 1H, 1x OCH₂CHCH₂), 4.69 (d, *J* = 3.2 Hz, 1H, H-1), 4.59 (d, *J* = 5.0 Hz, 1H, OH-3), 4.55 (d, *J* = 6.0 Hz, 1H, OH-2), 4.46 (d, *J* = 4.3 Hz, 1H, OH-4), 4.12 – 4.05 (m, 1H, 1x OCH₂CHCH₂), 3.97 – 3.90 (m, 1H, 1x OCH₂CHCH₂), 3.78 – 3.73 (m, 1H, H-5), 3.67 (q, *J* = 2.5, 1.4 Hz, 1H, H-4), 3.63 – 3.54 (m, 2H, H-2, H-3), 3.51 (dd, *J* = 10.2, 4.8 Hz, 1H, 1x H-6), 3.45 – 3.41 (m, 1H, 1x H-6), 3.41 – 3.35 (m, 2H, PhCH₂CH₂CH₂), 2.64 – 2.57 (m, PhCH₂CH₂CH₂), 1.82 – 1.72 (m, 2H, PhCH₂CH₂CH₂). ¹³C NMR (126 MHz, DMSO) δ 141.80 (ArC), 135.01 (OCH₂CHCH₂), 128.30 (2x ArCH), 128.27 (2x ArCH), 125.69 (ArCH), 116.21 (OCH₂CHCH₂), 98.38 (C-1), 70.04 (C-6), 69.53 (PhCH₂CH₂CH₂), 69.44 (C-5), 69.40 (C-3?), 69.29 (C-4), 68.24 (C-2?), 67.33 (OCH₂CHCH₂), 31.66 (PhCH₂CH₂CH₂), 31.11 (PhCH₂CH₂CH₂). HR-MS [C₁₈H₂₆O₆ + Na]⁺ calcd. 361.1622, found: 361.1616.

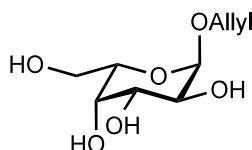
6-O-Phenoxypropyl- α -Allyl-L-Galactopyranoside (16b)



The reaction was performed in analogy to Koch *et al.*^[156] The isopropylidene protected L-galactopyranoside **15a** (140 mg, 0.35 mmol, 1 eq.) was dissolved in allyl alcohol and Dowex® 50WX4-50 (272 mg) was added. The mixture was refluxed overnight, cooled to room temperature, filtered, and the pure product precipitated from the filtrate. The supernatant was concentrated under reduced pressure and the remaining residue was purified by flash-chromatography (CH₂Cl₂/MeOH 95:5). The anomeric mixture was obtained as a light-brown solid (64.4 mg, 0.20 mmol, 57%) that was further purified by preparative HPLC (CH₃CN/H₂O gradient: 5–75% CH₃CN) to yield the final α -product **16b** (27 mg, 0.09 mmol, 26%). ¹H NMR (500 MHz, DMSO-*d*₆) δ 7.33 – 7.22 (m, 2H, 2x ArCH), 7.00 – 6.83 (m, 3H, 3x ArCH), 5.93 – 5.84 (m, 1H, OCH₂CHCH₂), 5.31 – 5.26 (m, 1H, 1x OCH₂CHCH₂), 5.13 – 5.08 (m, 1H, 1x OCH₂CHCH₂), 4.68 (d, *J* = 3.3 Hz, 1H, H-1), 4.57 (d, *J* = 5.2 Hz, 1H, OH-3), 4.53 (d, *J* = 6.1 Hz, 1H, OH-2), 4.46 (d, *J* = 4.4 Hz, 1H, OH-4), 4.11 – 4.04 (m, 2H, 1x OCH₂CHCH₂), 4.01 (t, *J* = 6.4 Hz, 2H, PhOCH₂CH₂CH₂), 3.95 – 3.87 (m, 1H, 1x OCH₂CHCH₂), 3.78 – 3.73 (m, 1H, H-5),

3.68 – 3.64 (m, 1H, H-4), 3.62 – 3.50 (m, 5H, H-2, H-3, 1x H-6, PhOCH₂CH₂CH₂), 3.48 – 3.42 (m, 1H, 1x H-6), 1.98 – 1.90 (m, 1H, PhOCH₂CH₂CH₂). ¹³C NMR (126 MHz, DMSO) δ 159.02 (ArC), 135.41 (OCH₂CHCH₂), 129.94 (2x ArCH), 120.91 (ArCH), 116.66 (OCH₂CHCH₂), 114.85 (2x ArCH), 98.81 (C-1), 70.68 (C-6), 69.88 (C-5, C-3), 69.76 (C-4), 68.70 (C-2), 67.74 (OCH₂CHCH₂), 67.49 (PhOCH₂CH₂CH₂), 64.84 (PhOCH₂CH₂CH₂), 29.67 (PhOCH₂CH₂CH₂). HR-MS [C₁₈H₂₆O₇ + Na]⁺ calcd. 377.1571, found: 377.1565.

Allyl-α-L-Galactopyranoside (17)



The FISCHER glycosylation was performed in analogy to Page *et al.*^[164] Allyl alcohol (6 mL) was cooled to 0 °C, and acetyl chloride (0.5 mL, 7 mmol, 2.5 eq.) was added dropwise. Then, L-galactose (0.5 g, 2.76 mmol, 1 eq.) was added and the solution was stirred for 4 h at 70 °C before reducing the temperature to 40 °C overnight. The next day, the reaction mixture was cooled to room temperature and the solvents were removed under reduced pressure. The mixture was purified by silica gel column chromatography (CH₂Cl₂/MeOH 9:1). The collected sugar was recrystallized twice from abs. EtOH resulting in the final product **17** (0.035 g, 7%). ¹H NMR (500 MHz, H₂O-d₂) δ 6.04 – 5.94 (m, 1H, (OCH₂CHCH₂), 5.37 (s, 1H), 5.32 – 5.23 (m, 1H, 1x OCH₂CHCH₂), 5.00 (d, *J* = 3.5 Hz, 1H, H-1), 4.27 – 4.20 (m, 1H, OCH₂CHCH₂), 4.12 – 4.04 (m, 1H, 1x OCH₂CHCH₂), 3.99 – 3.97 (m, 1H, H-4), 3.95 (t, *J* = 6.2 Hz, 1H, H-5), 3.89 – 3.79 (m, 2H, H-2, H-3), 3.75 – 3.72 (m, 2H, H-6). ¹³C NMR (126 MHz, H₂O-d₂) δ 133.55 (OCH₂CHCH₂), 118.16 (OCH₂CHCH₂), 97.48 (C-1), 70.93 (C-5), 69.45 (C-3?), 69.21 (C-4), 68.47 (OCH₂CHCH₂), 68.15 (C-2?), 61.15 (C-6). HR-MS [C₉H₁₆O₆ + Na]⁺ calcd. 243.0839, found: 243.0836.

Biophysical Validation

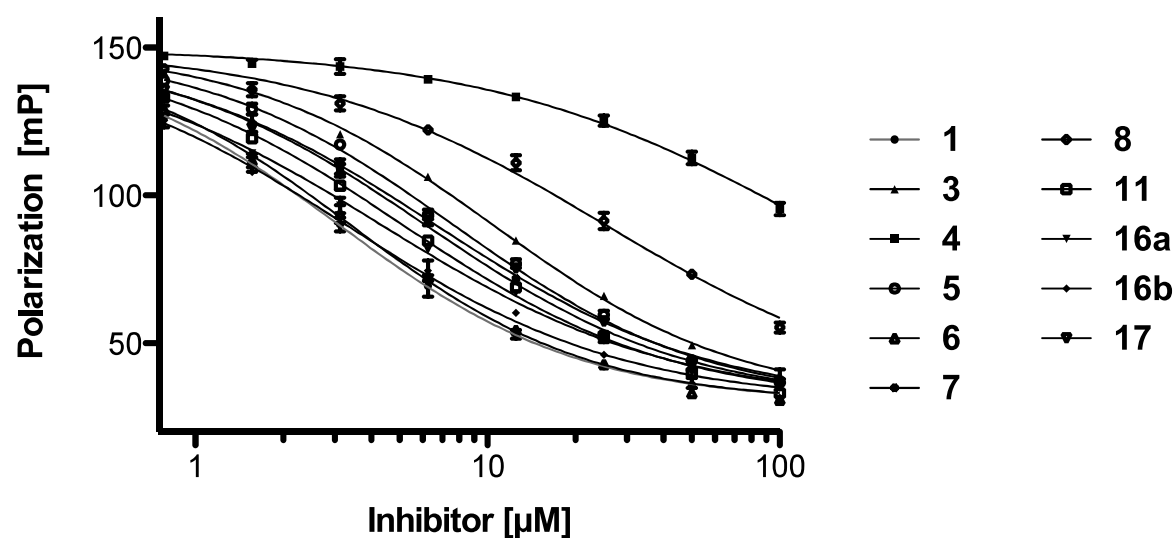


Figure S10: Results of the best *in-house* library BambL inhibitors together with the three synthesized galactosides **16a–16b** and **17** measured in the competitive binding assay based on fluorescence polarization. The positive control Me- α -L-Fuc (**1**) is highlighted in red.

Docking

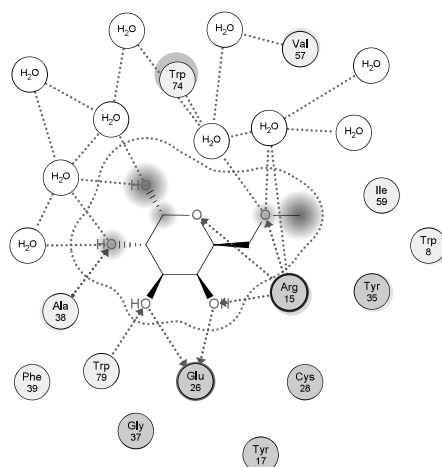


Figure S11: 2D ligand interaction map 6-O-methyl- α -L-galactoside showing a possible hydrogen bond between the ether function and Arg15 of BamBL visualized by MOE (purple = polar, purple red = acidic, purple blue = basic, green = greasy, blue arrow = backbone acceptor/ donor, green arrow = sidechain acceptor/ donor, blue = ligand exposure, light blue = receptor exposure, grey line = solvent contract).

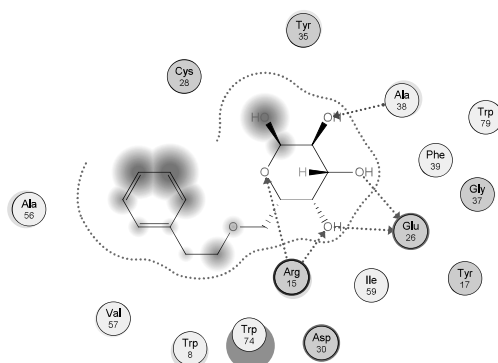


Figure S12: 2D ligand interaction map of the second best binding pose of the designed compound **12** in complex with BamBL (purple = polar, purple red = acidic, purple blue = basic, green = greasy, blue arrow = backbone acceptor/ donor, green arrow = sidechain acceptor/ donor, blue = ligand exposure, light blue = receptor exposure, grey line = solvent contract).

