

# **Biophysical Characterization of Carbohydrate-Lectin Interactions**

**Inauguraldissertation**

zur

Erlangung der Würde eines Doktors der Philosophie

vorgelegt der

Philosophisch-Naturwissenschaftlichen Fakultät

der Universität Basel

von

Pascal Severin Zihlmann

aus Malters, Luzern

Basel, 2018

Genehmigt von der Philosophisch-Naturwissenschaftlichen Fakultät

auf Antrag von:

Prof. Dr. Beat. Ernst  
Institut für Molekulare Pharmazie  
Universität Basel

Prof. Dr. Gerhard Klebe  
Institut für Pharmazeutische Chemie  
Philipps-Universität Marburg, Deutschland

Basel, den 23. Februar 2016

Prof. Dr. Jörg Schibler  
Dekan

### Acknowledgements

First and foremost, I would like to thank *Prof. Dr. Beat Ernst* for giving me the opportunity to work in his outstanding research group. Your vast experience in drug discovery and your willingness to share your knowledge with me made the last three years an enlightening experience. I highly appreciate all your contributions of time, ideas, encouragement and funding to make my doctorate possible. I truly enjoyed this time.

I further would like to thank *Prof. Dr. Gerhard Klebe* for accepting to be co-referee on my thesis-committee. This is a real honor for me as your talks and publications have been a great inspiration for my work.

I would like to thank *Prof. Dr. Timm Maier* for supporting me with your time and knowledge in my crystallization projects. I would like to point out the great achievements of *Dr. Roman Peter Jakob*. The crystal structures you solved were of inestimable value for my research projects.

A very special thanks goes to *Dr. Timothy Sharpe* for his invaluable support in biophysics. There is probably no biophysical question you can't answer (at least I was never able to pose one). You have been a real support and friend.

*Tobias Mühlethaler, Stefan Siegrist, and Michael Misev* earn my high appreciation for being the best master students I could have imagined. I'm really grateful for all the work you contributed to my thesis. I wish you all the best for your futures and your scientific careers.

Several people from the IMP deserve special thanks including *Anja Sigl* and *Christoph Sager* for the amazing atmosphere in our "Sunnestübli". You were precious friends and inspiring co-workers and I truly enjoyed the stimulating discussions we had. Another special "thank you" goes to *Dr. Said Rabbani* for sharing his extensive experience in molecular biology and for providing me with incredible amounts of FimH protein. I would like to thank *Claudia Huber* for reminding me to sign up for courses, for keeping me awake with coffee, and for being the good soul of the IMP.

## Acknowledgements

---

Furthermore, I would like to thank all *former members of the IMP* that my work is based on. Your amazing spadework has made this thesis possible. I'm especially grateful to *Dr. Roland Preston, Dr. Katrin Lemme, and Dr. Adam Zalewski* who initiated many of those fascinating projects I was involved with.

My time in Basel was made much more enjoyable due to the many friends in the IMP-family. One thing I always loved about the IMP was that every new member was immediately integrated without any exceptions. I remember those memorable trips into the Alps, the vivid discussions in the coffee corner, and the wonderful evenings we spent together during our "summer beer" events. I would like to thank *Brigitte, Philipp, Katja, Simon, Jacqueline, Deniz, Marleen, Meike, and Daniela* from the biology group. Huge thanks go also to all the chemists, namely *Bea, Xiaohua, Giulio, Olly, Priska, Kathi, Mirko, Norbert, Blijke, Lijuan, Fan, Rachel, Hélène, Maja, and Wojtek*. The members of the "Molecular Modeling" group from *Prof. Dr. Angelo Vedani* have been a source of friendships, as well as good advice and collaboration. Thank you, *Martin, Sameh, Oya, Joël, and Zhenquan*. I hope to spend much more time with you all.

Last but not least, I want to thank my family and friends and my wife for all of their love and encouragement. Thank you, *Maya and Bruno*, for believing in me and supporting me over more than three decades. Thank you, *Barbara*, for your faithful support during the last 10 years and for the wonderful relationship we have.

**Abstract**

Improved knowledge of the biological role of lectins has raised the demand for carbohydrate-based therapeutics in recent years. The potential market is estimated to be greater than 20 billion dollars. However, lectins are challenging drug targets due to the unique binding properties of their extensively hydroxylated carbohydrate ligands. Hydroxyl groups provide directionality and therefore specificity, but are penalized with high desolvation costs. Consequently, monovalent carbohydrate-lectin interactions tend to be rather weak, often in the millimolar range. Moreover, the polar character of carbohydrates creates large obstacles for drug application regarding oral availability and long-lasting plasma levels. The key to the successful development of carbohydrate-based drugs is the simultaneous optimization of carbohydrate lead structures in terms of pharmacodynamics and pharmacokinetics. To further enhance the success rate of carbohydrate-based drug candidates, the understanding of carbohydrate-lectin interactions on a molecular basis has to be improved. For this purpose, we combined structural information (X-ray crystallography and nuclear magnetic resonance spectroscopy), binding data (isothermal titration calorimetry, microscale thermophoresis, and fluorescence polarization assay) and computational methods (quantum mechanical calculations and molecular dynamics simulations) to explore the lectins FimH and E-selectin and their interaction with carbohydrates and mimetics thereof.

FimH is a virulence factor of uropathogenic *E. coli* located at the tip of the bacterial type 1 pili. It interacts with the mannosylated glycoprotein uroplakin 1a in the urothelial mucosa and thereby mediates adhesion to the bladder wall as the initial step of urinary tract infections (UTI). In *manuscript 1* we investigated the energy contribution to binding of the hydroxyl groups mediating the interaction between FimH and the carbohydrate moiety of its ligands. The rigidity of this bacterial lectin was demonstrated in *manuscript 2*, where we showed that the affinity of a septanose as a mannose mimic is reduced by a factor of 10, mainly due to its flexibility in solution and the consequent conformational restrictions upon binding. In *manuscript 3* we analyzed interactions between the tyrosine gate motif of FimH and the aglycones of different ligand classes. This motif (Tyr48, Tyr137) forms the entrance of the binding pocket and significantly contributes to binding affinity. In *manuscript 4* we explored 2-C-branched mannosides as a

## Abstract

---

novel family of FimH antagonists. In *manuscript 5*, a pharmacodynamically and pharmacokinetically optimized FimH antagonist was explored by oral application in a mouse model for UTI, resulting in a 1000-fold reduction of the bacterial load in the bladder. Finally, in *manuscript 6* we reanalyzed ITC results from our previous studies with the novel analytical tool kinITC, allowing the determination of kinetics in addition to the thermodynamics of binding. The hydrophobic aglycone turned out to be mainly responsible for guiding the antagonist to its binding site whereas the hydrogen bond network between the mannose moiety and the protein had predominantly an impact on the off-rate.

E-selectin is a lectin expressed on the surface of vascular endothelial cells and is involved in the recruitment of leukocytes to the site of inflammation. By interacting with the tetrasaccharide epitope sialyl Lewis<sup>x</sup>, E-selectin establishes the initial contact and enables leukocytes to roll along the endothelial surface. Whereas this process is a defense mechanism in case of infections and injuries, excessive extravasation of leukocytes can have deleterious consequences in case of numerous diseases with an inflammatory component, e.g. asthma, psoriasis or stroke. Thus, blocking the interaction of E-selectin with its physiological ligands is a promising strategy to suppress the inflammatory response at the beginning of the cascade. For reliable and materially efficient affinity measurements, we developed and evaluated a novel assay for E-selectin based on microscale thermophoresis technology in *manuscript 7*. In the subsequent manuscripts, we applied the microscale thermophoresis assay. In *manuscript 8* we performed a competitive library screen, whereby four promising small-molecule fragments were identified for further development towards a non-carbohydrate E-selectin antagonist. Finally, in *manuscript 9* we were able to improve the affinity of a sialyl Lewis<sup>x</sup> mimic to E-selectin by pre-organizing the acid in its bioactive conformation.

---

**Abbreviations**

3-D	Three-dimensional space
$[\alpha]_D^{20}$	Optical rotation
ABP	L-arabinose-binding protein
ABTS	2,2'-azino-bis[3-ethylbenzthiazoline-6-sulfonic acid]
Ac	Acetyl
AcOH	Acetic acid
AFM	Atomic force microscopy
Ala	Alanine
AM	1,5-Anhydro-D-mannitol
aq.	Aqueous
Ar	Aryl
Arg	Arginine
Asn	Asparagine
Asp	Aspartic acid
ATP	Adenosine triphosphate
AUC	Area under the curve
ax	Axial
BF	4-Biphenyl- $\alpha$ -D-mannopyranoside
Bicine	2-(Bis(2-hydroxyethyl)amino)acetic acid
BisTris	2-[Bis(2-hydroxyethyl)amino]-2-(hydroxymethyl)propane-1,3-diol
Bn	Benzyl
Boc	<i>tert</i> -butyloxycarbonyl
BSA	Bovine serum albumin
Bu	Butyl
Bz	Benzoyl
<i>c</i>	Wiseman parameter
$C_0$	Initial concentration
Caco-2 cells	Caucasian colon adenocarcinoma cells
calcd	Calculated
cat.	Catalyst / Catalytic amount
CD	Circular-dichroism
CD22	Cluster of differentiation 22 / Siglec-2
CES	Carboxylesterase
CFU	Colony forming units
C-HEGA	Cyclohexylbutanoyl- <i>N</i> -hydroxyethyl-D-glucamide
CIP	Ciprofloxacin
$CL_{tot}$	Total clearance
ConA	Concanavalin A
conc.	Concentration
Conf. I.	Confidence Interval
COPD	Chronic obstructive pulmonary disease
COSY	Correlation spectroscopy (NMR)
COX	Cyclooxygenase

## Abbreviations

---

$\Delta C_p$	Change in heat capacity
CRD	Carbohydrate recognition domain
Cryo-EM	Cryo-electron microscopy
CSP	Chemical shift perturbation
Cy(c)	Cyclohexyl
CYP	Cytochrome
Cys	Cysteine
$\delta$	Chemical shift
d	Doublet (NMR)
D	Distribution coefficient
$D$	Dielectric constant
D <sub>2</sub> O	Heavy water, deuterium oxide
DAMP	Damage-associated molecular pattern
DAST	Diethylaminosulfur trifluoride
DC-SIGN	Dendritic cell-specific ICAM-3 grabbing nonintegrin
DCE	1,2-Dichloroethane
DCM	Dichloromethane
DL	Detection limit
DMAP	4-Dimethylaminopyridine
DME	1,2-Dimethoxyethane
DMEM	Dulbecco's modified eagle medium
DMF	<i>N,N</i> -Dimethylformamide
DMSO	Dimethyl sulfoxide
DMTST	Dimethyl(methylthio)sulfonium triflate
DNA	Deoxyribonucleic acid
DsF	FimF donor strand
DsG	FimG donor strand
dsRNA	Double stranded RNA
DTBMP	2,6-Di- <i>tert</i> -butyl-4-methylpyridine
$\epsilon_0$	Vacuum permittivity
<i>E. coli</i>	<i>Escherichia coli</i>
e.g.	For example
EA	Ethyl acetate
EDTA	Ethylenediaminetetraacetic acid
ee	Enantiomeric excess
EGF	Epidermal growth factor
ELAM	Endothelial-leukocyte adhesion molecule
ELISA	Enzyme-linked immunosorbent assay
ELLA	Enzyme-linked lectin assay
eq	Equatorial
equiv	Equivalent
ESI	Electrospray ionization (-MS)
ESL-1	E-selectin ligand-1
Et	Ethyl
ETC	Equilibration time curve

---

EtOAc	Ethyl acetate
EtOH	Ethanol
FACS	Fluorescence-activated cell sorting
FCS	Fetal calf serum
FimH <sub>L</sub>	FimH lectin domain
FimH <sub>P</sub>	FimH pilin domain
FITC	Fluorescein isothiocyanate
FP	Fluorescence polarization
Fuc	Fucose
$\Delta G$	Change in free energy of binding
$\Delta G^{\circ}_{\text{int}}$	Change in intrinsic free energy of binding
$\Delta G^{\circ}$	Change in standard free energy of binding
$\Delta G^{\circ}_{\text{solv}}$	Change in solvation free energy of binding
Gal	Galactose
GdmCl	Guadinium chloride
GFP	Green fluorescent protein
Glc	Glucose
GlcNAc	<i>N</i> -Acetylglucosamine
Gln	Glutamine
Glu	Glutamic acid
Gly	Glycine
GlyCAM-1	Glycosylation-dependent cell adhesion molecule-1
GMP-140	Granule membrane protein-140
$\Delta H^{\circ}_{\text{ion}}$	Change in ionization enthalpy
$\Delta H^{\circ}$	Change in enthalpy
H bond	Hydrogen bond
h(rs)	Hour(s)
HA	Hemagglutination assay
hCAII	Human carbonic anhydrase II
HEPES	2-[4-(2-Hydroxyethyl)piperazin-1-yl]ethanesulfonic acid
HEV	High endothelial venule
HIA	Hemagglutination inhibition assay
His	Histidine
HIV	Human immunodeficiency virus
HM	<i>n</i> -Heptyl $\alpha$ -D-mannopyranoside
HMBC	Heteronuclear multiple bond correlation
HPLC	High performance liquid chromatography
HRMS	High-resolution mass spectrometry
HSQC	Heteronuclear single quantum coherence
Hz	Hertz
IBC	Intracellular bacterial community
$IC_{50}$	Half-maximal inhibitory concentration
ICAM-3	Intercellular adhesion molecule-3
IFD	Induced fit docking

## Abbreviations

---

IgG	Immunoglobulin G
IL-1	Interleukin-1
Ile	Isoleucine
IMP	Institute of Molecular Pharmacy
<i>i</i> PrOH	2-Propanol
IPTG	Isopropyl $\beta$ -D-1-thiogalactopyranoside
ITC	Isothermal titration calorimetry
iv	Intravenous
$K_A$	Equilibrium association constant
$K_D$	Equilibrium dissociation constant
$K_i$	Inhibitory constant
kinITC	Kinetic ITC
$k_{\text{off}}$	Dissociation rate constant
$k_{\text{on}}$	Association rate constant
L	Ligand
Lac	Lactic acid
LAM	Lymphocyte adhesion molecule
LB	Luria–Bertani
LC	Liquid chromatography (-MS)
LECAM	Leukocyte-endothelial cell adhesion molecule
LED	Light-emitting diode
LF	Ligand-free
LPS	Lipopolysaccharide
Lys	Lysine
m	Multiplet (NMR)
$m/z$	Mass-to-charge ratio
MAC <sub>90</sub>	Minimal concentration to inhibit 90% adhesion
MAAdCAM-1	Mucosal vascular addressin cell adhesion molecule-1
MAG	Myelin-associated glycoprotein / Siglec-4
Man	Mannose
MD	Molecular dynamics
Me	Methyl
MeCN	Acetonitrile
MeOH	Methanol
MFI	Mean fluorescence intensity
min(s)	Minute(s)
MM-GBSA	Molecular Mechanics - Generalized Born Surface Area
mol.	Molecular
MOPS	3-Morpholinopropane-1-sulfonic acid
MPLC	Medium-pressure liquid chromatography
mRNA	Messenger ribonucleic acid
MS	Mass spectrometry
MST	Microscale thermophoresis
MWCO	Molecular weight cut-off
<i>N</i>	Stoichiometry

---

n.b.	Not binding
n.d.	Not determined
NADPH	Nicotinamide adenine dinucleotide phosphate
Neu5Ac	<i>N</i> -Acetylneuraminic acid, sialic acid
NHS	<i>N</i> -Hydroxysuccinimide
Ni-NTA	Nickel-nitrilotriacetic acid
NIS	<i>N</i> -Iodosuccinimide
NMR	Nuclear magnetic resonance
NOE	Nuclear Overhauser effect
OD	Optical density
OH	Hydroxyl
P	Partitioning coefficient
P	Protein
p-TSOH	<i>p</i> -Toluenesulfonic acid
PAA	Polyacrylamide
PAMP	Pathogen-associated molecular pattern
PAMPA	Parallel artificial membrane permeation assay
$P_{app}$	Apparent permeability
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PD	Pharmacodynamics
PDB	Protein data bank
$P_e$	Effective permeability
PE	Petroleum ether
PEG	Polyethylene glycol
PES	Polyethersulfone
PGA-LM	Poly- $\gamma$ -glutamic acid low-molecular-weight polymer
Ph	Phenyl
Phe	Phenylalanine
PK	Pharmacokinetics
PL	Protein-ligand complex
po	Per os (oral)
PPB	Plasma protein binding
ppm	Parts per million
Pr	Propyl
Pro	Prolin
PRR	Pattern recognition receptor
PSGL-1	P-selectin glycoprotein ligand-1
PTFE	Poly(1,1,2,2-tetrafluoroethylene)
py	Pyridine
q	Quartet
QIR	Quiescent intracellular reservoir
QM	Quantum mechanics
QTOF	Quadrupole TOF MS
<i>R</i>	Universal gas constant

## Abbreviations

---

$R_f$	Retention factor
RLM	Rat liver microsome
RMSD	Root-mean-square deviation
RNA	Ribonucleic acid
RP	Reversed-phase
rt	Room temperature
$\Delta S^\circ$	Change in entropy
$\Delta S^\circ_{\text{conf}}$	Change in conformational entropy
$\Delta S^\circ_{\text{solv}}$	Change in solvation entropy
$\Delta S^\circ_{\text{trans/rot}}$	Change in translational and rotational entropy
s	Second or singlet (NMR)
SAR	Structure-activity relationship
SASA	Solvent accessible surface area
sat.	Saturated
SAXS	Small-angle X-ray scattering
sc	Subcutaneous
SCR	Short consensus repeat
SDS-PAGE	Sodium dodecyl sulfate - Polyacrylamide gel electrophoresis
sec(s)	Second(s)
Ser	Serine
Siglec	Sialic acid-binding immunoglobulin-like lectins
SKR	Structure-kinetic relationship
sLe <sup>a</sup>	Sialyl Lewis <sup>a</sup>
sLe <sup>x</sup>	Sialyl Lewis <sup>x</sup>
SNF	Swiss National Science Foundation
SPR	Surface plasmon resonance
STD	Saturation-transfer difference (NMR)
Std.	Standard
$\tau$	Residence time
t	Triplet (NMR)
$T$	Temperature
$t_{1/2}$	Dissociation half-life, residence time
TAPS	<i>N</i> -Tris(hydroxymethyl)methyl-3-aminopropanesulfonic acid
TBAB	Tetra- <i>n</i> -butylammonium bromide
TBAF	Tetra- <i>n</i> -butylammonium fluoride
TBAI	Tetra- <i>n</i> -butylammonium iodide
TBDMS	<i>tert</i> -Butyldimethylsilyl
TBDPS	<i>tert</i> -Butyldiphenylsilyl
TCDI	1,1'-Thiocarbonyldiimidazole
TEER	Transepithelial electrical resistance
Tf	Triflate
TFA	Trifluoroacetic acid
TfOH	Trifluoromethanesulfonic acid
Th	Thrombin
THF	Tetrahydrofuran, Oxolane

---

Thr	Threonine
TLC	Thin-layer chromatography
TMS	Trimethylsilyl
TMSOTf	Trimethylsilyl trifluoromethanesulfonate
TNF- $\alpha$	Tumor necrosis factor alpha
TOCSY	Total correlation spectroscopy
TOF	Time of flight
Tricine	<i>N</i> -(Tri(hydroxymethyl)methyl)glycine
Tris	Tris(hydroxymethyl)aminomethane
trNOE	Transfer NOE (-NMR)
Trp	Tryptophan
TS <sup>‡</sup>	Transition state
TSP	3-(Trimethylsilyl)propanoic acid
Tyr	Tyrosine
Tys	Sulfotyrosine
UPEC	Uropathogenic <i>E. coli</i>
UPK1a	Uroplakin 1a protein
UTC	Urinary tract infection
UV	Ultraviolet
Val	Valine
WT	Wild type

### Table of contents

<b>I. Introduction</b>	<b>1</b>
1. Molecular Interaction	3
1.1. Milestones in the research history of molecular interaction	3
1.2. Mathematical description of a molecular binding event	4
1.3. Driving forces of molecular interactions: Thermodynamics	5
1.3.1. Enthalpic driving forces	6
1.3.2. Entropic driving forces	10
1.3.3. Solvent effects	11
2. Evaluating Molecular Interaction	13
2.1. Methods to determine the structure of complexes	13
2.2. Methods to determine the inhibitory constant	14
2.3. Methods to determine the binding constants	15
3. Carbohydrate-Lectin Interactions	19
3.1. Carbohydrates	19
3.2. Lectins	19
3.2.1. Microbial lectins	20
3.2.2. Human lectins	20
3.3. Lectins and carbohydrates in drug discovery	21
3.4. Molecular basis of the regulation of carbohydrate-lectin affinity	22
3.4.1. Affinity regulation: Carbohydrates	23
3.4.2. Affinity regulation: Lectins	24
3.4.3. Multivalency	25
3.5. Thermodynamics of carbohydrate-lectin interactions	25
4. References	28
<b>II. The Bacterial Adhesin FimH</b>	<b>35</b>
1. Introduction	37
1.1. Urinary tract infection	37
1.2. Infection cycle of uropathogenic <i>Escherichia coli</i>	37
1.3. Structure	39
1.3.1. Type 1 pili and <i>fim</i> operon	39
1.3.2. The lectin FimH	40
1.3.3. FimH binding site	44
1.4. FimH antagonists	46
1.5. References	49

---

2. Manuscripts	53
2.1. Manuscript 1: High-affinity carbohydrate-lectin interactions: How nature makes it possible	53
2.2. Manuscript 2: The price of flexibility – A case study on septanoses as pyranose mimetics	95
2.3. Manuscript 3: The tyrosine gate of the bacterial lectin FimH: A conformational analysis by NMR Spectroscopy and X-ray crystallography.	107
2.4. Manuscript 4: 2-C-Branched mannosides as a novel family of FimH antagonists – Synthesis and biological evaluation	121
2.4. Manuscript 5: FimH antagonists: Bioisosteres to improve the in vitro and in vivo PK/PD Profile.	133
2.5. Manuscript 6: KinITC – One method supports both thermodynamic and kinetic SARs	155
<b>III. E-selectin</b>	<b>203</b>
1. Introduction	205
1.1. A brief overview of the selectin family members	205
1.2. Structural features	205
1.3. Selectin ligands	207
1.4. The role of selectins in inflammation	208
1.4.1. Inflammatory cascade	208
1.4.2. Role of selectins in disease	209
1.5. Structural binding properties	210
1.5.1. E-selectin	211
1.5.2. P-selectin	213
1.5.3. Catch-bond mechanism	214
1.6. Selectin antagonists	215
1.7. References	218
2. Manuscripts	
2.1. Manuscript 7: Microscale thermophoresis: A powerful opportunity for carbohydrate-based drug discovery	223
2.2. Manuscript 8: Fragment screening towards an orally available E-selectin antagonist using microscale thermophoresis	245
2.4. Manuscript 9: Is it a hydrophobic clash that pre-organizes sLe <sup>x</sup> mimics in the bioactive conformation for the interaction with E-selectin?	263



**Section I.**  
**Introduction**



## 1. Molecular Interaction

### 1.1. Milestones in the research history of molecular interaction

It has been more than 120 years since Emil Fischer discovered a fundamental underlying principle of molecular interaction.<sup>(1)</sup> He observed that methyl- $\alpha$ -D-glucoside but not methyl- $\beta$ -D-glucoside is hydrolyzed in the presence of the yeast extract invertin. In contrast, methyl- $\beta$ -D-glucoside but not its  $\alpha$ -anomer was hydrolyzed in the presence of the almond extract emulsin. He interpreted his findings with the “lock and key” principle:

“Um ein Bild zu gebrauchen, will ich sagen, dass Enzym und Glucosid wie Schloss und Schlüssel zu einander passen müssen, um eine chemische Wirkung auf einander ausüben zu können.”

The “lock and key” principal of Fischer and the finding of Paul Ehrlich that agents can only work when they are bound (“*corpora non agunt nisi fixata*”) provided the basis for modern pharmacology.<sup>(2,3)</sup> They allowed medicinal chemists to synthesize “keys” that activate (agonist) or inhibit (antagonist) the function of a given “lock” (protein).<sup>(4)</sup> Since then, the number of known “locks” has been steadily growing and nowadays includes proteins of various groups like e.g. enzymes, lectins, hormone receptors, ion channels, transporters, structural proteins or transcription regulators.

However, the static ‘key and lock’ principle could not explain why non-competitive inhibition can occur and why, for example, glucose but not water is phosphorylated in the binding pocket of hexokinase. In consequence, Daniel E. Koshland introduced the dynamic ‘induced-fit model’ in 1958.<sup>(5,6)</sup>

“a) a precise orientation of catalytic groups is required for enzyme action, b) the substrate causes an appreciable change in the three-dimensional relationship of the amino acids at the active site, and c) the changes in the protein structure caused by the substrate will bring the catalytic groups into the proper alignment, whereas a non-substrate will not.”

In recent years, the concept of ‘conformational selection’ introduced from Ruth Nussinov and co-workers<sup>(7)</sup> has become more and more popular as an alternative dynamic model. This model assumes a selection for a conformational subpopulation in a system of dynamically fluctuating protein and ligand species. It is in contrast to the ‘induced-fit model’, where the conformational change is thought to be a consequence of the interaction of protein and ligand. Both models might be combined to an initial conformational selection followed by a conformational adjustment.<sup>(8)</sup>

### 1.2. Mathematical description of a molecular binding event

The equations to describe a protein-ligand binding event is the law of mass action deduced by Guldberg and Waage in 1879.<sup>(9)</sup>



$$K = \frac{k_1}{k_{-1}} = \frac{[C]^c [D]^d}{[A]^a [B]^b} \quad (\text{eq. 2})$$

Equation 2 describes the ratio between the concentrations of the reactants ( $[A]$ ,  $[B]$ ) and products ( $[C]$ ,  $[D]$ ) at the equilibrium state of the reaction described in equation 1, where  $a$ ,  $b$ ,  $c$ , and  $d$  are the stoichiometric coefficients. The equilibrium constant  $K$  defines the extent of the reaction, whereas the rate is defined by the rate constant of the forward reaction ( $k_1$ ) and the reverse reaction ( $k_{-1}$ ). Equilibrium is reached when the rate of reaction in forward and reverse directions is equal.

The law of mass action can be used analogously to describe a protein-ligand binding interaction (equations 3 and 4).



$$K_A = \frac{k_{\text{on}}}{k_{\text{off}}} = \frac{[PL]}{[P][L]} = \frac{1}{K_D} \quad (\text{eq. 4})$$

$$t_{1/2} = \frac{\ln 2}{k_{\text{off}}} \quad (\text{eq. 5})$$

The equilibrium constants (binding constants) are called the association constant ( $K_A$ ) and the dissociation constant ( $K_D$ ). They describe the concentration in an equilibrium at which half of the protein in a system is complexed by a ligand. The rate constants are called the association rate constant  $k_{on}$  and the dissociation rate constants  $k_{off}$ . From  $k_{off}$  the half-life ( $t_{1/2}$ ) of the protein-ligand complex can be obtained (equation 5).

### 1.3. Driving forces of molecular interactions: Thermodynamics

A spontaneous interaction between ligand and receptor only takes place when the overall change of free energy of binding ( $\Delta G$ ) is negative.  $\Delta G$  of a reaction is determined by the nature of the reactants and their concentrations with  $R$  is the universal gas constant (8.314 J/mol K) and  $T$  being the absolute temperature (equation 6):<sup>(10)</sup>

$$\Delta G = \Delta G^\circ + RT \ln \frac{[P][L]}{[PL]} = \Delta G^\circ + RT \ln K_D \quad (\text{eq. 6})$$

$\Delta G^\circ$  is the change in free energy of binding under standard condition that is a reactant concentration of 1 M in solution. The system is at equilibrium when  $\Delta G = 0$ . Therefore, the relation between the standard free energy and the equilibrium constant can be derived from equations 7 and 8.

$$0 = \Delta G^\circ + RT \ln K_D \quad (\text{eq. 7})$$

$$\Delta G^\circ = RT \ln K_D = - RT \ln K_A \quad (\text{eq. 8})$$

$\Delta G^\circ$  is furthermore determined by the change of enthalpy ( $\Delta H^\circ$ ) and change of entropy ( $\Delta S^\circ$ ) in dependence of the absolute temperature ( $T$ ) (equation 9):

$$\Delta G^\circ = \Delta H^\circ - T\Delta S^\circ \quad (\text{eq. 9})$$

The change in entropy results from the sum of the changes in solvation entropy, conformational entropy, and the rigid body motions (translational and rotational entropy). The change in enthalpy reflects the change in bond interaction energy (covalent and non-covalent, including short and long-range electrostatics).

In drug discovery often a phenomenon called *enthalpy-entropy compensation* is observed, where an enthalpic improvement is linearly correlated with a loss in the entropy and *vice versa*.<sup>(11,12)</sup> A newly formed interaction between protein and ligand may lead to an improved enthalpy, but simultaneously constrains the conformational flexibility of protein and ligand, leading to an entropic penalty. It is furthermore accepted that water plays a key role in enthalpy-entropy compensation. When a binding event takes place in aqueous solution (solvent), both ligand and protein (solutes) shed their solvation shell in the contact area. The costs of this desolvation process have to be considered when calculating the binding energy. Hence, *the change in free energy of binding is the difference between the solute-solute and the solute-solvent interactions*. Depending on the polarity of the binding site and the ligand, this desolvation process is enthalpically more or less costly as hydrogen bonds between solute and solvent are to be broken.<sup>(13)</sup> However, this effect might again be partially compensated by a beneficial entropy term as water molecules are released to bulk.

### 1.3.1. Enthalpic driving forces

Enthalpic driving forces of a binding reaction are characterized by their release or consumption of heat. A negative enthalpy means the release of heat and contributes beneficially to the free energy of binding. *The nature of enthalpic interactions is electrostatic*. This means that they are formed *between charged or dipolar atoms* and are therefore subject to the Coulomb's law (equation 10):

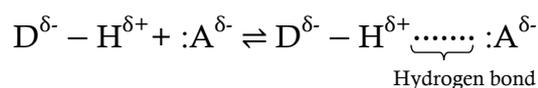
$$E = \frac{q_1 q_2}{4\pi\epsilon_0 r D} \quad (\text{eq. 10})$$

$E$  is the electrostatic energy,  $r$  is the distance,  $q_1$  and  $q_2$  are the charges of two interacting atoms,  $\epsilon_0$  is the vacuum permittivity, and  $D$  is the dielectric constant of the surrounding medium. Electrostatic interactions are therefore either attractive (oppositely charged) or repulsive (same charge). As described in Coulomb's law, the energy of an interaction does not only depend on the individual charges of the interacting, but also on their distance and the solvent it takes place in. The solvent of biological systems (water) has a dielectric constant of 80. In close proximity to a protein, the dielectric constants drops to values in the range of 20 and in protein cavities to 5 or lower. In a vacuum,  $D$  is defined

to be 1.<sup>(14,15)</sup> Hence, *the magnitude of the energy of an electrostatic interaction is increased in proximity to the protein due to the lower dielectric constant.* Furthermore, electrostatic energies vary according to the nature of the charges and their distance. Between ions with net charges, the energy falls off slowly with distance  $1/r$  (long-range electrostatics), while the energy of randomly oriented dipoles falls off rapidly with distance  $1/r^6$  (short-range electrostatics).<sup>(16)</sup>

## Hydrogen bonds

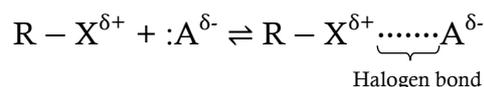
A hydrogen bond is a non-covalent interaction between a *hydrogen bond donor* ( $D^{\delta-}$  covalently bound to a hydrogen atom  $H^{\delta+}$ ) and a *hydrogen bond acceptor* with a free lone pair ( $:A^{\delta-}$ ). Both donor and acceptor are electronegative atoms, usually nitrogen or oxygen.



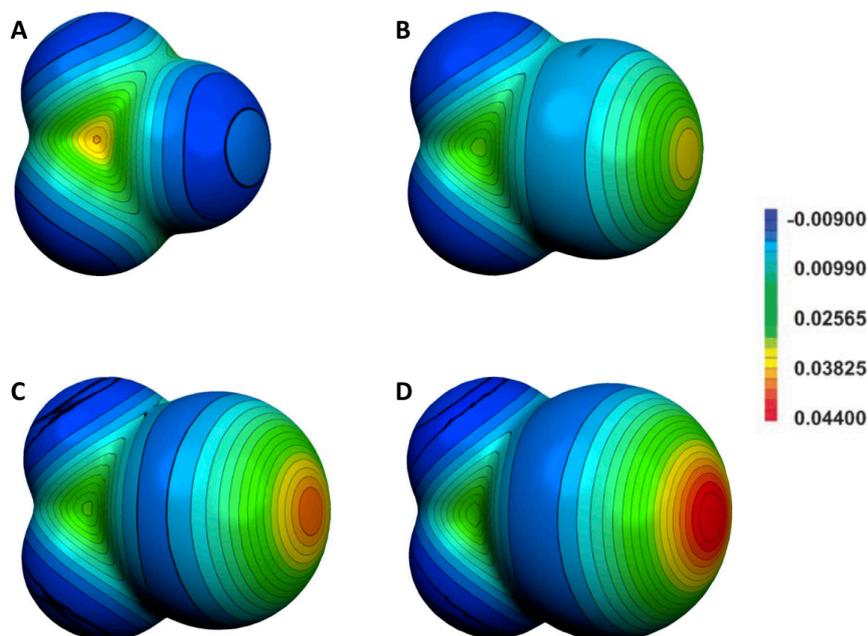
Hydrogen bonds are *electrostatic dipole-dipole interactions with covalent features*: the distance between the donor and acceptor is shorter than the sum of the van-der-Waals radii of the interacting atoms (but longer than a covalent bond) and their quality depends on *geometrical rules* (bond length between 2.5-3.2 Å; bond angle 130-180°). *Directionality provides specificity* and is therefore crucial for molecular recognition.<sup>(17)</sup> According to Jeffrey and Saenger (1991), a hydrogen bond contributes 1-4 kcal/mol to the free energy of binding.<sup>(18)</sup> A formal charge on one (charge-assisted hydrogen bond) or both (salt bridge) interaction partner(s) increases the energy of the bond.<sup>(19)</sup> However, hydrogen bonds are not restricted to the interaction between protein and ligand, but can also be formed with the solvent (water). Whether the disruption of these interactions during desolvation (desolvation penalty) is compensated by the newly formed interactions between ligand and protein depends on the specific case. However, even if hydrogen bonds do not always contribute significantly to the free energy of binding they might nonetheless be important for the specificity of a molecular binding event.<sup>(17)</sup>

## Halogen bonds

A halogen bond is built between a *covalently bound halogen atom* ( $R-X^{\delta+}$ ) ( $X$ =halogen atom) and a *halogen bond acceptor* ( $A^{\delta-}$ ). The halogen atom is the donor, while the acceptor is usually a lone pair of a nitrogen or an oxygen atom. However, also  $\pi$  electrons of an unsaturated conjugated system or a second halogen atom may act as an acceptor.<sup>(20)</sup>



Although often described as entirely negative, the electrons of halogens are distributed anisotropically (Figure 1). The outer lobe of the half-filled  $p$ -Orbital is electronically depleted and therefore forms a more negative region along the transaxial plane and a positive region ( $\sigma$ -hole) along the R-X axis. Hence, the optimal bond angle for a halogen bond is  $180^\circ$ .<sup>(21)</sup> In theory, all halogens are able to form halogen bonds, but the positive electrostatic potential and the bond energy increases with the polarizability of the halogen atom ( $F < Cl < Br < I$ ).<sup>(21)</sup> However, the electronegativity and a high level of  $sp$ -hybridization of fluorine cause an electron influx that neutralizes the  $\sigma$ -hole.<sup>(22)</sup>



**Figure 1. Molecular electrostatic potential of halogen atoms onto a 0.001 electrons Bohr<sup>-3</sup> isodensity surface in Hartrees ( $E_h$ ).** Chlorine (B), bromine (C) and iodine (D) but not fluorine (A) form a positively charged  $\sigma$ -hole. The covalently attached electron withdrawing  $CF_3$  group reinforcing the formal positive charge. (Picture from Clark *et al.* (2007) modified after Metrangolo *et al.* (2008); reprinted with permission from Springer Science + Business Media)<sup>(22,23)</sup>

Like hydrogen bonds, halogen bonds are highly directional, electrostatically driven non-covalent interactions. Halogen bonds are weaker than hydrogen bonds, but nevertheless still of importance in drug discovery because halogens are often used to provide solubility, metabolic stability and bioavailability to molecules.

### Van der Waals contacts

Van der Waals (VdW) interactions are non-specific electrostatic interactions formed between dipoles. Although one single contact is weak, a large number of interactions contribute significantly to the overall binding energy.<sup>(24-26)</sup> While polar molecules have permanent dipoles, non-polar molecules only have induced dipoles formed by short-termed random fluctuations in the distribution of the electron density. A distinction can be made between the interaction of a permanent dipole with an induced dipole (Debye force), two permanent dipoles (Keesom interaction) or two induced dipoles (London dispersion force). The VdW interactions are mathematically described by the 12-6 Lennard-Jones potential ( $V$ ) (equation 11),<sup>(27)</sup> where  $A$  and  $B$  are specific molecular constants (based on the VdW radii and the magnitude of attraction) of the interacting atoms and  $r$  describes their distance:

$$V = \frac{A}{r^{12}} - \frac{B}{r^6} \quad (\text{eq. 11})$$

$A/r^{12}$  represents the repulsive potential and  $B/r^6$  the attractive potential. The closer two atoms approach, the higher the attraction is (negative potential  $V$ ). Yet, when the atoms come too close together, the repulsion term rapidly increases due to overlapping electron clouds (positive potential  $V$ ). At  $V=0$ , when attraction and repulsion are in balance, the atoms are in VdW contact, which is the sum of the VdW radii of the two interacting atoms.<sup>(28)</sup>

### Stacking with $\pi$ systems

The amino acids phenylalanine, tryptophan, tyrosine, and histidine as well as many drug molecules contain  $\pi$ -systems. The conjugation of their p-orbitals enables delocalization of  $\pi$  electrons. This electron delocalization polarizes the aromatic ring resulting in two partially negative charged rings above and below the plane of the partially positive

charged atom ring (quadrupole).<sup>(29)</sup> Two aromatic systems mainly interact in a parallel or perpendicular configuration. The parallel arrangement is more favored when electron-withdrawing groups are introduced that weaken the quadrupole.<sup>(30)</sup> On the contrary, electron donating groups increase the quadrupole charge distribution and therefore promote the perpendicular constellation between two aromatic rings, but also the interaction with other formally positive charged groups (cation- $\pi$  interaction).<sup>(31)</sup> Furthermore, interactions with aliphatic groups and halogens are able to favorably contribute to the free energy of binding.<sup>(32)</sup>

### 1.3.2. Entropic driving forces

The entropy of a thermodynamic system is the measure of its disorder (motion). In contrast to the change in enthalpy, there is no direct measure of the change in entropy. However, since the  $K_A$  and  $\Delta H^\circ$  can be determined,  $\Delta S^\circ$  can be calculated according to equations 8 and 9. A positive change in entropy contributes beneficially to the change in free energy of binding. Like  $\Delta H^\circ$ ,  $\Delta S^\circ$  is an additive measure of different processes taking place during a binding event and it can be further dissected (equation 12).<sup>(33,34)</sup>

$$\Delta S^\circ = \Delta S^\circ_{\text{solv}} + \Delta S^\circ_{\text{conf}} + \Delta S^\circ_{\text{trans/rot}} \quad (\text{eq. 12})$$

The change in solvation entropy ( $\Delta S^\circ_{\text{solv}}$ ) depends on the hydration shell that alters upon the binding of a ligand to a protein. Water molecules are released to bulk water where they are assumed to gain motion. The hydration shell “melts” with increasing temperature and it completely passes to bulk water at a temperature of 385 K ( $T_R$ ).<sup>(35,36)</sup> Therefore, determination of the temperature dependence of the thermodynamic constants allows the calculation of  $\Delta S^\circ_{\text{solv}}$  (equation 13 and 14), where  $\Delta C_p$  is the change in heat capacity and  $T$  is the absolute temperature.

$$\Delta C_p = \frac{\partial \Delta H^\circ}{\partial T} \quad (\text{eq. 13})$$

$$\Delta S^\circ_{\text{solv, T}} = \Delta C_p \ln \left( \frac{T}{T_R} \right) \quad (\text{eq. 14})$$

The change in translational and rotational entropy ( $\Delta S_{\text{trans/rot}}^{\circ}$ ) is called the cratic term. It is a constant loss of energy due to the loss of degrees of freedom as upon complex formation when two rigid bodies (each being able to translate and rotate) join to one. Literature values for the cratic term range from -5.4 up to -62 kJ/mol (at 298.15 K).<sup>(37-39)</sup> However, the most often used change in entropy at room temperature is -10 kJ/mol calculated from equation 15, where  $R$  is the universal gas constant (8.314 J/mol K) and 55.6 in the divisor corresponds to the molarity of 1 liter water.<sup>(39,40)</sup>

$$\Delta S_{\text{trans/rot}}^{\circ} = R \ln \left( \frac{1}{55.6} \right) \quad (\text{eq. 15})$$

The change in conformational entropy ( $\Delta S_{\text{conf}}^{\circ}$ ) can therefore be calculated according to equation 12.  $\Delta S_{\text{conf}}^{\circ}$  is the change of conformational degrees of freedom of the bound protein-ligand complex compared to their unbound states in solvent.  $\Delta S_{\text{conf}}^{\circ}$  is usually unfavorable (negative) as the amino acid side chains involved in binding, as well as the rotational bonds of the ligand, are more restricted in a complex. Conformational changes of the protein (induced fit) may also contribute unfavorably to  $\Delta S_{\text{conf}}^{\circ}$ .

### 1.3.3. Solvent effects

To understand the role of water is crucial to understand binding. It is often neglected, but each water molecule has to be considered as an additional interaction partner when a protein-ligand interaction takes place in an aqueous environment. Therefore, the change in free energy of binding of a protein-ligand binding event contains the energy change of the whole system, including the *desolvation* of the interaction interfaces, newly formed interactions of the protein-ligand complex and *water reorganization*.

Liquid water favorably forms approx. three hydrogen bonds while retaining a certain amount of mobility.<sup>(18)</sup> However, water molecules in the first solvation layer with close contact to the solute might not be able to form three hydrogen bonds (enthalpically *frustrated water*) or be limited in their mobility (conformationally *trapped water*).<sup>(24,41-44)</sup> To release these water molecules to the bulk may be highly beneficial. Dunitz calculated an entropic gain of up to 2 kcal/mol (at 300 K) for the release of a trapped water molecule.<sup>(45)</sup> In terms of enthalpy, the release of a trapped water molecule might be less

## Section I. Introduction – Molecular Interaction

---

beneficial, as some of them play an important role in mediating interactions between protein and ligand (*structural water*).<sup>(32,46)</sup> Still, by far the most water molecules are loosely associated with the protein and can be displaced with minor effect to the overall free energy of binding. Hence, the solvation of an unbound protein binding site is predetermining the binding characteristics. While some water patterns promote binding, others request high desolvation costs and prevent strong interactions.

Furthermore, the properties of the solute are decisive for solvation, in particular, its polarity. The highly polar water molecules tend to form hydrogen bonds to hydrophilic solutes. Breaking these hydrogen bonds upon desolvation is enthalpically expensive, but is partially compensated by the increased mobility of the water molecules in the bulk. On the other hand, hydrophobic solutes do not interact via hydrogen bonds with the solvent but rather form a cavity disrupting the water structure. The concept of the *classical hydrophobic effect* proposed by Kauzman assumes a more ordered water structure surrounding the non-polar solute.<sup>(47-51)</sup> Desolvation upon binding is therefore accompanied by a gain of entropy. However, ITC measurements revealed a *non-classical hydrophobic effect* characterized by an enthalpy gain upon desolvation.<sup>(44,52-55)</sup> This effect is explained by the gain of electrostatic interactions between the non-polar solutes as well as by the optimized formation of hydrogen bonds between the water molecules in the bulk.

## 2. Evaluating Molecular Interaction

### 2.1. Methods to determine the structure of complexes

A full understanding of a molecular binding interaction requires not only knowledge of the energetics of binding, but also structural information about the bound and unbound states of the binding partners. Otherwise, crucial effects such as conformational changes and oligomerization might not be detected and can lead to a misinterpretation of binding data. Probably the most important source for high resolution structural information is *X-ray crystallography*, where the diffraction pattern of a molecular structure (e.g. lectin-carbohydrate complex) is gained by the exposure of its crystal to a beam of monochromatic X-rays.<sup>(56)</sup> The X-ray scatters at the atoms in the crystal and the emerging diffraction pattern can be transformed into 3-dimensional information regarding individual atom positions. However, information about complex structure from X-ray crystallography may be compromised by poor resolution or unfavorable crystal contacts that prevent binding or distort the structure of the bound state. Moreover, an X-ray structure is a static picture of a dynamic system and it is wise to consider other sources of structural information like small-angle X-ray scattering (SAXS),<sup>(57)</sup> cryo-electron microscopy (Cryo-EM),<sup>(58)</sup> nuclear magnetic resonance (NMR) spectroscopy, atomic force microscopy (AFM)<sup>(59,60)</sup> or homology modeling.<sup>(61)</sup> *Computational* approaches like molecular dynamics (MD) simulations or quantum mechanical (QM) calculations are very useful to evaluate questions of motion, conformation, solvation and energy composition.<sup>(62)</sup> The most powerful technique to obtain both structural information and binding data is *NMR spectroscopy* and it is, in contrast to X-ray crystallography, measured in solution. A range of informative NMR data such as chemical shifts and intra- or intermolecular nuclear Overhauser effects (NOE) can be combined to derive the 3D structure of a protein or protein-ligand complex.<sup>(63,64)</sup> Further experiments can be applied to elucidate the dynamics of a protein, ranging from fast backbone or side chain fluctuations to global conformational changes on a slow timescale.<sup>(65,66)</sup> Protein-ligand binding events can be detected and quantified by a broad range of experiments with either observation of the ligand signals (e.g. saturation transfer difference, transfer NOE or relaxation-based experiments)<sup>(67-69)</sup> or of the protein signals (chemical shift perturbation experiments, CSP).<sup>(70)</sup> CSP experiments are especially suited to the identification of unknown binding sites on a protein target.<sup>(71)</sup>

### 2.2. Methods to determine the inhibitory constant

Carbohydrate-lectin interactions are frequently weak, having  $K_D$  values in the high  $\mu\text{M}$  to  $\text{mM}$  range. Many techniques that are routinely used to quantify high-affinity interactions ( $K_D$  in the  $\mu\text{M}$  to  $\text{pM}$  range) are not applicable or practicable for measuring such lower affinities. Therefore, the next two sections will focus on techniques suitable for the measurement of the carbohydrate-lectin interactions.

Many competitive assay formats have been established to quantify carbohydrate-lectin interactions. The phenomenon of hemagglutination that led to the discovery of the first lectin by Stillmark in the 19<sup>th</sup> century, is still the most often used method.<sup>(72,73)</sup> The *hemagglutination assay (HA)* takes advantage of the heterogeneous glycosylation of erythrocytes and the multivalent binding properties of most lectins.<sup>(74,75)</sup> While the presence of anticoagulants red blood cells would segregate in suspension, the addition of a multivalent lectin cross-links the blood cells and creates a lattice structure. An assay format where the carbohydrate competes with erythrocytes for binding to the lectin allows the determination of a carbohydrate concentration that inhibits 50% of the agglutination ( $IC_{50}$ ) (*Hemagglutination inhibition assay, HIA*).  $IC_{50}$  values are helpful to compare and rank the inhibition potency of different ligands, but they are not suitable to compare inhibition concentrations measured with different methods or in different laboratories (e.g. when differently glycosylated erythrocytes are used). Furthermore, the HIA is subject to the numerous limitations, e.g. irreversible cross-linkage without formation of an equilibrium or multivalent ligands cross-linking the erythrocytes. These drawbacks are eliminated by the *enzyme-linked lectin assay (ELLA)*.<sup>(73)</sup> In the ELLA method, a polyvalent ligand is immobilized on a microtiter plate and a lectin-horseradish peroxidase conjugate is incubated with a serial dilution of competing ligand of interest. After equilibration, the supernatant containing the unbound ligand of interest and lectin is washed off and replaced by a solution of chromogenic peroxidase substrate. The protein bound on the microtiter plate can now be quantified using UV spectroscopy. This method can be modified, e.g. when the lectin is immobilized on a microtiter plate and the ligand is competing with a spectroscopically detectable glycopolymer (*Polymer-binding assay*).<sup>(76)</sup> However, the output of both assays is an  $IC_{50}$  value with all its drawbacks. A possible approach to solve these drawbacks is the fitting of the binding curves using an analytical solution to derive the inhibition constant ( $K_i$ ) for competitive binding derived

by Wang.<sup>(77)</sup> However, for both assays, HIA and ELLA, the Wang equation cannot be applied as the exact concentration of all participating molecules as well as the dissociation constant ( $K_D$ ) between the standard competitor and the lectin is not known. A competitive assay where the Wang equation can be applied is a technique based on *fluorescence polarization (FP)*. Here, the competing ligand (tracer) is linked to a fluorophore.<sup>(78)</sup> In steady-state fluorescence polarization measurements, polarized light is used to excite a fluorophore-containing tracer compound, and the degree of polarization of fluorescence emission is determined by measuring fluorescence through polarizers parallel and perpendicular to the axis of excitation polarization. The extent of depolarization of fluorescence emission depends upon the rotation of the fluorophore during its fluorescence lifetime. This depends upon the tumbling rate of the fluorophore, which is influenced by its hydrodynamic radius. Hence, a tracer bound to a protein exhibits slower tumbling compared to a free tracer in solution and has increased fluorescence polarization. Titrating a dilution series of protein against a constant concentration of tracer gives a binding curve from which it is possible to derive the exact  $K_D$  using a standard single-site binding model (e.g. the function given by Cooper<sup>(79)</sup>). Therefore, all constants are known to use the Wang equation for the competitive assay format.

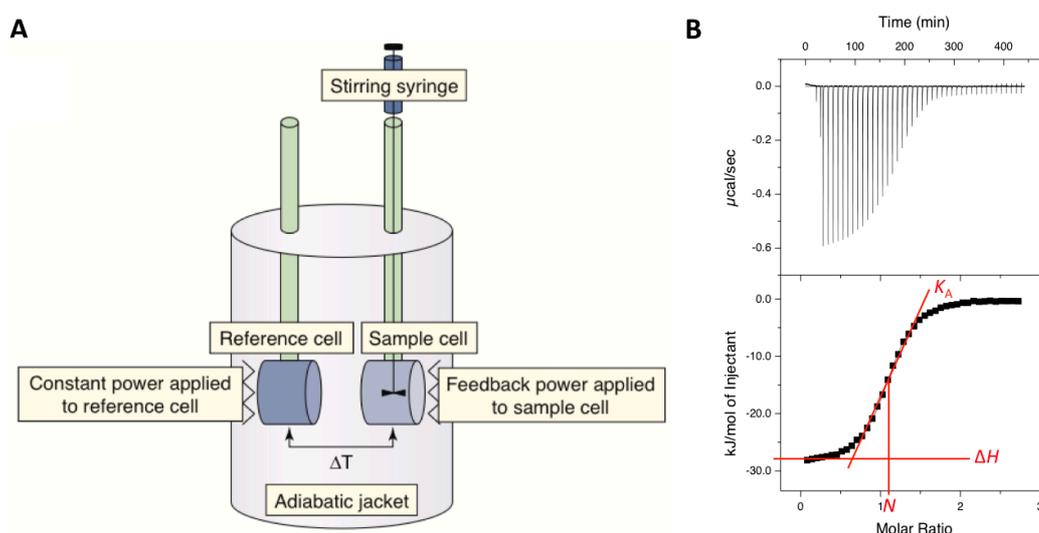
### 2.3. Methods to determine the binding constants

More reliable is the direct determination of the  $K_D$  without the complication of a competitive assay format, e.g. using *microscale thermophoresis (MST)*.<sup>(80)</sup> We routinely use MST to determine affinities of E-selectin ligands because the extremely small assay volume preserves valuable protein and ligand resources (*Manuscripts 7-9*). The assay is based on the phenomenon that molecules move along a temperature gradient, a phenomenon known as thermophoresis. The rate of thermophoresis is sensitive to changes in size, charge, and solvation of a molecule, which allows discriminating bound from unbound species.<sup>(81)</sup> For some proteins it is possible to follow thermophoresis by intrinsic tryptophan fluorescence, while for many others it is necessary to label one interaction partner with a fluorescent dye.

## Section I. Introduction – Evaluating Molecular Interaction

Another label-free technique to determine  $K_D$ , as well as the association rate constant ( $k_{on}$ ) and the dissociation rate constant ( $k_{off}$ ) is *surface plasmon resonance (SPR)*.<sup>(82)</sup> Either the ligand or the protein is immobilized on a functionalized gold surface. The backside of the gold layer reflects a light source through a prism to a detector. At a certain resonance angle, the light is absorbed by the electrons of the gold layer causing them to resonate. The angle at which this occurs depends upon the refractive index of the solution at the functionalized gold surface. A binding event at the surface changes that refractive index and is observable as a shift of the minimum in the reflectivity curve. Monitoring the change in resonance over time enables to derive the rate constants  $k_{on}$  and  $k_{off}$ .

The only technique to determine directly the thermodynamics of a binding reaction is *isothermal titration calorimetry (ITC)*.<sup>(83)</sup> ITC is the most frequently used technique throughout this thesis and is therefore explained in more detail. With one ITC experiment, it is possible to simultaneously measure the change in enthalpy ( $\Delta H^\circ$ ), the binding stoichiometry ( $N$ ) and  $K_A$ . According to a recent publication by Burnouf *et al.* even the rate constants  $k_{on}$  and  $k_{off}$  can be deduced from ITC measurements.<sup>(84)</sup> For a standard ITC experiment, a threefold surplus of ligand is stepwise titrated through a syringe into a cell containing the protein solution (Figure 2A). However, also an inverse experiment is possible.



**Figure 2. Illustrated mode of operation of an isothermal titration calorimeter.** (A) The small temperature changes are not measured directly, but the power supply to maintain a constant temperature in the sample cell is compared to a reference cell. (B) Enthalpogram of an ITC experiments. *Top:* The recorded change in heat ( $\mu\text{cal}/\text{sec}$ ) that is set free upon complex formation after stepwise injection of ligand. *Bottom:* By integrating the area under the curve of all injections, a binding curve is obtained that contains information on stoichiometry ( $N$ ), change in enthalpy ( $\Delta H^\circ$ ) and the association constant ( $K_A$ ). (2A is a picture from Holdgate and Ward (2005); reprinted with permission from Elsevier)<sup>(85)</sup>

The formation of a complex (equation 3) either consumes or releases heat, which can be monitored during an experiment by comparing the power supplied to heaters on the reference and sample cells to maintain them at equal temperature during the experiment. The energy released or consumed ( $q$ ) is associated with the change in complex concentration ( $\Delta[PL]$ ) and depends on the enthalpy change at the given temperature ( $\Delta H_{(T)}$ ) and the reaction volume of the sample cell ( $V$ ) (equation 16):

$$q = V\Delta H_{(T)}\Delta[PL] \quad (\text{eq. 16})$$

For a tight-binding ligand with an exothermic enthalpy of binding, the surplus of protein in the sample cell enables the entire injected ligand to bind at the beginning of the experiment. When the experiment proceeds, a steadily increasing number of binding sites are occupied leading to a decreased release of heat. When all binding sites are occupied only the heat of dilution is observed (Figure 2B). The integral heat ( $Q$ ) released or consumed after the  $i^{\text{th}}$  injection is calculated according to equation 17,

$$Q = N[P]V\Delta H_{(T)}\theta_i \quad (\text{eq. 17})$$

where  $[P]$  is the total protein concentration,  $\theta_i$  the fractional saturation and  $N$  the stoichiometry. Equation 18 is used to calculate the differential heat after the  $i^{\text{th}}$  injection ( $q_i$ ).

$$q_i = N[P]V\Delta H_{(T)}(\theta_i - \theta_{i-1}) \quad (\text{eq. 18})$$

Both, nonlinear fitting to the hyperbolic saturation curve of  $Q$  vs. the total ligand concentration  $[L]$  based on equation 17 or to the sigmoidal saturation curve of  $q_i$  vs.  $[L]$  based on equation 18 yield in the parameters  $K_A$ ,  $N$  and  $\Delta H^\circ$ . All three parameters can only be determined from one experiment within a certain range of the dimensionless Wiseman parameter ( $c$ ), depending on the affinity of the interacting molecules and the receptor concentration (equation 19).<sup>(86)</sup>

$$c = [P]K_A = [P] \frac{1}{K_D} \quad (\text{eq. 19})$$

## Section I. Introduction – Evaluating Molecular Interaction

---

An optimal  $c$ -value is within a range of 5 and 1000 and allows the determination of all three parameters.<sup>(87)</sup> When the  $c$ -value increases, the sigmoidal binding curve becomes steeper and insensitive to changes in  $K_A$ , while the fitting of the parameters  $\Delta H^\circ$  and  $N$  is still reliable. For the determination of the binding constants of high-affinity ligands, a competitive displacement experiment has to be performed.<sup>(88)</sup> Because lower  $c$ -values require higher concentrations of both, ligand and protein<sup>(89)</sup> problems with solubility and material consumption may occur. Thus, using a larger surplus of ligand to completely saturate the protein within one experiment, allows the determination of  $K_A$  and  $\Delta H^\circ$  using a two-parameter fit with a fixed stoichiometry. With this experimental setup, the accuracy of the fitted parameters depends entirely on the accurate determination of both ligand and protein concentration.

### 3. Carbohydrate-Lectin Interactions

#### 3.1. Carbohydrates

Carbohydrates (saccharides, sugars) are the product of photosynthesis in plants. They are composed of carbon, oxygen and hydrogen atoms ( $C_n(H_2O)_n$ ) and cyclize by either via an aldehyde or a keto group with a hydroxyl of the distal carbons. 5-membered rings are called furanoses, while 6-membered rings are called pyranoses. Carbohydrates are a *major source of energy* for all living organisms and their covalent linkage from polymers (polysaccharides) allows the spatially efficient *storage of this energy* (e.g. starch, glycogen). Due to their ability to form large cross-linked, rigid polymers, they also fulfill important *structural functions* (e.g. cellulose, chitin). Furthermore, the surface of cells is coated with a broad variety of branched polysaccharides (*glycan*) forming a layer of carbohydrates (*glycocalyx*). The glycans are either linked to proteins (glycoproteins, proteoglycans) or lipids (glycolipids). One example of the clinical relevance of glycans is that differential glycosylation of red blood cells determines the human blood groups.<sup>(90)</sup> The glycosylation of a protein may also alter its structure, folding, function or solubility.<sup>(91)</sup> Moreover, glycans are able to promote or inhibit intermolecular binding sterically, but also specifically. The *coding capacity of polysaccharides is enormous*: With the ten mammalian monosaccharides<sup>(92)</sup> glycosidically linked  $\alpha$  or  $\beta$  billions of oligosaccharides can be formed exhibiting different three-dimensional structures with distinct binding specificities. Finally, post-translational modifications (e.g. acetylation, sulfation, methylation, phosphorylation) even further increase the coding capacity of carbohydrates.<sup>(93)</sup>

#### 3.2. Lectins

When carbohydrates are encoding information, *lectins are the structures to read this 'glycocode'* by specific binding. Lectins are defined as *carbohydrate-binding proteins of non-immunological origin and without enzymatic activity*. The discovery of lectins leads back to 19<sup>th</sup> century when Stillmark isolated a protein extract of seeds of *Ricinus communis* and found it to agglutinate with animal erythrocytes.<sup>(72,94)</sup> The connection between carbohydrates and lectins was made when the binding of erythrocytes to Concanavalin A

(a lectin isolated from *Canavalia ensiformis*) could be inhibited by specific saccharides.<sup>(95)</sup> However, not only plants but a broad variety of species, such as bacteria, viruses, and animals were found to express lectins with a wide spectrum of functions.<sup>(96)</sup>

### 3.2.1. Microbial lectins

Many microorganisms make use of lectins for the initial attachment to the cell-surface glycan of host cells.<sup>(97)</sup> Due to this function, these lectins are also termed ‘adhesins’ and are important virulence factors of pathogens. Similar to plant lectins, microbial lectins were often detected due to their ability to agglutinate erythrocytes (hemagglutination). The first detected and still best characterized microbial lectin is hemagglutinin (HA), which was detected on the surface of the influenza virus in the 1950s by Alfred Gottschalk.<sup>(98)</sup> The first bacterial adhesins were discovered in the 1970s.<sup>(99,100)</sup> They were found to be associated with filamentous protein structures at the surface of bacteria (pili). The most prominent representatives to date are FimH, a mannose-binding lectin at the tip of the type 1 pili and PapG, a galabiose binding lectin at the tip of P pili of uropathogenic *E. coli*. FimH and PapG are virulence factors for urinary tract infections (see Section II).

### 3.2.2. Human lectins

The role of lectins in humans is more versatile. In 1988, Kurt Drickamer classified calcium ion-dependent lectins as C-type lectins and lectin with free thiols as S-type lectins.<sup>(101)</sup> Among the best characterized *C-type* lectins are the collectins<sup>(102)</sup> (stimulators of the innate immune system by recognizing glycans of pathogens), DC-SIGN<sup>(103)</sup> (a lectin on dendritic cells that binds to oligomannose-type-N-glycan present in the envelope of viruses, such as HIV and Hepatitis C) and the selectins<sup>(104)</sup> (cell-adhesion molecules that mediate the initial steps of leukocyte trafficking). Function and structure of the selectin family are described in more detail in Section III. In the same way that today lectins are being classified according to structural features of their ‘carbohydrate recognition domain’ (CRD), the earlier S-type lectins were divided into galectins and P-type lectins. *P-type lectins* are intracellular mannose 6-phosphate receptors and are responsible for the routing of lysozymal enzymes.<sup>(105)</sup> *Galectins* are a very diverse family of lectins with a broad spectrum of functions such as cellular growth and apoptosis regulation, cell-matrix

interactions, chemotaxis, immune regulation, and the promotion of pro- and anti-inflammatory responses.<sup>(106)</sup> A further group of lectins are the *I-type* lectins which are members of the immunoglobulin superfamily.<sup>(107)</sup> The most prominent subgroup of the *I-type*, the *Siglecs*, share sialic acid as their common binding motif and fulfill widespread functions such as the maintenance of myelinated axons (Siglec-4 = MAG), or the prevention of an overactivation and development of autoimmune reactions of B-cells (Siglec-2 = CD22).<sup>(108)</sup>

### 3.3. Lectins and carbohydrates in drug discovery

Carbohydrates have been used as drugs since 1930 when the polysaccharide heparin was clinically prescribed as an antithrombotic.<sup>(109)</sup> A further early example of a carbohydrate-based drug is hyaluronic acid used as a vitreous substitution during eye surgery in the 1950s.<sup>(110)</sup> However, both heparin and hyaluronic acid were isolated from natural sources, as opposed to chemically synthesized what is probably one of the reasons why not more carbohydrate-based drugs made their way to clinical application. *Compared to other drug molecules, the synthesis of oligosaccharides is more complex* due to multiple protection and deprotection steps, numerous stereocenters and a lack of general methods for routine preparation.<sup>(111)</sup> However, recent developments are beginning to address these problems.<sup>(112)</sup> The first synthetically produced heparin (Fondaparinux, ARIXTRA, 2002) may exemplify this progress.<sup>(113)</sup>

On the other hand, progress in molecular biology has improved the knowledge of the biological role of lectins in cell regulation and recognition at a molecular level and this has raised the demand for new drugs and therapeutics. Of special interest for drug discovery are microbial adhesins (e.g. FimH, PapG) in prospect of an anti-adhesion therapy of bacterial infections, as well as extracellular human lectins (e.g. C-type, galectins, I-type lectins). An excellent overview of promising lectin drug targets and marketed carbohydrate-based drugs was written by Beat Ernst and John L. Magnani (2009) and is recommended for more detailed information.<sup>(114)</sup> In the year 2010, the potential market for carbohydrate-based drugs was estimated to be greater than 20 billion dollars a year.<sup>(115)</sup> Still, *at the time of writing, no lectin antagonist has made it to the market.* Nevertheless, the development of lectin antagonists is in progress and several drug

candidates are in clinical evaluation. One candidate that has to be highlighted in this context is rivipansel (GMI-1070, Glycomimetics Inc.),<sup>(116)</sup> which is currently in clinical phase III. Rivipansel is a pan-selectin antagonist that mimics the natural binding motif sialyl Lewis<sup>x</sup> and gives hope to patients suffering from sickle cell disease (for more details see Section III).

In contrast to lectin antagonists, *several inhibitors of glycosidases are already marketed*. Successful pharmaceuticals are the viral neuraminidase inhibitors oseltamivir (Tamiflu, Roche)<sup>(117)</sup> and zanamivir (Relenza, GlaxoSmithKline),<sup>(118)</sup> which mimic the structure of sialic acid in the transition state of the enzymatic reaction. Other prominent examples are  $\alpha$ -glucosidase inhibitors like miglitol (Glyset, Pfizer),<sup>(119)</sup> acarbose (Glucobay, Bayer),<sup>(120)</sup> and voglibose (Glustat, Takeda)<sup>(121)</sup> that are used for the treatment of type II diabetes by preventing the digestion of carbohydrates. Miglustat (Zavesca, Actelion)<sup>(122)</sup> inhibits glucosylceramide synthase and is prescribed for the treatment of type 1 Gaucher disease. The key to the successful carbohydrate-based drugs is to overcome the *pharmacokinetic drawbacks* inherently linked to native carbohydrates. Their polarity has to be systematically reduced by substitution of the hydroxyl groups not necessary for the biological activity. Subsequent optimization in terms of pharmacodynamics and pharmacokinetics (e.g. bioavailability, metabolic stability) give the *glycomimetic drug* an advantage over its natural competitor. An important prerequisite to further improve the success rate of carbohydrate-based drug candidates is a better understanding of carbohydrate-lectin interactions on a molecular basis.

### 3.4. Molecular basis of the regulation of carbohydrate-lectin affinity

Despite the structural diversity and diverse origins, lectins exhibit some common features probably related to their common interaction partners.<sup>(123)</sup> *Fast association and dissociation kinetics* and therefore short half-lives of the carbohydrate-lectin complexes characterize highly dynamic systems.<sup>(124)</sup> Furthermore, carbohydrates generally bind to their target structures with relatively *weak affinities*, often in the milli- to micromolar range.<sup>(73,125)</sup> This seems of particular importance under physiological conditions, as it allows the regulation of an interaction with higher resolution on different levels (Table 1).

**Table 1. The four levels of regulation of carbohydrate – lectin affinity.** From monomer to interaction interface to scaffold to cluster. Inspired by Gabius *et al.* (2011).<sup>(126)</sup>

Level	Carbohydrate		Lectin	
	Structure	Affinity regulation	Structure	Affinity regulation
Monomer	Monosaccharide	Substitution; bidentate interactions; cooperative hydrogen bonding; stacking	Amino acids	Charged interactions; cooperative hydrogen bonding; solvation; stacking
Interaction interface	Oligosaccharide	Conformational flexibility; Solvation	Binding site	Shape; dielectric constant; preorganization; solvation; structural ions
Scaffold	Glycan	Branching	Protein	Catch bond
Cluster	Glycocalyx	Multiple glycans; microdomains	Multivalency	Multiple lectins; lectin oligomers; multiple binding sites on one lectin

### 3.4.1. Affinity regulation: Carbohydrates

Not surprisingly, hydroxyl groups are the dominant feature when carbohydrates interact with lectins and at the same time, the main reason for the weak affinity. *Desolvation* of carbohydrates is enthalpically extremely costly because hydroxyl groups form strong interactions with the surrounding water molecules.<sup>(73)</sup> Hence, to compensate for this desolvation penalty a hydroxyl group needs to form strong *hydrogen bonds* with a lectin. Therefore, with complex hydrogen bond networks between protein and ligand, high selectivity and affinity can be reached. Moreover, several *substitutions* on carbohydrates (e.g. amine- and carboxyl groups, sulfation) are also involved in hydrogen bonding and *charged groups* can further increase the interaction energy. Hydroxyl groups frequently act simultaneously as both, donor and acceptor of a hydrogen bond (*bidentate hydrogen bonds*).<sup>(127)</sup> This is beneficial in three ways: first, a single desolvated group forming multiple interactions in the bound state allows more efficient compensation of the desolvation penalty for that group, because the penalty is paid once but compensated multiply. Second, freezing a rotational bond is penalized by entropic costs. Thus, only the first interaction is entropically costly, while for further interactions the hydroxyl group is already pre-organized. Third, the hydrogen bonds strengthen each other due to a phenomenon called ‘*cooperative hydrogen bonding*’.<sup>(18,128-130)</sup> Donation of a hydrogen bond by a hydroxyl group increases the electron density (partial charge) on the oxygen lone pair and improves its H-bond acceptance quality and *vice versa*. A less known feature of carbohydrate-lectin interactions is *C-H... $\pi$  stacking*.<sup>(131)</sup> Partially positive charged

hydrophobic C-H patches on the sugar surface are able to stack against the quadrupole of a  $\pi$  system of an aromatic amino acid side chain (Trp, Tyr, Phe). The enormous importance of the diverse linkage possibilities for the specificity of carbohydrate binding was already mentioned before. However, they also play a role in affinity regulation. Affinity to a lectin is obviously increased by increasing the maximum number of contacts, but also by affecting the *conformational flexibility/rigidity* of the oligosaccharide (e.g. core fucosylation).<sup>(126)</sup>

### 3.4.2. Affinity regulation: Lectins

On the monomer level (Table 1), similarities between monosaccharides and amino acids are detected: hydrogen bonds between carbohydrates and proteins are often *charge assisted and bidentate*.<sup>(127)</sup> Therefore, it is not surprising that the (at physiological pH) negatively charged aspartic and glutamic acid, as well as the positively charged arginine, are over-represented in the binding sites of lectins.<sup>(127,132)</sup> All of them are able to form bidentate hydrogen bonds. Furthermore, the presence of the aromatic amino acids tryptophan and tyrosine (*stacking*) is above the average.<sup>(132)</sup> On the level of the interaction interface (binding site), the evolution of lectins has resulted in two different adaptations to bind carbohydrates with physiologically acceptable affinities: Small and deep binding pockets binding monosaccharides (group I), or large and shallow grooves binding oligosaccharides (group II).<sup>(128)</sup> Deep buried binding pockets have the advantage of a lower *dielectric constant*, which increases the energy of electrostatic interactions according to Coulomb's law (equation 10). Hence, group I lectin often provide higher affinities and are predominantly observed in plant, viral, and bacterial lectins (adhesins), whereas group II lectins are highly specific and predominant among animal lectins. A well *pre-organized* binding site thereby facilitates higher affinity binding as it reduces the conformational entropy penalty upon binding.<sup>(125,133)</sup> However, conformational changes of the lectin might also be beneficial for the interaction lifetime despite a concomitant entropy penalty. An example is the so-called '*catch-bond behavior*' inducing a slower dissociation rate of a bound ligand by a conformational change.<sup>(134,135)</sup> A further adaptation to enhance affinity is the presence of salt-bridges with *strategically presented*  $Ca^{2+}$  ions (C-type lectins).

### 3.4.3. Multivalency

Despite the evolutionarily optimized interactions of lectins and glycans, low affinity remains a problem for some applications where firm adhesion is required. Nature's way to bridge the gap is 'multivalency'.<sup>(136)</sup> Multiple receptors bind to multiple ligands. There are, however, different types of organization clusters: a multi-antennary glycan, different glycans on one macromolecule or different glycans on different macromolecules (microdomains) may interact with different individual lectins, a lectin with multiple binding sites or lectin oligomers.<sup>(126,137)</sup> The binding strength of a multivalent interaction is not equal to the sum of the single binding affinities. It is influenced by structural conditions, but also by an increase of local concentration in proximity to a binding event. The increase of binding strength provided by multivalent interaction is described by the term 'avidity'.

### 3.5. Thermodynamics of carbohydrate-lectin interactions

Intuitively, one would assume that the abundant hydroxyl groups, which determine many properties of carbohydrate-lectin interactions, would also make a dominant contribution to the thermodynamics of binding. However, this is not the case.<sup>(125,133,138)</sup> A reaction in aqueous solution is too varied for a general trend to be discerned.<sup>(125)</sup> Nevertheless, Eric J. Toone (1994) recognized three recurrent trends when comparing different thermodynamic profiles of carbohydrate-lectin interactions: (i) a more negative change in enthalpy than in free energy of binding (enthalpy driven); (ii) a strong linear enthalpy-entropy compensatory behavior; (iii) a small and negative heat capacity (< 400 J/mol K). The negative enthalpy might be attributed to the hydrogen bond network and reorganization of disordered water near the protein surface. Still, it is somewhat surprising that only a few cases of entropy-driven carbohydrate-lectin interactions are described considering that lectin binding sites are well pre-organized and preferably bind the bioactive conformation of their carbohydrate ligand (*Manuscript I*).<sup>(133,139)</sup> However, it has to be noted that both, the enthalpy driven binding and enthalpy-entropy compensation, are not solely features of carbohydrates binding to lectins, but commonly observed phenomena of interactions in aqueous solution.<sup>(11,133)</sup> Hence, heat capacity is the only exclusive parameter related to carbohydrate-lectin interactions and reflects the reorganization of solvent.<sup>(140)</sup> Small  $\Delta C_p$  values indicate minor

solvent reorganization processes and are observed for the burial of polar surface area.<sup>(141)</sup> To quantify solvent reorganization for carbohydrate-lectin interactions, Chervenak and co-workers evaluated binding reactions in light ( $H_2O$ ) and heavy ( $D_2O$ ) water.<sup>(142)</sup> Solvent reorganization was found to contribute about 25-100% to the observed enthalpy, while the binding free energy remained unchanged due to compensatory effects in the entropy term.