X-ray Crystallography

Table S9: X-ray data collection and refinement statistics of Bambl-3, Bambl-6, Bambl-11, Bambl-16b, Bambl-17 complexes.

Crystallization	
Protein: Bambl (8–9 mg mL ⁻¹) in 20 mM tris-HCl pH = 7.5, 150 mM NaCl; + 1 mM 17, 1.25% DMSO; 1 + 1 µL 100 mM NaOAc pH = 4.6, 26% PEG 2K MME; cryoprotectant: same buffer + 28% PEG 2K MME	
Data collection	
Beamtime: 10.12.2021 at Proxima 2, Synchrotron SOLEIL, crystal form: broken rod	
Wavelength [Å]	0.9801
Resolution range [Å]	45.48 –1.4 (1.42–1.4)
Space group	P 21 21 21
Unit cell dimensions	45.178 47.392 98.835 90.00 90.00 90.00
Total number of observations	538279 (23244)
Total number unique	42662 (2082)
Multiplicity	12.6 (11.2)
Completeness (%)	100.0 (99.9)
Mean I/sigma(I)	29.0 (6.5)
Wilson B-factor	12.68
R-merge (within I+/I-)	0.049 (0.302)
R-meas (within I+/I-)	0.053 (0.334)
R-pim (within I+/I-)	0.020 (0.139)
CC1/2	0.999 (0.966)
Refinement	
Reflections used in refinement	42592 (4180)
Reflections used for R-free	2083 (198)
R-work	0.153
R-free	0.180
Rms BondLength [Å]	0.0153
Rms BondAngle [°]	1.77
Rms ChirVolume [°]	0.0839
Ramachandran favored [%]	97.65
Ramachandran allowed [%]	2.35
Ramachandran outliers [%]	0.00
Rotamer outliers [%]	0.00
Clashscore	1.80
Total number of non-hydrogen atoms...	2471
...macromolecules	2015
...ligands	45
...solvent	411

Protein residues	261
Average B-factor... [\AA^2]	14.52
...macromolecules [\AA^2]	12.5
...ligands [\AA^2]	18.63
...solvent [\AA^2]	23.99

Statistics for the highest-resolution shell are shown in parentheses.

Crystallization

Protein: BamBL (8–9 mg mL⁻¹) in 20 mM tris-HCl pH = 7.5, 150 mM NaCl; + 2.5 mM 11; 1 + 1 μ L 200 mM (NH₄)₂SO₄, 30% PEG 2K MME (= SG1 1-45), 100 mM NaOAc pH = 4.6

Data collection

Beamtime: 09.06.2021 at Proxima 2, Synchrotron SOLEIL, **crystal form:** plates

Wavelength [\AA]	0.9801
Resolution range [\AA]	37.75–1.65 (1.68–1.65)
Space group	C 1 2 1
Unit cell dimensions	76.059 43.92 74.315 90.00 113.579 90.00
Total number of observations	184382 (9262)
Total number unique	27257 (1331)
Multiplicity	6.8 (7.0)
Completeness (%)	100.0 (100.0)
Mean I/sigma(I)	8.4 (3.1)
Wilson B-factor	15.83
R-merge (within I+/I-)	0.141 (0.526)
R-meas (within I+/I-)	0.168 (0.621)
R-pim (within I+/I-)	0.090 (0.328)
CC1/2	0.991 (0.868)

Refinement

Reflections used in refinement	25974 (2678)
Reflections used for R-free	1282 (122)
R-work	0.1665
R-free	0.2173
Rms BondLength [\AA]	0.016
Rms BondAngle [$^\circ$]	1.71
Rms ChirVolume [$^\circ$]	0.0813
Ramachandran favored [%]	96.84
Ramachandran allowed [%]	100
Ramachandran outliers [%]	0.00
Rotamer outliers [%]	1.06
Clashscore	3.40
Total number of non-hydrogen atoms...	2383
...macromolecules	1961

...ligands	120
...solvent	302
Protein residues	261
Average B-factor... [Å ²]	18.67
...macromolecules [Å ²]	17.16
...ligands [Å ²]	21.37
...solvent [Å ²]	27.45

Statistics for the highest-resolution shell are shown in parentheses.

Crystallization

Protein: BamBL (8–9 mg mL⁻¹) in 20 mM tris-HCl pH = 7.5, 150 mM NaCl; + 2.5 mM **3**, 5% DMSO; 1 + 1 µL 4.3 M NaCl, 100 mM HEPES pH = 7.5 (= SG1 2-48)

Data collection

Beamtime: 09.06.2021 at Proxima 2, Synchrotron SOLEIL, **crystal form:** plates

Wavelength [Å]	0.9801
Resolution range [Å]	43.84–1.95 (2.00–1.95)
Space group	C 2 2 21
Unit cell dimensions	44.268 76.77 131.527 90.00 90.00 90.00
Total number of observations	120955 (8107)
Total number unique	16777 (1145)
Multiplicity	7.2 (7.1)
Completeness (%)	100.0 (100.0)
Mean I/sigma(I)	10.7 (3.4)
Wilson B-factor	17.4
R-merge (within I+/I-)	0.119 (0.608)
R-meas (within I+/I-)	0.139 (0.711)
R-pim (within I+/I-)	0.070 (0.366)
CC1/2	0.997 (0.905)

Refinement

Reflections used in refinement	15943 (1651)
Reflections used for R-free	793 (62)
R-work	0.1713
R-free	0.2411
Rms BondLength [Å]	0.0153
Rms BondAngle [°]	1.729
Rms ChirVolume [°]	0.0903
Ramachandran favored [%]	97.25
Ramachandran allowed [%]	100.0
Ramachandran outliers [%]	0.00
Rotamer outliers [%]	0.00
Clashscore	4.48

Total number of non-hydrogen atoms...	2311
...macromolecules	1947
...ligands	118
...solvent	246
Protein residues	261
Average B-factor... [Å ²]	19.26
...macromolecules [Å ²]	18.37
...ligands [Å ²]	19.37
...solvent [Å ²]	26.24

Statistics for the highest-resolution shell are shown in parentheses.

Crystallisation

Protein: BamBL (8–9 mg mL⁻¹) in 20 mM tris-HCl pH = 7.5, 150 mM NaCl; + 2.5 mM **6**; 1 + 1 µL 24% PEG 4K, 10% *i*PrOH (= SG1 1-6)

Data collection

BamBL + **6**: **Beamtime**: 09.06.2021 at Proxima 2, Synchrotron SOLEIL, **crystal form**: rod from rod cluster

Wavelength [Å]	0.980118
Resolution range [Å]	43.22–1.83 (1.87–1.83)
Space group	P 1 21 1
Unit cell dimensions	50.39 84.096 62.127 90.00 91.239 90.00
Total number of observations	208611 (12979)
Total number unique	44741 (2721)
Multiplicity	4.7 (4.8)
Completeness (%)	98.1 (97.2)
Mean I/sigma(I)	10.7 (3.2)
Wilson B-factor	15.9
R-merge (within I+/I-)	0.087 (0.491)
R-meas (within I+/I-)	0.112 (0.653)
R-pim (within I+/I-)	0.07 (0.398)
CC1/2	0.996 (0.809)

Refinement

Reflections used in refinement	42536 (4378)
Reflections used for R-free	2217 (190)
R-work	0.164
R-free	0.197
Rms BondLength [Å]	0.0161
Rms BondAngle [°]	1.6970
Rms ChirVolume [°]	97.64
Ramachandran favored [%]	100
Ramachandran allowed [%]	97.6
Ramachandran outliers [%]	0.00

rotamer outliers [%]	0.54
Clashscore	3.29
Total number of non-hydrogen atoms...	4560
...macromolecules	3905
...ligands	214
...solvent	441
Protein residues	521
Average B-factor... [Å ²]	19.12
...macromolecules [Å ²]	17.99
...ligands [Å ²]	22.57
...solvent [Å ²]	27.34

Statistics for the highest-resolution shell are shown in parentheses.

Crystallisation

BambL + **16b**: **Beamtime**: 10.12.2021 at Proxima 2, Synchrotron SOLEIL, **crystal form**: parallelepiped

Data collection

Protein: BambL (8–9 mg/mL) in 20 mM tris-HCl pH = 7.5, 150 mM NaCl; + 2.5 mM **16b**, 5% DMSO; 1 + 1 µL 20% PEG 10K, 100 mM HEPES pH = 7.5, cryoprotectant: + 10% ethyleneglycol

Wavelength [Å]	0.9801
Resolution range [Å]	47.58–1.35 (1.37–1.35)
Space group	P 21 21 21
Unit cell dimensions	45.56 47.585 97.211 90.00 90.00 90.00
Total number of observations	595648 (30811)
Total number unique	47278 (2381)
Multiplicity	12.6 (12.9)
Completeness (%)	100.0 (100.0)
Mean I/sigma(I)	21.9 (3.5)
Wilson B-factor	15.3
R-merge (within I+/I-)	0.058 (0.645)
R-meas (within I+/I-)	0.063 (0.705)
R-pim (within I+/I-)	0.025 (0.279)
CC1/2	0.999 (0.937)

Refinement

Reflections used in refinement	44831 (4619)
Reflections used for R-free	2377 (251)
R-work	0.142
R-free	0.188
Rms BondLength [Å]	0.016
Rms BondAngle [°]	1.75
Rms ChirVolume [°]	0.0902
Ramachandran favored [%]	96.85

Ramachandran allowed [%]	100.0
Ramachandran outliers [%]	0.00
Rotamer outliers [%]	0.50
Clashscore	2.34
Total number of non-hydrogen atoms...	2369
...macromolecules	2012
...ligands	25
...solvent	332
Protein residues	260
Average B-factor... [Å ²]	18.92
...macromolecules [Å ²]	17.20
...ligands [Å ²]	24.37
...solvent [Å ²]	28.96

Statistics for the highest-resolution shell are shown in parentheses.

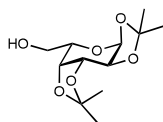
Table S10: 2D protein-ligand interactions maps of the different BambL-3, BambL-6, BambL-11, BambL-17, BambL-16b complexes in the *inter/intra* carbohydrate binding sites. (purple = ligand bond, brown = non-ligand bond, green = hydrogen bond and its length, reddish-brown eyelash similar bows = non-ligand residues involved in hydrophobic contact(s), reddish-brown lines on atoms = corresponding atoms involved in hydrophobic contact(s).