A further method to obtain characteristic information of carbohydrate-lectin interactions is the systematic substitution of hydroxyl groups of the ligand (*Manuscript 1*). Many studies with monodeoxy analogs were published, but only in a few cases (Table 2) the thermodynamics of binding upon a loss of a hydroxyl group involved in a hydrogen bond could be determined. The most comprehensive monodeoxy studies were carried out by Fred Brewer and co-workers who investigated the binding of concanavalin A to deoxy analogs of 3,6-di-O-( $\alpha$ -D-mannopyranosyl)-D-mannose.<sup>(143)</sup> His research group extended the investigations to other (all structurally and thermodynamically related) *Leguminosae* lectins,<sup>(144,145)</sup> supported thermodynamic data with structural information,<sup>(146)</sup> and determined the solvent reorganization effect in  $D_2O$ .<sup>(147)</sup> The measurements in  $D_2O$  revealed that the solvation of deoxy ligands alters significantly with the loss of a hydroxyl group. These desolvation costs were considered in the study Daranas *et al.* (2004)<sup>(148)</sup> who concluded that the intrinsic free energy of D-galactose binding to the L-Arabinose-binding protein is remarkably constant (30 kJ/mol), independent on the number of hydrogen bonds formed. In their publication on the thermodynamics of deoxy analogs of 3,6-di-O-( $\alpha$ -D-mannopyranosyl)-D-mannose binding to artocarpin, Surolia and co-workers<sup>(149)</sup> pointed out the nonlinearity of the enthalpy and free energy contributions to binding of individual hydroxyl groups. For example, the sum of all  $\Delta\Delta H^\circ$  values for monodeoxy analogs (295.4 kJ/mol) was almost 7 times greater than the  $\Delta H^\circ$  of the unmodified reference trimannoside (-44.8 kJ/mol). They speculated that the loss in enthalpy and free energy upon the loss of a hydroxyl group represents different contributions of the solvent and the protein and not only those related to the loss of the hydrogen bond. All of the mentioned lectins (the plant lectins concanavalin A and artocarpin and the bacterial lectin L-Arabinose binding protein) share the similarity of deeply buried binding pocket. Hence, the thermodynamic fingerprint for the loss of a

**Table 2. Overview of thermodynamic parameters of lectins binding to monodeoxy ligands.** Concanavalin A is representatively shown for the *Leguminosae* lectin family, which all exhibit closely related thermodynamics of binding.

Protein / Ligand / Temperature	Deoxy analog	$\Delta\Delta G^\circ$ [kJ/mol]	$\Delta\Delta H^\circ$ [kJ/mol]	$-\Delta T\Delta S^\circ$ [kJ/mol]	Interaction	Ref
Concanavalin A / 3,6-di-O-( $\alpha$ -D- mannopyranosyl)-D- mannose / 300 K	Man3-3H	5.5	14.2	-8.7	2	(143,144)
	Man3-4H	4.2	8.8	-4.6	2	
	Man6-3H	6.3	13.4	-7.1	1	
	Man6-4H	7.4	11.3	-3.9	2	
	Man6-6H	7.0	11.7	-4.7	2	
	ManC-2H	3.6	4.2	-0.6	1	
	ManC-4H	5.9	9.6	-3.7	1	
Artocarpin / 3,6-di-O-( $\alpha$ -D- mannopyranosyl)-D- mannose / 293 K	Man1-3H	4.2	33.8	-29.7	1	(149)
	Man1-4H	6.4	39.0	-32.6	1	
	Man1-6H	n.b.	n.b.	n.b.	3	
	Man2-2H	6.8	29.5	-22.7	2	
L-Arabinose binding protein / D-Galactose / 308 K	Gal-1H	22.6	32.0	-9.0	1	(148)
	Gal-2H	15.1	34.0	-18.0	2 (1 via H <sub>2</sub> O)	
	Gal-3H	24.4	38.0	-13.0	3	
	Gal-4H	n.b.	n.b.	n.b.	2	
	Gal-6H	6.3	2.0	5.0	2 (1 via H <sub>2</sub> O)	
Calreticulin / Glc $\alpha$ 1-3Man- $\alpha$ 1- 2Man / 303 K	Glc1-6H	-1.1	22.3	-23.4	1 via H <sub>2</sub> O	(150)
	Glc1-3H	8.7	0.9	7.8	1	
	Man2-4H	0.9	0.1	0.8	1	
	Man2-6H	7.4	-3.5	10.9	1	
	Man3-4H	7.2	-8.7	16.0	1	
	Man3-3H	0.5	-0.4	0.9	1 via H <sub>2</sub> O	

hydroxyl group is also very uniform: the loss in enthalpy exceeds the loss in free energy due to a partial compensation by the entropy. In contrast, the thermodynamic fingerprint of monodeoxy analogs of Glc $\alpha$ 1-3Man- $\alpha$ 1-2Man binding to the shallow binding site of the human chaperone calreticulin is more heterogeneous.<sup>(150)</sup> The loss of 6-OH of glucose results in a gain of free energy of binding due to an overcompensating gain in entropy, while the loss of the hydrogen bonds formed by 4-OH on the central mannose moiety and 3-OH on the terminal mannose moiety have barely an effect on the binding properties. Interestingly, the removal of 6-OH of the central mannose unit and 4-OH on the terminal mannose unit leads to a decreased free energy of binding due to an unbeneficial entropy term. Unfortunately, the authors of the study do not comment on these unusual findings. In summary, the comparison of conservatively modified ligands (or proteins) (e.g.  $\Delta\Delta H^\circ$ ,  $\Delta\Delta G^\circ$  and  $-\Delta T\Delta S^\circ$  values) allow a more meaningful interpretation of binding characteristics than the bare cross-comparison of thermodynamic fingerprints of lectins binding carbohydrates

### 4. References

1. Fischer, E. (1894). Einfluss der Configuration auf die Wirkung der Enzyme. *Ber. Dtsch. Chem. Ges.*, 27(3): 2985-2993.
2. Ehrlich, P. (1913). Chemotherapeutics: Scientific Principles, Methods, and Results. *The Lancet*, 182(4694): 445-451.
3. Bosch, F. & Rosich, L. (2008). The contributions of Paul Ehrlich to pharmacology: a tribute on the occasion of the centenary of his Nobel Prize. *Pharmacology*, 82(3): 171-179.
4. Zalewski, A. (2013). *In pursuit of a novel UTI treatment strategy - an "in silico" study of the FimH adhesin*. PhD Dissertation, University of Basel, Basel.
5. Koshland, D.E. (1958). Application of a Theory of Enzyme Specificity to Protein Synthesis. *Proc. Natl. Acad. Sci. U. S. A.*, 44(2): 98-104.
6. Koshland, D.E. (1995). The Key–Lock Theory and the Induced Fit Theory. *Angew. Chem. Int. Ed.*, 33(23-24): 2375-2378.
7. Tsai, C.J., Kumar, S., Ma, B., & Nussinov, R. (1999). Folding funnels, binding funnels, and protein function. *Protein Sci.*, 8(6): 1181-1190.
8. Csermely, P., Palotai, R., & Nussinov, R. (2010). Induced fit, conformational selection and independent dynamic segments: an extended view of binding events. *Trends Biochem. Sci.*, 35(10): 539-546.
9. Guldberg, C.M. & Waage, P. (1879). Ueber die chemische Affinität. *J. Prakt. Chem.*, 19(1): 69-114.
10. Gibbs, J.W. (1871). A Method of Geometrical Representation of the Thermodynamic Properties of Substances by Means of Surfaces. *Transactions of the Connecticut Academy of Arts and Sciences*, The Connecticut Academy of Arts and Sciences. pp. 382-404.
11. Dunitz, J.D. (1995). Win some, lose some: enthalpy-entropy compensation in weak intermolecular interactions. *Chem. Biol.*, 2(11): 709-712.
12. Lumry, R. & Rajender, S. (1970). Enthalpy-entropy compensation phenomena in water solutions of proteins and small molecules: a ubiquitous property of water. *Biopolymers*, 9(10): 1125-1227.
13. Cabani, S., Gianni, P., Mollica, V., & Lepori, L. (1981). Group contributions to the thermodynamic properties of non-ionic organic solutes in dilute aqueous solution. *J. Solution Chem.*, 10(8): 563-595.
14. Perutz, M.F. (1978). Electrostatic Effects in Proteins. *Science*, 201(4362): 1187-1191.
15. Li, L., Li, C., Zhang, Z., & Alexov, E. (2013). On the Dielectric "Constant" of Proteins: Smooth Dielectric Function for Macromolecular Modeling and Its Implementation in DelPhi. *J. Chem. Theory Comput.*, 9(4): 2126-2136.
16. Fersht, A. (1999). Forces between molecules, and binding energies. *Structure and mechanism in protein science: a guide to enzyme catalysis and protein folding*, W. H. Freeman and Company, New York. pp. 324-348.
17. Bissantz, C., Kuhn, B., & Stahl, M. (2010). A Medicinal Chemist's Guide to Molecular Interactions. *J. Med. Chem.*, 53(14): 5061-5084.
18. Jeffrey, G.A. & Saenger, W. (1991). *Hydrogen Bonding in Biological Structures*. Springer-Verlag, Berlin.
19. Perrin, C.L. & Nielson, J.B. (1997). "Strong" Hydrogen Bonds in Chemistry and Biology. *Annu. Rev. Phys. Chem.*, 48(1): 511-544.
20. Metrangolo, P., Neukirch, H., Pilati, T., & Resnati, G. (2005). Halogen bonding based recognition processes: A world parallel to hydrogen bonding. *Acc. Chem. Res.*, 38(5): 386-395.
21. Politzer, P., Lane, P., Concha, M.C., Ma, Y.G., & Murray, J.S. (2007). An overview of halogen bonding. *J. Mol. Model.*, 13(2): 305-311.
22. Clark, T., Hennemann, M., Murray, J., & Politzer, P. (2007). Halogen bonding: the  $\sigma$ -hole. *J. Mol. Model.*, 13(2): 291-296.
23. Metrangolo, P., Meyer, F., Pilati, T., Resnati, G., & Terraneo, G. (2008). Halogen bonding in supramolecular chemistry. *Angew. Chem. Int. Ed.*, 47(33): 6114-6127.
24. Gohlke, H. & Klebe, G. (2002). Approaches to the Description and Prediction of the Binding Affinity of Small-Molecule Ligands to Macromolecular Receptors. *Angew. Chem. Int. Ed.*, 41(15): 2644-2676.
25. Dunitz, J.D. & Gavezzotti, A. (1999). Attractions and Repulsions in Molecular Crystals: What Can Be Learned from the Crystal Structures of Condensed Ring Aromatic Hydrocarbons? *Acc. Chem. Res.*, 32(8): 677-684.

26. Smith, R.D., Engdahl, A.L., Dunbar Jr, J.B., & Carlson, H.A. (2012). Biophysical limits of protein–ligand binding. *J. Chem. Inf. Model.*, 52(8): 2098-2106.
27. Jones, J.E. (1924). On the determination of molecular fields. II. From the equation of state of a gas. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, The Royal Society, pp. 463-477.
28. Bondi, A. (1964). Van der Waals Volumes and Radii. *J. Phys. Chem.*, 68(3): 441-451.
29. Hunter, C.A. & Sanders, J.K.M. (1990). The nature of  $\pi$ - $\pi$  interactions. *J. Am. Chem. Soc.*, 112(14): 5525-5534.
30. Raju, R.K., Bloom, J.W., An, Y., & Wheeler, S.E. (2011). Substituent effects on non-covalent interactions with aromatic rings: insights from computational chemistry. *ChemPhysChem*, 12(17): 3116-3130.
31. Dougherty, D.A. (1996). Cation- $\pi$  interactions in chemistry and biology: a new view of benzene, Phe, Tyr, and Trp. *Science*, 271(5246): 163-168.
32. Bronowska, A.K. (2011). *Thermodynamics of ligand-protein interactions: implications for molecular design*. INTECH Open Access Publisher. pp. 1-48.
33. Baker, B.M. & Murphy, K.P. (1997). Dissecting the energetics of a protein-protein interaction: the binding of ovomucoid third domain to elastase. *J. Mol. Biol.*, 268(2): 557-569.
34. Cedervik, F.H., *et al.* (2007). Thermodynamic analysis of allosamidin binding to a family 18 chitinase. *Biochemistry*, 46(43): 12347-12354.
35. Murphy, K.P., Privalov, P.L., & Gill, S.J. (1990). Common features of protein unfolding and dissolution of hydrophobic compounds. *Science*, 247(4942): 559.
36. Baldwin, R.L. (1986). Temperature dependence of the hydrophobic interaction in protein folding. *Proc. Natl. Acad. Sci. U. S. A.*, 83(21): 8069-8072.
37. Chervenak, M.C. & Toone, E.J. (1995). Calorimetric Analysis of the Binding of Lectins with Overlapping Carbohydrate-Binding Ligand Specificities. *Biochemistry*, 34(16): 5685-5695.
38. Williams, D.H., Stephens, E., O'Brien, D.P., & Zhou, M. (2004). Understanding Noncovalent Interactions: Ligand Binding Energy and Catalytic Efficiency from Ligand-Induced Reductions in Motion within Receptors and Enzymes. *Angew. Chem. Int. Ed.*, 43(48): 6596-6616.
39. Murphy, K.P., Xie, D., Thompson, K.S., Amzel, L.M., & Freire, E. (1994). Entropy in biological binding processes: estimation of translational entropy loss. *Proteins: Struct. Funct. Bioinform.*, 18(1): 63-67.
40. Irudayam, S.J. & Henchman, R.H. (2009). Entropic Cost of Protein–Ligand Binding and Its Dependence on the Entropy in Solution. *J. Phys. Chem. B*, 113(17): 5871-5884.
41. Lam, P., *et al.* (1994). Rational design of potent, bioavailable, nonpeptide cyclic ureas as HIV protease inhibitors. *Science*, 263(5145): 380-384.
42. Ladbury, J.E. (1996). Just add water! The effect of water on the specificity of protein-ligand binding sites and its potential application to drug design. *Chem. Biol.*, 3(12): 973-980.
43. Poornima, C.S. & Dean, P.M. (1995). Hydration in drug design. 2. Influence of local site surface shape on water binding. *J. Comput. Aided Mol. Des.*, 9(6): 513-520.
44. Jencks, W.P. (1987). *Catalysis in chemistry and enzymology*. Courier Corporation.
45. Dunitz, J.D. (1994). The entropic cost of bound water in crystals and biomolecules. *Science*, 264(5159): 670-670.
46. Poornima, C. & Dean, P. (1995). Hydration in drug design. 3. Conserved water molecules at the ligand-binding sites of homologous proteins. *J. Comput. Aided Mol. Des.*, 9(6): 521-531.
47. Kauzmann, W. (1959). Some Factors in the Interpretation of Protein Denaturation. *Adv. Protein Chem.*, eds C.B. Anfinsen MLAKB & John TE, Academic Press. Vol. Volume 14, pp. 1-63.
48. Ben-Naim, A. (1987). *Solvation thermodynamics*. Plenum Press, New York. p. 246.
49. Ben-Naim, A. (1980). *Hydrophobic interactions*. Plenum Press, New York. p. 311.
50. Dill, K.A. (1990). Dominant forces in protein folding. *Biochemistry*, 29(31): 7133-7155.
51. Silverstein, K.A.T., Haymet, A.D.J., & Dill, K.A. (1998). A Simple Model of Water and the Hydrophobic Effect. *J. Am. Chem. Soc.*, 120(13): 3166-3175.
52. Meyer, E.A., Castellano, R.K., & Diederich, F. (2003). Interactions with aromatic rings in chemical and biological recognition. *Angew. Chem. Int. Ed.*, 42(11): 1210-1250.
53. Syme, N.R., Dennis, C., Phillips, S.E., & Homans, S.W. (2007). Origin of heat capacity changes in a “nonclassical” hydrophobic interaction. *ChemBioChem*, 8(13): 1509-1511.
54. Muller, N. (1990). Search for a realistic view of hydrophobic effects. *Acc. Chem. Res.*, 23(1): 23-28.
55. Blokzijl, W. & Engberts, J.B.F.N. (1993). Hydrophobic Effects. Opinions and Facts. *Angew. Chem. Int. Ed.*, 32(11): 1545-1579.

56. Turnbull, P.A. & Emsley, P. (2013). Studying Protein–Ligand Interactions Using X-Ray Crystallography. *Protein-Ligand Interactions*, Methods in Molecular Biology, eds Williams MA & Daviter T, Humana Press. Vol. 1008, pp. 457-478.
57. Bernadó, P., Mylonas, E., Petoukhov, M.V., Blackledge, M., & Svergun, D.I. (2007). Structural characterization of flexible proteins using small-angle X-ray scattering. *J. Am. Chem. Soc.*, 129(17): 5656-5664.
58. Liao, M.F., Cao, E.H., Julius, D., & Cheng, Y.F. (2013). Structure of the TRPV1 ion channel determined by electron cryo-microscopy. *Nature*, 504(7478): 107-112.
59. Müller, D.J. & Dufrene, Y.F. (2008). Atomic force microscopy as a multifunctional molecular toolbox in nanobiotechnology. *Nat. Nanotechnol.*, 3(5): 261-269.
60. Binnig, G., Quate, C.F., & Gerber, C. (1986). Atomic force microscope. *Phys. Rev. Lett.*, 56(9): 930-933.
61. Schwede, T., Kopp, J., Guex, N., & Peitsch, M.C. (2003). SWISS-MODEL: an automated protein homology-modeling server. *Nucleic Acids Res.*, 31(13): 3381-3385.
62. Jorgensen, W.L. (2004). The many roles of computation in drug discovery. *Science*, 303(5665): 1813-1818.
63. Sattler, M., Schleucher, J., & Griesinger, C. (1999). Heteronuclear multidimensional NMR experiments for the structure determination of proteins in solution employing pulsed field gradients. *Prog. Nucl. Magn. Reson. Spectrosc.*, 34(2): 93-158.
64. Grzesiek, S. & Sass, H.-J. (2009). From biomolecular structure to functional understanding: new NMR developments narrow the gap. *Curr. Opin. Struct. Biol.*, 19(5): 585-595.
65. Kleckner, I.R. & Foster, M.P. (2011). An introduction to NMR-based approaches for measuring protein dynamics. *BBA-Proteins Proteom.*, 1814(8): 942-968.
66. Mittermaier, A.K. & Kay, L.E. (2009). Observing biological dynamics at atomic resolution using NMR. *Trends Biochem. Sci.*, 34(12): 601-611.
67. Cala, O., Guillièrre, F., & Krimm, I. (2014). NMR-based analysis of protein–ligand interactions. *Anal. Bioanal. Chem.*, 406(4): 943-956.
68. Meyer, B. & Peters, T. (2003). NMR spectroscopy techniques for screening and identifying ligand binding to protein receptors. *Angew. Chem. Int. Ed.*, 42(8): 864-890.
69. Peng, J.W., Moore, J., & Abdul-Manan, N. (2004). NMR experiments for lead generation in drug discovery. *Prog. Nucl. Magn. Reson. Spectrosc.*, 44(3): 225-256.
70. Williamson, M.P. (2013). Using chemical shift perturbation to characterise ligand binding. *Prog. Nucl. Magn. Reson. Spectrosc.*, 73: 1-16.
71. Shuker, S.B., Hajduk, P.J., Meadows, R.P., & Fesik, S.W. (1996). Discovering high-affinity ligands for proteins: SAR by NMR. *Science*, 274(5292): 1531-1534.
72. Stillmark, H. (1888). *Über Ricin ein giftiges Ferment aus den Samen von Ricinus comm. L. und einige anderen Euphorbiaceen*. Inaugural-Dissertation, Kaiserliche Universität zu Dorpat, Tartu.
73. Burkhalter, N.F., et al. (2008). Protein-Carbohydrate Interaction: Fundamental Considerations. *Carbohydrates in Chemistry and Biology*, Wiley-VCH Verlag GmbH. pp. 863-914.
74. Christensen, T. & Toone, E.J. (2003). Calorimetric evaluation of protein-carbohydrate affinities. *Methods Enzymol.*, 362: 486.
75. Abgottspon, D., et al. (2010). Development of an aggregation assay to screen FimH antagonists. *J. Microbiol. Methods*, 82(3): 249-255.
76. Rabbani, S., Jiang, X., Schwardt, O., & Ernst, B. (2010). Expression of the carbohydrate recognition domain of FimH and development of a competitive binding assay. *Anal. Biochem.*, 407(2): 188-195.
77. Wang, Z.-X. (1995). An exact mathematical expression for describing competitive binding of two different ligands to a protein molecule. *FEBS Lett.*, 360(2): 111-114.
78. Kleeb, S., et al. (2015). FimH antagonists: bioisosteres to improve the in vitro and in vivo PK/PD profile. *J. Med. Chem.*, 58(5): 2221-2239.
79. Cooper, A. (2004). Thermodynamics and interactions. *Biophys. Chem.*, ed Cooper A, The Royal Society of Chemistry. 2nd Ed, Vol. 16, pp. 99-122.
80. Jerabek-Willemsen, M., Wienken, C.J., Braun, D., Baaske, P., & Duhr, S. (2011). Molecular interaction studies using microscale thermophoresis. *Assay Drug Dev. Technol.*, 9(4): 342-353.
81. Duhr, S. & Braun, D. (2006). Why molecules move along a temperature gradient. *Proc. Natl. Acad. Sci. U. S. A.*, 103(52): 19678-19682.
82. Guo, X. (2012). Surface plasmon resonance based biosensor technique: a review. *J. Biophotonics*, 5(7): 483-501.

83. Perozzo, R., Folkers, G., & Scapozza, L. (2004). Thermodynamics of protein–ligand interactions: history, presence, and future aspects. *J. Recept. Signal Transduct.*, 24(1-2): 1-52.
84. Burnouf, D., *et al.* (2012). kinITC: a new method for obtaining joint thermodynamic and kinetic data by isothermal titration calorimetry. *J. Am. Chem. Soc.*, 134(1): 559-565.
85. Holdgate, G.A. & Ward, W.H. (2005). Measurements of binding thermodynamics in drug discovery. *Drug Discov. Today*, 10(22): 1543-1550.
86. Wiseman, T., Williston, S., Brandts, J.F., & Lin, L.-N. (1989). Rapid measurement of binding constants and heats of binding using a new titration calorimeter. *Anal. Biochem.*, 179(1): 131-137.
87. Broecker, J., Vargas, C., & Keller, S. (2011). Revisiting the optimal cvalue for isothermal titration calorimetry. *Anal. Biochem.*, 418(2): 307-309.
88. Sigurskjold, B.W. (2000). Exact analysis of competition ligand binding by displacement isothermal titration calorimetry. *Anal. Biochem.*, 277(2): 260-266.
89. Turnbull, W.B. & Daranas, A.H. (2003). On the value of c: can low affinity systems be studied by isothermal titration calorimetry? *J. Am. Chem. Soc.*, 125(48): 14859-14866.
90. Watkins, W.M. & Morgan, W.T.J. (1952). Neutralization of the anti-H agglutinin in eel serum by simple sugars. *Nature*, 169(4307): 825-826.
91. Helenius, A. & Aebi, M. (2004). Roles of N-linked glycans in the endoplasmic reticulum. *Annu. Rev. Biochem.*, 73(1): 1019-1049.
92. Ohtsubo, K. & Marth, J.D. (2006). Glycosylation in Cellular Mechanisms of Health and Disease. *Cell*, 126(5): 855-867.
93. Yu, H. & Chen, X. (2007). Carbohydrate post-glycosylational modifications. *Org. Biomol. Chem.*, 5(6): 865-872.
94. Hartmann, M. & Lindhorst, T.K. (2011). The Bacterial Lectin FimH, a Target for Drug Discovery – Carbohydrate Inhibitors of Type 1 Fimbriae-Mediated Bacterial Adhesion. *Eur. J. Org. Chem.*, 2011(20-21): 3583-3609.
95. Sumner, J.B. & Howell, S.F. (1936). Identification of Hemagglutinin of Jack Bean with Concanavalin A. *J. Bacteriol.*, 32(2): 227-237.
96. Varki, A., Etzler, M.E., Cummings, R.D., & Esko, J.D. (2009). Discovery and Classification of Glycan-Binding Proteins. *Essentials of Glycobiology*, eds Varki A, Cummings RD, Esko JD, Freeze HH, Stanley P, Bertozzi CR, Hart GW, & Etzler ME, Cold Spring Harbor (NY). 2nd Ed.
97. Mouricout, M. (1997). Interactions between the enteric pathogen and the host. An assortment of bacterial lectins and a set of glycoconjugate receptors. *Adv. Exp. Med. Biol.*, 412: 109-123.
98. Gottschalk, A. (1959). Chemistry of virus receptors. *The viruses: biochemical, biological and biophysical properties*, eds Burnet FM & Stanley WM, Academic Press, Inc., New York, NY. pp. 51-61.
99. Ofek, I., Mirelman, D., & Sharon, N. (1977). Adherence of Escherichia coli to human mucosal cells mediated by mannose receptors. *Nature*, 265(5595): 623-625.
100. Old, D.C. (1972). Inhibition of the interaction between fimbrial haemagglutinins and erythrocytes by D-mannose and other carbohydrates. *J. Gen. Microbiol.*, 71(1): 149-157.
101. Drickamer, K. (1988). Two distinct classes of carbohydrate-recognition domains in animal lectins. *J. Biol. Chem.*, 263(20): 9557-9560.
102. Gupta, G. & Surolia, A. (2007). Collectins: sentinels of innate immunity. *Bioessays*, 29(5): 452-464.
103. Geijtenbeek, T.B.H., *et al.* (2000). DC-SIGN, a dendritic cell-specific HIV-1-binding protein that enhances trans-infection of T cells. *Cell*, 100(5): 587-597.
104. McEver, R.P. (2002). Selectins: lectins that initiate cell adhesion under flow. *Curr. Opin. Cell Biol.*, 14(5): 581-586.
105. Dahms, N.M. & Hancock, M.K. (2002). P-type lectins. *BBA-Gen. Subjects*, 1572(2-3): 317-340.
106. Leffler, H., Carlsson, S., Hedlund, M., Qian, Y., & Poirier, F. (2004). Introduction to galectins. *Glycoconj. J.*, 19(7-9): 433-440.
107. Angata, T. & Brinkman-Van der Linden, E.C.M. (2002). I-type lectins. *BBA-Gen. Subjects*, 1572(2-3): 294-316.
108. Crocker, P.R., Paulson, J.C., & Varki, A. (2007). Siglecs and their roles in the immune system. *Nat. Rev. Immunol.*, 7(4): 255-266.
109. Lever, R., Mulloy, B., & Page, C.P. (2012). *Heparin-a century of progress*. Springer Science & Business Media.
110. Necas, J., Bartosikova, L., Brauner, P., & Kolar, J. (2008). Hyaluronic acid (hyaluronan): a review. *Vet. Med. Czech*, 53(8): 397-411.
111. Cipolla, L., *et al.* (2010). Discovery and design of carbohydrate-based therapeutics. *Expert Opin. Drug Discov.*, 5(8): 721-737.

112. Boltje, T.J., Buskas, T., & Boons, G.J. (2009). Opportunities and challenges in synthetic oligosaccharide and glycoconjugate research. *Nat. Chem.*, 1(8): 611-622.
113. Petitou, M. & van Boeckel, C.A. (2004). A synthetic antithrombin III binding pentasaccharide is now a drug! What comes next? *Angew. Chem. Int. Ed.*, 43(24): 3118-3133.
114. Ernst, B. & Magnani, J.L. (2009). From carbohydrate leads to glycomimetic drugs. *Nat. Rev. Drug Discov.*, 8(8): 661-677.
115. Bielik, A. & Zaia, J. (2010). Historical Overview of Glycoanalysis. *Functional Glycomics, Methods in Molecular Biology*, ed Li J, Humana Press. Vol. 600, pp. 9-30.
116. Chang, J., *et al.* (2010). GMI-1070, a novel pan-selectin antagonist, reverses acute vascular occlusions in sickle cell mice. *Blood*, 116(10): 1779-1786.
117. Kim, C.U., *et al.* (1997). Influenza neuraminidase inhibitors possessing a novel hydrophobic interaction in the enzyme active site: design, synthesis, and structural analysis of carbocyclic sialic acid analogues with potent anti-influenza activity. *J. Am. Chem. Soc.*, 119(4): 681-690.
118. von Itzstein, M., *et al.* (1993). Rational design of potent sialidase-based inhibitors of influenza virus replication. *Nature*, 363(6428): 418-423.
119. Campbell, L.K., Baker, D.E., & Campbell, R.K. (2000). Miglitol: assessment of its role in the treatment of patients with diabetes mellitus. *Ann. Pharmacother.*, 34(11): 1291-1301.
120. Truscheit, E., *et al.* (1981). Chemistry and Biochemistry of Microbial  $\alpha$ -Glucosidase Inhibitors. *Angew. Chem. Int. Ed.*, 20(9): 744-761.
121. Chen, X.L., Zheng, Y.G., & Shen, Y.C. (2006). Voglibose (Basen (R) AO-128), one of the most important alpha-glucosidase inhibitors. *Curr. Med. Chem.*, 13(1): 109-116.
122. Weinreb, N.J., *et al.* (2005). Guidance on the use of miglustat for treating patients with type 1 Gaucher disease. *Am. J. Hematol.*, 80(3): 223-229.
123. del Carmen Fernandez-Alonso, M., *et al.* (2012). Protein-carbohydrate interactions studied by NMR: from molecular recognition to drug design. *Curr. Protein Pept. Sci.*, 13(8): 816-830.
124. Scharenberg, M., *et al.* (2014). Kinetic Properties of Carbohydrate–Lectin Interactions: FimH Antagonists. *ChemMedChem*, 9(1): 78-83.
125. Toone, E.J. (1994). Structure and energetics of protein-carbohydrate complexes. *Curr. Opin. Struct. Biol.*, 4(5): 719-728.
126. Gabius, H.-J., André, S., Jiménez-Barbero, J., Romero, A., & Solís, D. (2011). From lectin structure to functional glycomics: principles of the sugar code. *Trends Biochem. Sci.*, 36(6): 298-313.
127. Quijcho, F.A. (1986). Carbohydrate-binding proteins: tertiary structures and protein-sugar interactions. *Annu. Rev. Biochem.*, 55: 287-315.
128. Vyas, N.K. (1991). Atomic features of protein-carbohydrate interactions. *Curr. Opin. Struct. Biol.*, 1(5): 732-740.
129. López de la Paz, M., *et al.* (2002). Carbohydrate Hydrogen-Bonding Cooperativity – Intramolecular Hydrogen Bonds and Their Cooperative Effect on Intermolecular Processes – Binding to a Hydrogen-Bond Acceptor Molecule. *Eur. J. Org. Chem.*, 2002(5): 840-855.
130. Frank, H.S. & Wen, W.-Y. (1957). Ion-solvent interaction. Structural aspects of ion-solvent interaction in aqueous solutions: a suggested picture of water structure. *Discuss. Faraday Soc.*, 24(0): 133-140.
131. Asensio, J.L., Ardá, A., Cañada, F.J., & Jiménez-Barbero, J. (2013). Carbohydrate–Aromatic Interactions. *Acc. Chem. Res.*, 46(4): 946-954.
132. Taroni, C., Jones, S., & Thornton, J.M. (2000). Analysis and prediction of carbohydrate binding sites. *Protein Eng.*, 13(2): 89-98.
133. Ambrosi, M., Cameron, N.R., & Davis, B.G. (2005). Lectins: tools for the molecular understanding of the glycode. *Org. Biomol. Chem.*, 3(9): 1593-1608.
134. Thomas, W.E., Trintchina, E., Forero, M., Vogel, V., & Sokurenko, E.V. (2002). Bacterial adhesion to target cells enhanced by shear force. *Cell*, 109(7): 913-923.
135. Marshall, B.T., *et al.* (2003). Direct observation of catch bonds involving cell-adhesion molecules. *Nature*, 423(6936): 190-193.
136. Rini, J.M. (1995). Lectin structure. *Annu. Rev. Biophys. Biomol. Struct.*, 24: 551-577.
137. Reynolds, M. & Pérez, S. (2011). Thermodynamics and chemical characterization of protein-carbohydrate interactions: The multivalency issue. *C. R. Chim.*, 14(1): 74-95.
138. Dam, T.K. & Brewer, C.F. (2002). Thermodynamic studies of lectin-carbohydrate interactions by isothermal titration calorimetry. *Chem. Rev.*, 102(2): 387-429.
139. Roldós, V., Cañada, F.J., & Jiménez-Barbero, J. (2011). Carbohydrate–Protein Interactions: A 3D View by NMR. *ChemBioChem*, 12(7): 990-1005.

140. Sturtevant, J.M. (1977). Heat capacity and entropy changes in processes involving proteins. *Proc. Natl. Acad. Sci. U. S. A.*, 74(6): 2236-2240.
141. Spolar, R.S., Ha, J.H., & Record, M.T. (1989). Hydrophobic effect in protein folding and other noncovalent processes involving proteins. *Proc. Natl. Acad. Sci. U. S. A.*, 86(21): 8382-8385.
142. Chervenak, M.C. & Toone, E.J. (1994). A Direct Measure of the Contribution of Solvent Reorganization to the Enthalpy of Binding. *J. Am. Chem. Soc.*, 116(23): 10533-10539.
143. Gupta, D., Dam, T.K., Oscarson, S., & Brewer, C.F. (1997). Thermodynamics of lectin-carbohydrate interactions. Binding of the core trimannoside of asparagine-linked carbohydrates and deoxy analogs to concanavalin A. *J. Biol. Chem.*, 272(10): 6388-6392.
144. Dam, T.K., Oscarson, S., & Brewer, C.F. (1998). Thermodynamics of binding of the core trimannoside of asparagine-linked carbohydrates and deoxy analogs to Dioclea grandiflora lectin. *J. Biol. Chem.*, 273(49): 32812-32817.
145. Dam, T.K., *et al.* (2000). Thermodynamic binding studies of lectins from the diocleinae subtribe to deoxy analogs of the core trimannoside of asparagine-linked oligosaccharides. *J. Biol. Chem.*, 275(21): 16119-16126.
146. Rozwarski, D.A., Swami, B.M., Brewer, C.F., & Sacchettini, J.C. (1998). Crystal structure of the lectin from Dioclea grandiflora complexed with core trimannoside of asparagine-linked carbohydrates. *J. Biol. Chem.*, 273(49): 32818-32825.
147. Dam, T.K., Oscarson, S., Sacchettini, J.C., & Brewer, C.F. (1998). Differential solvation of "core" trimannoside complexes of the Dioclea grandiflora lectin and concanavalin A detected by primary solvent isotope effects in isothermal titration microcalorimetry. *J. Biol. Chem.*, 273(49): 32826-32832.
148. Daranas, A.H., Shimizu, H., & Homans, S.W. (2004). Thermodynamics of binding of D-galactose and deoxy derivatives thereof to the L-arabinose-binding protein. *J. Am. Chem. Soc.*, 126(38): 11870-11876.
149. Rani, P.G., Bachhawat, K., Reddy, G.B., Oscarson, S., & Surolia, A. (2000). Isothermal titration calorimetric studies on the binding of deoxytrimannoside derivatives with artocarpin: implications for a deep-seated combining site in lectins. *Biochemistry*, 39(35): 10755-10760.
150. Gupta, G., Gemma, E., Oscarson, S., & Surolia, A. (2008). Defining substrate interactions with calreticulin: an isothermal titration calorimetric study. *Glycoconj. J.*, 25(8): 797-802.



## **Section II.**

# **The Bacterial Adhesin FimH**



### 1. Introduction

#### 1.1. Urinary tract infection

Urinary tract infections (UTI) are among the most prevalent bacterial infections worldwide causing high medical costs.<sup>(1,2)</sup> During these infections, gut bacteria colonize in the urethra and ascend to the bladder (lower urinary tract infection, cystitis) and the kidneys (upper urinary tract infection, pyelonephritis). 50% of women experience a symptomatic UTI at least once during their lifetime.<sup>(2)</sup> About 25% of the UTIs recur within 6 months caused by a genotypically similar bacteria strain.<sup>(3)</sup> Typical patients are young, sexually active women and elderly women with a catheter or diabetes.<sup>(4)</sup> The standard therapy is a three-day antibiotic treatment. However, rising numbers of bacterial resistance to antibiotics and high reoccurrence rates reveal the need for an alternative treatment.<sup>(5)</sup>

#### 1.2. Infection cycle of uropathogenic *Escherichia coli*

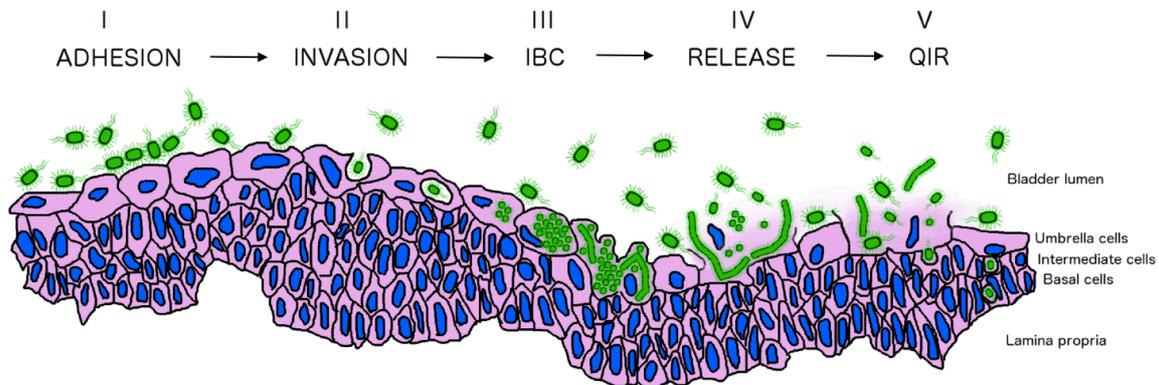
*Escherichia coli* (*E. coli*) is a usually mutual, gram-negative, facultatively anaerobic bacterium of the human gut flora. Nevertheless, some serotypes of *E. coli* are pathogenic, among them uropathogenic *E. coli* (UPEC). UPEC causes up to 90% of all UTIs<sup>(6)</sup> and typically expresses adhesins allowing the attachment to the urinary tract epithelium. The adhesion is the initial step<sup>(7)</sup> of an infection cycle,<sup>(8)</sup> which increases the persistence of UPEC in the bladder:

- 1. Adhesion:** UPEC are able to express type 1 pili with a carbohydrate binding lectin (FimH) at their tip. This enables the interaction with uroplakin 1a (UPK1a), which is an integral component of the luminal membrane of urothelial cells.<sup>(9)</sup> UPK1a is highly mannosylated with N-linked carbohydrates and therefore presents mannose carbohydrates to bladder lumen. FimH specifically binds to these mannosides. The interaction with the bladder epithelium prevents UPEC from being flushed out with the bulk flow of urine.
- 2. Invasion:** Exocytosis of specialized vesicles (fusiform vesicles) regulates the bladder surface when the volume of the bladder increases due to the accumulation of

## Section II. The Bacterial Adhesin FimH – Introduction

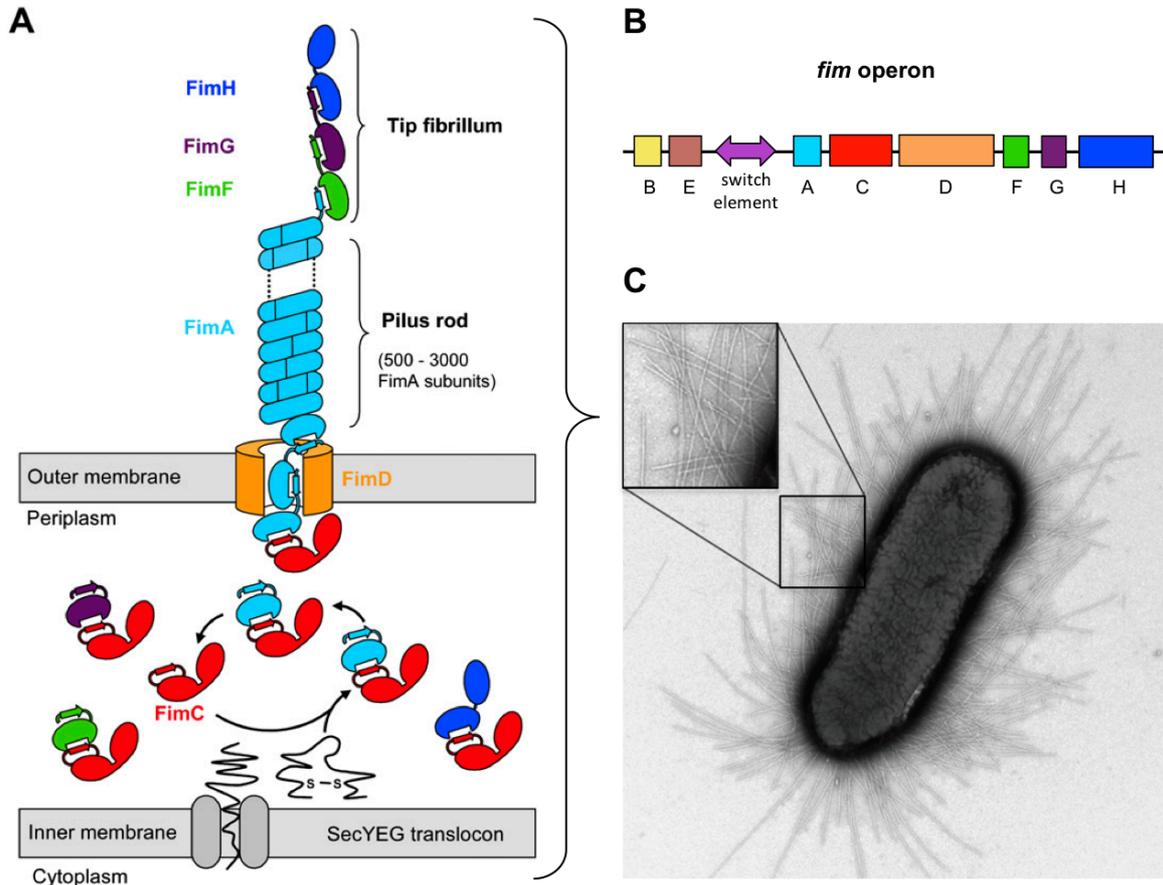
urine.<sup>(10-12)</sup> Bladder relaxation and the consequent internalization of fusiform vesicle allow attached UPEC to invade bladder cells.

- 3. Intracellular bacterial communities (IBC):** Once internalized, *E. coli* bacteria start at an early stage to replicate and build IBCs.<sup>(13,14)</sup> After approx. 6-8 hours they change their morphology from the normal rod-shape to a coccoid form and reduce their replication time. Furthermore, they build a polysaccharide biofilm protecting the IBC from immune responses.<sup>(15)</sup>
- 4. Release:** To re-enter the bladder lumen, the bacteria change again their morphology from the coccoid form (0.7  $\mu\text{m}$ ) to a long filamentous shape (70  $\mu\text{m}$ ) to burst the urothelial cell. Back in the lumen they are able to change to their normal rod-shaped morphology and start expressing type 1 pili.<sup>(16,17)</sup>
- 5. Quiescent intracellular reservoir (QIR):** Burst urothelial cells increase the accessibility to subjacent intermediate and basal urothelial cells. In these cells, the *E. coli* remains in a latent state. These bacteria build a QIR and persist over a longer period of time and are most likely responsible for the high reoccurrence rate of UTI within 6 months after an infection.<sup>(18,19)</sup>



**Figure 1. Infection cycle of uropathogenic *E. coli* (UPEC) in the lower urinary tract.** Adhesion is established by an interaction between the type 1 pili of UPEC and a mannosylated glycoprotein on the bladder epithelium. Subsequent invasion and intracellular bacterial colonies (IBC) with biofilm formation facilitate amplification in the absence of immune measures and urine flush. The subsequent release and formation of quiescent intracellular reservoirs (QIR) in deeper lying tissue cells increase the persistence of UPEC in the bladder.

## 1.3. Structure

1.3.1. Type 1 pili and the *fim* operon

**Figure 2. Structure of type 1 pili of UPEC.** Schematic representation of the assembly of a type 1 pilus (A) and the genetic cluster of the *fim* operon (B) encoding for the structural Fim proteins. *fimB*, *fimE*, and the invertible switch element regulate the expression of type 1 pili. (C) Electron microscopic picture of a UPEC expressing type 1 pili. (Images adapted from Puorger *et al.* (2008)<sup>(20)</sup> and Hahn *et al.* (2002)<sup>(21)</sup>; reprinted with permission from Elsevier).

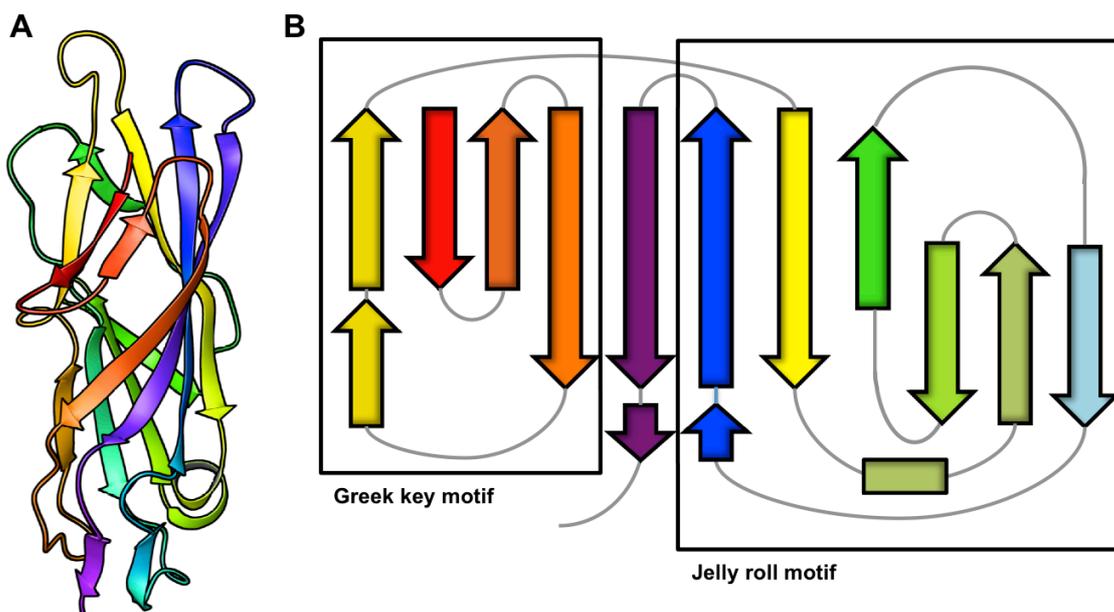
The type 1 pilus is a hetero-oligomeric protein complex protruding from the cell wall of UPEC (Figure 2A). The rod-like shape is formed by ~500 to 3000 helically assembled copies of FimA, followed by one unit of FimF, FimG, and FimH at the distal end of the pilus.<sup>(21,22)</sup> The terminal lectin domain of FimH is responsible for the adhesion to the mannosylated glycans of uroplakin 1a (UPK1a) on the bladder epithelium. The incomplete immunoglobulin-like fold (pilin-fold) of FimA, FimF, FimG, and FimH is lacking a C-terminal  $\beta$ -strand that is crucial for the protein stability.<sup>(23)</sup> Before pilus assembly by the usher  $\beta$ -barrel pore protein FimD in the outer bacterial membrane, the FimC chaperone protein provides the missing  $\beta$ -strand in the periplasmic space (donor

## Section II. The Bacterial Adhesin FimH – Introduction

strand complementation).<sup>(24-26)</sup> However, for the assembly of a stable pilus, a donor strand exchange is required. Therefore, the subunits (except the terminal subunit FimH) possess a prolonged N-terminal  $\beta$ -strand that complements for the lacking C-terminal  $\beta$ -strand of the neighboring protein.

Genes of the Type 1 pilus proteins are encoded on the *fim* operon (Figure 2B). The expression of the structural genes in the *fim* gene cluster (*fimA-fimH*) can be regulated by phase-variable inversion of a 314 bp DNA switch element containing the promoter region of the *fimA* gene.<sup>(27)</sup> This inversion is catalyzed by two upstream-encoded recombinases *fimB* and *fimE*.<sup>(28)</sup> The recombinases are sensitive to external factors such as pH, temperature, and osmolarity and allow an adaption to the environment.<sup>(29,30)</sup>

### 1.3.2. The lectin FimH



**Figure 3. Topology of FimH<sub>L</sub>.** The immunoglobulin-like fold of the FimH<sub>L</sub> (FimH lectin domain) (A) is mainly composed of antiparallel  $\beta$ -strands (arrows) forming a greek key and a jelly roll motif (B).

The tip of type 1 pili is formed by the FimH lectin. The 29 kDa protein consists of 279 amino acids forming two immunoglobulin-like domains: an N-terminal lectin domain (FimH<sub>L</sub>) (residues 1-156) linked via three amino acids to a C-terminal pilin domain (FimH<sub>P</sub>) (residues 160-279). As stated before, FimH is unstable due to the missing  $\beta$ -strand. However, the separated FimH<sub>L</sub> (PDB code: 1AUU) (Figure 3) and the full-

length FimH protein in complex with the chaperone protein FimC (PDB code: 1QUN) or a FimG/FimF peptide donor strand (DsG/DsF) (PDB code: 4XOD/4XO9) are stable.

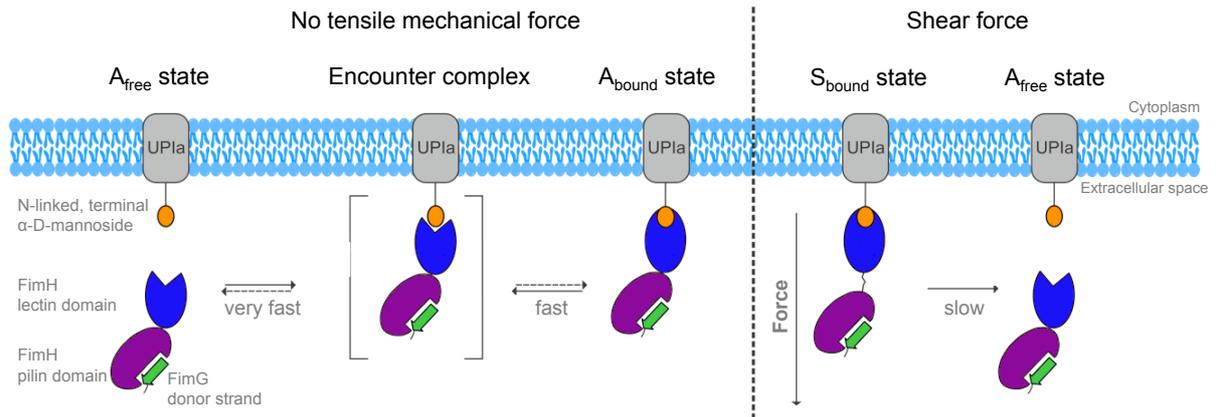
In 2002, Sokurenko and co-workers discovered a special property of FimH: type 1 pili expressing *E. coli* improve their adhesion to erythrocytes under flow conditions.<sup>(31)</sup> This discovery led to the proposal of a two-state catch-bond binding model allosterically regulated by FimH<sub>P</sub>.<sup>(32,33)</sup> In the so-called ‘low-affinity state’ FimH<sub>P</sub> and FimH<sub>L</sub> are in close contact, whereas shear force induces a domain separation referred as ‘high-affinity state’. This mechanism is of biological relevance as it prevents the bacteria from being flushed out by the urinary bulk flow. Mutations that weakened the contact between FimH<sub>L</sub> and FimH<sub>P</sub> increased the affinity of interaction with mannosides.<sup>(34)</sup> This is in good agreement with the fact that most naturally occurring mutations are in proximity to the interdomain interface.<sup>(35)</sup> Hence, allosteric mutations enable adaption to various conditions by regulating the binding properties without changing the binding specificity to the natural ligand.

A recent publication from Sauer *et al.* (2015) revealed further insights into structural and kinetic properties of FimH.<sup>(36)</sup> An additional state and consequently a new nomenclature to the previously introduced two-state catch-bond mechanism was proposed (Figure 4). Without shear force, FimH<sub>L</sub> and FimH<sub>P</sub> remain associated. Via an encounter complex, the transition between the associated unbound ( $A_{\text{free}}$ ) and the associated bound ( $A_{\text{bound}}$ ) state is highly dynamic. This allows the bacterium to maintain a certain amount of mobility on the urothelium in the absence of flow, while in the presence of flow shear force separates FimH<sub>L</sub> and FimH<sub>P</sub> ( $S_{\text{bound}}$ ) providing firm adhesion.

## Section II. The Bacterial Adhesin FimH – Introduction

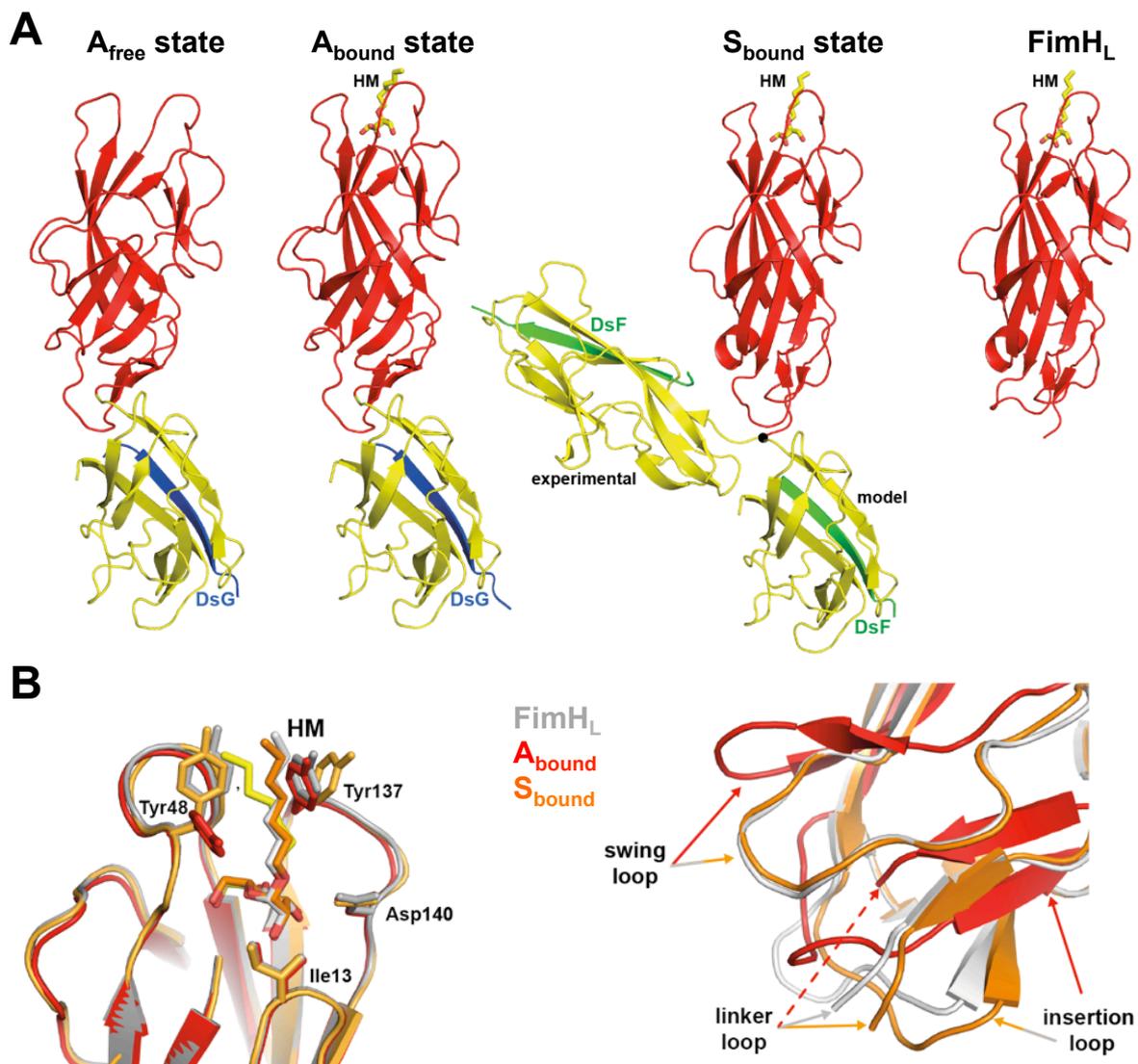
**Table 1. Overview of the FimH structures deposited in the Protein Data Bank (PDB).** Resolution for X-ray structures is given in Ångström (Å). DsG/DsF = Donor strand from FimG/FimF.

PDB-Code	Construct	Ligand	Resolution	Remarks	Ref
1QUN	FimH-FimC	apo	2.8		(23)
1KIU	FimH-FimC	methyl- $\alpha$ -D-mannopyranoside	3.0	Q133N	(37)
1KLF	FimH-FimC	$\alpha$ -D-mannopyranoside	2.8		(37)
3JWN	FimH-FimF-FimG-FimC	apo	2.7	V27A N70S S78N	(38)
3RFZ	FimD-FimH-FimC	apo	2.8		(39)
4J3O	FimD-FimH-FimF-FimG-FimC	apo	3.8		(40)
4AUU	FimH <sub>L</sub>	apo	1.6		(41)
3ZPD	FimH <sub>L</sub>	apo		Solution NMR	(42)
2VCO	FimH <sub>L</sub>	Man <sub>5</sub> -GlcNAc <sub>2</sub>	2.1		(43)
4CA4	FimH <sub>L</sub>	heptyl- $\alpha$ -D-mannopyranoside	2.8	Y48A	-
4XOA	FimH-DsG	heptyl- $\alpha$ -D-mannopyranoside	2.5	V27A N70S S78N	(36)
4XOB	FimH-DsF	heptyl- $\alpha$ -D-mannopyranoside	3.0	V27A N70S S78N	(36)
4XO8	FimH <sub>L</sub>	heptyl- $\alpha$ -D-mannopyranoside	1.7	V27A N70S S78N	(36)
4XO9	FimH-DsG	apo	1.1	V27A N70S S78N	(36)
1UWF, 1TR7, 3MCY, 4AUU, 4AV0, 4AV4, 4AV5, 4AVH, 4AVI, 4AVJ, 4AVK, 4ATT, 4AUJ, 3ZL1, 3ZL2, 4LOV, 4BUQ, 4CSS, 4CST, 4X50, 4X5P, 4X5Q, 4X5R	FimH <sub>L</sub>	$\alpha$ -D-mannopyranosides with varying aglycones	1.0 - 2.9		(41,42,44-49)



**Figure 4. Model for a three-state catch-bond mechanism of FimH.** Without shear force, FimH<sub>L</sub> and FimH<sub>P</sub> remain associated. The transition between the associated bound ( $A_{\text{bound}}$ ) and the associated unbound ( $A_{\text{free}}$ ) state remains highly dynamic via an encounter complex. In the presence of shear force, FimH<sub>L</sub> and FimH<sub>P</sub> become separated ( $S_{\text{bound}}$ ). Image adapted from Sauer *et al.* 2015.<sup>(36)</sup>

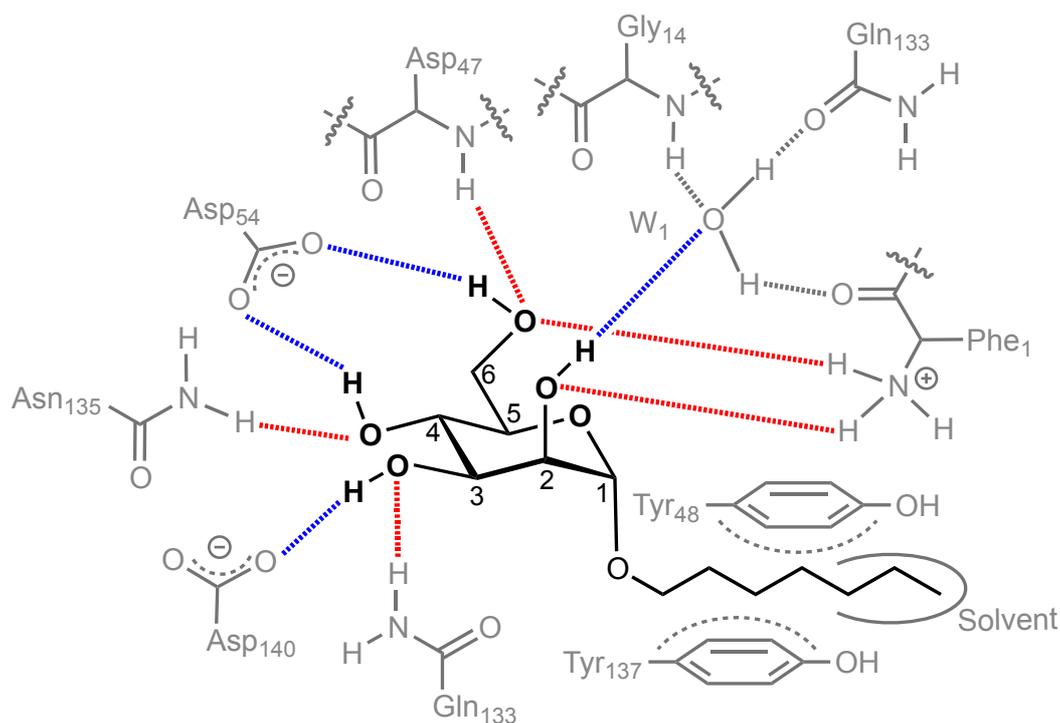
Conformational changes between  $A_{\text{free}}$  and  $A_{\text{bound}}$  are restricted to the ligand binding site and conformational changes between  $A_{\text{bound}}$  and  $S_{\text{bound}}$  are restricted to the interdomain interface of the lectin domain (Figure 5A). The isolated FimH<sub>L</sub> construct represents the  $S_{\text{bound}}$  state. Compared to FimH<sub>L</sub>, the binding affinity of  $A_{\text{bound}}$  decreases by a factor of 30 and the presence of FimH<sub>P</sub> may be described as a negative allosteric regulator. However, the binding site of all bound states ( $A_{\text{bound}}$ ,  $S_{\text{bound}}$ , and FimH<sub>L</sub>) is nearly identical (Figure 5B). Hence, FimH<sub>L</sub> is a valid model for the molecular investigation of ligands interacting with the FimH binding site.



**Figure 5. Conformational similarities and differences of the FimH states.** (A) 4 crystal structures represent a three-state model with catch-bond behavior: In the absence of shear force and ligand ( $A_{\text{free}}$ ; PDB Code: 4XO9), in the absence of shear force and the presence of ligand ( $A_{\text{bound}}$ ; PDB codes: 4XOA/4XOB) and in the presence of shear force and ligand ( $S_{\text{bound}}$ ; PDB code: 4XO8).  $FimH_L$  represents the  $S_{\text{bound}}$  state. (B) Structural differences between  $FimH_L$ ,  $A_{\text{bound}}$ , and  $S_{\text{bound}}$  in the binding site (left) and the interdomain region (right). Image adapted from Sauer *et al.* 2015.<sup>(36)</sup>

### 1.3.3. FimH binding site

Many co-crystal structures (Table 1) of FimH and  $\alpha$ -D-mannopyranosides revealed a rather deep and charged cavity with a hydrophobic ridge at the entrance (Figure 6). *Manuscript 1* “High-affinity carbohydrate-lectin interaction: How nature makes it possible” discusses properties of the FimH binding site in depth with a focus on hydroxyl groups. In brief, the interaction between FimH and mannose ligands is a highly optimized network of 9 hydrogen bonds (Figure 6).



**Figure 6. Schematic representation of the FimH binding pocket.** From the ligand accepted (red) and donated (blue) hydrogen bonds are drawn according to X-ray crystal structure of *n*-heptyl α-D-mannopyranoside in complex with FimH<sub>L</sub> (PDB code 4BUQ).

No other monosaccharide was found to bind to FimH with an affinity comparable to that of mannose.<sup>(44)</sup> One hydrogen bond is formed with the amino acid side chains of the residues Gln133, Asp140, Asn135 and with the protein backbone of residue Asp47. Multiple hydrogen bonds are formed with the negatively charged amino acid side chain of Asp54 (2 direct bonds) and with the positively charged N-terminus (2 direct and 1 bridged bond via the structural water molecule W<sub>1</sub>). The amino acids forming charged interactions contribute the most to the free energy of binding.<sup>(50)</sup> All hydrogen bonds have a reasonable geometry (angles and hydrogen bond length). The hydrophobic ridge at the entrance of the cavity is formed by two parallel-aligned tyrosines (Tyr48, Tyr137), the so-called ‘tyrosine gate’ (Figure 5B, Figure 6). While the cavity provides specificity, the tyrosine gate is the key element for the optimization of monovalent α-D-mannopyranosides derivatives.<sup>(45,48,51-53)</sup> The interaction between different aglycones and the tyrosine gate is described in more detail in *Manuscript 3*: “The Tyrosine Gate of the Bacterial Lectin FimH: A Conformational Analysis by NMR Spectroscopy and X-ray Crystallography”.<sup>(49)</sup>

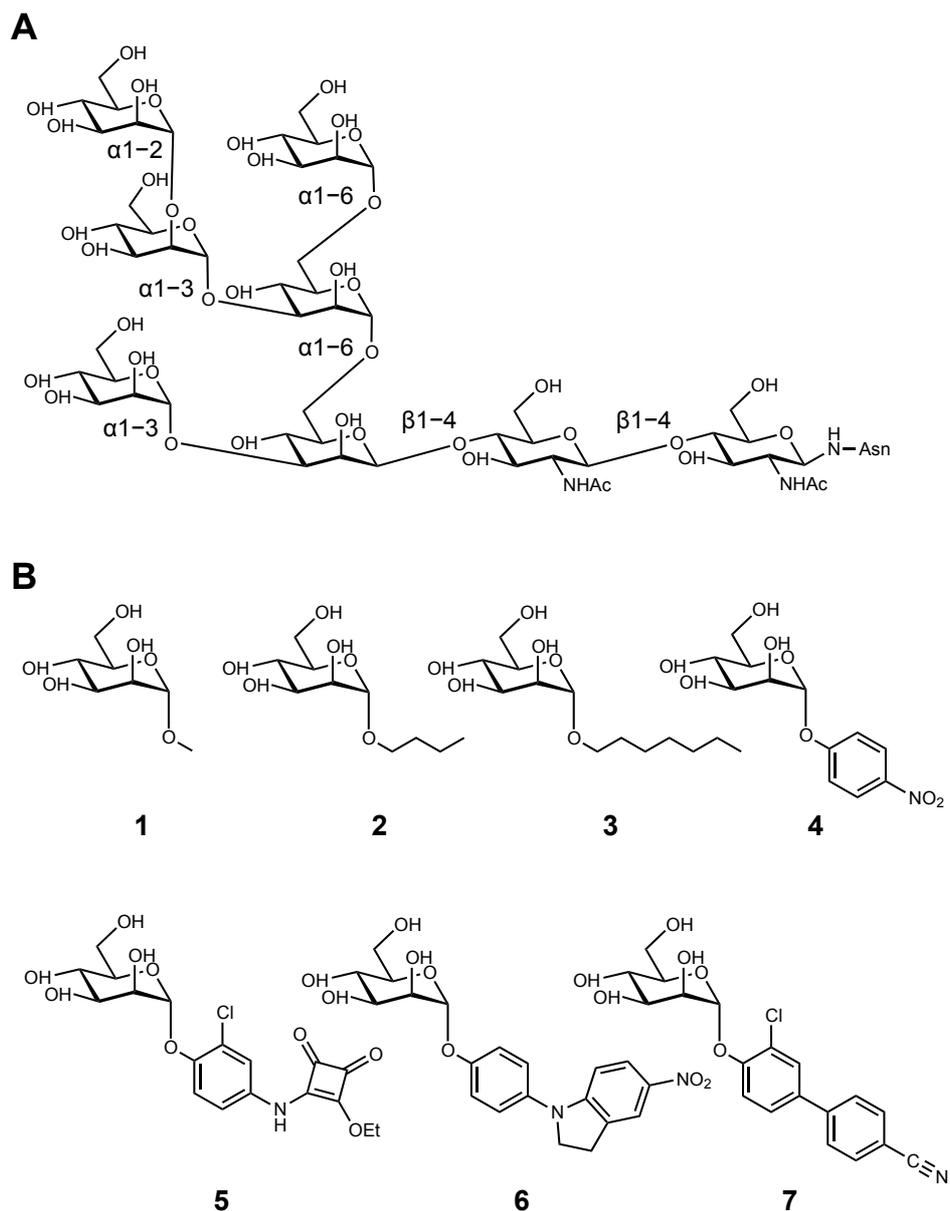
### 1.4. FimH antagonists

UTI patients could benefit from a long-term anti-adhesion treatment considering that 25% suffer from recurrent infections within 6 months, and emerging antibiotic resistances in UPEC isolates are creating increasing problems.<sup>(5)</sup> Still, many properties are required to launch a FimH antagonist as a drug for UTI: the affinity of a successful candidate has to exceed the affinity of FimH to the natural target, needs to be orally available, and must be specific to the receptor.

An excellent review of the history of FimH antagonists was published by Hartmann and Lindhorst in 2011 and is recommended for a more detailed overview.<sup>(54)</sup> In brief, the research interest in type 1 pili started in the 1970s when Sharon found *n*-methyl  $\alpha$ -D-mannopyranoside (1) (Figure 7B) to inhibit hemagglutination.<sup>(55)</sup> In the 1980s the research was mainly focused on the natural ligand of FimH. On urothelial cells, mannose is presented on high-mannose N-linked glycans, consisting of five (Man<sub>5</sub>-GlcNAc<sub>2</sub>) to nine (Man<sub>9</sub>-GlcNAc<sub>2</sub>) mannose residues. While on Man<sub>9</sub>-GlcNAc<sub>2</sub> glycans all terminal mannose residues are  $\alpha$ 1-2 linked, Man<sub>5</sub>-GlcNAc<sub>2</sub> presents  $\alpha$ 1-2,  $\alpha$ 1-3 and  $\alpha$ 1-6 linked mannoses to the bladder lumen (Figure 7A). FimH was found to have the highest affinity towards  $\alpha$ 1-3 linked mannosides.<sup>(56)</sup> These results were later confirmed and rationalized by Wellens *et al.* 2008,<sup>(43)</sup> who concluded that the central mannose stacks with the aromatic ring of Tyr48 and thereby increases the affinity to FimH.

Investigations with natural mannosides and the knowledge about the cluster effect<sup>(57)</sup> of type 1 pili led to the development of multivalent glycomimetics.<sup>(58)</sup> Indeed, multiple presentations of  $\alpha$ -D-mannosides significantly increased the affinity compared to monovalent ligands. However, this effect is still not fully understood. Multivalent constructs are unlikely to bind more than one FimH binding site as they are still too small in size to crosslink multiple lectins. A possible explanation is the increase of the local concentration of  $\alpha$ -D-mannosides and therefore statistically favored binding. The major drawback to this class of FimH ligands are their pharmacokinetic properties. Due to their polarity and large size, they are very unlikely to permeate through membranes to achieve oral bioavailability.

The more promising approach in terms of oral bioavailability is the design of monovalent high-affinity ligands. A crystal structure solved by Bouckaert *et al.* 2005 unexpectedly contained *n*-butyl  $\alpha$ -D-mannopyranoside (**2**) in the binding pocket of FimH<sub>L</sub> (Figure 7B) stemming from the bacterial growth medium (PDB Code: 1UWF).<sup>(44)</sup> Further investigations revealed that long-chain alkyl-mannosides strongly increase the affinity by interacting with the tyrosine gate. The highest affinity among alkyl-mannosides was determined for *n*-heptyl  $\alpha$ -D-mannopyranoside (**3**), binding about 65 times stronger than methyl mannoside (**1**).<sup>(44)</sup> Since then, rational design approaches were guided by a broad variety of co-crystal structures (Table 1). In terms of affinity improvement and pharmacokinetic properties, aryl aglycones turned out to be beneficial over alkyl moieties. Already Sharon in 1987 reported *p*-nitrophenyl  $\alpha$ -D-mannopyranoside (**4**) to be a high-affinity FimH binder.<sup>(55)</sup> Indeed, these results could be confirmed and **4** was shown to have a comparable affinity for FimH<sub>L</sub> as *n*-heptyl mannoside (**3**).<sup>(44)</sup>  $\pi$ - $\pi$  stacking with Tyr48 could be further improved by substitution of the phenyl ring in *para* position; e.g. the squaric acid moiety (**5**) was shown to increase the affinity by an additional factor of 10 compared to **4**.<sup>(59,60)</sup> and an indolinyphenyl derivative (**6**) improved affinity as well as the pharmacokinetic properties, which were tested in mice and revealed an improved renal elimination profile after intravenous application.<sup>(51)</sup> The affinity of biphenyl  $\alpha$ -D-mannopyranosides was found to be optimized by substituents in *ortho* position of the phenyl ring adjacent to the anomeric oxygen (e.g. Cl, Me, CF<sub>3</sub>) and electron withdrawing substituents in *para* and *meta* position on the terminal ring of the biphenyl aglycone (e.g. carboxylic acid, cyano group).<sup>(45,52)</sup> Oral availability was finally achieved with 3'-chloro-4'-( $\alpha$ -D-mannopyranosyloxy)biphenyl-4-carbonitrile (**7**) using a phosphate buffered saline (PBS) formulation containing 5% DMSO and 1% Tween 80 (*Manuscript 5*).<sup>(48)</sup> This FimH antagonist binds with a 1.3 nM affinity to FimH<sub>L</sub> and reduces the bacterial load in the bladder of mice as efficaciously as a comparable dose of the antibiotic ciprofloxacin, which is used for standard treatment of a UTI.



**Figure 7: Overview of FimH ligand structures.** (A)  $\text{Man}_5\text{-GlcNAc}_2$  is a naturally occurring polysaccharide in the glycan of uroplakin 1a on urothelial cells. FimH binds to  $\alpha$ 1-3 linked mannose with the highest affinity. (B) Milestones in the development of monovalent FimH antagonists towards a highly affine and orally bioavailable ligand (7).