Compound	<i>Intra</i> Carbohydrate Binding Site	<i>Inter</i> Carbohydrate Binding Site
α - Fucopyranosyl benzamide (3)		

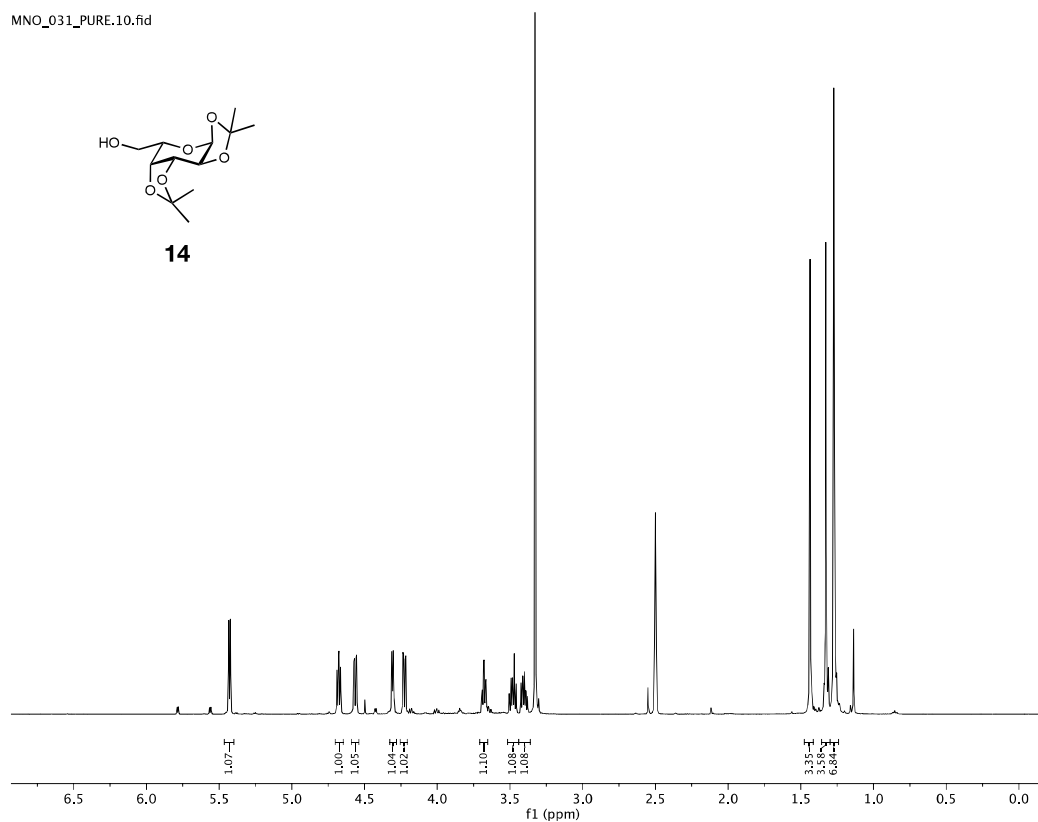
Compound	Intra Carbohydrate Binding Site	Inter Carbohydrate Binding Site
β - Fucopyranosyl -methyl 2- thiophene- sulfonamide (6)		
1-C-(3- phenylthio)- propyl- β -D- mannopyrano- side (12)		
Allyl- α -L- galactopyrano- -side (17)		No ligand observed
6-O-Phenoxy- α -allyl-L- galactopyrano- -side (16b)		No ligand observed

Spectra

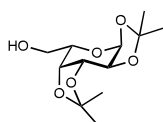
MNO_031_PURE.10.fid



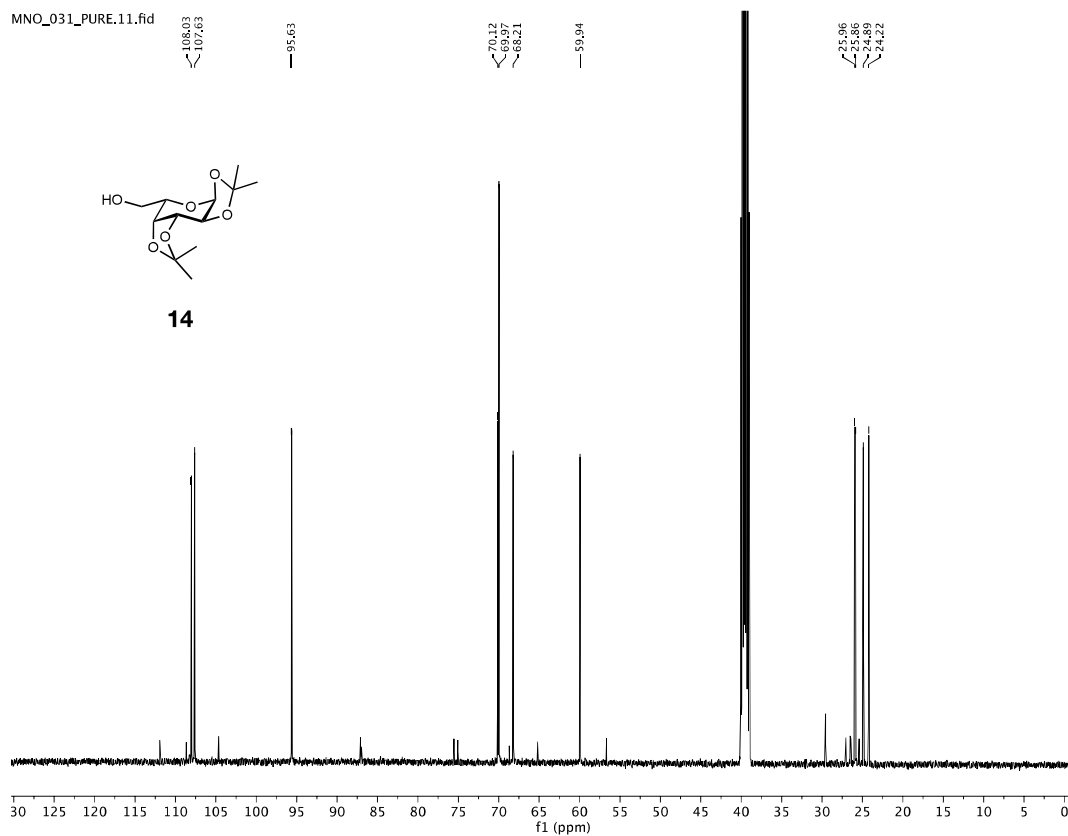
14



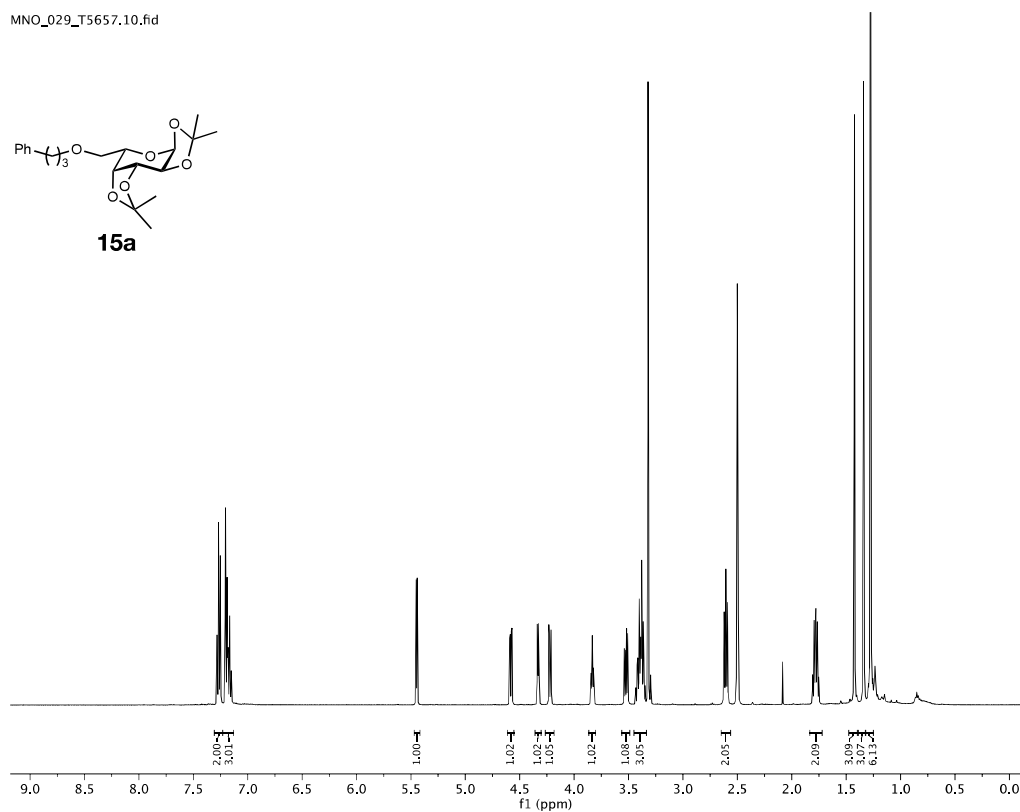
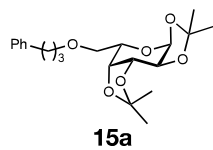
MNO_031_PURE.11.fid



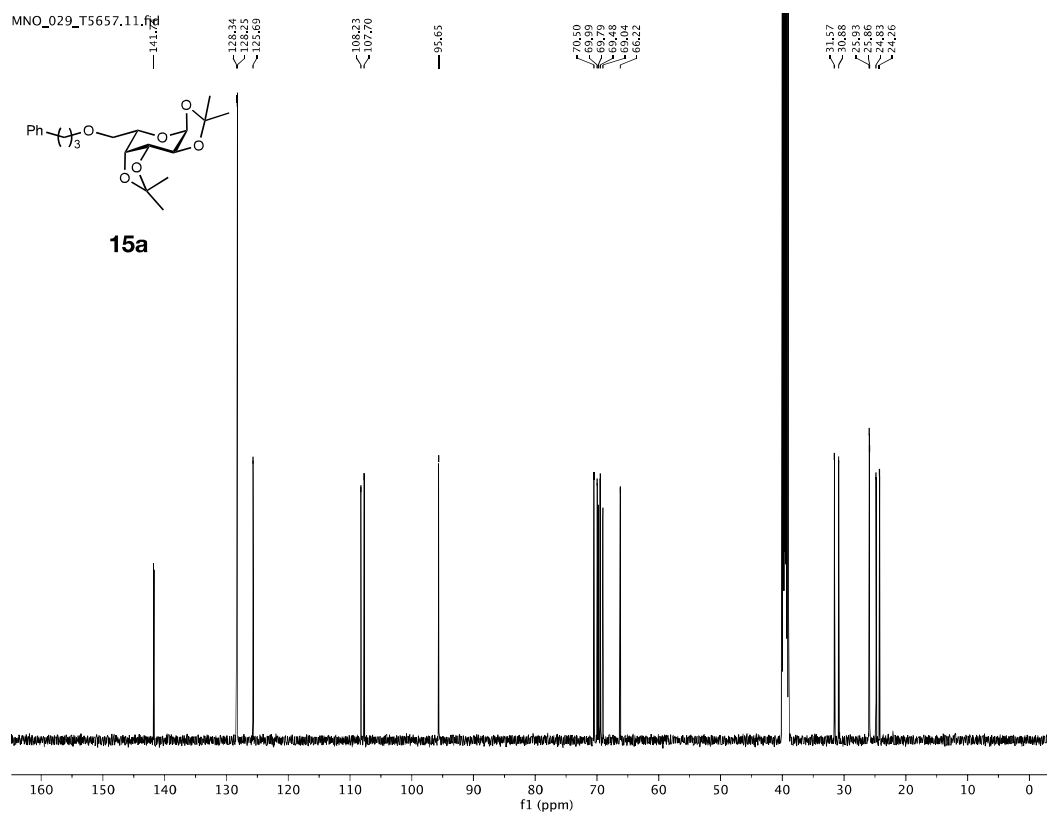
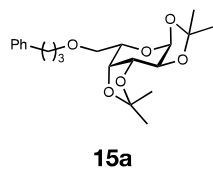
14