## 1.5. References

1. Foxman, B. (2002). Epidemiology of urinary tract infections: Incidence, morbidity, and economic costs. *Am. J. Med.*, 113: 5-13.
2. Foxman, B., Barlow, R., D'Arcy, H., Gillespie, B., & Sobel, J.D. (2000). Urinary tract infection: self-reported incidence and associated costs. *Ann. Epidemiol.*, 10(8): 509-515.
3. Hooton, T.M. (2001). Recurrent urinary tract infection in women. *Int. J. Antimicrob. Agents*, 17(4): 259-268.
4. Fihn, S.D. (2003). Acute Uncomplicated Urinary Tract Infection in Women. *N. Engl. J. Med.*, 349(3): 259-266.
5. Sanchez, G.V., Master, R.N., Karlowsky, J.A., & Bordon, J.M. (2012). In Vitro Antimicrobial Resistance of Urinary Escherichia coli Isolates among US Outpatients from 2000 to 2010. *Antimicrob. Agents Chemother.*, 56(4): 2181-2183.
6. Ronald, A. (2002). The etiology of urinary tract infection: Traditional and emerging pathogens. *Am. J. Med.*, 113: 14-19.
7. Ofek, I., Hasty, D., Abraham, S., & Sharon, N. (2002). Role of Bacterial Lectins in Urinary Tract Infections. *Genes and Proteins Underlying Microbial Urinary Tract Virulence*, Advances in Experimental Medicine and Biology, eds Emoödy L, Pál T, Hacker J, & Blum-Oehler G, Springer US. Vol. 485, pp. 183-192.
8. Jorgensen, I. & Seed, P.C. (2012). How to Make It in the Urinary Tract: A Tutorial by Escherichia coli. *PLoS Path.*, 8(10): e1002907.
9. Wu, X.R., Sun, T.T., & Medina, J.J. (1996). In vitro binding of type 1-fimbriated Escherichia coli to uroplakins Ia and Ib: relation to urinary tract infections. *Proc. Natl. Acad. Sci. U. S. A.*, 93(18): 9630-9635.
10. Bishop, B.L., et al. (2007). Cyclic AMP-regulated exocytosis of Escherichia coli from infected bladder epithelial cells. *Nat. Med.*, 13(5): 625-630.
11. Duncan, M.J., Li, G., Shin, J.S., Carson, J.L., & Abraham, S.N. (2004). Bacterial penetration of bladder epithelium through lipid rafts. *J. Biol. Chem.*, 279(18): 18944-18951.
12. Chen, Y., et al. (2003). Rab27b is associated with fusiform vesicles and may be involved in targeting uroplakins to urothelial apical membranes. *Proc. Natl. Acad. Sci. U. S. A.*, 100(24): 14012-14017.
13. Anderson, G.G., Dodson, K.W., Hooton, T.M., & Hultgren, S.J. (2004). Intracellular bacterial communities of uropathogenic Escherichia coli in urinary tract pathogenesis. *Trends Microbiol.*, 12(9): 424-430.
14. Rosen, D.A., Hooton, T.M., Stamm, W.E., Humphrey, P.A., & Hultgren, S.J. (2007). Detection of intracellular bacterial communities in human urinary tract infection. *PLoS Med.*, 4(12): e329.
15. Anderson, G.G., et al. (2003). Intracellular bacterial biofilm-like pods in urinary tract infections. *Science*, 301(5629): 105-107.
16. Andersen, T.E., et al. (2012). Escherichia coli uropathogenesis in vitro: invasion, cellular escape, and secondary infection analyzed in a human bladder cell infection model. *Infect. Immun.*, 80(5): 1858-1867.
17. Justice, S.S., et al. (2004). Differentiation and developmental pathways of uropathogenic Escherichia coli in urinary tract pathogenesis. *Proc. Natl. Acad. Sci. U. S. A.*, 101(5): 1333-1338.
18. Mysorekar, I.U. & Hultgren, S.J. (2006). Mechanisms of uropathogenic Escherichia coli persistence and eradication from the urinary tract. *Proc. Natl. Acad. Sci. U. S. A.*, 103(38): 14170-14175.
19. Schwartz, D.J., Chen, S.L., Hultgren, S.J., & Seed, P.C. (2011). Population Dynamics and Niche Distribution of Uropathogenic Escherichia coli during Acute and Chronic Urinary Tract Infection. *Infect. Immun.*, 79(10): 4250-4259.
20. Puorger, C., et al. (2008). Infinite kinetic stability against dissociation of supramolecular protein complexes through donor strand complementation. *Structure*, 16(4): 631-642.
21. Hahn, E., et al. (2002). Exploring the 3D Molecular Architecture of Escherichia coli Type 1 Pili. *J. Mol. Biol.*, 323(5): 845-857.
22. Jones, C.H., et al. (1995). FimH adhesin of type 1 pili is assembled into a fibrillar tip structure in the Enterobacteriaceae. *Proc. Natl. Acad. Sci. U. S. A.*, 92(6): 2081-2085.
23. Choudhury, D., et al. (1999). X-ray structure of the FimC-FimH chaperone-adhesin complex from uropathogenic Escherichia coli. *Science*, 285(5430): 1061-1066.

## Section II. The Bacterial Adhesin FimH – Introduction

---

24. Sauer, F.G., *et al.* (2000). Chaperone-assisted pilus assembly and bacterial attachment. *Curr. Opin. Struct. Biol.*, 10(5): 548-556.
25. Knight, S.D., Berglund, J., & Choudhury, D. (2000). Bacterial adhesins: structural studies reveal chaperone function and pilus biogenesis. *Curr. Opin. Chem. Biol.*, 4(6): 653-660.
26. Hung, D.L. & Hultgren, S.J. (1998). Pilus biogenesis via the chaperone/usher pathway: An integration of structure and function. *J. Struct. Biol.*, 124(2-3): 201-220.
27. Sohanpal, B.K., Kulasekara, H.D., Bonnen, A., & Blomfield, I.C. (2001). Orientational control of fimE expression in Escherichia coli. *Mol. Microbiol.*, 42(2): 483-494.
28. Klemm, P. (1986). Two regulatory fim genes, fimB and fimE, control the phase variation of type 1 fimbriae in Escherichia coli. *EMBO J.*, 5(6): 1389-1393.
29. Schwan, W.R., Lee, J.L., Lenard, F.A., Matthews, B.T., & Beck, M.T. (2002). Osmolarity and pH growth conditions regulate fim gene transcription and type 1 pilus expression in uropathogenic Escherichia coli. *Infect. Immun.*, 70(3): 1391-1402.
30. Gally, D.L., Bogan, J.A., Eisenstein, B.I., & Blomfield, I.C. (1993). Environmental regulation of the fim switch controlling type 1 fimbrial phase variation in Escherichia coli K-12: effects of temperature and media. *J. Bacteriol.*, 175(19): 6186-6193.
31. Thomas, W.E., Trintchina, E., Forero, M., Vogel, V., & Sokurenko, E.V. (2002). Bacterial Adhesion to Target Cells Enhanced by Shear Force. *Cell*, 109(7): 913-923.
32. Thomas, W., *et al.* (2006). Catch-Bond Model Derived from Allostery Explains Force-Activated Bacterial Adhesion. *Biophys. J.*, 90(3): 753-764.
33. Yakovenko, O., *et al.* (2008). FimH Forms Catch Bonds That Are Enhanced by Mechanical Force Due to Allosteric Regulation. *J. Biol. Chem.*, 283(17): 11596-11605.
34. Aprikian, P., *et al.* (2007). Interdomain interaction in the FimH adhesin of Escherichia coli regulates the affinity to mannose. *J. Biol. Chem.*, 282(32): 23437-23446.
35. Sokurenko, E.V., *et al.* (1998). Pathogenic adaptation of Escherichia coli by natural variation of the FimH adhesin. *Proc. Natl. Acad. Sci. U. S. A.*, 95(15): 8922-8926.
36. Sauer, M.M., *et al.* (2015). Catch-bond mechanism of the bacterial adhesin FimH. *Nat. Struct. Biol.*, submitted.
37. Hung, C.S., *et al.* (2002). Structural basis of tropism of Escherichia coli to the bladder during urinary tract infection. *Mol. Microbiol.*, 44(4): 903-915.
38. Le Trong, I., *et al.* (2010). Structural basis for mechanical force regulation of the adhesin FimH via finger trap-like beta sheet twisting. *Cell*, 141(4): 645-655.
39. Phan, G., *et al.* (2011). Crystal structure of the FimD usher bound to its cognate FimC-FimH substrate. *Nature*, 474(7349): 49-53.
40. Geibel, S., Procko, E., Hultgren, S.J., Baker, D., & Waksman, G. (2013). Structural and energetic basis of folded-protein transport by the FimD usher. *Nature*, 496(7444): 243-246.
41. Wellens, A., *et al.* (2012). The tyrosine gate as a potential entropic lever in the receptor-binding site of the bacterial adhesin FimH. *Biochemistry*, 51(24): 4790-4799.
42. Vanwetswinkel, S., *et al.* (2014). Study of the Structural and Dynamic Effects in the FimH Adhesin upon  $\alpha$ -d-Heptyl Mannose Binding. *J. Med. Chem.*, 57(4): 1416-1427.
43. Wellens, A., *et al.* (2008). Intervening with urinary tract infections using anti-adhesives based on the crystal structure of the FimH-oligomannose-3 complex. *PLoS One*, 3(4): e2040.
44. Bouckaert, J., *et al.* (2005). Receptor binding studies disclose a novel class of high-affinity inhibitors of the Escherichia coli FimH adhesin. *Mol. Microbiol.*, 55(2): 441-455.
45. Han, Z., *et al.* (2010). Structure-Based Drug Design and Optimization of Mannoside Bacterial FimH Antagonists. *J. Med. Chem.*, 53(12): 4779-4792.
46. Roos, G., *et al.* (2013). Validation of Reactivity Descriptors to Assess the Aromatic Stacking within the Tyrosine Gate of FimH. *ACS Med. Chem. Lett.*, 4(11): 1085-1090.
47. Brument, S., *et al.* (2013). Thiazolylaminomannosides as potent antiadhesives of type 1 pilated Escherichia coli isolated from Crohn's disease patients. *J. Med. Chem.*, 56(13): 5395-5406.
48. Kleeb, S., *et al.* (2015). FimH antagonists: bioisosteres to improve the in vitro and in vivo PK/PD profile. *J. Med. Chem.*, 58(5): 2221-2239.
49. Fiege, B., *et al.* (2015). The Tyrosine Gate of the Bacterial Lectin FimH: A Conformational Analysis by NMR Spectroscopy and X-ray Crystallography. *ChemBioChem*, 16(8): 1235-1246.
50. Zalewski, A. (2013). *In pursuit of a novel UTI treatment strategy - an "in silico" study of the FimH adhesin.* PhD Dissertation, University of Basel, Basel.
51. Jiang, X., *et al.* (2012). Antiadhesion Therapy for Urinary Tract Infections—A Balanced PK/PD Profile Proved To Be Key for Success. *J. Med. Chem.*, 55(10): 4700-4713.

52. Pang, L., *et al.* (2012). FimH antagonists: structure-activity and structure-property relationships for biphenyl alpha-D-mannopyranosides. *ChemMedChem*, 7(8): 1404-1422.
53. Schwardt, O., *et al.* (2011). Design, synthesis and biological evaluation of mannosyl triazoles as FimH antagonists. *Bioorg. Med. Chem.*, 19(21): 6454-6473.
54. Hartmann, M. & Lindhorst, T.K. (2011). The Bacterial Lectin FimH, a Target for Drug Discovery – Carbohydrate Inhibitors of Type 1 Fimbriae-Mediated Bacterial Adhesion. *Eur. J. Org. Chem.*, 2011(20-21): 3583-3609.
55. Sharon, N. (1987). Bacterial lectins, cell-cell recognition and infectious disease. *FEBS Lett.*, 217(2): 145-157.
56. Firon, N., Ofek, I., & Sharon, N. (1983). Carbohydrate specificity of the surface lectins of *Escherichia coli*, *Klebsiella pneumoniae*, and *Salmonella typhimurium*. *Carbohydr. Res.*, 120(0): 235-249.
57. Lundquist, J.J. & Toone, E.J. (2002). The Cluster Glycoside Effect. *Chem. Rev.*, 102(2): 555-578.
58. Lindhorst, T.K., Kieburg, C., & Krallmann-Wenzel, U. (1998). Inhibition of the type 1 fimbriae-mediated adhesion of *Escherichia coli* to erythrocytes by multiantennary alpha-mannosyl clusters: The effect of multivalency. *Glycoconj. J.*, 15(6): 605-613.
59. Rabbani, S., Jiang, X., Schwardt, O., & Ernst, B. (2010). Expression of the carbohydrate recognition domain of FimH and development of a competitive binding assay. *Anal. Biochem.*, 407(2): 188-195.
60. Sperling, O., Fuchs, A., & Lindhorst, T.K. (2006). Evaluation of the carbohydrate recognition domain of the bacterial adhesin FimH: Design, synthesis and binding properties of mannoside ligands. *Org. Biomol. Chem.*, 4(21): 3913-3922.



## Manuscript 1

### High-Affinity Carbohydrate-Lectin Interactions: How Nature Makes it Possible

Pascal Zihlmann,<sup>#</sup> Xiaohua Jiang,<sup>#</sup> Christoph P. Sager,<sup>#</sup> Brigitte Fiege,<sup>#</sup>  
Roman P. Jakob,<sup>§</sup> Stefan Siegrist,<sup>#</sup> Adam Zalewski,<sup>#</sup> Said Rabbani,<sup>#</sup> Deniz Eriş,<sup>#</sup>  
Marleen Silbermann,<sup>#</sup> Lijuan Pang,<sup>#</sup> Tobias Mühlethaler,<sup>#</sup> Timothy Sharpe,<sup>†</sup>  
Timm Meier<sup>§</sup> and Beat Ernst<sup>\*</sup>

<sup>#</sup>University of Basel, Institute of Molecular Pharmacy,  
Klingelbergstr. 50, 4056 Basel, Switzerland

<sup>§</sup>University of Basel, Institute of Structural Biology,  
Klingelbergstr. 70, 4056 Basel, Switzerland

<sup>†</sup>University of Basel, Biophysics Facility,  
Klingelbergstr. 70, 4056 Basel, Switzerland

<sup>\*</sup>Corresponding author.

Tel.: 0041 (0)61 267 15 51; Fax: 0041 (0)61 267 15 52;  
E-mail: beat.ernst@unibas.ch

#### Contributions of Pascal Zihlmann:

- Manuscript preparation
- FP experiments
- Thermodynamic and kinetic profiling by ITC



### Abstract

Lectins belong to the most challenging targets in drug discovery due to the unique binding properties of their polyhydroxylated carbohydrate ligands. Whereas the hydroxyl groups provide directionality and specificity, the high desolvation costs of carbohydrates are the origin for their notoriously low affinities.

Nonetheless, some lectins with high affinity to monovalent carbohydrate ligands have been reported.<sup>(1-6)</sup> One of these rare examples is the bacterial lectin FimH, located at the tip of the pili of uropathogenic *E. coli* (UPEC). It mediates adhesion to the mannosylated glycoproteins uroplakin Ia on urothelial host cells. By combining computational methods (QM, MD simulations) with structural information (X-ray, NMR) and binding data (FP, ITC, kinITC) the complex and cooperative hydrogen bond network formed by mannoside ligands interacting with FimH was elucidated. Deoxy- and deoxy-halogeno derivatives of *n*-heptyl  $\alpha$ -D-mannoside (**1**) reveal that the loss individual hydroxyl groups not only leads to a decrease of the association rate ( $k_{on}$ ), but also to an increased dissociation rate ( $k_{off}$ ) and as a result to a dramatic drop of affinity ( $K_D$ ). Furthermore, a comparison of thermodynamic profiles obtained by isothermal titration calorimetry (ITC) indicates that the loss in affinity (corresponding to a  $\Delta\Delta G^\circ$  of 15-21 kJ/mol per hydroxyl group) originates from unfavorable enthalpy contributions partly compensated by an entropic gain.

### Introduction

It is generally accepted that hydrogen bonds (H-bonds) provide directionality and therefore specificity to ligand-receptor interactions, whereas hydrophobic interactions, although rather unspecific, predominantly contribute to binding energies.<sup>(7)</sup> Since carbohydrate-lectin interactions are mainly based on H-bond formation, they are characterized by a high degree of specificity, but often suffer from a lack of affinity.<sup>(8)</sup> The high specificity allows carbohydrates to fulfill their broad biological tasks, such as signal transduction,<sup>(9,10)</sup> cell recognition<sup>(11,12)</sup> or cell adhesion.<sup>(13)</sup> Furthermore, considering their fundamental importance in numerous disease-related processes, carbohydrate mimetic drugs offer potential new therapeutic applications.<sup>(14)</sup> However, the polar character of carbohydrates creates pharmacokinetic challenges related to oral availability, plasma half-life or renal excretion. Moreover, tight interactions with lectins seem to be against the nature of most carbohydrates, mainly due to the high desolvation costs related to their numerous hydroxyl groups. Toone *et al.* appraised these limitations of carbohydrate-lectin interactions as “fundamental, severe, and likely insurmountable”<sup>(7)</sup> and Hopkins *et al.* regarded the likelihood of modulating a lectin with an orally available small molecule drug to be very low.<sup>(15)</sup> Therefore, to analyze and solve the structural drawbacks common to carbohydrates is of fundamental importance when therapeutic applications are envisaged.

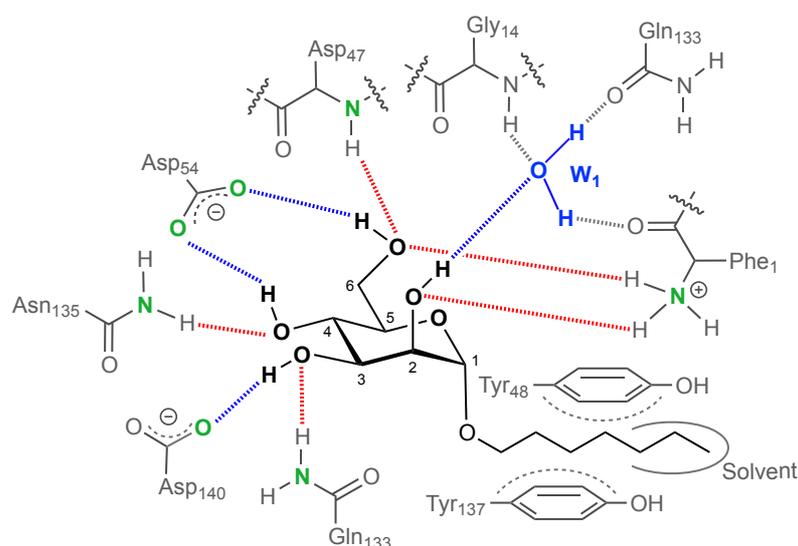
In the present study, the bacterial lectin FimH - one of the rare examples of a lectin undergoing high-affinity interactions with carbohydrates - is analyzed and the role of the individual hydroxyl groups in terms of thermodynamic and kinetic contribution to binding is studied. FimH, a virulence factor of uropathogenic *E. coli* (UPEC), is located at the tip of bacterial type 1 pili.<sup>(16,17)</sup> By interacting with the urothelial glycoprotein uroplakin Ia, it mediates the bacterial adhesion to the bladder wall as the initial step in urinary tract infections (UTI). A high-affinity interaction between FimH and the oligomannosides of the host's uroplakin Ia is a prerequisite to prevent UPEC to be washed out of the bladder by the urinary bulk flow.<sup>(18,19)</sup>

With the goal to reveal their individual contributions of the various hydroxyl groups to binding, the thermodynamic and kinetic properties of the reference compound *n*-heptyl

$\alpha$ -D-mannopyranoside (1) were compared with those of derivatives deoxygenated in the 2-, 3-, 4- or 6-position and those where the hydroxyl groups were replaced by halogens. By combining structural data from X-ray crystallography and solution NMR experiments with computational methods (quantum mechanics, QM) as well as thermodynamic (ITC) and kinetic data (kinITC) the individual contributions of the various hydroxyl groups to binding were analyzed. Therewith, our study completes previous thermodynamic studies with deoxygenated carbohydrates.<sup>(6,20-29)</sup>

## Results and Discussion

**The mannose-binding pocket of FimH (Figure 1).** The interaction of  $\alpha$ -D-mannosides with the lectin domain of FimH (FimH<sub>LD</sub>) has been extensively studied. The pronounced loss of affinity induced by the replacement of the D-mannose moiety by other hexoses, e.g. D-glucose or D-galactose,<sup>(30)</sup> corroborates the importance of the extended hydrogen-bond network which can be established by mannose (Figure 1). As a result, D-mannose exhibits a for carbohydrate-lectin interactions remarkable micromolar affinity of 2.3  $\mu$ M.<sup>24</sup> Furthermore, alkyl or aryl aglycones can establish beneficial hydrophobic interactions with the so-called tyrosine gate (residues Tyr48 and Tyr137) forming the entrance to the mannose-binding pocket.<sup>(31-35)</sup> This effect leads to a further 100-fold improvement of affinity as documented by *n*-heptyl  $\alpha$ -D-mannoside (**1**,  $K_D$ : 22 nM).<sup>(36)</sup>



Acceptor	Donor	Distance [Å]	Angle Donor	Angle Acceptor
O (2-OH)	N (N-terminus)	2.70	161.8	127.2
O (H <sub>2</sub> O)	O (2-OH)	2.68	134.8	112.1
O (3-OH)	N (Gln133)	3.02	173.0	123.0
O (Asp140)	O (3-OH)	2.78	155.8	119.1
O (4-OH)	N (Asn135)	2.95	155.8	114.8
O (Asp54)	O (OH-4)	2.60	165.6	140.6
O (6-OH)	N (Asp47)	2.96	159.2	115.0
O (6-OH)	N (N-terminus)	2.76	165.6	116.9
O (Asp54)	O (6-OH)	2.56	173.6	120.6

**Figure 1. Two-dimensional schematic representation of *n*-heptyl  $\alpha$ -D-mannoside (**1**, HM) binding to FimH according to X-ray crystallography (PDB code 4XO8).** H-bonds donated by **1** are shown in blue, those accepted in red, the structural water W1 is highlighted in blue; to avoid overlaps, Gln133 was drawn twice. In the table, angles and distances from co-crystal structure FimH<sub>LD</sub> in complex with *n*-heptyl  $\alpha$ -D-mannoside (**1**). (PDB 4XO8). Hydrogen atoms were energy minimized with an OPLS\_2005 force field.

The extraordinary high affinity of this carbohydrate-lectin interaction originates from a multitude of factors. First, the interaction of  $\alpha$ -D-mannopyranosides with FimH<sub>LD</sub> (e.g. PDB code 4XO8, 4CST, 4BUQ) is characterized by a total of 9 hydrogen bonds (Figure 1). They exhibit optimal geometries related to distance, donor and acceptor angles (see Table in Figure 1) with the high degree of complementarity necessary for tight binding.<sup>(29,37)</sup> According to MD simulations, the charge-assisted hydrogen bonds formed by Asp54 (acceptor) as well as by the positively charged N-terminus (donor) provide the largest energy contributions.<sup>(38)</sup> Furthermore, compared to a solvent-exposed binding site, electrostatic interactions in the *deeply buried binding pocket* of FimH benefit from a much lower dielectric constant and thus render increased contribution to binding (Table S1).<sup>(39)</sup>

Second, the loss of *rotational freedom* of each hydroxyl group is penalized by entropic costs. However, since all hydroxyl groups of the mannose moiety form multiple hydrogen bonds, entropic costs arise only for the first interaction of a given hydroxyl group. Subsequent interactions from the same hydroxyl group do not induce additional entropic penalty.

Third, the high *desolvation costs* of 26 kJ/mol associated with one hydroxyl group<sup>(40)</sup> might be lowered by perturbed water molecules in the solvation shell of the interaction surfaces due to their polyamphiphilic character.<sup>(29)</sup> However, the high desolvation penalty cannot be compensated by the formation of single hydrogen bond.<sup>(41)</sup> Desolvation costs ( $\Delta G^{\circ}_{\text{solv}}$ ) for the ligand *n*-heptyl  $\alpha$ -D-mannoside and deoxy derivatives thereof were calculated (AMSOL 7.1) and considered for the calculations of the bond energies. The complete desolvation of *n*-heptyl  $\alpha$ -D-mannoside was calculated to cost 38.8 kJ/mol, which is almost as much as the entire binding energy between *n*-heptyl  $\alpha$ -D-mannoside and FimH<sub>LD</sub> ( $\Delta G^{\circ} = -43.7$  kJ/mol) (Table 2). Hence, forming multiple interactions allows compensating the high desolvation penalty, which has to be paid only once, more efficiently.

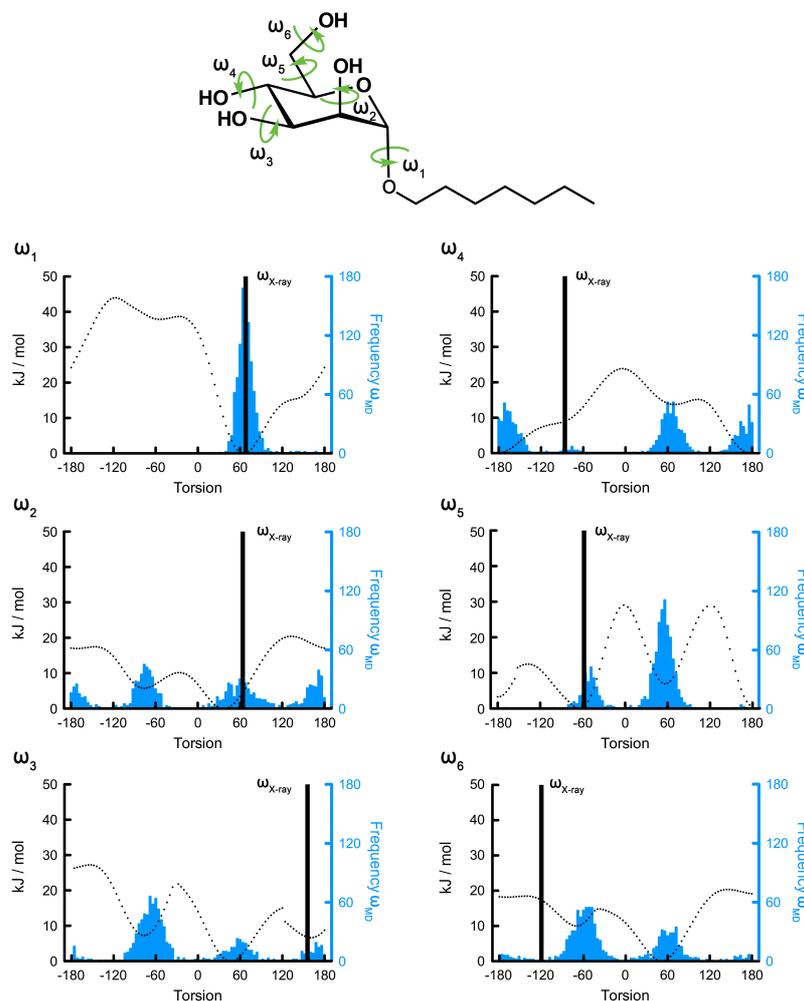
Fourth, H-bonds strengthen each other due to a phenomenon called '*cooperative hydrogen bonding*'.<sup>(42-45)</sup> When a hydroxyl group acts as hydrogen bond donor, the electron density on its oxygen lone pair is slightly increased and therefore its hydrogen bond accepting properties are improved. Since every hydroxyl group of the mannose moiety is involved

in at least two hydrogen bond, the formation of a *cooperative hydrogen bond network* is favored. The structural water (Figure 1,  $W_1$ ) introduces positive cooperativity by interacting with the backbone of Phe1 and thereby increasing the positive charge of the N-terminus that forms charged hydrogen bonds with 2-OH and 6-OH. Opposed to that, negative cooperativity is presumed between 6-OH and 2-OH and between 4-OH and 6-OH, which both have to share a charge on their interaction partners (Asp54, N-terminus). Positive cooperativity is furthermore occurring when multiple amino acids of one protein loop form interactions with the ligand, i.e. observed in the FimH binding site between Asp140 and Asn135 interacting with 3-OH and 4-OH, respectively. This is favorable because the interactions stabilize each other and the entropic costs for the loss of the loop mobility are redundant.<sup>(46,47)</sup> It is therefore delicate to investigate contributions of a single hydroxyl group in a cooperative system as a pyranose, as any change might affect the whole H-bond network.<sup>(48)</sup> Consequently, to attribute the energy loss to the removal of a certain hydroxyl group and to exclude changes of the whole interaction system we carefully considered structural changes observed by NMR and X-ray.

**Protein pre-organization.** In a previous study, we exemplified the rigidity of FimH<sub>LD</sub>, which accommodates also 7-membered ring analogs of D-mannose without any conformational adaption.<sup>(49)</sup> In order to demonstrate the rigidity of FimH<sub>LD</sub>, the “pseudo-apo” structure (PDB-Code: 4AUU, containing thioethanol in the binding site) and the crystal structure co-crystallized with *n*-heptyl  $\alpha$ -D-mannoside (**1**, HM) (PDB-Code: 4XO8) were compared. An RMSD of 0.266 considering the 7 heavy atoms directly interacting with the mannose moiety (see Figure 1, heavy atoms highlighted in green) revealed that ligand binding-induced only minor rearrangements.

**Ligand pre-organization (Figure 2).** In order to derive information regarding the degree of pre-organization of *n*-heptyl  $\alpha$ -D-mannoside (**1**) in the bioactive conformation, torsion angles of solution structures obtained by molecular dynamics simulations ( $\omega_{MD}$ ) were compared with its structure ( $\omega_{X-ray}$ ) in complex with FimH<sub>LD</sub> (PDB-Code: 4XO8). To classify the energy profile of each hydroxyl group, they were rotated stepwise around their torsion angles ( $\omega_1$ - $\omega_6$ ) while calculating the energies of the resulting structures by quantum mechanics (QM). The solution conformations are broadly distributed because they are separated only by low energy barriers. The conformational energy in the bound

state was found to be more favorable than the averaged energies of the solution conformations (Table 1) in four out of six analyzed dihedral angles ( $\omega_{1-3, 5}$ ).

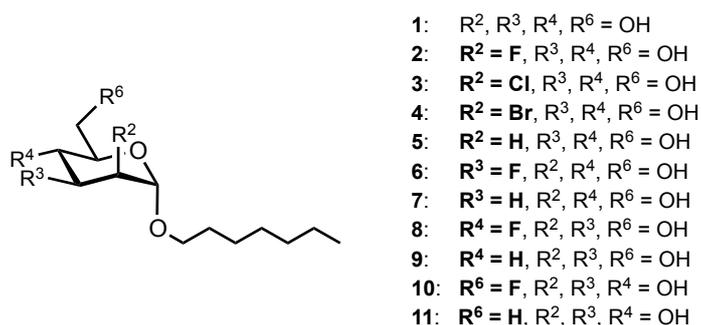


**Figure 2.** Calculated energy as a function of the torsion angles of methyl  $\alpha$ -D-mannopyranoside. Dotted lines correspond to the energy of a given torsion angle. Left y-axis: calculated energy in kJ/mol, right y-axis: frequency of solution conformation from MD simulation, x-axis: torsion angle, conformation of heptyl  $\alpha$ -D-mannopyranoside co-crystallized with FimH-CRD (PDB-Code: 4XO8) shown as a black bar, frequency distribution of solution conformation shown in blue bars.

**Table 1.** Comparing the energy of the torsion angles  $\omega_{1-6}$  of *n*-heptyl  $\alpha$ -D-mannopyranoside (**1**) in solution and bound to FimH<sub>LD</sub>. Torsion angles ( $\omega_{X-ray}$ ) according to the co-crystal structure (PDB-code: 4XO8) and the corresponding calculated energy of the bound conformation in comparison with the averaged energy of the solution conformations over the course of a 9.6 ns MD simulation.

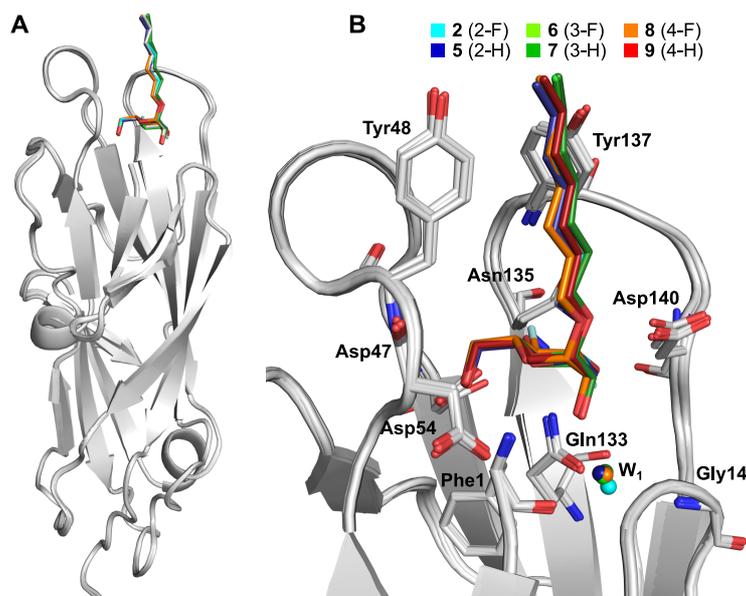
Torsion angles	Solution conformation	Bound conformation		Gain (-) /loss (+) of conformational energy upon binding
	Average Energy [kJ/mol]	Torsion angle ( $\omega_{X-ray}$ )	Energy [kJ/mol]	Energy [kJ/mol]
$\omega_1$	1.5	68.5°	0.1	- 1.4
$\omega_2$	9.9	64.7°	4.6	- 5.3
$\omega_3$	8.8	155.4°	6.8	- 2.0
$\omega_4$	6.8	-86.9°	9.0	+ 2.2
$\omega_5$	8.2	-57.4°	0.4	- 7.8
$\omega_6$	9.4	-119.8°	17.3	+ 7.9

**Synthesis of deoxy- and deoxy-halogeno-derivatives of *n*-heptyl  $\alpha$ -D-mannosides.** To study the contribution of the individual hydroxyl groups, a series of deoxy- and deoxy-halogeno-derivatives of *n*-heptyl  $\alpha$ -D-mannoside (**1**) were synthesized (Figure 3). Since fluorine shares a comparable polarity and a close isosteric relationship with oxygen but is unable to form H-bonds,<sup>(50,51)</sup> a comparison of the parent mannoside **1** with fluorine analogs will allow discriminating between the contribution of H-bonds and other electrostatic interactions. Because the micromolar affinity of unsubstituted D-Man would lead to substantial protein consumption for the thermodynamic analysis, derivatives with an *n*-heptyl aglycone leading to a significantly improved affinity for FimH<sub>LD</sub> were used (D-Mannose:  $K_D = 2.3 \mu\text{M}$ ; *n*-heptyl  $\alpha$ -D-mannoside (**1**):  $K_D = 22 \text{ nM}$ ).<sup>(30,36)</sup> This enabled the study of even those deoxy- and deoxy-halogeno-derivatives that have severely diminished affinities compared to the unmodified mannoside. The synthesis of compounds **2-4**, **6**, **7** and **9-11** is summarized the Supporting Information, while the syntheses of **1**, **5** and **8** have been published recently.<sup>(36)</sup>



**Figure 3.** Structure of *n*-heptyl  $\alpha$ -D-mannoside (**1**) and the deoxy- and deoxy-halogeno-derivatives **2 - 11**.

**Structural analysis of deoxy- (**5**, **7** & **9**) and deoxy-fluoro-derivatives (**2**, **6** & **8**) of *n*-heptyl  $\alpha$ -D-mannoside (**1**) bound to FimH<sub>LD</sub>.** To analyze how the substitution of an individual hydroxyl group influences ligand and lectin conformation, the structures **2** & **5 - 9** co-crystallized with FimH<sub>LD</sub> were determined applying conditions we previously reported for *n*-heptyl  $\alpha$ -D-mannoside (**1**).<sup>(32,36)</sup> Resolutions between 1.9 and 2.1 Å (except, **7** at 3.0 Å) and crystals in two different space groups were obtained (Table S2).



**Figure 4. Superimposition of co-crystal structures of FimH<sub>LD</sub> with *n*-heptyl  $\alpha$ -D-mannoside and six deoxy- and deoxy-fluoro derivatives.** A) The parent *n*-heptyl  $\alpha$ -D-mannoside (1) (PDB code 4XO8) and all proteins are shown in grey, while the deoxy- (5, 7 & 9) and the deoxy-fluoro-derivatives (2, 6 & 8) are colored as indicated. B) Important binding pocket residues and the structural water (W<sub>1</sub>) are shown as sticks and spheres, respectively.

Superposition of the crystal structures revealed RMSD values for the protein backbone between 0.11 and 0.20 Å (Figure 4A) and almost identical binding pockets (Figure 4B). All binding modes were identical, except for the 2-deoxy-mannoside 5 which, however, was slightly tilted (Figure S1). In all crystal structures, the structural water W<sub>1</sub> mediates the H-bond interaction of the 2-hydroxyl group with the backbone amides of Phe1 and Gly14, and as well as to the side chain of Gln133 (Figure 1). In case of the 2-deoxy-2-fluoro-derivative 2, this water molecule is located significantly closer to the protein (Figure 4B).

To analyze the dynamics of the hydrogen bond network of the various complexes <sup>1</sup>H, <sup>15</sup>N-HSQC NMR spectra were recorded. Whereas <sup>1</sup>H chemical shifts of backbone and side chain amides report on H-bond formations, <sup>15</sup>N shifts respond sensitively to changes in the dihedral angles of the protein backbone and side chains.<sup>(52)</sup> We measured <sup>1</sup>H, <sup>15</sup>N-HSQC fingerprint spectra of FimH<sub>LD</sub> in presence of 2 and 5-10 and assigned the signals on the basis of chemical shift proximity to the spectrum with unmodified *n*-heptyl  $\alpha$ -D-mannoside (1). Because of low affinity, mannoside 10 (6-F) and 11 (6-H) modified in 6-position of the mannose moiety could not be evaluated by ITC (Table 2). From the <sup>1</sup>H, <sup>15</sup>N-HSQC NMR spectrum of 10 (6-F) it was apparent that 10 (6-F) binds very weakly

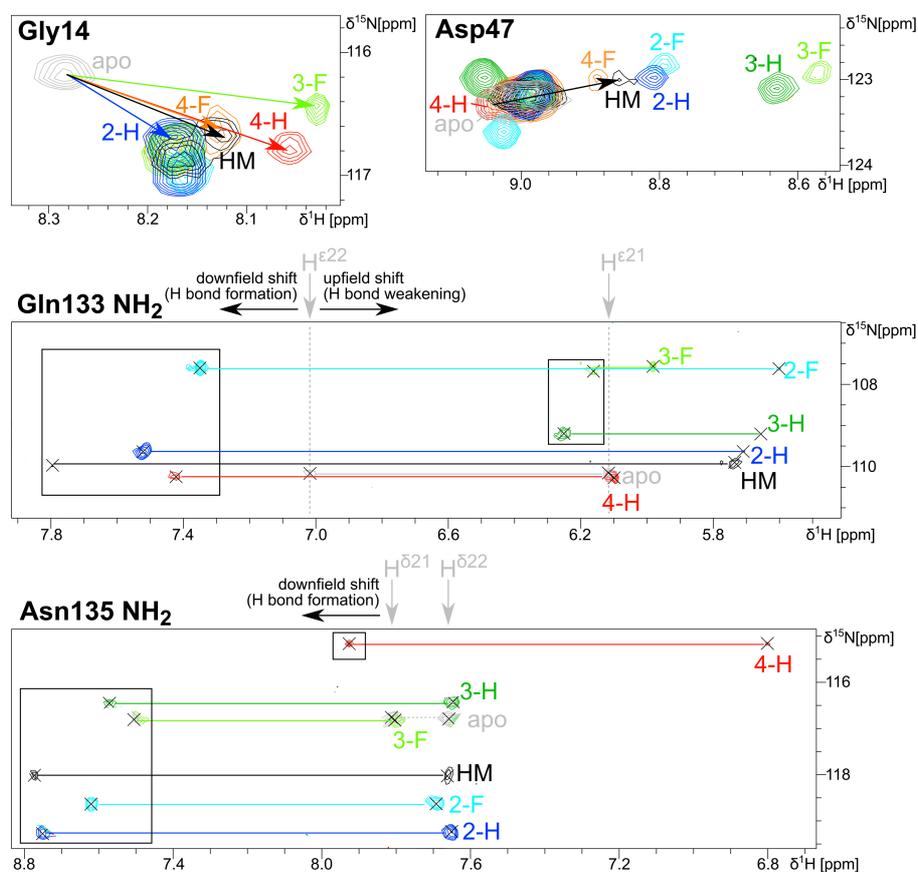
to FimH<sub>LD</sub>, as even at high ligand concentration (20 mM) no chemical shift perturbations were observed. Since the affinity of **11** (6-H) was expected to be even lower than for **10** (6-F), **11** was not subjected to NMR experiments. For all other deoxy- and deoxy-fluoro mannosides, chemical shift perturbation of residues in the binding pocket indicated specific interactions with the protein, while no chemical shift changes of residues remote from the binding pocket were observed (Figure S2; Figure S3).

Downfield shifts indicate new H-bond formation or strengthening, while upfield shifts indicate H-bond weakening or disruption. It was of particular interest to observe chemical shift changes of signals of backbone amide and side chain signals that function as H-bond donors to the ligand's OH groups, i.e. of Gly14 H<sup>N</sup> (W<sub>1</sub> H-bond to 2-OH), Asp47 H<sup>N</sup> (to 6-OH), Gln133 H<sup>ε22</sup> (to 3-OH) and Asn135 H<sup>δ21</sup> (to 4-OH) (Figure 5).

The chemical shifts of Gly14 H<sup>N</sup> are almost identical for **5** (2-H), **6** (3-F), **8** (4-F) and **9** (4-H), indicating a similar orientation of W<sub>1</sub>, even though for **5** (2-H) the interacting hydroxyl group does not exist (Figure 5). The slight proton upfield shift of Gly14 H<sup>N</sup> with all ligands indicates that the coordination of mannose weakens the H-bond of this residue to W<sub>1</sub>. In agreement with this, the data suggest the strongest Gly14-water interaction for **5** (2-H), for which W<sub>1</sub> is more “free” to coordinate with Gly14. For **2** (2-F), Gly14 cannot be assigned and may be shifted significantly due to the close fluorine atom. Indeed, in the co-crystal structure with **2** (2-F), W<sub>1</sub> is more buried (Figure 4B) which may contribute to the large chemical shift change of Gly14. The relative chemical shift changes of Asp47 H<sup>N</sup> suggest that compared to *n*-heptyl α-D-mannoside the direct H-bond to 6-OH is weakened (upfield shift) for the ligand **6** (3-F) and **7** (3-H), strengthened for **9** (4-H) whereas no significant changes were observed for **2** (2-F), **5** (2-H) and **8** (4-F) (Figure 5). The H<sup>ε22</sup> of the Gln133 amide side chain signal is shifted downfield by almost 0.8 ppm upon addition of *n*-heptyl α-D-mannoside indicative for the H-bond formation to 3-OH (Figure 5). For **2** (2-F), **5** (2-H) and **9** (4-H), similar downfield shifts demonstrate H-bond formation to 3-OH, although the smaller shifts (0.33–0.50 ppm) suggest slightly weaker H-bonds. Importantly, the Gln133 H<sup>ε22</sup> signals with **6** (3-F) and **7** (3-H) show dramatic upfield shifts [ca. –1.64 ppm relative to *n*-heptyl α-D-mannoside (**1**)] as direct evidence for the absence of the corresponding H-bond. Similarly, Asn135 H<sup>δ21</sup> is strongly shifted downfield by 0.70 to 0.96 ppm upon addition of *n*-heptyl α-D-mannoside (**1**), **2** (2-F), **5**

(2-H), 6 (3-F) and 7 (3-H) as a consequence of H-bond formation to 4-OH. With 9 (4-H), a relative upfield shift of this signal reports on the absence of the corresponding H-bond. For 8 (4-F), the Asn135 side chain signals could not be assigned.

In summary, the structural information obtained by NMR and X-ray is in good agreement and the conformational changes of protein and ligand are remarkably small upon the loss of relevant interactions. This qualifies FimH<sub>LD</sub> as a model to investigate contributions of distinct hydroxyl groups by deoxy- and deoxy-halogeno derivatives.



**Figure 5.**  $^1\text{H}$ ,  $^{15}\text{N}$ -HSQC spectral regions of binding pocket residues directly involved in H-bond donation to the mannosyl moiety. Boxes in case of side chain signals of Gln133 and Asn135 indicate relative downfield and upfield proton shifts, respectively. Peaks not of interest were faded-out to improve the clarity.

**Thermodynamic analysis of deoxy- and deoxy-halogeno-derivatives of *n*-heptyl  $\alpha$ -D-mannoside (1) (Table 2).** Measuring interaction by ITC is valuable in several respects. First, it is a label-free method that allows the determination of the change in enthalpy of

binding ( $\Delta H^\circ$ ), association constant ( $K_A$ ) and stoichiometry ( $N$ ) in one experiment, and second, recent progress in data analysis now allows the deviation of kinetic rate constants for association and dissociation from the same raw data. ITC experiments (Table 2) were carried out at 25 °C in 10 mM HEPES buffer at pH 7.4 containing 150 mM NaCl. Measurements in other buffering systems revealed a significant heat of ionization ( $\Delta H^\circ_{\text{ion}}$ ) originating from a partial proton transfer from the solvent to the N-terminus of the FimH<sub>LD</sub> protein (Figure S4, Table S3).<sup>(53,54)</sup> By measuring the enthalpy of *n*-heptyl  $\alpha$ -D-mannoside (**1**) binding to FimH<sub>LD</sub> in different buffers at pH 7.4 (Cacodylate, HEPES, Tris), pH 8.5 (Bicine, Tricine, TAPS, Tris) and pH 9.0 (Bicine; TAPS, Tris), the pK<sub>a</sub> of the N-terminus was determined to be approximately 8.3. This proton transfer is endothermic and therefore the intrinsic enthalpy is underestimated applying standard experiment conditions. However, the relevant  $\Delta\Delta$ -values are constant and we decided to omit correction for the heat of ionization in this publication to maintain the comparability with earlier publications.

**Table 2. Change in thermodynamic parameters of FimH<sub>LD</sub> binding to *n*-heptyl  $\alpha$ -D-mannoside (**1**) and deoxy- (**5**, **7** & **9**) and deoxy-halogeno derivatives (**2-4**, **6** & **8**) thereof.** All values are relative to the absolute values of mannoside **1**. Measurements were carried out at 298.15 K in 10 mM HEPES buffer containing 150 mM NaCl at pH 7.4. Protein and ligand concentrations, confidence intervals and stoichiometries are part of the supplementary information (Table S5). Affinities from a fluorescence polarization (FP) assay are shown for comparison.

	FP: $K_D$ [ $\mu\text{M}$ ]	ITC: $K_D$ [ $\mu\text{M}$ ]	$\Delta G^\circ$ [kJ/mol]	$\Delta H^\circ$ [kJ/mol]	$-T\Delta S^\circ$ [kJ/mol]
<i>n</i> -heptyl mannoside ( <b>1</b> )	0.028 <sup>(32)</sup>	0.022	-43.7	-50.5	6.7
			$\Delta\Delta G^\circ$ [kJ/mol]	$\Delta\Delta H^\circ$ [kJ/mol]	$-T\Delta\Delta S^\circ$ [kJ/mol]
<b>2</b> (2-F)	0.44	0.53	7.9	22.7	-14.8
<b>3</b> (2-Cl)	2.26	1.34	10.2	25.9	-15.8
<b>4</b> (2-Br)	4.24	1.83	11.0	29.6	-18.7
<b>5</b> (2-H)	13.95	9.77	15.1	31.1	-16.0
<b>6</b> (3-F)	0.45	1.19	9.9	19.1	-9.2
<b>7</b> (3-H)	23.35	19.88	16.9	25.7	-8.8
<b>8</b> (4-F)	135.2	103.43	21.0	29.9	-9.0
<b>9</b> (4-H)	101.2	88.40	20.6	30.6	-10.0
<b>10</b> ( <b>6-F</b> )	n.d.	>1000			
<b>11</b> ( <b>6-H</b> )	n.d.	>1000			

**Deoxygenation and substitution of 2-OH of mannose.** Substitutions in the 2-position of the mannose moiety have the lowest impact on the free energy of binding [e.g.  $\Delta\Delta G^\circ = 15.1$  kJ/mol for **5** (2-H)]. However, 2-deoxy mannoside **5** suffers from the largest loss in enthalpy of the whole series relative to **1** ( $\Delta\Delta H^\circ = 31.1$  kJ/mol), which is partially

compensated by a gain in entropy ( $-T\Delta\Delta S^\circ = -16.0$  kJ/mol). Overall, this leads to a 440-fold loss in affinity. The unusually high enthalpic loss originates from the excellent H-bond of the 2-OH group with the positively charged N-terminus. Furthermore, structural analysis by X-ray indicated that the mannose moiety of **5** (2-H) is slightly twisted relative to the parent *n*-heptyl  $\alpha$ -D-mannoside (**1**), weakening the H-bond network and thus increasing the flexibility of the ligand in the binding pocket (Figure S1). However, the twist could not be confirmed by NMR data, as all binding pocket residues have nearly identical chemical shifts compared to **1**. Phe1 might indeed be more flexible due to the absence of the H-bond to 2-OH, but unfortunately the N-terminal Phe1 signal cannot be observed directly in the NMR experiments. A further insight into the binding mode was obtained when the 2-hydroxyl group was replaced by halogens. It was expected that the electrostatic interaction exerted by the former oxygen atom can be maintained, whereas obviously the H-bonds formed with W1 and Phe<sub>1</sub> are lost. The increasing electronegativity of bromine < chlorine < fluorine induces a repulsion of the structural water W<sub>1</sub> (Figure 4B), while the interaction with the positively charged N-terminus becomes more beneficial. The difference between the loss of all electrostatic interactions for **5** (2-H) ( $\Delta\Delta G^\circ = 15.1$  kJ/mol) and the loss of only the H-bonds of **2** (2-F) ( $\Delta\Delta G^\circ = 7.9$  kJ/mol) leads to the conclusion that the *permanent dipole-buried charge interaction* contributes approximately 7.2 kJ/mol to the free energy of binding.

**Deoxygenation and substitution of 3-OH of mannose.** Deoxygenation of the 3-position of *n*-heptyl  $\alpha$ -D-mannoside (**1**) ( $\rightarrow$  **7**) leads to a substantial loss in enthalpy ( $\Delta\Delta H^\circ = 25.7$  kJ/mol), which is only partially compensated by an entropy gain ( $-T\Delta\Delta S^\circ = -8.8$  kJ/mol). As a result, a 900-fold loss in affinity was observed. Compared to **7** (3-H), the introduction of a fluoro substituent [ $\rightarrow$  **6** (3-F)] results in an enthalpic gain ( $\Delta\Delta H^\circ = -6.6$  kJ/mol), as well as a small entropic advantage ( $-T\Delta\Delta S^\circ = -0.4$  kJ/mol). Interestingly, crystal structures of *n*-heptyl  $\alpha$ -D-mannoside (**1**) and the 3-fluoro derivative **6** co-crystallized with FimH<sub>LD</sub> (PDB QQ & PDB QQ) reveal repulsion between the fluorine atom and the negative charge on Asp140 and consequently an increased distance (0.5 Å) between the interacting heavy atoms. However, Asp140 is part of a loop and its perpendicular position to the ligand allow for a rearrangement with minor impact on binding energy. The difference between the intrinsic loss of all electrostatic interactions for **7** (3-H) ( $\Delta\Delta G^\circ = 16.9$  kJ/mol) and the loss of only the H-bonds to Gln133 and

Asp140 of **6** (3-F) ( $\Delta\Delta G^\circ = 9.9$  kJ/mol) discloses a contribution of approximately 7.0 kJ/mol for the *charge-permanent dipole interaction* and 9.9 kJ/mol for the H-bonds to the free energy of binding of 3-OH. The close contact of two complementary surfaces in the absence of solvent water may be an explanation for the large contribution of the charge-dipole interaction.<sup>(55-57)</sup>

**Deoxygenation and substitution of 4-OH of mannose.** The 4-deoxy derivative **9** (4-H) suffers from a 4'000-fold loss of affinity. Compared to the effect of the substitution of the 3-hydroxyl group, the affinity is reduced by another factor of 4, although number and type of H-bond interactions of 3-OH and 4-OH are comparable. However, C-4 is buried more deeply in the pocket than C-3, where a lower dielectric constant and consequently an increased energy for electrostatic interactions are expected.<sup>(58)</sup> Interestingly, in contrast to the 2- or 3-position, the introduction of a fluorine substituent in the 4-position does not improve binding energy compared to the 4-deoxy derivative. Most likely, the electronegative fluoride induces a repulsion of the negatively charged Asp54, which is part of a rigid  $\beta$ -sheet. Crystal structures reveal a change of the orientation between 4-OH and 4-F (dihedral angle HO4-C4-C5-O =  $-174.9^\circ$  and F4-C4-C5-O =  $-167.3^\circ$ ), which supports this assumption. Therefore, the loss of binding energy upon the removal of 4-OH [ $\rightarrow$  **9** (4-H),  $\Delta\Delta G^\circ = 20.6$  kJ/mol] is almost identical resulting from the loss of the hydrogen bonds [ $\rightarrow$  **8** (4-F),  $\Delta\Delta G^\circ = 21.0$  kJ/mol].

**Deoxygenation and substitution of 6-OH of mannose.** Finally, modifications in the 6-position of the mannose moiety have the most severe impact on the binding energy. Competitive ITC titrations with *n*-heptyl mannoside (**1**) when FimH<sub>LD</sub> was preincubated with a 300-fold excess of **10** (6-F) or **11** (6-H) did neither induce a change in binding affinity nor enthalpy. Furthermore, direct titration with a 6 mM solution of mannoside **10** (6-F) did not result in a signal change in the ITC isotherm. We conclude that the binding affinity of **10** (6-F) and **11** (6-H) can be expected to be below 1 mM. The 6-hydroxy group forms 3 excellent hydrogen bonds, whereof two are charge assisted (N-terminus, Asp54). Considering the reduced desolvation costs and number of rotational bonds of derivatives deoxygenated or fluorated in the 6-position when compared to mannoside **1**, an approximate 30'000-fold loss in affinity is surprising and reveals the tremendous importance of this deeply buried hydroxyl group.

**Kinetic analysis of *n*-heptyl deoxy- and deoxy-halogeno- $\alpha$ -D-mannosides (Table 3).**

With kinetic isothermal titration calorimetry (kinITC) kinetic data are accessible from ITC measurements.<sup>(59)</sup> It measures the time that the differential power curve takes to return to baseline after an injection of ligand solution that is a function binding kinetics. Narrower peaks at the beginning and the end of the titration, and wider peaks around the inflection point yield a bell-shaped equilibration time curve (ETC) that can be analyzed to determine dissociation rate constant  $k_{\text{off}}$ . The association rate constant  $k_{\text{on}}$  is calculated from  $k_{\text{off}}$  and the equilibrium constant ( $K_D = k_{\text{off}} / k_{\text{on}}$ ) (Table 3).

**Table 3. Kinetic binding parameters for the interaction of FimH<sub>LD</sub> with deoxy- and deoxy-halogeno derivatives of *n*-heptyl  $\alpha$ -D-mannoside (1) determined by kinITC.** Confidence intervals of the fitted parameters  $k_{\text{on}}$ ,  $k_{\text{off}}$ ,  $K_D$ , and the response time are part of the supplementary information (Table S7). Relative changes ( $r k_{\text{on}}$ ,  $r k_{\text{off}}$ ) are compared to *n*-heptyl  $\alpha$ -D-mannoside (1).

Compound	Method	$k_{\text{on}}$ [M <sup>-1</sup> s <sup>-1</sup> ]	1/ $r k_{\text{on}}$ n	$k_{\text{off}}$ [s <sup>-1</sup> ]	$r k_{\text{off}}$	$t_{1/2}$ [min]
1 (HM)	ITC	3.32*10 <sup>4</sup>	1.0	7.27*10 <sup>-4</sup>	1.0	15.90
2 (2-F)	ITC	2.04*10 <sup>4</sup>	1.6	1.02*10 <sup>-2</sup>	14.0	1.14
3 (2-Cl)	ITC	8.90*10 <sup>3</sup>	3.7	1.19*10 <sup>-2</sup>	16.4	0.97
4 (2-Br)	ITC	7.90*10 <sup>3</sup>	4.2	1.45*10 <sup>-2</sup>	19.9	0.80
5 (2-H)	ITC	5.35*10 <sup>3</sup>	6.2	4.13*10 <sup>-2</sup>	56.8	0.28
5 (2-H)	SPR	1.05*10 <sup>4</sup>	3.1	5.08*10 <sup>-2</sup>	71.1	0.23
6 (3-F)	ITC	1.66*10 <sup>4</sup>	2.0	1.98*10 <sup>-2</sup>	27.3	0.58
7 (3-H)	ITC	6.50*10 <sup>3</sup>	5.1	1.29*10 <sup>-1</sup>	177.9	0.09
8 (4-F)	ITC	1.00*10 <sup>3</sup>	33.2	1.03*10 <sup>-1</sup>	142.3	0.11
9 (4-H)	ITC	9.56*10 <sup>2</sup>	34.7	8.45*10 <sup>-1</sup>	116.3	0.14

Binding kinetics for **5** (2-H) determined either by surface plasmon resonance (SPR) or kinITC-ETC are in excellent agreement and exhibited deviations smaller than a factor of 2 (Table 3). In contrast to other carbohydrate-lectin interactions, FimH<sub>LD</sub> is characterized by unusually slow dissociation rates.<sup>(60)</sup> The reduced affinity of all deoxy- and deoxy-halogeno derivatives is mainly resulting from increased  $k_{\text{off}}$  values, leading to substantially reduced complex half-lives ( $t_{1/2} = \ln 2 / k_{\text{off}}$ ). For example, while  $t_{1/2}$  of *n*-heptyl  $\alpha$ -D-mannoside is more than 16 minutes, it is reduced to less than a minute for all deoxy derivatives. Although the removal of a hydroxyl group in general has a smaller effect on the  $k_{\text{on}}$  than  $k_{\text{off}}$  values, the association rates deliver an interesting insight into the binding process. Especially the most buried and charge-assisted interaction of 4-OH turns out to be important for the association of **9** (4-H), since its removal leads to a more than 30-fold reduction of the  $k_{\text{on}}$  value. A much smaller effect was observed for **5** (2-H) and **7** (3-H). An interesting trend can be observed for halogen substituents in 2-position, where

the decreasing electronegativity (fluorine > chlorine > bromine) affects the  $k_{\text{on}}$  values to a significantly greater extent than  $k_{\text{off}}$  values. However, we can not rule out that increasing atomic radii could have a steric effect on the formation of the protein-ligand complexes. In summary, We hypothesize that changes in the short-range electrostatic interactions of FimH-mannoside complexes (e.g. hydrogen bonds, dipole-dipole interactions) mainly affect the dissociation rate, while medium-range electrostatics, such as the dipole-charged interaction of 2-OH with the N-terminus, are additionally of importance for the association rate as described before for protein-protein interactions.<sup>(61-64)</sup>

## Conclusion

We demonstrated how the lectin FimH applies a broad spectrum of strategies to overcome the intrinsically low binding affinities of a carbohydrate ligand. In particular, quantum mechanically derived torsional barriers of a mannoside ligand were calculated and the solution conformations were compared to the bound conformation obtained from crystal structures. FimH<sub>LD</sub> turned out to be well pre-organized and to bind hydroxyl groups mainly in low energy orientations.

Furthermore, it was possible to analyze the contributions of three hydroxyl groups of **1** (HM) and its deoxy- and deoxy-halogeno derivatives by ITC thermodynamically and kinetically by considering their structural binding properties. As expected, the removal of a hydroxyl group was accompanied by a loss of enthalpy ( $\Delta\Delta H^\circ = 26 - 31$  kJ/mol) and a smaller gain in entropy ( $-T\Delta\Delta S^\circ = -9 - -16$  kJ/mol). Hence, contributions to  $\Delta G^\circ$  could be determined as 15 kJ/mol for 2-OH close to the protein surface, 17 kJ/mol for 3-OH with a medium position, and 21 kJ/mol for the most buried hydroxyl group 4-OH. We observed a general tendency that deeper buried hydroxyl groups contribute more to the overall binding affinity of mannoside ligands. This effect was most stark for removal of the hydroxyl group in position C-6, resulting in the total loss of binding upon. Furthermore,

It furthermore became clear that the contribution of a single hydroxyl group is larger than the sum of the loss of its H-bonds and van der Waals interactions. Summing up the  $\Delta G^\circ$  values of the hydroxyl groups 2-4 (without including the essential 6-OH) yields a cumulative contribution of 53 kJ/mol which exceeds the total free binding energy  $\Delta G^\circ$  of *n*-heptyl  $\alpha$ -D-mannoside. The explanation is that H-bonds build a cross-linked network in which only the cooperative interplay between multiple H-bonds is stabilizing the protein-ligand complex. Consequently, the loss of one single hydroxyl group (6-OH) can prevent a molecule from binding, although the remaining hydroxyl groups could form 6 hydrogen bonds that are structurally identical to those observed for the unmodified ligand.

Moreover, by performing an analysis with kinITC, additional kinetic values were obtained. As expected, the loss of hydroxyl groups had a stronger influence on  $k_{\text{off}}$  than

on  $k_{\text{on}}$ . However, the short-range electrostatic H-bonds were found to be less important for  $k_{\text{on}}$ , while the medium-range electrostatic interactions with the positively charged N-terminus (2-OH) and the negatively charged Asp54 (4-OH) were of great importance for an increased association rate.

The detailed dissection of mannose binding to FimH<sub>LD</sub> enhances our general understanding of carbohydrate binding, and give insights how nature successfully developed high affinity carbohydrate-protein interactions. We believe, this knowledge will be of great importance for the future design and development of glycomimetic drug candidates.

### **Acknowledgements**

We thank Prof. Stephan Grzesiek (University of Basel) for the access to a 600 MHz NMR spectrometer. We furthermore appreciate the support by Eva Muñoz for the use of kinITC (AFFINImeter software). The authors gratefully acknowledge for a scholarship of B.F. by the German Academic Exchange Service (DAAD).

## References

1. Thomson, J., Liu, Y., Sturtevant, J.M., & Quijcho, F.A. (1998). A thermodynamic study of the binding of linear and cyclic oligosaccharides to the maltodextrin-binding protein of *Escherichia coli*. *Biophys. Chem.*, 70(2): 101-108.
2. Gupta, G., Gemma, E., Oscarson, S., & Surolia, A. (2008). Defining substrate interactions with calreticulin: an isothermal titration calorimetric study. *Glycoconj. J.*, 25(8): 797-802.
3. Kapoor, M., *et al.* (2004). Mutational analysis provides molecular insight into the carbohydrate-binding region of calreticulin: pivotal roles of tyrosine-109 and aspartate-135 in carbohydrate recognition. *Biochemistry*, 43(1): 97-106.
4. Dam, T.K., Roy, R., Das, S.K., Oscarson, S., & Brewer, C.F. (2000). Binding of multivalent carbohydrates to concanavalin A and Dioclea grandiflora lectin. Thermodynamic analysis of the "multivalency effect". *J. Biol. Chem.*, 275(19): 14223-14230.
5. Quijcho, F.A. (1993). Probing the atomic interactions between proteins and carbohydrates. *Biochem. Soc. Trans.*, 21(2): 442-448.
6. Daranas, A.H., Shimizu, H., & Homans, S.W. (2004). Thermodynamics of binding of D-galactose and deoxy derivatives thereof to the L-arabinose-binding protein. *J. Am. Chem. Soc.*, 126(38): 11870-11876.
7. Burkhalter, N.F., Dimick, S.M., & Toone, E.J. (2008). Protein-Carbohydrate Interaction: Fundamental Considerations. *Carbohydrates in Chemistry and Biology*, eds Ernst B, Hart GW, & Sinaý P, Wiley-VCH Verlag GmbH. pp. 863-914.
8. Lis, H. & Sharon, N. (1998). Lectins: Carbohydrate-specific proteins that mediate cellular recognition. *Chem. Rev.*, 98(2): 637-674.
9. Aplin, A.E., Howe, A., Alahari, S.K., & Juliano, R.L. (1998). Signal transduction and signal modulation by cell adhesion receptors: the role of integrins, cadherins, immunoglobulin-cell adhesion molecules, and selectins. *Pharmacol. Rev.*, 50(2): 197-263.
10. Lasky, L.A. (1992). Selectins: Interpreters of Cell-Specific Carbohydrate Information During Inflammation. *Science*, 258(5084): 964-969.
11. Sharon, N. & Lis, H. (1995). Lectins--proteins with a sweet tooth: functions in cell recognition. *Essays in biochemistry*, 30: 59-75.
12. Weis, W.I. & Drickamer, K. (1996). Structural basis of lectin-carbohydrate recognition. *Annual review of biochemistry*, 65: 441-473.
13. Ley, K., Laudanna, C., Cybulsky, M.I., & Nourshargh, S. (2007). Getting to the site of inflammation: the leukocyte adhesion cascade updated. *Nature reviews. Immunology*, 7(9): 678-689.
14. Ernst, B. & Magnani, J.L. (2009). From carbohydrate leads to glycomimetic drugs. *Nature reviews. Drug discovery*, 8(8): 661-677.
15. Hopkins, A.L. & Groom, C.R. (2002). The druggable genome. *Nature reviews Drug discovery*, 1(9): 727-730.
16. Hartmann, M. & Lindhorst, T.K. (2011). The Bacterial Lectin FimH, a Target for Drug Discovery - Carbohydrate Inhibitors of Type 1 Fimbriae-Mediated Bacterial Adhesion. *Eur. J. Org. Chem.*, (20-21): 3583-3609.
17. Abgottspson, D. & Ernst, B. (2012). In vivo Evaluation of FimH Antagonists - A Novel Class of Antimicrobials for the Treatment of Urinary Tract Infection. *Chimia*, 66(4): 166-169.
18. Connell, H., *et al.* (1996). Type 1 fimbrial expression enhances *Escherichia coli* virulence for the urinary tract. *Proc. Natl. Acad. Sci. U. S. A.*, 93(18): 9827-9832.
19. Wellens, A., *et al.* (2008). Intervening with urinary tract infections using anti-adhesives based on the crystal structure of the FimH–oligomannose-3 complex. *PLoS One*, 3(4): e2040.
20. Rani, P.G., Bachhawat, K., Reddy, G.B., Oscarson, S., & Surolia, A. (2000). Isothermal titration calorimetric studies on the binding of deoxytrimannoside derivatives with artocarpin: implications for a deep-seated combining site in lectins. *Biochemistry*, 39(46): 14364.
21. Swaminathan, C.P., Gupta, D., Sharma, V., & Surolia, A. (1997). Effect of substituents on the thermodynamics of D-galactopyranoside binding to winged bean (*Psophocarpus tetragonolobus*) basic lectin. *Biochemistry*, 36(43): 13428-13434.
22. Schwarz, F.P., Misquith, S., & Surolia, A. (1996). Effect of substituent on the thermodynamics of D-glucopyranoside binding to concanavalin A, pea (*Pisum sativum*) lectin and lentil (*Lens culinaris*) lectin. *Biochem. J.*, 316 ( Pt 1): 123-129.

23. Gupta, D., Dam, T.K., Oscarson, S., & Brewer, C.F. (1997). Thermodynamics of lectin-carbohydrate interactions. Binding of the core trimannoside of asparagine-linked carbohydrates and deoxy analogs to concanavalin A. *J. Biol. Chem.*, 272(10): 6388-6392.
24. Dam, T.K., Oscarson, S., & Brewer, C.F. (1998). Thermodynamics of binding of the core trimannoside of asparagine-linked carbohydrates and deoxy analogs to Dioclea grandiflora lectin. *J. Biol. Chem.*, 273(49): 32812-32817.
25. Dam, T.K., *et al.* (2011). Fine specificities of two lectins from *Cymbosema roseum* seeds: a lectin specific for high-mannose oligosaccharides and a lectin specific for blood group H type II trisaccharide. *Glycobiology*, 21(7): 925-933.
26. Winter, H.C., Oscarson, S., Slattegard, R., Tian, M., & Goldstein, I.J. (2005). Banana lectin is unique in its recognition of the reducing unit of 3-O-beta-glucosyl/mannosyl disaccharides: a calorimetric study. *Glycobiology*, 15(10): 1043-1050.
27. Solis, D., Romero, A., Kaltner, H., Gabius, H.J., & DiazMaurino, T. (1996). Different architecture of the combining site of the two chicken galectins revealed by chemical mapping studies with synthetic ligand derivatives. *J. Biol. Chem.*, 271(22): 12744-12748.
28. Clarke, C., *et al.* (2001). Involvement of water in carbohydrate-protein binding. *J. Am. Chem. Soc.*, 123(49): 12238-12247.
29. Lemieux, R.U. (1996). How water provides the impetus for molecular recognition in aqueous solution. *Acc. Chem. Res.*, 29(8): 373-380.
30. Bouckaert, J., *et al.* (2005). Receptor binding studies disclose a novel class of high-affinity inhibitors of the *Escherichia coli* FimH adhesin. *Mol. Microbiol.*, 55(2): 441-455.
31. Han, Z., *et al.* (2010). Structure-Based Drug Design and Optimization of Mannoside Bacterial FimH Antagonists. *J. Med. Chem.*, 53(12): 4779-4792.
32. Kleeb, S., *et al.* (2015). FimH antagonists: bioisosteres to improve the in vitro and in vivo PK/PD profile. *J. Med. Chem.*, 58(5): 2221-2239.
33. Pang, L., *et al.* (2012). FimH antagonists: structure-activity and structure-property relationships for biphenyl alpha-D-mannopyranosides. *ChemMedChem*, 7(8): 1404-1422.
34. Schwardt, O., *et al.* (2011). Design, synthesis and biological evaluation of mannosyl triazoles as FimH antagonists. *Bioorg. Med. Chem.*, 19(21): 6454-6473.
35. Jiang, X., *et al.* (2012). Antiadhesion therapy for urinary tract infections--a balanced PK/PD profile proved to be key for success. *J. Med. Chem.*, 55(10): 4700-4713.
36. Fiege, B., *et al.* (2015). The Tyrosine Gate of the Bacterial Lectin FimH: A Conformational Analysis by NMR Spectroscopy and X-ray Crystallography. *ChemBioChem*, 16(8): 1235-1246.
37. Steiner, T. (2002). The hydrogen bond in the solid state. *Angew. Chem. Int. Ed.*, 41(1): 48-76.
38. Zalewski, A. (2013). *In pursuit of a novel UTI treatment strategy - an "in silico" study of the FimH adhesin*. Diss Phil -Nat Univ Basel, 2013 - Ref : A Vedani, J Bouckaert, University of Basel, Basel.
39. Quijcho, F.A. (1986). Carbohydrate-binding proteins: tertiary structures and protein-sugar interactions. *Annu. Rev. Biochem.*, 55: 287-315.
40. Cabani, S., Gianni, P., Mollica, V., & Lepori, L. (1981). Group contributions to the thermodynamic properties of non-ionic organic solutes in dilute aqueous solution. *J. Solution Chem.*, 10(8): 563-595.
41. Vedani, A. & Dunitz, J.D. (1985). Lone-pair directionality in hydrogen-bond potential functions for molecular mechanics calculations: the inhibition of human carbonic anhydrase II by sulfonamides. *J. Am. Chem. Soc.*, 107(25): 7653-7658.
42. Vyas, N.K. (1991). Atomic features of protein-carbohydrate interactions. *Curr. Opin. Struct. Biol.*, 1(5): 732-740.
43. Jeffrey, G.A. & Saenger, W. (1991). *Hydrogen Bonding in Biological Structures*. Springer-Verlag, Berlin.
44. López de la Paz, M., *et al.* (2002). Carbohydrate Hydrogen-Bonding Cooperativity – Intramolecular Hydrogen Bonds and Their Cooperative Effect on Intermolecular Processes – Binding to a Hydrogen-Bond Acceptor Molecule. *Eur. J. Org. Chem.*, 2002(5): 840-855.
45. Frank, H.S. & Wen, W.-Y. (1957). Ion-solvent interaction. Structural aspects of ion-solvent interaction in aqueous solutions: a suggested picture of water structure. *Discuss. Faraday Soc.*, 24(0): 133-140.
46. Vorov, O.K., Livesay, D.R., & Jacobs, D.J. (2011). Nonadditivity in conformational entropy upon molecular rigidification reveals a universal mechanism affecting folding cooperativity. *Biophys. J.*, 100(4): 1129-1138.
47. Vorov, O.K., Livesay, D.R., & Jacobs, D.J. (2008). Conformational entropy of an ideal cross-linking polymer chain. *Entropy*, 10(3): 285-308.