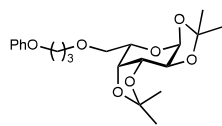
MNO_029_T5657.10.fhd



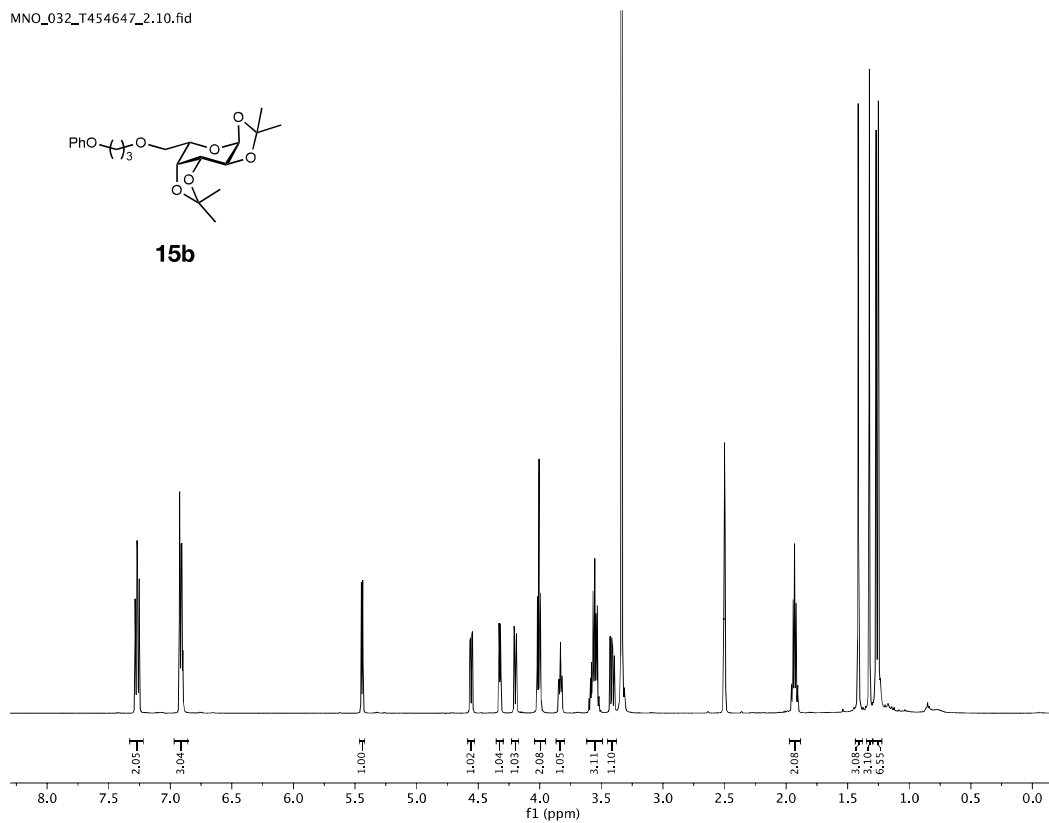
MNO_029_T5657.11.fhd



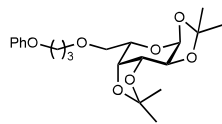
MNO_032_T454647_2.10.fid



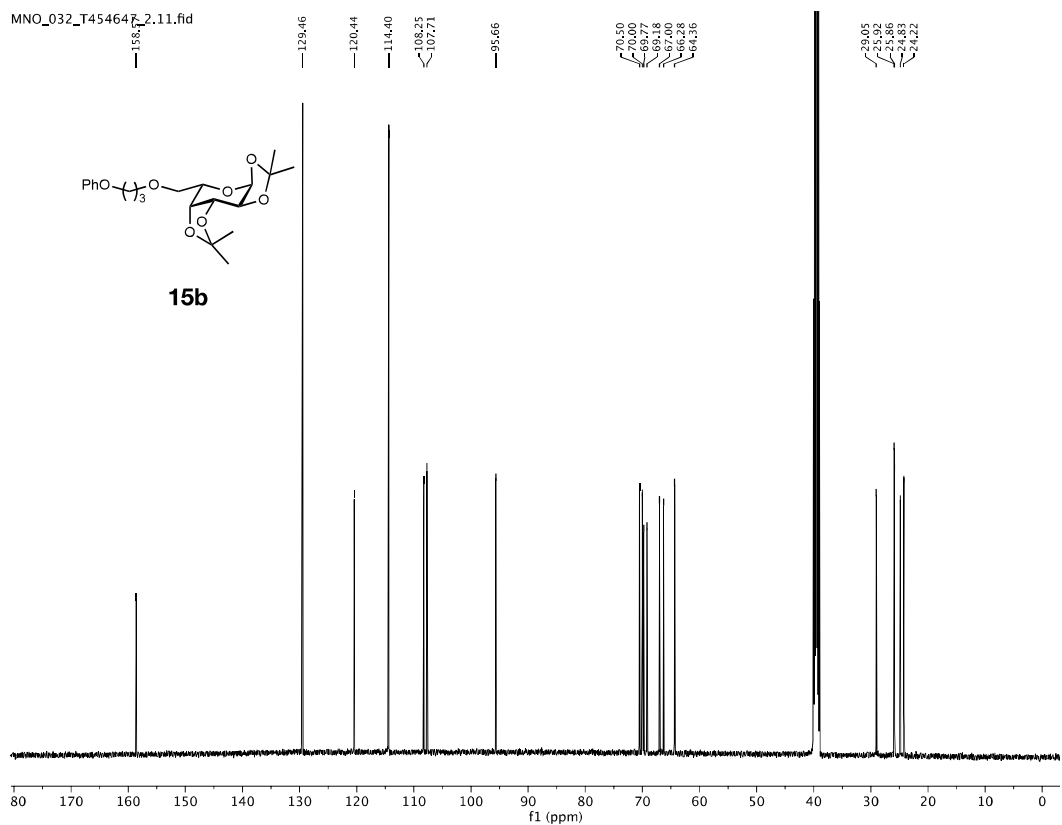
15b



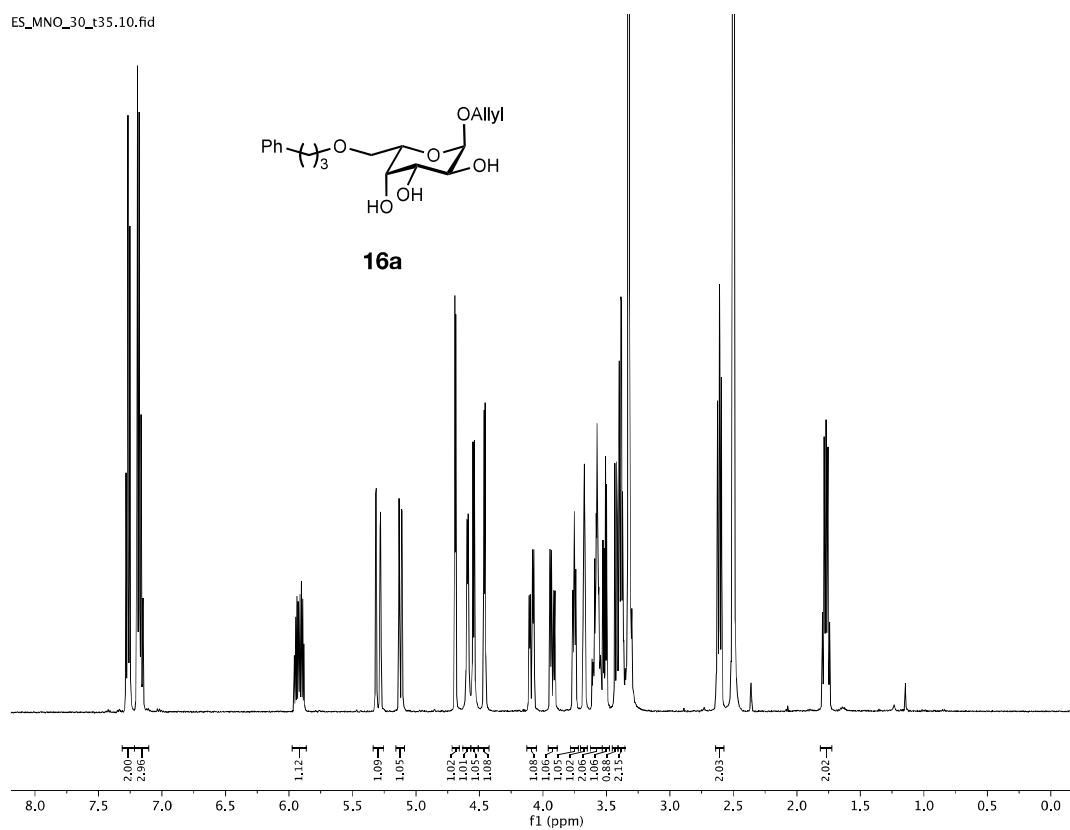
MNO_032_T454647_2.11.fid



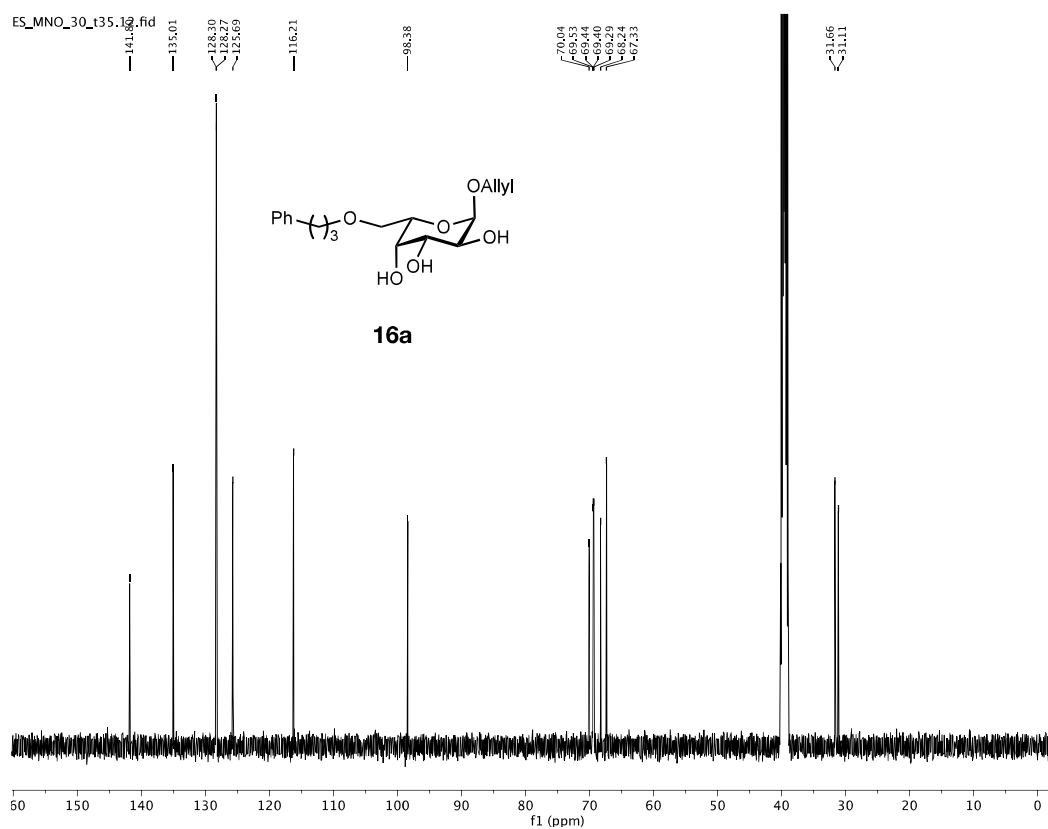
15b



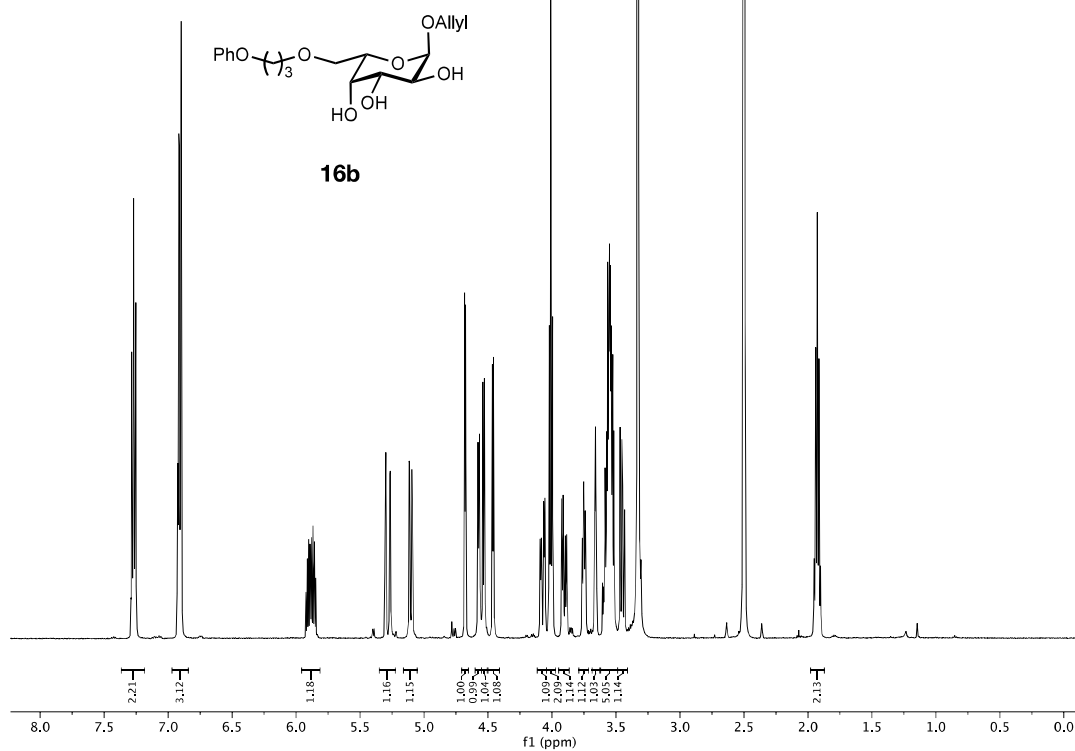
ES_MNO_30_t35.10.fid



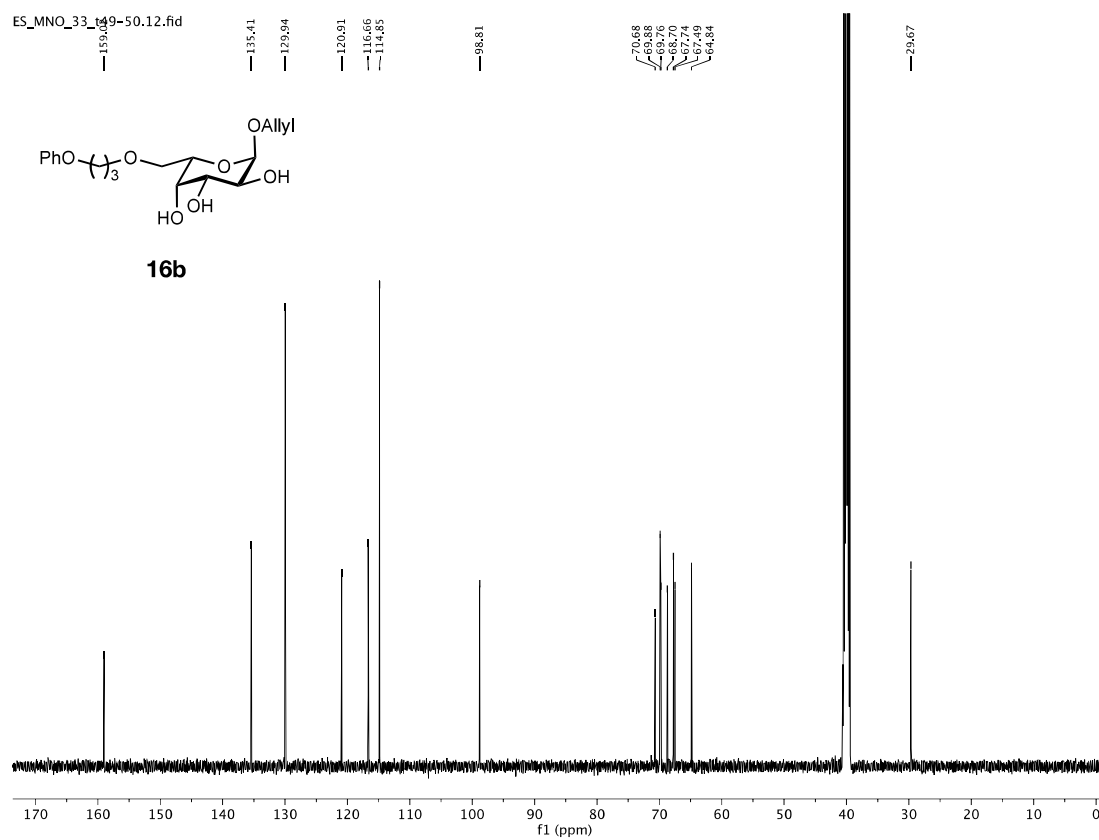
ES_MNO_30_t35.10.fid



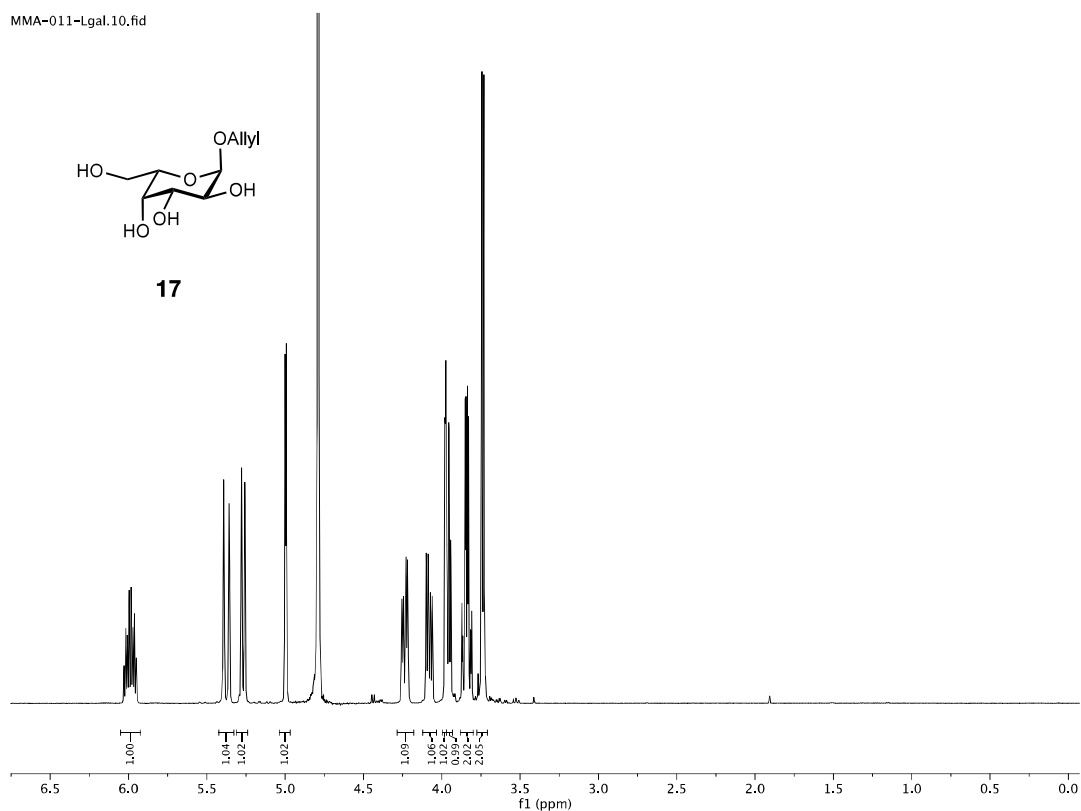
ES_MNO_33_t49-50.10.fid



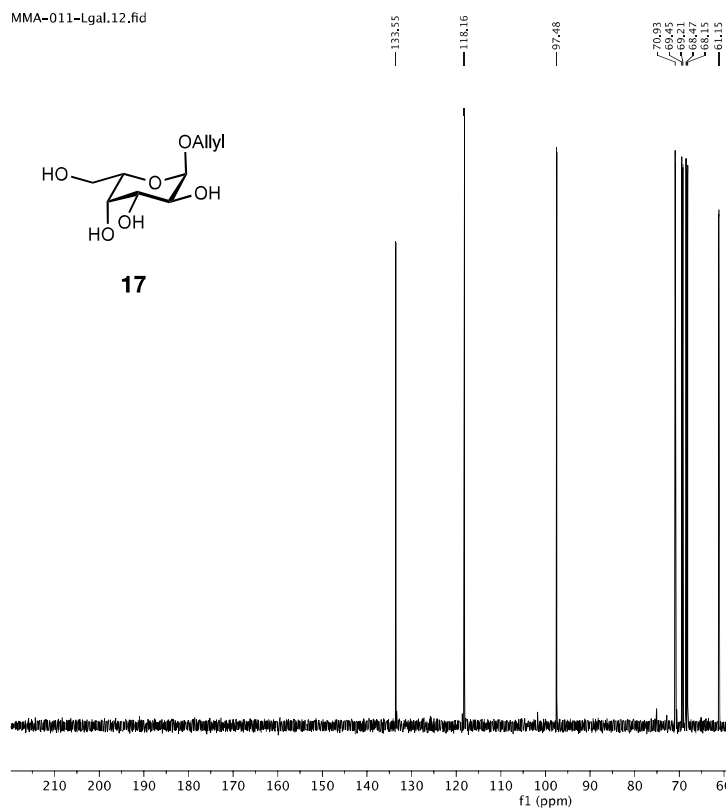
ES_MNO_33_t49-50.12.fid



MMA-011-Lgal.10.fid



MMA-011-Lgal.12.fid



References

- [1] E. D. Brown, G. D. Wright, *Nature* **2016**, 529, 336–343.
- [2] M. Hutchings, A. Truman, B. Wilkinson, *Curr. Opin. Microbiol.* **2019**, 51, 72–80.
- [3] D. C. Domínguez, S. M. Meza-Rodriguez, *Pharm. Pers. Care Prod. Waste Manag. Treat. Technol. Emerg. Contam. Micro Pollut.* **2019**, 383–408.
- [4] S. Santajit, N. Indrawattana, *Biomed Res. Int.* **2016**, 2016, 8.
- [5] World Health Organization, “New report calls for urgent action to avert antimicrobial resistance crisis,” can be found under <https://www.who.int/news/item/29-04-2019-new-report-calls-for-urgent-action-to-avert-antimicrobial-resistance-crisis> [03.09.2021], **2019**.
- [6] G. Kappor, S. Saigal, A. Elongavan, *J. Anaesthesiol. Clin. Pharmacol.* **2017**, 33, 300–305.
- [7] J. H. Leitão, *Int. J. Mol. Sci.* **2020**, 21, 1–6.
- [8] S. Divakar, M. Lama, K. Asad U., *Antimicrob. Resist. Infect. Control* **2019**, 3, 1–10.
- [9] T. E. P. Kimkes, M. Heinemann, *FEMS Microbiol. Rev.* **2019**, 44, 106–122.
- [10] J. Böhning, A. Dobbelstein, N. Sulkowski, K. Eilers, A. von Kügelgen, A. K. Tarafder, V. Alva, A. Filloux, T. A. M. Bharat, *bioRxiv* **2022**, DOI 10.1134/s0320972519050051.
- [11] H. C. Flemming, J. Wingender, *Nat. Rev. Microbiol.* **2010**, 8, 623–633.
- [12] V. I. Francis, E. C. Stevenson, S. L. Porter, *FEMS Microbiol. Lett.* **2017**, 364, 1–22.
- [13] C. R. Armbruster, C. K. Lee, J. Parker-Gilham, J. de Anda, A. Xia, K. Zhao, K. Murakami, B. S. Tseng, L. R. Hoffman, F. Jin, et al., *Elife* **2020**, 9, 1–29.
- [14] M. Klausen, A. Aaes-Jørgensen, S. Molin, T. Tolker-Nielsen, *Mol. Microbiol.* **2003**, 50, 61–68.
- [15] S. Helaine, E. Kugelberg, *Trends Microbiol.* **2014**, 22, 417–424.
- [16] T. Rasamiravaka, Q. Labtani, P. Duez, M. El Jaziri, *Biomed Res. Int.* **2015**, 2015, 1–17.
- [17] E. Olivares, S. Badel-Berchoux, C. Provot, G. Prévost, T. Bernardi, F. Jehl, *Front. Microbiol.* **2020**, 10, 2894.
- [18] I. Keren, N. Kaldalu, A. Spoering, Y. Wang, K. Lewis, *FEMS Microbiol. Lett.* **2004**, 230, 13–18.

- [19] T. Stalder, E. Top, *npj Biofilms Microbiomes* **2016**, 2, 16022.
- [20] T. B. Rasmussen, M. Givskov, *Int. J. Med. Microbiol.* **2006**, 296, 149–161.
- [21] S. Elborn, E. Vallieres, *Adv. Genomics Genet.* **2014**, 4, 161–172.
- [22] S. Malhotra, D. Hayes, D. J. Wozniak, *Clin. Microbiol. Rev.* **2019**, 32, e00138–18.
- [23] V. Venkatakrisnan, M. Thaysen-Andersen, S. C. A. Chen, H. Nevalainen, N. H. Packer, *Glycobiology* **2015**, 25, 88–100.
- [24] A. A. Belaaouaj, Kwang Sik Kim, S. D. Shapiro, *Science (80-.)*. **2000**, 289, 1185–1187.
- [25] A. M. Cantin, D. Hartl, M. W. Konstan, J. F. Chmiel, *J. Cyst. Fibros.* **2015**, 14, 419–430.
- [26] J. Laval, A. Ralhan, D. Hartl, *Biol. Chem.* **2016**, 397, 485–496.
- [27] M. A. Khan, Z. S. Ali, N. Sweezey, H. Grasemann, N. Palaniyar, *Genes (Basel)*. **2019**, 10, 183–206.
- [28] M. W. Konstan, M. Berger, *Pediatr. Pulmonol.* **1997**, 24, 137–142.
- [29] E. Faure, K. Kwong, D. Nguyen, *Front. Immunol.* **2018**, 9, 1–10.
- [30] 2019 Annual Data Report, “Cystic Fibrosis Foundation Patient Registry,” can be found under <https://www.cff.org/Research/Researcher-Resources/Patient-Registry/2019-Patient-Registry-Annual-Data-Report.pdf> [17.09.2021], **2019**.
- [31] P. W. Whitby, K. B. Carter, K. L. Hatter, J. J. LiPuma, T. L. Stull, *J. Clin. Microbiol.* **2000**, 38, 2962–2965.
- [32] G. Zahariadis, M. H. Levy, J. L. Burns, *Can. J. Infect. Dis.* **2003**, 14, 123–125.
- [33] N. Renders, H. Verbrugh, A. Van Belkum, *Infect. Genet. Evol.* **2001**, 1, 29–39.
- [34] L. Eberl, B. Tümmler, *Int. J. Med. Microbiol.* **2004**, 294, 123–131.
- [35] G. Döring, P. Flume, H. Heijerman, J. S. Elborn, *J. Cyst. Fibros.* **2012**, 11, 461–479.
- [36] M. E. Condren, M. D. Bradshaw, *J. Pediatr. Pharmacol. Ther.* **2013**, 18, 8–13.
- [37] M. Lopes-Pacheco, *Front. Pharmacol.* **2020**, 10, 1–29.
- [38] I. Francolini, C. Vuotto, A. Piozzi, G. Donelli, *Apmis* **2017**, 125, 392–417.
- [39] S. Wagner, R. Sommer, S. Hinsberger, C. Lu, R. W. Hartmann, M. Empting, A. Titz, *J. Med. Chem* **2016**, 59, 5929–5969.