48. Baum, B., *et al.* (2010). Non-additivity of Functional Group Contributions in Protein–Ligand Binding: A Comprehensive Study by Crystallography and Isothermal Titration Calorimetry. *J. Mol. Biol.*, 397(4): 1042-1054.
49. Sager, C.P., *et al.* (2018). The price of flexibility - a case study on septanoses as pyranose mimetics. *Chemical Science*.
50. Bondi, A. (1964). Van der Waals Volumes and Radii. *J. Phys. Chem.*, 68(3): 441-451.
51. Hoffmann, M. & Rychlewski, J. (2002). When, in the context of drug design, can a fluorine atom successfully substitute a hydroxyl group? *Int. J. Quantum Chem*, 89(4): 419-427.
52. Williamson, M.P. (2013). Using chemical shift perturbation to characterise ligand binding. *Prog. Nucl. Magn. Reson. Spectrosc.*, 73: 1-16.
53. Goldberg, R.N., Kishore, N., & Lennen, R.M. (2002). Thermodynamic Quantities for the Ionization Reactions of Buffers. *J. Phys. Chem. Ref. Data*, 31(2): 231-370.
54. Baker, B.M. & Murphy, K.P. (1996). Evaluation of linked protonation effects in protein binding reactions using isothermal titration calorimetry. *Biophys. J.*, 71(4): 2049-2055.
55. Malham, R., *et al.* (2005). Strong solute-solute dispersive interactions in a protein-ligand complex. *J. Am. Chem. Soc.*, 127(48): 17061-17067.
56. Barratt, E., *et al.* (2005). Van der Waals interactions dominate ligand-protein association in a protein binding site occluded from solvent water. *J. Am. Chem. Soc.*, 127(33): 11827-11834.
57. Yang, L., Adam, C., Nichol, G.S., & Cockroft, S.L. (2013). How much do van der Waals dispersion forces contribute to molecular recognition in solution? *Nat. Chem.*, 5(12): 1006-1010.
58. Gohlke, H. & Klebe, G. (2002). Approaches to the Description and Prediction of the Binding Affinity of Small-Molecule Ligands to Macromolecular Receptors. *Angew. Chem. Int. Ed.*, 41(15): 2644-2676.
59. Burnouf, D., *et al.* (2012). kinITC: a new method for obtaining joint thermodynamic and kinetic data by isothermal titration calorimetry. *J. Am. Chem. Soc.*, 134(1): 559-565.
60. Scharenberg, M., *et al.* (2014). Kinetic Properties of Carbohydrate–Lectin Interactions: FimH Antagonists. *ChemMedChem*, 9(1): 78-83.
61. Chari, R., Jerath, K., Badkar, A., & Kalonia, D. (2009). Long- and Short-Range Electrostatic Interactions Affect the Rheology of Highly Concentrated Antibody Solutions. *Pharm. Res.*, 26(12): 2607-2618.
62. Darling, R.J., *et al.* (2002). Glycosylation of Erythropoietin Affects Receptor Binding Kinetics: Role of Electrostatic Interactions. *Biochemistry*, 41(49): 14524-14531.
63. Schreiber, G., Haran, G., & Zhou, H.X. (2009). Fundamental Aspects of Protein–Protein Association Kinetics. *Chem. Rev.*, 109(3): 839-860.
64. Pan, A.C., Borhani, D.W., Dror, R.O., & Shaw, D.E. (2013). Molecular determinants of drug–receptor binding kinetics. *Drug Discov. Today*, 18(13): 667-673.

## Supporting Information

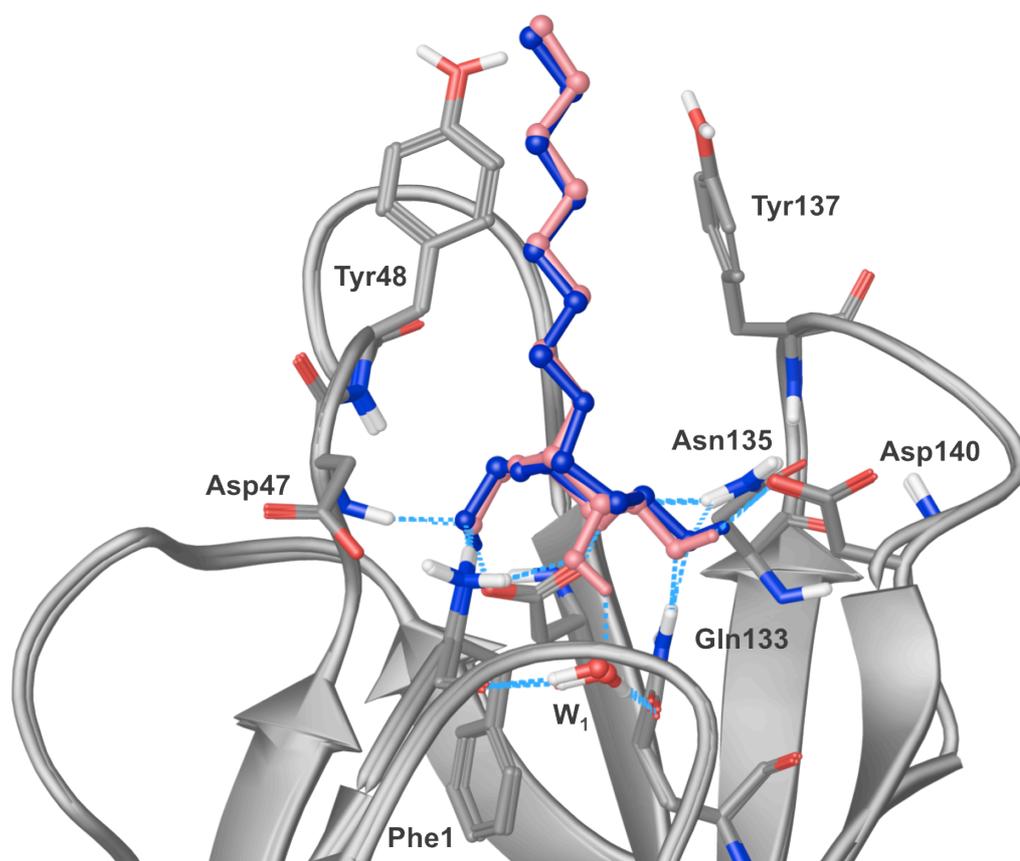
**Table S1. Calculated surface accessible solvation area.** SASA was calculated in a bound and unbound state of HM and its deoxy and deoxy-fluoro derivatives. Values are given in Å<sup>2</sup>.

	1 (HM)	2 (2-F)	5 (2-H)	6 (3-F)	7 (3-H)	8 (4-F)	9 (4-H)
Ligand donor_unbound	52	38	38	35	37	40	40
Ligand donor_bound	0	0	0	0	0	0	0
Ligand donor_delta	52	38	38	35	37	40	40
<b>Ligand donor_delta %</b>	<b>100%</b>						
Ligand acceptor_unbound	123	99	101	98	100	98	100
Ligand acceptor_bound	1	0	3	0	2	0	0
Ligand acceptor_delta	123	99	98	98	98	98	100
<b>Ligand acceptor_delta %</b>	<b>99%</b>	<b>100%</b>	<b>97%</b>	<b>100%</b>	<b>98%</b>	<b>100%</b>	<b>100%</b>
Ligand total_unbound	552	547	541	548	544	548	544
Ligand total_bound	137	136	135	135	173	134	111
Ligand total_delta	415	410	405	412	372	414	432
<b>Ligand total_delta %</b>	<b>75%</b>	<b>75%</b>	<b>75%</b>	<b>75%</b>	<b>68%</b>	<b>76%</b>	<b>80%</b>
Receptor total_unbound	532	531	542	506	438	531	598
Receptor total_bound	337	335	352	321	255	331	405
Receptor total_delta	195	197	190	185	183	201	192
<b>Receptor total_delta %</b>	<b>37%</b>	<b>37%</b>	<b>35%</b>	<b>37%</b>	<b>42%</b>	<b>38%</b>	<b>32%</b>

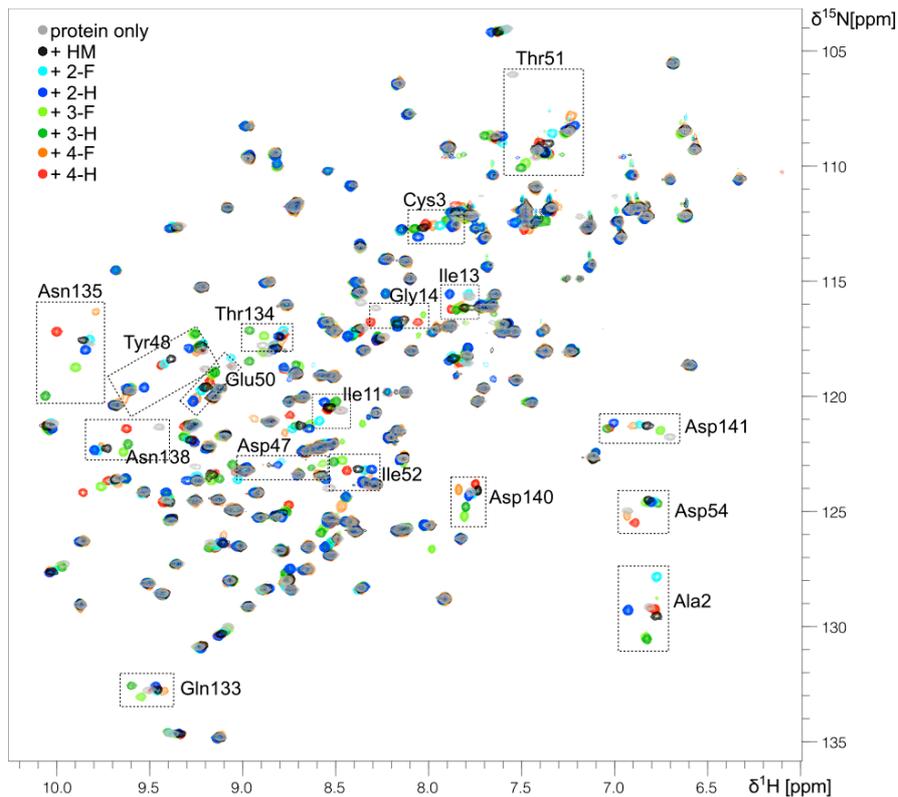
Table S2. Statistics on diffraction data and refinement of FimH<sub>LD</sub> and its ligand complexes.

PDB Identifier	FimH <sub>LD</sub> 5 (2-H)		FimH <sub>LD</sub> 7 (3-H)		FimH <sub>LD</sub> 9 (4-H)		FimH <sub>LD</sub> 2 (2-F)		FimH <sub>LD</sub> 6 (3-F)		FimH <sub>LD</sub> 8 (4-F)	
	5L4T	5L4V	5L4T	5L4V	5L4X	5L4U	5L4U	5L4W	5L4Y			
Wavelength (Å)	1.00001	1.00004	1.00003	1.00001	1.00001	1.00001	1.00001	1.00001	1.00001	1.00001	1.00001	1.00001
Resolution range (Å)	55.3 - 1.90 (2.01 - 1.90)* P 21 2 21	41.8 - 3.0 (3.17 - 3.0)* P 1 21 1	32.2 - 1.90 (1.98 - 1.90)* P 21 21 21	32.2 - 1.90 (1.98 - 1.90)* P 21 21 21	55.65 - 2.10 (2.33 - 2.1)* P 21 21 21	61.08 61.38 95.63 90 90 90	63.55 68.30 96.1 90 90 90	44.83 95.34 70.80 90 105.6 90	44.83 95.34 70.80 90 105.6 90	68.1 - 1.90 (2.01 - 1.90)* P 1 21 1	68.1 - 1.90 (2.01 - 1.90)* P 1 21 1	55.9 - 1.90 (2.01 - 1.90)* P 21 21 21
Space group	67.76 68.57 96.11 90 90 90	44.83 95.34 70.80 90 105.0 90	61.08 61.38 95.63 90 90 90	61.08 61.38 95.63 90 90 90	63.55 68.30 96.1 90 90 90	44.83 95.34 70.80 90 105.6 90	44.83 95.34 70.80 90 105.6 90	44.83 95.34 70.80 90 105.6 90	44.83 95.34 70.80 90 105.6 90	63.23 68.66 96.17 90 90 90	63.23 68.66 96.17 90 90 90	63.23 68.66 96.17 90 90 90
Total reflections	195676 (30395)	51948 (7481)	214352 (12961)	214352 (12961)	159603 (26006)	144207 (16535)	144207 (16535)	144207 (16535)	144207 (16535)	200752 (23165)	200752 (23165)	200752 (23165)
Unique reflections	35539 (5439)	11564 (1748)	35551 (2143)	35551 (2143)	24926 (4026)	40710 (4983)	40710 (4983)	40710 (4983)	40710 (4983)	33190 (4876)	33190 (4876)	33190 (4876)
Multiplicity	5.5 (5.5)	4.4 (4.2)	6.0 (6.0)	6.0 (6.0)	6.4 (6.4)	3.5 (3.3)	3.5 (3.3)	3.5 (3.3)	3.5 (3.3)	6.0 (4.7)	6.0 (4.7)	6.0 (4.7)
Completeness (%)	98.5 (98.7)	98.6 (93.9)	99.9 (99.9)	99.9 (99.9)	99.4 (99.0)	90.0 (71.2)	90.0 (71.2)	90.0 (71.2)	90.0 (71.2)	98.5 (94.7)	98.5 (94.7)	98.5 (94.7)
Mean I/sigma(I)	7.0 (1.5)	4.5 (1.3)	8.8 (2.3)	8.8 (2.3)	8.9 (2.2)	13.8 (2.2)	13.8 (2.2)	13.8 (2.2)	13.8 (2.2)	9.2 (2.0)	9.2 (2.0)	9.2 (2.0)
Wilson B-factor	25.3	31.0	16.6	16.6	29.1	26.6	26.6	26.6	26.6	23.3	23.3	23.3
R-meas	0.175 (1.445)	0.301 (1.13)	0.062 (0.72)	0.062 (0.72)	0.075 (0.614)	0.101 (0.757)	0.101 (0.757)	0.101 (0.757)	0.101 (0.757)	0.172 (0.83)	0.172 (0.83)	0.172 (0.83)
CC1/2	0.995 (0.707)	0.860 (0.520)	0.997 (0.925)	0.997 (0.925)	0.998 (0.876)	0.997 (0.699)	0.997 (0.699)	0.997 (0.699)	0.997 (0.699)	0.997 (0.841)	0.997 (0.841)	0.997 (0.841)
R-work	0.205 (0.32)	0.248 (0.337)	0.178 (0.295)	0.178 (0.295)	0.205 (0.224)	0.174 (0.268)	0.174 (0.268)	0.174 (0.268)	0.174 (0.268)	0.194 (0.228)	0.194 (0.228)	0.194 (0.228)
R-free	0.221 (0.361)	0.276 (0.361)	0.205 (0.306)	0.205 (0.306)	0.236 (0.264)	0.202 (0.287)	0.202 (0.287)	0.202 (0.287)	0.202 (0.287)	0.229 (0.284)	0.229 (0.284)	0.229 (0.284)
RMS(bonds)	0.004	0.011	0.006	0.006	0.009	0.010	0.010	0.010	0.010	0.009	0.009	0.009
RMS(angles)	0.96	1.6	1.09	1.09	1.11	1.13	1.13	1.13	1.13	1.11	1.11	1.11
Ramachandran favored (%)	97.2	98.1	97.4	97.4	98.7	98.1	98.1	98.1	98.1	97.2	97.2	97.2
Ramachandran outliers (%)	0	0	0	0	0	0	0	0	0	0	0	0
Clashscore	1.2	1.5	2.7	2.7	1.3	1.0	1.0	1.0	1.0	0.4	0.4	0.4

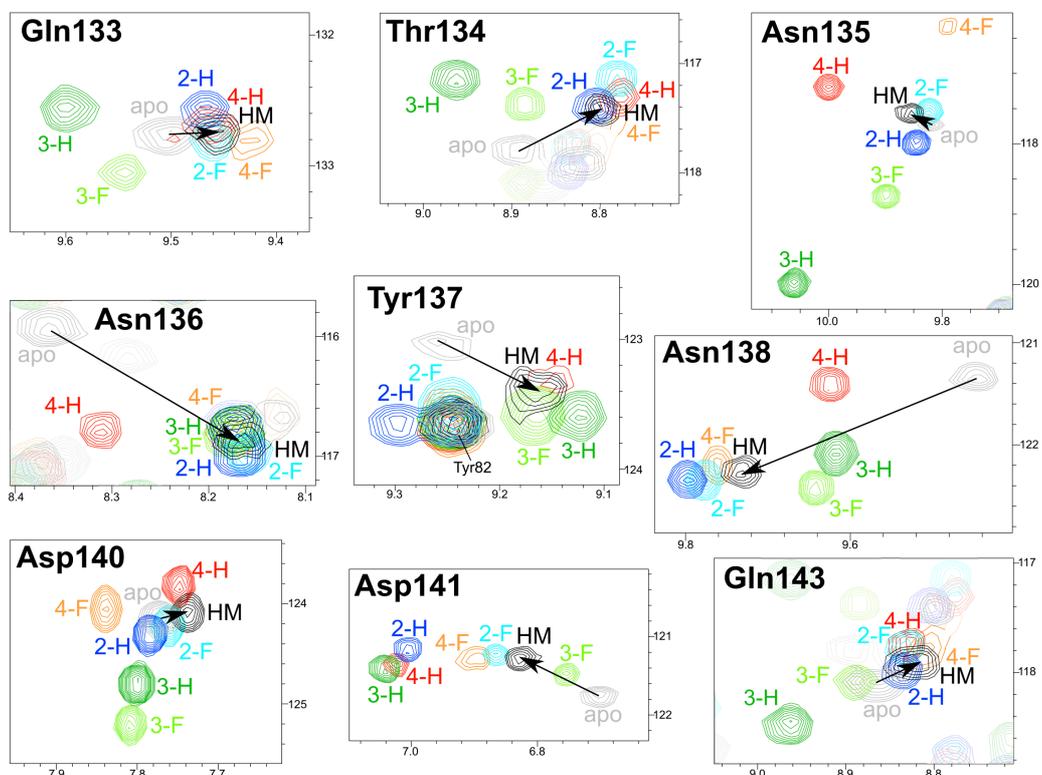
\*The values in parentheses correspond to the highest resolution shell



**Figure S1. Superimposition of X-ray co-crystal structures of FimH<sub>LD</sub> with HM and 2-H.** Both proteins are shown in grey, the ligands are shown in pink **1** (HM) and blue **5** (2-H). The missing contact of 2-OH to the structural water molecule (W<sub>1</sub>) and the positively charged N-terminus of **5** slightly tilts its mannose moiety, which potentially affects the interactions of the remaining hydroxyl groups to the protein. However, the structural changes are within the error of the measurement.



**Figure S2.** Overlap of  $^1\text{H}$ ,  $^{15}\text{N}$ -HSQC spectra of FimH in absence of ligand (grey) and in presence of 1 (HM, black), 2 (2-F, cyan), 5 (2-H, blue), 6 (3-F, light green), 7 (3-H, green), 8 (4-F, orange) and 9 (4-H, red). The spectrum in presence of 10 (6-F) is not shown due to the absence of any chemical shift changes.

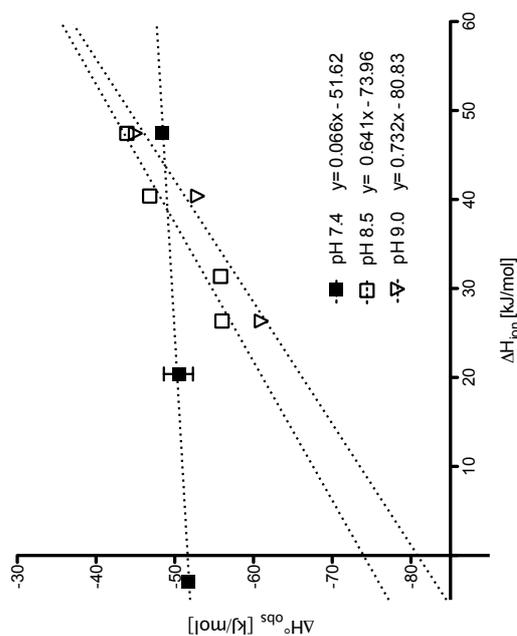


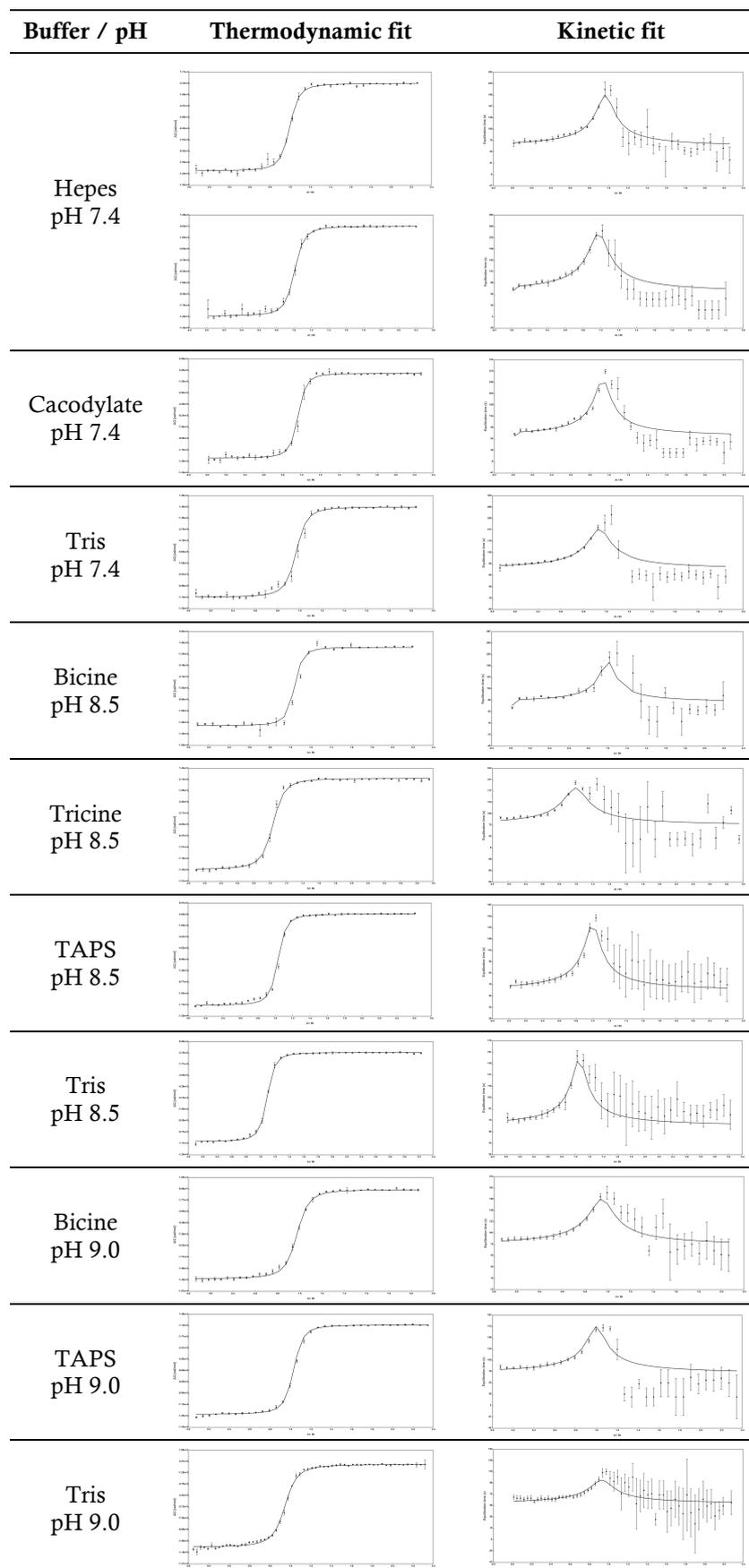
**Figure S3.**  $^1\text{H}$ ,  $^{15}\text{N}$ -HSQC spectral regions of the backbone signals of Gln133 to Gln143.

**Table S3. Thermodynamic parameters of compound 1 (HM) binding to FimH with varying buffers and pH.** Measurements were carried out at 298.15 K in 10 mM buffer containing 150 mM NaCl. The confidence interval is given in parentheses.

Buffer	pH	Ligand [ $\mu\text{M}$ ]	Protein [ $\mu\text{M}$ ]	$K_D$ [nM]	$\Delta G^{\circ}_{\text{obs}}$ [kJ/mol]	$\Delta H^{\circ}_{\text{obs}}$ [kJ/mol]	$-T\Delta S^{\circ}$ [kJ/molK]	$N$	$c$ -value
HEPES	7.4	100	10.0 / 8.6	21.9 (16.8 - 27.3)	-43.7	-50.5 (-48.6 - -52.3)	6.7	0.97/0.98	518/351
Cacodylate	7.4	100	9.1	15.4 (13.4 - 17.4)	-44.6	-51.7 (-51.5 - -52.0)	7.2	0.93	591
Tris	7.4	100	10.1	22.2 (19.3 - 25.1)	-43.7	-48.4 (-48.2 - -48.7)	4.8	0.94	455
Bicine	8.5	150	14.0	18.6 (15.6 - 21.6)	-44.1	-56.0 (-55.9 - -56.2)	11.9	1.00	754
Tricine	8.5	150	10.5	46.8 (43.4 - 50.2)	-41.8	-55.8 (-55.6 - -56.0)	13.9	1.00	224
TAPS	8.5	150	13.0	25.7 (24.0 - 27.5)	-43.3	-46.8 (-46.7 - -46.8)	3.4	1.00	506
Tris	8.5	150	10.5	24.8 (22.2 - 27.4)	-43.4	-43.9 (-43.7 - -44.0)	0.5	1.05	423
Bicine	9.0	100	10.0	26.7 (24.8 - 28.6)	-43.2	-61.0 (-60.8 - -61.2)	17.8	0.95	375
TAPS	9.0	150	13.2	24.9 (23.6 - 26.2)	-43.4	-52.9 (-52.8 - -53.0)	9.5	1.00	530
Tris	9.0	200	20.0	76.3 (73.2 - 79.4)	-40.6	-45.0 (-44.8 - -45.1)	4.3	0.93	262

**Figure S4. pH-dependent protonation of the FimH N-terminus.** Heat of ionization ( $\Delta H^{\circ}_{\text{ion}}$ ) is buffer dependent (Cacodylate, -3 kJ/mol; HEPES, +20.4 kJ/mol; Bicine, +26.34 kJ/mol; Tricine +31.37 kJ/mol; TAPS +40.40 kJ/mol; Tris, +47.45 kJ/mol). The enthalpy ( $\Delta H^{\circ}_{\text{obs}}$ ) of HM binding to FimH was measured by ITC in different buffers at pH 7.4, pH 8.5 and pH 9.0. The proton uptake was calculated according to the equation  $\Delta H^{\circ}_{\text{obs}} = \Delta H^{\circ}_{\text{int}} + nH^+ \cdot \Delta H^{\circ}_{\text{ion}}$ , where the steepness of the slope reflects the number of protons taken up or released by this binding event ( $nH^+$ ) and  $\Delta H^{\circ}_{\text{int}}$  is the intrinsic enthalpy of binding. 0.07 protons are taken up from the solvent at pH 7.4, 0.64 at pH 8.5 and 0.73 at pH 9.0. According to the Henderson-Hasselbalch equation this corresponds to a  $pK_a$  between 8.25 and 8.55 for the unbound FimH protein. The N-terminus is the only residue in the FimH binding pocket that is able to take up a proton in this range of pH.



**Table S4.** Thermodynamic and kinetic fits of *n*-heptyl  $\alpha$ -D-mannopyranoside mannoside (1) binding to FimH<sub>LD</sub> in varying buffers and pH. Measurements were analyzed using AFFINImeter software.

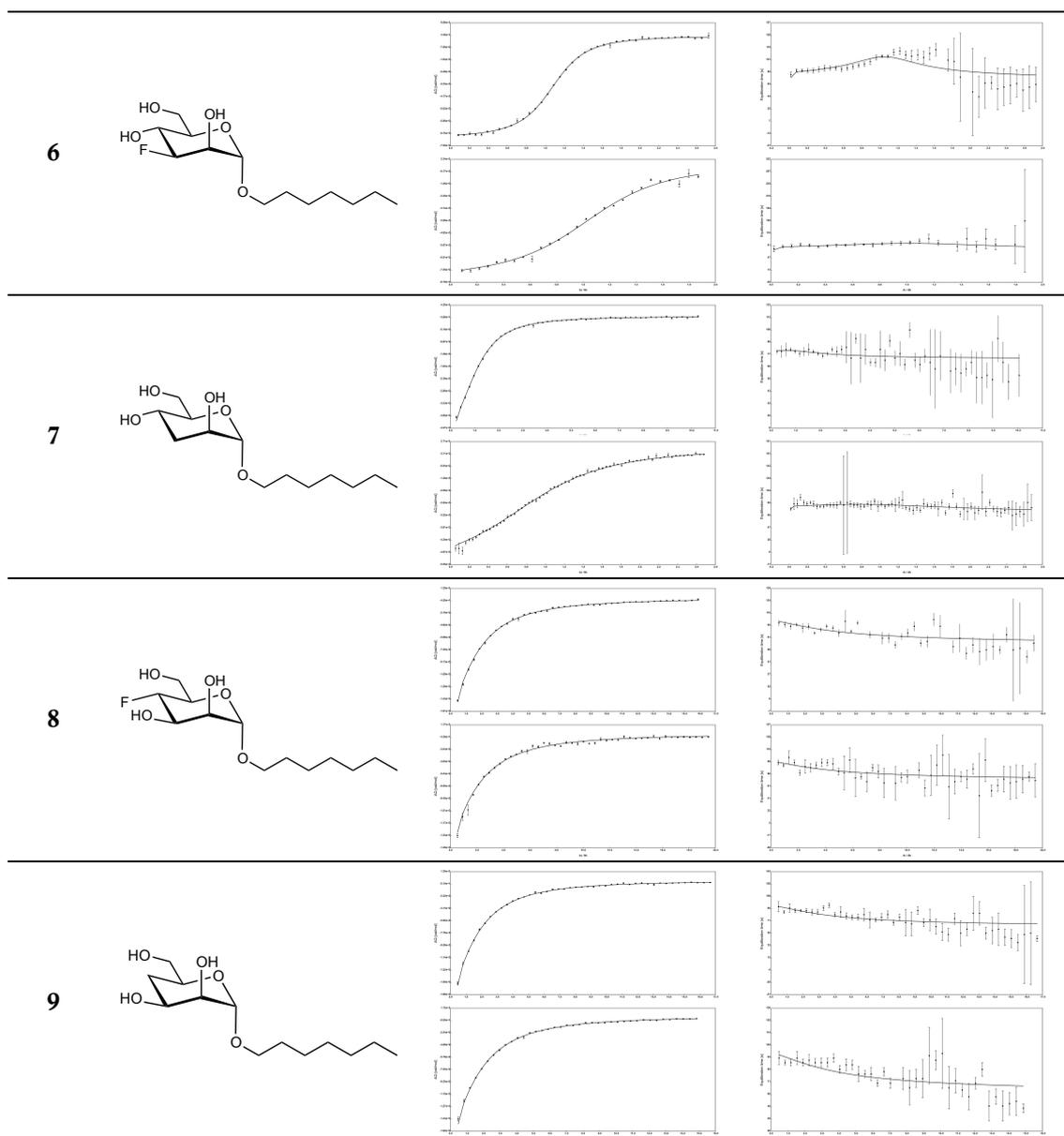
**Table S5. Thermodynamic parameters of HM and deoxy and deoxy-halogeno derivatives thereof binding to FimH.** Measurements were carried out at 298.15 K in 10mM HEPES buffer adjusted to pH 7.4 containing 150 mM NaCl. Two independent measurements of each compound were performed and analyzed using AFFINImeter software. The confidence interval of each measurement is given in parentheses. An asterisk indicates a fixed stoichiometry ( $N=1$ ).

Ligand	Ligand [ $\mu$ M]	Protein [ $\mu$ M]	$K_D$ [ $\mu$ M]	$\Delta G^\circ$ [kJ/mol]	$\Delta H^\circ$ [kJ/mol]	$-T\Delta S^\circ$ [kJ/molK]	$N$	$c$ -value
1 (HM_1)	100	10.0	0.019 (0.017 - 0.022)	-44.0	-48.8 (-48.6 - -49.0)	4.8	0.97	518
1 (HM_2)	100	8.6	0.025 (0.022 - 0.027)	-43.4	-52.1 (-51.9 - -52.3)	8.7	0.98	351
1 (HM)			<b>0.022</b>	<b>-43.7</b>	<b>-50.5</b>	<b>6.7</b>		
2 (2-F_1)	500	25.5	0.499 (0.486 - 0.512)	-36.0	-28.0 (-27.9 - -28.1)	-8.0	1.03	51
2 (2-F_2)	500	26.2	0.565 (0.545 - 0.585)	-35.7	-27.5 (-27.4 - -27.6)	-8.2	1.02	46
2 (2-F)			<b>0.532</b>	<b>-35.8</b>	<b>-27.7</b>	<b>-8.1</b>		
3 (2-Cl_1)	1000	40.7	1.678 (1.645 - 1.712)	-33.0	-24.9 (-24.8 - -25.0)	-8.1	1.00	24
3 (2-Cl_2)	800	40.2	0.995 (0.977 - 1.013)	-34.3	-24.1 (-24.1 - -24.2)	-10.1	1.00	40
3 (2-Cl)			<b>1.337</b>	<b>-33.5</b>	<b>-24.5</b>	<b>-9.0</b>		
4 (2-Br_1)	1000	50.0	2.144 (2.106 - 2.182)	-32.4	-21.1 (-21.1 - -21.2)	-11.2	1.00	23
4 (2-Br_2)	800	39.4	1.517 (1.462 - 1.571)	-33.2	-20.5 (-20.4 - -20.6)	-12.7	1.00	26
4 (2-Br)			<b>1.830</b>	<b>-32.7</b>	<b>-20.8</b>	<b>-11.9</b>		
5 (2-H_1)	1000	50.4	7.709 (7.628 - 7.790)	-29.2	-18.4 (-18.4 - -18.5)	-10.7	1.00	6.5
5 (2-H_2)	1000	50.1	11.833 (11.717 - 11.949)	-28.1	-20.3 (-20.2 - -20.3)	-7.9	1.00	4.2
5 (2-H)			<b>9.771</b>	<b>-28.6</b>	<b>-19.3</b>	<b>-9.2</b>		
6 (3-F_1)	450	30.6	0.938 (0.923 - 0.954)	-34.4	-29.3 (-29.2 - -29.4)	-5.1	1.08	33
6 (3-F_2)	250	16.4	1.449 (1.405 - 1.493)	-33.3	-33.4 (-33.2 - -33.6)	0.1	1.10	11
6 (3-F)			<b>1.194</b>	<b>-33.8</b>	<b>-31.4</b>	<b>-2.4</b>		
7 (3-H_1)	2000	47.0	20.901 (20.749 - 21.053)	-26.7	-25.2 (-25.1 - -25.2)	-1.5	1.00	2.2
7 (3-H_2)	1000	77.1	18.855 (18.674 - 19.036)	-27.0	-24.4 (-24.3 - -24.4)	-2.6	1.00	4.1
7 (3-H)			<b>19.878</b>	<b>-26.8</b>	<b>-24.8</b>	<b>-2.1</b>		
8 (4-F_1)	4000	47.9	91.940 (90.726 - 93.154)	-23.0	-19.3 (-19.2 - -19.4)	-3.7	1.00	0.5
8 (4-F_2)	4000	42.1	114.91 (113.45 - 116.37)	-22.5	-21.7 (-21.6 - -21.8)	-0.8	1.00	0.4
8 (4-F)			<b>103.425</b>	<b>-22.7</b>	<b>-20.5</b>	<b>-2.2</b>		
9 (4-H_1)	4000	48.0	85.698 (85.124 - 86.272)	-23.2	-19.6 (-19.5 - -19.6)	-3.7	1.00	0.6
9 (4-H_2)	4000	47.3	91.106 (90.446 - 91.766)	-23.1	-20.1 (-20.1 - -20.2)	-3.0	1.00	0.5
9 (4-H)			<b>88.402</b>	<b>-23.1</b>	<b>-19.8</b>	<b>-3.3</b>		

**Table S6:** Structure, and thermodynamic and kinetic fits of *n*-heptyl  $\alpha$ -D-mannopyranoside (1) and its deoxy- and deoxy-halogeno derivatives (2-9) binding to FimH<sub>LD</sub>. All measurements were performed at 25°C in HEPES buffer adjusted to pH 7.4, containing 150 mM NaCl. Two independent experiments were carried out for each ligand and analyzed using AFFINImeter software.

Cpd	Structure	Thermodynamic fit	Kinetic fit
1			
2			
3			
4			
5			

## Section II. The Bacterial Adhesin FimH – Manuscript 1



**Table S7. Kinetic parameters of HM and deoxy- and deoxy-halogeno derivatives thereof binding to FimH.** The kinetic analysis was performed using the kinITC tool implemented in the ITC data analysis software AFFINImeter. An asterisk marks the measurements that were excluded from further analysis due to poor quality of the kinetic fit. Confidence intervals are given in parentheses.