- [40] H. P. Hauber, M. Schulz, A. Pforte, D. Mack, P. Zabel, U. Schumacher, *Int. J. Med. Sci.* **2008**, 5, 371–376.
- [41] M. A. Sartim, S. V. Sampaio, *J. Venom. Anim. Toxins Incl. Trop. Dis.* **2015**, 21,

- 1–11.
- [42] H. Stillmark, **1888**, Inaug. Diss Dorpat.
 - [43] W. C. Boyd, E. Shapleigh, *Science* (80-). **1954**, 119, 419.
 - [44] S. H. Barondes, *Trends Biochem. Sci.* **1988**, 13, 480–482.
 - [45] D. Avichezer, D. J. Katcoff, N. C. Garber, N. Gilboa-Garber, *J. Biol. Chem.* **1992**, 267, 23023–23027.
 - [46] K. Winzer, C. Falconer, N. C. Garber, S. P. Diggle, M. Camara, P. Williams, *J. Bacteriol.* **2000**, 182, 6401–6411.
 - [47] A. Imberty, M. Wimmerová, E. P. Mitchell, N. Gilboa-Garber, *Microbes Infect.* **2004**, 6, 221–228.
 - [48] N. Gilboa-Garber, D. Sudakevitz, M. Sheffi, R. Sela, C. Levene, *Glycoconj. J.* **1994**, 11, 414–417.
 - [49] S. P. Diggle, R. E. Stacey, C. Dodd, M. Cámara, P. Williams, K. Winzer, *Environ. Microbiol.* **2006**, 8, 1095–1104.
 - [50] D. Tielker, S. Hacker, R. Loris, M. Strathmann, J. Wingender, S. Wilhelm, F. Rosenau, K. E. Jaeger, *Microbiology* **2005**, 151, 1313–1323.
 - [51] R. U. Kadam, M. Bergmann, M. Hurley, D. Garg, M. Cacciarini, M. A. Swiderska, C. Nativi, M. Sattler, A. R. Smyth, P. Williams, et al., *Angew. Chem. Int. Ed.* **2011**, 50, 10631–10635.
 - [52] T. Eierhoff, B. Bastian, R. Thuenauer, J. Madl, A. Audfray, S. Aigal, S. Juillot, G. E. Rydell, S. Müller, S. de Bentzmann, et al., *Proc. Natl. Acad. Sci. U. S. A.* **2014**, 111, 12895–12900.
 - [53] C. Sabin, E. P. Mitchell, M. Pokorná, C. Gautier, J. P. Uille, M. Wimmerová, A. Imberty, *FEBS Lett.* **2006**, 580, 982–987.
 - [54] D. Passos Da Silva, M. L. Matwichuk, D. O. Townsend, C. Reichhardt, D. Lamba, D. J. Wozniak, M. R. Parsek, *Nat. Commun.* **2019**, 10, 2183.
 - [55] A. M. Boukerb, A. Decor, S. Ribun, R. Tabaroni, A. Rousset, L. Commin, S. Buff, A. Doléans-Jordheim, S. Vidal, A. Varrot, et al., *Front. Microbiol.* **2016**, 7, 1–16.
 - [56] J. Adam, M. Pokorná, C. Sabin, E. P. Mitchell, A. Imberty, M. Wimmerová, *BMC Struct. Biol.* **2007**, 7, 1–13.
 - [57] J. Klockgether, A. Munder, J. Neugebauer, C. F. Davenport, F. Stanke, K. D. Larbig, S. Heeb, U. Schöck, T. M. Pohl, L. Wiehlmann, et al., *J. Bacteriol.* **2010**, 192, 1113–1121.
 - [58] H. Mikkelsen, R. McMullan, A. Filloux, *PLoS One* **2011**, 6, e29113.

- [59] L. Wiehlmann, G. Wagner, N. Cramer, B. Siebert, P. Gudowius, G. Morales, T. Köhler, C. Van Delden, C. Weinl, P. Slickers, et al., *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104*, 8101–8106.
- [60] R. Sommer, S. Wagner, A. Varrot, C. M. Nycholat, A. Khaledi, S. Häussler, J. C. Paulson, A. Imberty, A. Titz, *Chem. Sci.* **2016**, *7*, 4990–5001.
- [61] 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.
- [62] L. Chiarini, A. Bevivino, C. Dalmastri, S. Tabacchioni, P. Visca, *Trends Microbiol.* **2006**, *14*, 277–286.
- [63] 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.
- [64] T. Dingjan, A. Imberty, S. Pérez, E. Yuriev, P. A. Ramsland, *Front. Pharmacol.* **2017**, *8*, 1–16.
- [65] I. Wilhelm, E. Levit-Zerdoun, J. Jakob, S. Villringer, M. Frensch, R. Übelhart, A. Landi, P. Müller, A. Imberty, R. Thuenauer, et al., *Sci. Signal.* **2019**, *12*, 1–18.
- [66] C. Ligeour, L. Dupin, A. Angeli, G. Vergoten, S. Vidal, A. Meyer, E. Souteyrand, J. J. Vasseur, Y. Chevolot, F. Morvan, *Org. Biomol. Chem.* **2015**, *13*, 11244–11254.
- [67] N. Galanos, E. Gillon, A. Imberty, S. E. Matthews, S. Vidal, *Org. Biomol. Chem.* **2016**, *14*, 3476–3481.
- [68] S. Kuhaudomlarp, L. Cerofolini, S. Santarsia, E. Gillon, S. Fallarini, G. Lombardi, M. Denis, S. Giuntini, C. Valori, M. Fragai, et al., *Chem. Sci.* **2020**, *11*, 12662–12670.
- [69] World Health Organization, “WHO publishes list of bacteria for which new antibiotics are urgently needed,” can be found under www.who.int/news/item/27-02-2017-who-publishes-list-of-bacteria-for-which-new-antibiotics-are-urgently-needed, **2017**.
- [70] A. W. MUSSER, P. R. BEAMER, *J. Indiana State Med. Assoc.* **1961**, *54*, 1627–1634.
- [71] J. Horcajada, Milagro Montero, A. Oliver, L. Sorlí, Sònia Luque, S. Gómez-Zorrilla, N. Benito, S. Grau, *Clin. Microbiol. Rev.* **2019**, *32*, 1–52.
- [72] M. Empting, *Nachrichten aus der Chemie* **2018**, *66*, 290–294.
- [73] W. M. Dunne, *Clin. Microbiol. Rev.* **2002**, *15*, 155–166.

- [74] A. L. Spoering, K. Lewis, *J. Bacteriol.* **2001**, *183*, 6746–6751.
- [75] S. Zheng, T. Eierhoff, S. Aigal, A. Brandel, R. Thuenauer, S. de Bentzmann, A. Imberty, W. Römer, *Biochim. Biophys. Acta - Mol. Cell Res.* **2017**, *1864*, 1236–1245.
- [76] A. M. Boukerb, A. Rousset, N. Galanos, J. B. Méar, M. Thépaut, T. Grandjean, E. Gillon, S. Cecioni, C. Abderrahmen, K. Faure, et al., *J. Med. Chem.* **2014**, *57*, 10275–10289.
- [77] I. Bucior, J. Abbott, Y. Song, M. A. Matthay, J. N. Engel, *Am. J. Physiol. - Lung Cell. Mol. Physiol.* **2013**, *305*, 1–22.
- [78] D. Hauck, I. Joachim, B. Frommeyer, A. Varrot, B. Philipp, H. M. Möller, A. Imberty, T. E. Exner, A. Titz, *ACS Chem. Biol.* **2013**, *8*, 1775–1784.
- [79] R. Sommer, K. Rox, S. Wagner, D. Hauck, S. S. Henrikus, S. Newsad, T. Arnold, T. Ryckmans, M. Brönstrup, A. Imberty, et al., *J. Med. Chem.* **2019**, *62*, 9201–9216.
- [80] R. Sommer, S. Wagner, K. Rox, A. Varrot, D. Hauck, E. C. Wamhoff, J. Schreiber, T. Ryckmans, T. Brunner, C. Rademacher, et al., *J. Am. Chem. Soc.* **2018**, *140*, 2537–2545.
- [81] S. Cecioni, A. Imberty, S. Vidal, *Chem. Rev.* **2015**, *115*, 525–561.
- [82] J. Rodrigue, G. Ganne, B. Blanchard, C. Saucier, D. Giguère, T. C. Shiao, A. Varrot, A. Imberty, R. Roy, *Org. Biomol. Chem.* **2013**, *11*, 6906–6918.
- [83] R. U. Kadam, M. Bergmann, D. Garg, G. Gabrieli, A. Stocker, T. Darbre, J. L. Reymond, *Chem. - A Eur. J.* **2013**, *19*, 17054–17063.
- [84] I. Joachim, S. Rikker, D. Hauck, D. Ponader, S. Boden, R. Sommer, L. Hartmann, A. Titz, *Org. Biomol. Chem.* **2016**, *14*, 7933–7948.
- [85] E. Zahorska, S. Kuhaudomlarp, S. Minervini, S. Yousaf, M. Lepsik, T. Kinsinger, A. K. H. Hirsch, A. Imberty, A. Titz, *Chem. Commun.* **2020**, *56*, 8822–8825.
- [86] G. Yu, A. C. Vicini, R. J. Pieters, *J. Org. Chem.* **2019**, *84*, 2470–2488.
- [87] A. Novoa, T. Eierhoff, J. Topin, A. Varrot, S. Barluenga, A. Imberty, W. Römer, N. Winssinger, *Angew. Chem. Int. Ed. Int. Ed.* **2014**, *53*, 8885–8889.
- [88] C. W. Murray, M. L. Verdonk, *J. Comput. Aided. Mol. Des.* **2002**, *16*, 741–753.
- [89] S. Wagner, D. Hauck, M. Hoffmann, R. Sommer, I. Joachim, R. Müller, A. Imberty, A. Varrot, A. Titz, *Angew. Chem. Int. Ed.* **2017**, *56*, 16559–16564.
- [90] E. Siebs, E. Shanina, S. Kuhaudomlarp, P. da Silva Figueiredo Celestino Gomes, C. Fortin, P. H. Seeberger, D. Rognan, C. Rademacher, A. Imberty, A.

- Titz, *ChemBioChem* **2022**, 23, e202100563.
- [91] S. Kuhaudomlarp, E. Siebs, E. Shanina, J. Topin, I. Joachim, P. da Silva Figueiredo Celestino Gomes, A. Varrot, D. Rognan, C. Rademacher, A. Imberty, et al., *Angew. Chem. Int. Ed. Int. Ed.* **2021**, 60, 8104–8114.
- [92] E. Shanina, S. Kuhaudomlarp, E. Siebs, F. F. Fuchsberger, M. Denis, F. P. da S. C. Gomes, M. H. Clausen, P. H. Seeberger, D. Rognan, A. Titz, et al., *Commun. Chem.* **2022**, 5, 64.
- [93] C. J. Carrano, K. N. Raymond, *J. Am. Chem. Soc.* **1979**, 101, 5401–5404.
- [94] E. A. Dertz, J. Xu, A. Stintzi, K. N. Raymond, *J. Am. Chem. Soc.* **2006**, 128, 22–23.
- [95] P. Saha, B. S. Yeoh, X. Xiao, R. M. Golonka, S. Kumarasamy, M. Vijay-Kumar, *Biochem. Pharmacol.* **2019**, 168, 71–81.
- [96] T. Sato, K. Yamawaki, *Clin. Infect. Dis.* **2019**, 69, S538–S543.
- [97] “FDA Approves Fetroja,” can be found under <https://www.drugs.com/newdrugs/fda-approves-fetroja-cefiderocol-complicated-urinary-tract-infections-5104.html>, **2019**.
- [98] “FDA Approves Fetroja (cefiderocol) for the Treatment of Hospital-acquired Bacterial Pneumonia and Ventilator-associated Bacterial Pneumonia,” can be found under <https://www.drugs.com/newdrugs/fda-approves-fetroja-cefiderocol-complicated-urinary-tract-infections-5104.html>, **2020**.
- [99] D. D. Truong, *Clin. Interv. Aging* **2009**, 4, 109–113.
- [100] I. Katsaiti, J. Nixon, *J. Parkinsons. Dis.* **2018**, 8, 217–231.
- [101] R. Ceravolo, P. Piccini, D. L. Bailey, K. M. Jorga, H. Bryson, D. J. Brooks, *Synapse* **2002**, 43, 201–207.
- [102] R. U. Kadam, D. Garg, J. Schwartz, R. Visini, M. Sattler, A. Stocker, T. Darbre, J.-L. Reymond, *ACS Chem. Biol.* **2013**, 8, 1925–1930.
- [103] I. Joachim, S. Rikker, D. Hauck, D. Ponader, S. Boden, R. Sommer, L. Hartmann, A. Titz, *Org. Biomol. Chem.* **2016**, 14, 7933–7948.
- [104] E. Shanina, E. Siebs, H. Zhang, D. Varón Silva, I. Joachim, A. Titz, C. Rademacher, *Glycobiology* **2021**, 31, 159–165.
- [105] M. Gelin, V. Delfosse, F. Allemand, F. Hoh, Y. Sallaz-Damaz, M. Pirocchi, W. Bourguet, J. L. Ferrer, G. Labesse, J. F. Guichou, *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2015**, 71, 1777–1787.
- [106] B. Blanchard, A. Nurisso, E. Hollville, C. Tétaud, J. Wiels, M. Pokorná, M.

- Wimmerová, A. Varrot, A. Imberty, *J. Mol. Biol.* **2008**, 383, 837–853.
- [107] W. Kabsch, *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2010**, 66, 125–132.
- [108] P. R. Evans, *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2011**, 67, 282–292.
- [109] A. J. McCoy, *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2006**, 63, 32–41.
- [110] P. Emsley, K. Cowtan, *Acta Crystallogr. Sect. D* **2004**, 60, 2126–2132.
- [111] G. N. Murshudov, P. Skubák, A. A. Lebedev, N. S. Pannu, R. A. Steiner, R. A. Nicholls, M. D. Winn, F. Long, A. A. Vagin, *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2011**, 67, 355–367.
- [112] F. Long, R. A. Nicholls, P. Emsley, S. Gražulis, A. Merkys, A. Vaitkus, G. N. Murshudov, *Acta Crystallogr. Sect. D Struct. Biol.* **2017**, 73, 112–122.
- [113] V. B. Chen, W. B. Arendall, J. J. Headd, D. A. Keedy, R. M. Immormino, G. J. Kapral, L. W. Murray, J. S. Richardson, D. C. Richardson, *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2010**, 66, 12–21.
- [114] M. Miethke, M. Pieroni, T. Weber, M. Brönstrup, P. Hammann, L. Halby, P. B. Arimondo, P. Glaser, B. Aigle, H. B. Bode, et al., *Nat. Rev. Chem.* **2021**, 5, 726–749.
- [115] Y. Carmeli, N. Troillet, G. M. Eliopoulos, M. H. Samore, *Antimicrob. Agents Chemother.* **1999**, 43, 1379–1382.
- [116] M. B. Calvert, V. R. Jumde, A. Titz, *Beilstein J. Org. Chem.* **2018**, 14, 2607–2617.
- [117] C. Chemani, A. Imberty, S. De Bentzmann, M. Pierre, M. Wimmerová, B. P. Guery, K. Faure, *Infect. Immun.* **2009**, 77, 2065–2075.
- [118] T. Eierhoff, B. Bastian, R. Thuenauer, J. Madl, A. Audfray, S. Aigal, S. Juillot, G. E. Rydell, S. Muller, S. de Bentzmann, et al., *Proc. Natl. Acad. Sci.* **2014**, 111, 12895–12900.
- [119] R. Thuenauer, A. Landi, A. Trefzer, S. Altmann, S. Wehrum, T. Eierhoff, B. Diedrich, J. Dengjel, A. Nyström, A. Imberty, et al., *Am. Soc. Microbiol.* **2020**, 11, e03260-19.
- [120] J. Meiers, E. Siebs, E. Zahorska, A. Titz, *Curr. Opin. Chem. Biol.* **2019**, 53, 51–67.
- [121] G. Cioci, E. P. Mitchell, C. Gautier, M. Wimmerová, D. Sudakevitz, S. Pérez, N. Gilboa-Garber, A. Imberty, *FEBS Lett.* **2003**, 555, 297–301.
- [122] J. Rodrigue, G. Ganne, B. Blanchard, C. Saucier, D. Giguère, T. C. Shiao, A. Varrot, A. Imberty, R. Roy, *Org. Biomol. Chem.* **2013**, 11, 6906–6918.

- [123] F. Pertici, N. J. de Mol, J. Kemmink, R. J. Pieters, *Chem. Eur. J.* **2013**, *19*, 16923–16927.
- [124] E. Mitchell, C. Houles, D. Sudakevitz, M. Wimmerova, C. Gautier, S. Pérez, A. M. Wu, N. Gilboa-Garber, A. Imberty, *Nat. Struct. Biol.* **2002**, *9*, 918–921.
- [125] R. Sommer, T. E. Exner, A. Titz, *PLoS One* **2014**, *9*, 1–22.
- [126] A. Hofmann, R. Sommer, D. Hauck, J. Stifel, I. Göttker-Schnetmann, A. Titz, *Carbohydr. Res.* **2015**, *412*, 34–42.
- [127] R. Sommer, D. Hauck, A. Varrot, S. Wagner, A. Audfray, A. Prestel, H. M. Möller, A. Imberty, A. Titz, *ChemistryOpen* **2015**, *4*, 756–767.
- [128] P. E. Allegretti, E. A. Castro, J. J. P. Furlong, *J. Mol. Struct. THEOCHEM* **2000**, *499*, 121–126.
- [129] F. Damkaci, P. DeShong, *J. Am. Chem. Soc.* **2003**, *125*, 4408–4409.
- [130] D. Hauck, V. R. Jumde, C. J. Crawford, A. Titz, in *Carbohydr. Chem.*, **2021**, pp. 17–22.
- [131] M. Andreini, M. Anderluh, A. Audfray, A. Bernardi, A. Imberty, *Carbohydr. Res.* **2010**, *345*, 1400–1407.
- [132] O. Šulák, G. Cioci, E. Lameignère, V. Balloy, A. Round, I. Gutsche, L. Malinovská, M. Chignard, P. Kosma, D. F. Aubert, et al., *PLoS Pathog.* **2011**, *7*, e1002238.
- [133] R. Bermeo Malo, Conception, Synthèse et Évaluation de Glycocomposés Dirigés Contre BC2L-C, Université Grenoble Alpes, **2021**.
- [134] G. M. Morris, H. Ruth, W. Lindstrom, M. F. Sanner, R. K. Belew, D. S. Goodsell, A. J. Olson, *J. Comput. Chem.* **2009**, *30*, 2785–2791.
- [135] D. Weininger, *J. Chem. Inf. Comput. Sci.* **1988**, *28*, 31–36.
- [136] O. Korb, T. Stützel, T. E. Exner, *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)* **2006**, *4150*, 247–258.
- [137] E. P. Mitchell, C. Sabin, L. Šnajdrová, M. Pokorná, S. Perret, C. Gautier, C. Hofr, N. Gilboa-Garber, J. Koča, M. Wimmerová, et al., *Proteins Struct. Funct. Genet.* **2005**, *58*, 735–746.
- [138] E. Gasteiger, A. Gattiker, C. Hoogland, I. Ivanyi, R. D. Appel, A. Bairoch, *Nucleic Acids Res.* **2003**, *31*, 3784–3788.
- [139] Iegrandp, J. Aishima, CV-GPhL, **2017**, DOI 10.5281/ZENODO.837886.
- [140] M. D. Winn, C. C. Ballard, K. D. Cowtan, E. J. Dodson, P. Emsley, P. R. Evans, R. M. Keegan, E. B. Krissinel, A. G. W. Leslie, A. McCoy, et al., *Acta Crystallogr.*

- Sect. D Biol. Crystallogr.* **2011**, 67, 235–242.
- [141] A. J. McCoy, R. W. Grosse-Kunstleve, P. D. Adams, M. D. Winn, L. C. Storoni, R. J. Read, *J. Appl. Crystallogr.* **2007**, 40, 658–674.
- [142] P. Emsley, B. Lohkamp, W. G. Scott, K. Cowtan, *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2010**, 66, 486–501.
- [143] M. Martinucci, E. Roscetto, V. D. Iula, A. Votsi, M. R. Catania, E. De Gregorio, *Lett. Appl. Microbiol.* **2016**, 62, 221–229.
- [144] A. Isles, I. Macluskay, M. Corey, R. Gold, C. Prober, P. Fleming, H. Levison, *J. Pediatr.* **1984**, 104, 206–210.
- [145] B. J. Lang, S. D. Aaron, W. Ferris, P. C. Hebert, N. E. Macdonald, *Am. J. Respir. Crit. Care Med.* **2000**, 162, 2241–2245.
- [146] A. D. Rhim, L. Stoykova, M. C. Glick, T. F. Scanlin, *Glycoconj. J.* **2001**, 18, 649–659.
- [147] R. Marchetti, L. Malinowska, E. Lameignère, L. Adamova, C. De Castro, G. Cioci, C. Stanetty, P. Kosma, A. Molinaro, M. Wimmerova, et al., *Glycobiology* **2012**, 22, 1387–1398.
- [148] O. Šulák, G. Cioci, M. Delia, M. Lahmann, A. Varrot, A. Imberty, M. Wimmerová, *Structure* **2010**, 18, 59–72.
- [149] G. Beshr, Biochemical Characterization and Ligand Identification of Pseudomonas Aeruginosa Lectins and Their Orthologs in Other Pathogenic Bacteria, **2018**.
- [150] P. Sýkorová, J. Novotná, G. Demo, G. Pompidor, E. Dubská, J. Komárek, E. Fajdiarová, J. Houser, L. Hároníková, A. Varrot, et al., *Int. J. Biol. Macromol.* **2020**, 152, 1113–1124.
- [151] M. Frensch, C. Jäger, P. F. Müller, A. Tadić, I. Wilhelm, S. Wehrum, B. Diedrich, B. Fischer, A. V. Meléndez, J. Dengjel, et al., *Cell. Mol. Life Sci.* **2021**, 78, 8165–8186.
- [152] C. Ligeour, A. Audfray, E. Gillon, A. Meyer, N. Galanos, S. Vidal, J. J. Vasseur, A. Imberty, F. Morvan, *RSC Adv.* **2013**, 3, 19515–19524.
- [153] T. Dingjan, É. Gillon, A. Imberty, S. Pérez, A. Titz, P. A. Ramsland, E. Yuriev, *J. Chem. Inf. Model.* **2018**, 58, 1976–1989.
- [154] R. T. Lee, T. L. Hsu, S. K. Huang, S. L. Hsieh, C. H. Wong, Y. C. Lee, *Glycobiology* **2011**, 21, 512–520.
- [155] J. Wang, L. Yao, *Sci. Rep.* **2019**, 9, 1–9.

- [156] S. Koch, D. Schollmeyer, H. Löwe, H. Kunz, *Chem. Eur. J.* **2013**, 19, 7020–7041.
- [157] T. F. Headen, C. A. Howard, N. T. Skipper, M. A. Wilkinson, D. T. Bowron, A. K. Soper, *J. Am. Chem. Soc.* **2010**, 132, 5735–5742.
- [158] C. R. Martinez, B. L. Iverson, *Chem. Sci.* **2012**, 3, 2191–2201.
- [159] D. F. Sticke, L. G. Presta, K. A. Dill, G. D. Rose, *J. Mol. Biol.* **1992**, 226, 1143–1159.
- [160] G. N. Murshudov, P. Skubák, A. A. Lebedev, N. S. Pannu, R. A. Steiner, R. A. Nicholls, M. D. Winn, F. Long, A. A. Vagin, *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2011**, 67, 355–367.
- [161] A. C. Wallace, R. A. Laskowski, J. M. Thornton, *Protein Eng.* **1995**, 8, 127–134.
- [162] J. Meiers, E. Zahorska, T. Röhrig, D. Hauck, S. Wagner, A. Titz, *J. Med. Chem.* **2020**, 63, 11707–11724.
- [163] H. E. Gottlieb, V. Kotlyar, A. Nudelman, *J. Org. Chem.* **1997**, 62, 7512–7515.
- [164] P. C. Bulman Page, Y. Chan, J. Liddle, M. R. J. Elsegood, *Tetrahedron* **2014**, 70, 7283–7305.