Compound	$K_D$ [ $\mu\text{M}$ ] AFFINImeter	$k_{\text{on}}$ [ $\text{M}^{-1}\text{s}^{-1}$ ]	$k_{\text{off}}$ [ $\text{s}^{-1}$ ]	$t_{1/2}$ [min]	Response time [s]
1 (HM_1)	0.019 (0.017 – 0.021)	$3.87 \cdot 10^4$	$7.47 \cdot 10^{-4}$ ( $6.10 \cdot 10^{-5}$ )	15.46	12.89
1 (HM_2)	0.026 (0.024 – 0.028)	$2.88 \cdot 10^4$	$7.06 \cdot 10^{-4}$ ( $5.73 \cdot 10^{-5}$ )	16.37	11.97
1 (HM)	<b>0.022</b>	<b><math>3.32 \cdot 10^4</math></b>	<b><math>7.27 \cdot 10^{-4}</math></b>	<b>15.90</b>	
2 (2-F_1)	0.50 (0.49 – 0.51)	$2.04 \cdot 10^4$	$1.02 \cdot 10^{-2}$ ( $1.22 \cdot 10^{-3}$ )	1.14	13.61
2 (2-F_2)*	0.56 (0.54 – 0.59)	$9.43 \cdot 10^4$	$5.33 \cdot 10^{-2}$ ( $3.13 \cdot 10^{-2}$ )	0.22	15.03
2 (2-F)	<b>0.50</b>	<b><math>2.04 \cdot 10^4</math></b>	<b><math>1.02 \cdot 10^{-2}</math></b>	<b>1.14</b>	
3 (2-Cl_1)	1.68 (1.64 – 1.71)	$6.84 \cdot 10^3$	$1.15 \cdot 10^{-2}$ ( $1.50 \cdot 10^{-3}$ )	1.01	11.77
3 (2-Cl_2)	0.99 (0.98 – 1.01)	$1.24 \cdot 10^4$	$1.23 \cdot 10^{-2}$ ( $1.42 \cdot 10^{-3}$ )	0.94	12.61
3 (2-Cl)	<b>1.34</b>	<b><math>8.90 \cdot 10^3</math></b>	<b><math>1.19 \cdot 10^{-2}</math></b>	<b>0.97</b>	
4 (2-Br_1)	2.14 (2.11 – 2.18)	$1.02 \cdot 10^4$	$2.19 \cdot 10^{-2}$ ( $3.49 \cdot 10^{-3}$ )	0.53	13.65
4 (2-Br_2)	1.52 (1.46 – 1.57)	$4.62 \cdot 10^3$	$7.00 \cdot 10^{-3}$ ( $5.01 \cdot 10^{-4}$ )	1.65	8.87
4 (2-Br)	<b>1.83</b>	<b><math>7.90 \cdot 10^3</math></b>	<b><math>1.45 \cdot 10^{-2}</math></b>	<b>0.80</b>	
5 (2-H_1)	7.71 (7.63 – 7.79)	$5.35 \cdot 10^3$	$4.13 \cdot 10^{-2}$ ( $1.17 \cdot 10^{-2}$ )	0.28	13.15
5 (2-H_2)*	11.83 (11.72 – 11.95)	$4.21 \cdot 10^4$	$4.98 \cdot 10^{-1}$ ( $8.21 \cdot 10^{-1}$ )	0.02	14.58
5 (2-H)	<b>7.71</b>	<b><math>5.35 \cdot 10^3</math></b>	<b><math>4.13 \cdot 10^{-2}</math></b>	<b>0.28</b>	
6 (3-F_1)	0.94 (0.92 – 0.95)	$1.14 \cdot 10^4$	$1.07 \cdot 10^{-2}$ ( $8.04 \cdot 10^{-4}$ )	1.08	12.82
6 (3-F_2)	1.45 (1.16 – 1.22)	$2.00 \cdot 10^4$	$2.90 \cdot 10^{-2}$ ( $8.16 \cdot 10^{-3}$ )	0.40	11.89
6 (3-F)	<b>1.19</b>	<b><math>1.66 \cdot 10^4</math></b>	<b><math>1.98 \cdot 10^{-2}</math></b>	<b>0.58</b>	
7 (3-H_1)	20.90 (20.75 – 21.05)	$8.85 \cdot 10^3$	$1.85 \cdot 10^{-1}$ ( $5.66 \cdot 10^{-2}$ )	0.06	13.57
7 (3-H_2)	18.86 (18.67 – 19.04)	$3.90 \cdot 10^3$	$7.36 \cdot 10^{-2}$ ( $1.86 \cdot 10^{-2}$ )	0.16	13.35
7 (3-H)	<b>19.88</b>	<b><math>6.50 \cdot 10^3</math></b>	<b><math>1.29 \cdot 10^{-1}</math></b>	<b>0.09</b>	
8 (4-F_1)	91.94 (90.73 – 93.15)	$1.12 \cdot 10^3$	$1.03 \cdot 10^{-1}$ ( $1.17 \cdot 10^{-2}$ )	0.11	13.49
8 (4-F_2)	114.91 (113.45 – 116.37)	$9.04 \cdot 10^2$	$1.04 \cdot 10^{-1}$ ( $1.48 \cdot 10^{-2}$ )	0.11	12.68
8 (4-F)	<b>103.43</b>	<b><math>1.00 \cdot 10^3</math></b>	<b><math>1.03 \cdot 10^{-1}</math></b>	<b>0.11</b>	
9 (4-H_1)	85.70 (85.12 – 86.27)	$1.01 \cdot 10^3$	$8.65 \cdot 10^{-2}$ ( $8.33 \cdot 10^{-3}$ )	0.13	10.85
9 (4-H_2)	91.11 (90.45 – 91.77)	$9.05 \cdot 10^2$	$8.25 \cdot 10^{-2}$ ( $7.34 \cdot 10^{-3}$ )	0.14	10.88
9 (4-H)	<b>88.40</b>	<b><math>9.56 \cdot 10^2</math></b>	<b><math>8.45 \cdot 10^{-2}</math></b>	<b>0.14</b>	

### Experimental Part

**Protein preparation.** FimH<sub>LD</sub> from *E.coli* K-12 strain was expressed with a C-terminal thrombin cleavage site and a 6His-tag (FimH<sub>LD</sub>-Th-6His, 173 residues) following a previously published protocol.<sup>(1)</sup> Briefly, the clone containing the FimH<sub>LD</sub> construct was expressed in the protease-deficient *E.coli* HM 125 strain at 30°C and 180 rpm in M9 minimal medium supplemented with 100 µg/mL ampicillin. The protein expression was induced by 1 mM IPTG at an OD<sub>600</sub> of 0.8. The cells were further cultivated for 16 hrs, harvested by centrifugation for 20 min at 5'000 rpm and 4°C. The pellet was resuspended in lysis buffer containing 50 mM Tris pH 7.4, 150 mM NaCl, 5 mM EDTA and 1 mg/mL polymyxin B sulfate. The supernatant containing the periplasmic extract was dialyzed against sodium phosphate buffer and purified on Ni-NTA columns. The protein was finally dialyzed against assay buffer containing 20 mM HEPES pH 7.4, 150 mM NaCl and 1 mM CaCl<sub>2</sub>. For long time storage the protein was frozen at -80°C. For the production of uniformly <sup>15</sup>N-labeled FimH<sub>LD</sub>-Th-6His for NMR experiments, *E.coli* HM125 was cultivated in M9 minimal medium containing 1 g/L <sup>15</sup>NH<sub>4</sub>Cl (CortecNet, France) as the sole source of nitrogen. The labeled protein was purified as described above and dialyzed against 20 mM phosphate buffer pH 7. The exact molecular weight (18860.2 Da) was determined by mass spectrometry.

**Fluorescence polarization assay.** Competitive FP assays were essentially performed as described previously [1]. Briefly, a serial dilution of unlabeled competitor was titrated into constant concentrations of **GNFP4** (final concentration 5 nM) and FimH<sub>LD</sub> (final concentration 15 nM) in the presence of 5% DMSO. All solutions were prepared in a buffer containing 20 mM HEPES (pH 7.4), 150 mM NaCl and 50 µg/ml BSA. Mixtures were incubated for 24 hours in a final volume of 100 µl per well in black, flat bottom, NBS-treated, 96-well microtiter plates (Corning, U.S.A.). Competitor *K<sub>D</sub>*s were determined via the displacement of **GNFP4**. The associated decrease in fluorescence polarization was recorded at 528 nm (excitation at 485 nm) through appropriately oriented polarizing filters. Resulting binding isotherms were fit to an equilibrium competition function<sup>(2)</sup> and analyzed using Prism (GraphPad Software, U.S.A.). The *K<sub>D</sub>* of **GNFP4** and FimH was defined as constraint and set to 1.7 nM during curve fitting.<sup>(3)</sup>

**Co-crystallization, data processing, and structure refinement.** For crystallization, FimH<sub>LD</sub> (residues 1-158) at a final concentration of 12 mg/mL (ca. 0.8 mM) with a threefold molar excess of ligand (2.5 mM) in 20 mM HEPES buffer pH 7.4. Crystals were grown in sitting-drop vapor diffusion at 4°C, 12°C and 20°C in 0.2 M (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 0.1 M HEPES pH 7 and 25-30% PEG3350. Plate-like crystals appeared after 2 weeks, were cryopreserved by addition of 20% glycerol (v/v) and flash-cooled with liquid nitrogen. Data were collected at the SLS beamlines X06DA and X06SA of the Swiss Light Source (Paul Scherrer Institute, Switzerland) and indexed, integrated and scaled with XDS.<sup>(4,5)</sup> Structures were solved by molecular replacement with PHASER<sup>(6)</sup> using the FimH<sub>LD</sub>-*n*-heptyl  $\alpha$ -D-mannopyranoside complex (PDB code 4XO8) as search model. The structures were built using the COOT software<sup>(7)</sup> and periodically refined with the PHENIX and Buster-TNT software.<sup>(8,9)</sup> Geometric restraints for the ligands were generated with PRODRG<sup>(10)</sup> and Molprobity<sup>(11)</sup> was used for validation. The atomic coordinates have been deposited in the RCSB Protein Data Bank and are available under the accession code 5L4T, 5L4U, 5L4V, 5L4W, 5L4X, and 5L4Y, respectively.

**RMSD calculation.** RMSD was calculated from the FimH<sub>LD</sub> apo structure (PDB-Code: 4AUU) and the co-crystal structure in complex with **1** (HM) (PDB-Code: 4XO8). From both structures, the heavy atoms involved in ligand binding (Phe1: N from N-terminus; Asp47: N from backbone; Asp54: 2xO from side chain; Gln133: N from side chain; Asn135: N from side chain; Asp140: O from side chain) were compared using Schrödinger Suite 2014-4.

**Ab initio calculations.** The lowest energy conformer of *n*-heptyl  $\alpha$ -D-mannoside from a conformational analysis with MacroModel 10.6<sup>(12)</sup> using the OPLS 2005 force field was subjected to geometrical optimization and energy calculation using the density functional theory (DFT) with the B3LYP functional and the 6-31G(d,p) basis-set in the gas-phase as implemented in Gaussian 09.<sup>(13)</sup> The torsional potential was determined using a relaxed energy potential surface scan with a 2° step size. Vibrational frequency calculations were carried out to confirm these minima. No imaginary frequencies were found.

**Molecular dynamics simulations.** Molecular dynamics simulations were carried out using Desmond<sup>(14-18)</sup> and the OPLS 2005 force field. Default parameters were applied unless stated otherwise. TIP3P was selected as water model and a physiological salt concentration (0.15 M) was added to the protein-complex simulation. An energy barrier of 5 kcal/mol restricted backbone movement. The MD simulation for the solution conformation was run for 9.6 ns, whereas the protein-complex simulation was run for 4.8 ns. The energies of the per-residue interactions were calculated from 1000 extracted MD frames using Prime.<sup>(19-21)</sup>

**NMR experiments.** <sup>1</sup>H, <sup>15</sup>N-HSQC NMR experiments were measured at 298 K on a Bruker Avance III 600 MHz NMR spectrometer equipped with a 5 mm TXI room temperature probe head. Samples contained 120 of <sup>15</sup>N-labeled FimH<sub>LD</sub> in 20 mM phosphate buffer pH 7.4 in water with 7% D<sub>2</sub>O. Deoxy- and deoxy-fluoro derivatives were solved in H<sub>2</sub>O at 10 to 20 mM concentrations and added stepwise up to 2- to 5-fold molar excess. In case of **10** (6-F), the required amount of ligand was added directly as a lyophilized powder to the protein. NMR spectra were acquired and processed with Topspin 3.2 (Bruker BioSpin, Switzerland) and analyzed with CcpNmr Analysis (version 2.2).<sup>(22)</sup> The backbone assignment of FimH<sub>LD</sub> was available from previous studies.<sup>(23)</sup> Combined chemical shift differences,  $\Delta\delta_{AV}$ , between free and ligand-bound protein signals were calculated as in equation 1<sup>(24)</sup>

$$\Delta\delta_{AV} = \sqrt{(\Delta\delta^1H^N)^2 + (0.2\Delta\delta^{15})^2} \quad (\text{eq. 1})$$

**Isothermal titration calorimetry.** ITC experiments were performed at 25°C using a VP-ITC (Malvern Instruments, Worcestershire, UK) with an injection volume between 4 and 10  $\mu$ l, a reference power of 10  $\mu$ cal/sec, a stirring speed of 307 rpm, high feedback, a spacing time between 360 and 600 seconds and a filter period of 2 seconds. Preceding the measurements, FimH<sub>LD</sub>-Th-His6 was dialyzed against 10 mM of the experimental buffer containing 150 mM NaCl. Two independent experiments with **1** (HM) and deoxy- and deoxy-halogeno derivatives thereof were performed in a HEPES buffer at pH 7.4. Protonation experiments were carried out in different buffers at pH 7.4, pH 8.5 and pH 9.0 (Cacodylate pH 7.4; Tris-HCl 7.4, 8.5 and 9.0; Bicine pH 8.5 and 9.0; Tricine pH 8.5; TAPS pH 8.5 and 9.0). Ligand and protein were dissolved in the same buffer. Protein concentration was determined by NanoDrop ND-1000 Spectrophotometer

(Thermo Scientific, MA, USA) using an extinction coefficient of  $24,180 \text{ M}^{-1} \text{ cm}^{-1}$ .<sup>(25)</sup> The ligand and protein concentrations used for the titrations are given in the Supporting Information (Table S3). Baseline adjustment and peak integration to determine the thermodynamic parameters  $K_A$  (association constant) and  $\Delta H^\circ$  (change in enthalpy), the kinetic parameters  $k_{\text{on}}$  (association rate constant) and  $k_{\text{off}}$  (dissociation rate constant) as well as  $N$  (stoichiometry), were performed using the fully automated analysis software AFFINImeter.<sup>(26,27)</sup> For the evaluation of the weak binding ligands **7** (3-H), **8** (4-F) and **9** (4-H) it was necessary to fix the stoichiometry ( $N=1$ ) for fitting.  $\Delta G^\circ$  (free energy of binding) and  $\Delta S^\circ$  (change in entropy) were calculated from equation 2

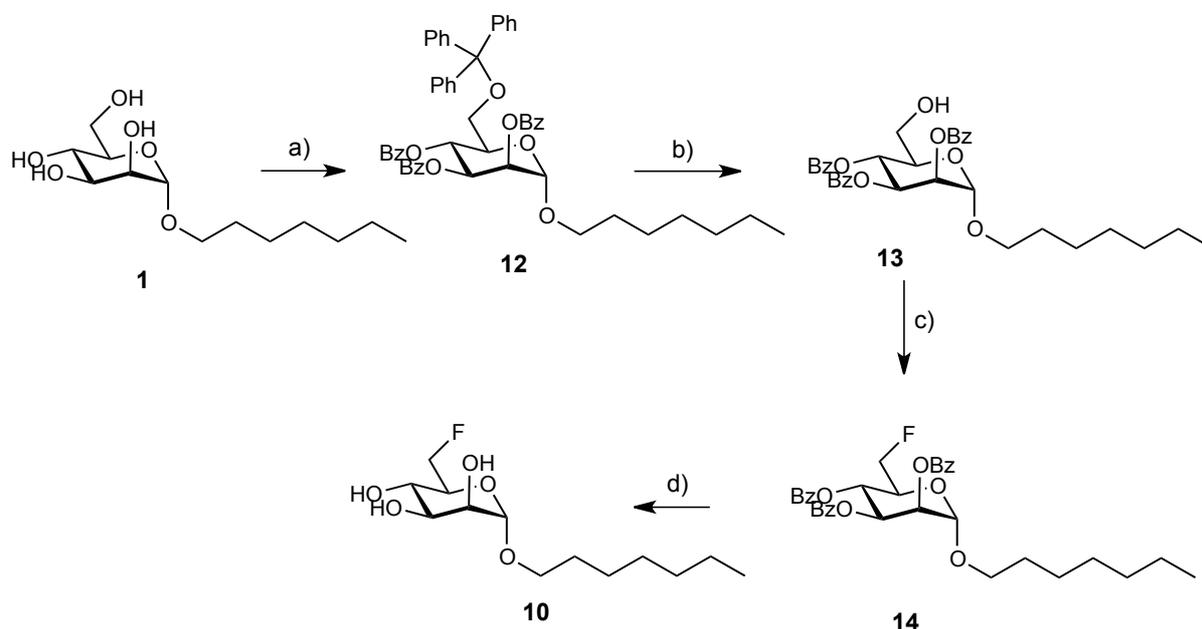
$$\Delta G^\circ = \Delta H^\circ - T\Delta S^\circ = -RT \ln K_A \quad (\text{eq. 2})$$

with  $T$  being the absolute temperature and  $R$  the universal gas constant ( $8.314 \text{ J/molK}$ ).

## Synthesis

**General methods.** Commercially available reagents were purchased from Aldrich, Merck or Alfa Aesar. Methanol was dried by distillation from sodium methoxide. Dichloromethane ( $\text{CH}_2\text{Cl}_2$ ) was dried by filtration through  $\text{Al}_2\text{O}_3$  (Fluka, basic; 0.05-0.15 mm). Toluene was dried by distillation from sodium/benzophenone. Optical rotations were measured at 20°C on a Perkin Elmer 341 polarimeter with a path length of 1 dm. Concentrations are given in g/100 mL. NMR spectra were obtained on a Bruker Avance 500 UltraShield spectrometer at 500.13 MHz ( $^1\text{H}$ ) or 125.76 MHz ( $^{13}\text{C}$ ). Chemical shifts are given in ppm and were calibrated on residual solvent peaks or to tetramethyl silane as internal standard. Multiplicities are specified as s (singlet), d (doublet), dd (doublet of a doublet), t (triplet), q (quartet) or m (multiplet). Assignment of the  $^1\text{H}$  and  $^{13}\text{C}$  NMR spectra was achieved using 2D methods (COSY, HSQC). ESI mass spectra were recorded on a Waters micromass ZQ instrument. High-resolution mass spectra were obtained on an ESI Bruker Daltonics micrOTOF spectrometer equipped with a TOF hexapole detector. Reactions were monitored by TLC using glass plates coated with silica gel 60 F<sub>254</sub> and visualized by using UV light and/or by charring with a molybdate solution (a 0.02 M solution of ammonium cerium sulfate dihydrate and ammonium molybdate tetrahydrate in aqueous 10%  $\text{H}_2\text{SO}_4$ ) with heating to 140°C for 5 min. Column chromatography was performed on a CombiFlash Companion (ISCO, Inc.) using RediSep normal phase disposable flash columns (silica gel). Reversed phase chromatography was performed on LiChroprep®RP-18 (Merck, 40- 63  $\mu\text{m}$ ).

### Synthesis of 1 (6-F) (*n*-Heptyl 6-Fluoro-6-deoxy- $\alpha$ -D-mannopyranoside)



**Scheme S1.** a) i. Trityl chloride, pyridine, DMAP, 80°C, overnight, ii. benzoyl chloride, two steps in one pot, (98%); b)  $\text{FeCl}_3$ , DCM, rt, overnight, (87%); c) DAST, DCM, rt, overnight, (45%); d) 0.5 M  $\text{CH}_3\text{ONa}/\text{CH}_3\text{OH}$ , (80%).

#### *n*-Heptyl 2,3,5-tri-*O*-benzoyl-6-*O*-triphenylmethyl- $\alpha$ -D-mannopyranoside (12)<sup>(28)</sup>

To a solution of 1 (310 mg, 1.114 mmol) in pyridine (4.0 mL) was added  $\text{TrCl}$  (388 mg, 1.392 mmol) and catalytic amount of 4-dimethylaminopyridine (DMAP). The mixture

was stirred at 80°C overnight and then cooled down to 0°C. To the above reaction mixture was added a premixed solution of BzCl (0.52 mL) in pyridine (0.5 mL) dropwise. The reaction mixture was stirred at 50°C overnight and then poured into ice-cold water, extracted with EtOAc, the organic layer was washed with aqueous NaHCO<sub>3</sub>, water, brine and dried over Na<sub>2</sub>SO<sub>4</sub>. The solvent was removed in vacuo and the residue was purified by flash chromatography on silica gel (PE-EE, 10:1-6:1) to afford the desired compound as a white solid (642 mg, 98%). <sup>1</sup>H-NMR (500 MHz, CDCl<sub>3</sub>): δ 8.16 (d, *J* = 7.5 Hz, 2H), 7.83 (d, *J* = 8.0 Hz, 2H), 7.73 (d, *J* = 8.0 Hz, 2H), 7.62 (t, *J* = 7.5 Hz, 1H), 7.50-7.40 (m, 10H), 7.31 (t, *J* = 8.0 Hz, 2H), 7.26 (t, *J* = 7.5 Hz, 2H), 7.16-7.07 (m, 9H), 6.01 (t, *J* = 10.0 Hz, 1H, H-4), 5.79 (dd, *J* = 10.0, 3.0 Hz, 1H, H-3), 5.67 (m, 1H, H-2), 5.12 (s, 1H, H-1), 4.18 (m, 1H, H-5), 3.86 (dt, *J* = 9.5, 6.5 Hz, 1H, H-OCH<sub>2</sub>C<sub>6</sub>H<sub>13</sub>), 3.58 (dt, *J* = 9.5, 7.0 Hz, 1H, H-OCH<sub>2</sub>C<sub>6</sub>H<sub>13</sub>), 3.37 (dd, *J* = 10.5, 2.0 Hz, 1H, H-6a), 3.27 (dd, *J* = 10.5, 5.0 Hz, 1H, H-6b), 1.71 (m, 2H), 1.43-1.32 (m, 8H), 0.90 (t, *J* = 7.0 Hz, 3H). <sup>13</sup>C-NMR (125 MHz, CDCl<sub>3</sub>): δ 165.70, 165.54, 165.11 (3xCO), 143.75, 133.38, 133.01, 132.96, 129.94, 129.74, 129.65, 129.59, 129.36, 129.26, 129.62, 128.60, 128.23, 128.13, 127.66, 126.80, 97.48 (C-1), 86.54, 70.97 (C-2), 70.66 (C-3), 70.31 (C-5), 68.36 (OCH<sub>2</sub>C<sub>6</sub>H<sub>13</sub>), 67.08 (C-4), 62.31 (C-6), 31.82, 29.45, 29.14, 26.16, 22.64, 14.12 (OCH<sub>2</sub>C<sub>6</sub>H<sub>13</sub>). ESI-MS Calcd for [M+Na]<sup>+</sup>, 855.35, found 855.28.

#### ***n*-Heptyl 2,3,5-tri-O-benzoyl- $\alpha$ -D-mannopyranoside (13)<sup>(29)</sup>**

To a solution of **12** (583 mg, 0.7 mmol) in DCM (10 mL) was added FeCl<sub>3</sub> (hygroscopic solid) (284 mg, 1.75 mmol) at rt. The reaction mixture was stirred at rt overnight, diluted with DCM, washed with brine, dried over Na<sub>2</sub>SO<sub>4</sub>. The solvent was removed in vacuo and the residue was purified by flash chromatography on silica gel (PE-EE, 10:1-4:1) to afford the desired compound as a white solid (359 mg, 87%). [ $\alpha$ ]<sub>D</sub><sup>20</sup> = -126.79 (c = 0.60, DCM). <sup>1</sup>H-NMR (500 MHz, CDCl<sub>3</sub>): δ 8.10 (d, *J* = 7.5 Hz, 2H), 7.98 (d, *J* = 7.5 Hz, 2H), 7.82 (dd, *J* = 8.0, 1.0 Hz, 2H), 7.62 (t, *J* = 7.5 Hz, 1H), 7.55-7.48 (m, 3H), 7.45-7.38 (m, 3H), 7.26 (m, 2H), 5.99 (dd, *J* = 10.0, 3.5 Hz, 1H, H-3), 5.82 (t, *J* = 10.0 Hz, 1H, H-4), 5.66 (dd, *J* = 3.0, 1.5 Hz, 1H, H-2), 5.09 (d, *J* = 1.0 Hz, 1H, H-1), 4.06 (m, 1H, H-5), 3.85-3.75 (m, 3H, H-6a, H-6b, H-OCH<sub>2</sub>C<sub>6</sub>H<sub>13</sub>), 3.54 (dt, *J* = 9.5, 6.5 Hz, 1H, H-OCH<sub>2</sub>C<sub>6</sub>H<sub>13</sub>), 2.63 (m, 1H, 6-OH), 1.68 (m, 2H), 1.42-1.25 (m, 8H), 0.91 (t, *J* = 7.0 Hz, 3H). <sup>13</sup>C-NMR (125 MHz, CDCl<sub>3</sub>): δ 166.56, 165.57, 165.49 (3xCO), 133.66, 133.52, 133.16, 129.91, 129.68, 129.33, 129.14, 128.73, 128.62, 128.50, 128.29, 97.68 (C-1), 70.82 (C-2), 70.77 (C-5), 69.69 (C-3), 68.67 (OCH<sub>2</sub>C<sub>6</sub>H<sub>13</sub>), 67.39 (C-4), 61.39 (C-6), 31.77, 29.38, 29.08, 26.08, 22.64, 14.11 (OCH<sub>2</sub>C<sub>6</sub>H<sub>13</sub>). HR-MS Calcd. for C<sub>34</sub>H<sub>38</sub>O<sub>9</sub> [M+Na]<sup>+</sup>, 613.2414, found 613.2420.

#### ***n*-Heptyl 2,3,5-tri-O-benzoyl-6-deoxy-6-fluoro- $\alpha$ -D-mannopyranoside (14)**

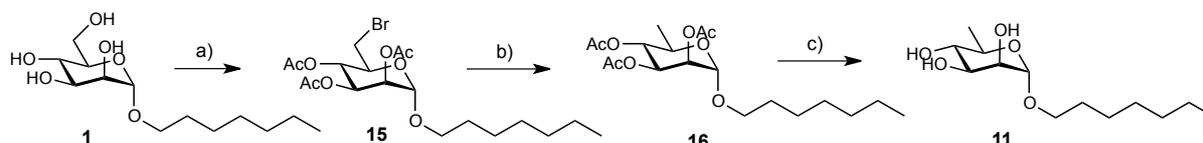
To a solution of **13** (125 mg, 0.2116 mmol) in dry DCM (0.5 mL) was added DAST (0.44 mL) at 0°C, then stirred at rt overnight. After cooling to 0°C, the reaction mixture was quenched with methanol, concentrated, the residue was purified by flash chromatography on silica gel using PE-EE (9:1-6:1) to afford the desired compound as colorless oil (56 mg, 45%). [ $\alpha$ ]<sub>D</sub><sup>20</sup> = -126.76 (c = 0.61, DCM). <sup>1</sup>H-NMR (500 MHz, CDCl<sub>3</sub>): δ 8.10 (m, 2H), 7.97 (m, 2H), 7.83 (m, 2H), 7.62 (t, *J* = 7.0 Hz, 1H), 7.55-7.48 (m, 3H), 7.45-7.38 (m, 3H), 7.26 (m, 2H), 5.92 (dd, *J* = 10.0, 3.0 Hz, 1H, H-3), 5.88 (t, *J* = 10.0 Hz, 1H, H-4), 5.67 (dd, *J* = 3.0, 1.5 Hz, 1H, H-2), 5.09 (d, *J* = 1.5 Hz, 1H, H-1), 4.64 (ddd, *J* = 47.5, 10.5, 4.5 Hz, 1H, H-6a), 4.60 (ddd, *J* = 47.0, 10.5, 2.5 Hz, 1H, H-6b), 4.28 (m, 1H, H-5), 3.82 (dt, *J* = 9.5, 6.5 Hz, 1H, H-OCH<sub>2</sub>C<sub>6</sub>H<sub>13</sub>), 3.57 (dt, *J* = 9.5, 6.5 Hz, 1H, H-OCH<sub>2</sub>C<sub>6</sub>H<sub>13</sub>), 1.71 (m, 2H), 1.37 (m, 8H), 0.91 (t, *J* = 7.0 Hz, 3H). <sup>13</sup>C-NMR

(125 MHz, CDCl<sub>3</sub>):  $\delta$  165.56, 165.51, 165.45 (3xCO), 133.52, 133.17, 129.91, 129.80, 129.70, 129.28, 129.05, 128.86, 128.60, 128.47, 128.29, 97.59 (C-1), 81.75 (d,  $J$  = 173.88 Hz, C-6), 70.55 (C-2), 69.94 (C-3), 69.59 (d,  $J$  = 19.0 Hz, C-5), 68.81 (OCH<sub>2</sub>C<sub>6</sub>H<sub>13</sub>), 66.36 (d,  $J$  = 9.0 Hz, C-4), 31.77, 29.36, 29.06, 26.06, 22.63, 14.10 (OCH<sub>2</sub>C<sub>6</sub>H<sub>13</sub>). HR-MS Calcd for C<sub>34</sub>H<sub>37</sub>FO<sub>8</sub> [M+Na]<sup>+</sup>, 615.2370, found 615.2373.

### *n*-Heptyl 6-deoxy-6-fluoro- $\alpha$ -D-mannopyranoside (10)

To a solution of *n*-heptyl **14** (44.8 mg, 0.0756 mmol) in dry methanol (1.0 mL) was added 0.5 M CH<sub>3</sub>ONa/MeOH (80  $\mu$ L) at rt. The reaction mixture was stirred overnight, then neutralized HOAc, concentrated to dryness. The residue was purified by flash chromatography on silica gel (DCM:MeOH 15:1 to 10:1) to afford **10** as a colorless syrup (17 mg, 80%).  $[\alpha]_D^{20}$  = +98.8 (c = 0.28, DCM). <sup>1</sup>H-NMR (500 MHz, CD<sub>3</sub>OD):  $\delta$  4.73 (d,  $J$  = 1.5 Hz, 1H, H-1), 4.65 (m, 1H, H-6a), 4.55 (m, 1H, H-6b), 3.79 (dd,  $J$  = 3.0, 2.0 Hz, 1H, H-2), 3.72-3.61 (m, 4H, H-OCH<sub>2</sub>C<sub>6</sub>H<sub>13</sub>, H-3, H-4, H-5), 3.42 (dt,  $J$  = 9.5, 6.5 Hz, 1H, H-OCH<sub>2</sub>C<sub>6</sub>H<sub>13</sub>), 1.59 (m, 2H), 1.32 (m, 8H), 0.91 (t,  $J$  = 7.0 Hz, 3H, CH<sub>3</sub>). <sup>13</sup>C-NMR (125 MHz, CD<sub>3</sub>OD):  $\delta$  101.74 (C-1), 84.23 (d,  $J$  = 169.88 Hz, C-6), 73.35 (d,  $J$  = 17.75 Hz, C-5), 72.63 (C-3), 72.10 (C-2), 68.71 (OCH<sub>2</sub>C<sub>6</sub>H<sub>13</sub>), 67.48 (d,  $J$  = 7.25 Hz, C-4), 32.99, 30.60, 30.21, 27.32, 23.69, 14.42 (OCH<sub>2</sub>C<sub>6</sub>H<sub>13</sub>). <sup>19</sup>F-NMR (376 MHz, CD<sub>3</sub>OD):  $\delta$  -230.46 (m). ESI-MS Calcd for [M+Na]<sup>+</sup>, 303.16, found 303.07. HR-MS Calcd. for C<sub>13</sub>H<sub>25</sub>FO<sub>5</sub> [M+Na]<sup>+</sup>, 303.1578, found 303.1575.

### Synthesis of **54** (6-H) (*n*-Heptyl 6-deoxy- $\alpha$ -D-mannopyranoside)



**Scheme S2.** a) i. CBr<sub>4</sub>, Ph<sub>3</sub>P, pyr, 0-65°C, 3 h, ii. Ac<sub>2</sub>O/pyr, DMAP, two steps (81%); b) Bu<sub>3</sub>SnH/toluene, reflux, overnight, (80%); c) CH<sub>3</sub>ONa/CH<sub>3</sub>OH, (66%).

### *n*-Heptyl 2,3,4-tri-O-acetyl-6-bromo-6-deoxy- $\alpha$ -D-mannopyranoside (15)

To a solution of **1** (110 mg, 0.359 mmol) in pyridine (3 mL) was added triphenylphosphine (189 mg, 0.72 mmol) in one portion at 0°C, then added CBr<sub>4</sub> (179 mg, 0.54 mmol) in portions at 0°C. The reaction mixture was stirred at 0°C for 0.5 h, then was heated at 65°C for 3 h. After cooling, methanol (2 mL) was added dropwise and the mixture was concentrated to dryness. The residue was acetylated with Ac<sub>2</sub>O (1 mL), pyridine (2 mL) and DMAP (2 mg). The reaction mixture was stirred at rt for 3 h and then concentrated, the residue was purified by flash chromatography on silica gel (PE-EA 6:1-4:1) to give the desired product (136 mg, 81%) as a colorless syrup.  $[\alpha]_D^{20}$  = +51.60 (c = 0.89, CHCl<sub>3</sub>); <sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>):  $\delta$  5.33 (m, 1H, H-3), 5.21 (d,  $J$  = 1.0 Hz, 1H, H-2), 5.16 (t,  $J$  = 10.0 Hz, 1H, H-4), 4.81 (s, 1H, H-1), 3.97 (m, 1H, H-5), 3.76 (m, 1H, OCH<sub>2</sub>C<sub>6</sub>H<sub>13</sub>), 3.42 (m, 3H, H-6a, H-6b, OCH<sub>2</sub>C<sub>6</sub>H<sub>13</sub>), 2.13, 2.06, 1.98 (3xs, 9H, 3xCOCH<sub>3</sub>), 1.59 (m, 2H), 1.30 (m, 8H, OC<sub>2</sub>H<sub>4</sub>C<sub>4</sub>H<sub>8</sub>CH<sub>3</sub>), 0.87 (t,  $J$  = 6.5 Hz, 3H, OC<sub>2</sub>H<sub>4</sub>C<sub>4</sub>H<sub>8</sub>CH<sub>3</sub>); <sup>13</sup>C NMR (125 MHz, CDCl<sub>3</sub>):  $\delta$  170.06, 169.84 (3xCOCH<sub>3</sub>), 97.28 (C-1), 69.88 (C-5), 69.67 (C-2), 68.91 (C-3), 68.84 (C-4), 68.41 (OCH<sub>2</sub>C<sub>6</sub>H<sub>13</sub>), 31.44 (C-6), 31.68, 29.16, 28.98, 25.99, 22.57, 14.05 (OCH<sub>2</sub>C<sub>6</sub>H<sub>13</sub>), 20.86, 20.74, 20.66 (3xCOCH<sub>3</sub>); ESI-MS Calcd for [M+Na+2]<sup>+</sup>, 491.13, found 491.11. HR-MS Calcd. for C<sub>19</sub>H<sub>31</sub>BrO<sub>8</sub> [M+Na]<sup>+</sup>, 489.1100, found 489.1099.

***n*-Heptyl 2,3,4-tri-O-acetyl-6-deoxy- $\alpha$ -D-mannopyranoside (16)**

A solution of Bu<sub>3</sub>SnH (0.26 mL, 0.96 mmol) in toluene (2.5 mL) was added dropwise to a gently refluxing solution of **15** (300 mg, 0.64 mmol) in toluene (4 mL) over 10 min and the mixture was refluxing overnight. After cooling, the mixture was concentrated and the residue was purified by flash chromatography on silica gel (PE-EA 3:1-3:2) to give the desired compound (198 mg, 80%) as a colorless oil.  $[\alpha]_D^{20} = +52.72$  ( $c = 0.30$ , MeOH); <sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>):  $\delta$  5.30 (dd,  $J = 10.0, 3.5$  Hz, 1H, H-3), 5.22 (dd,  $J = 3.5, 1.5$  Hz, 1H, H-2), 5.05 (t,  $J = 10.0$  Hz, 1H, H-4), 4.71 (s, 1H, H-1), 3.87 (m, 1H, H-5), 3.65 (dt,  $J = 9.5, 6.5$  Hz, 1H, OCH<sub>2</sub>C<sub>6</sub>H<sub>13</sub>), 3.40 (dt,  $J = 9.5, 6.5$  Hz, 1H, OCH<sub>2</sub>C<sub>6</sub>H<sub>13</sub>), 2.15, 2.05, 1.98 (3xs, 9H, 3xCOCH<sub>3</sub>), 1.57 (m, 2H, OCH<sub>2</sub>CH<sub>2</sub>C<sub>5</sub>H<sub>11</sub>), 1.30 (m, 8H, OC<sub>2</sub>H<sub>4</sub>C<sub>4</sub>H<sub>8</sub>CH<sub>3</sub>), 1.21 (d,  $J = 6.0$  Hz, 3H, H-6), 0.88 (t,  $J = 7.0$  Hz, 3H, OC<sub>2</sub>H<sub>4</sub>C<sub>4</sub>H<sub>8</sub>CH<sub>3</sub>); <sup>13</sup>C NMR (125 MHz, CDCl<sub>3</sub>):  $\delta$  170.21, 170.05, 170.00 (3xCOCH<sub>3</sub>), 97.38 (C-1), 71.23 (C-4), 70.02 (C-2), 69.18 (C-3), 68.23 (OCH<sub>2</sub>C<sub>6</sub>H<sub>13</sub>), 66.16 (C-5), 17.40 (C-6), 31.73, 29.30, 29.01, 26.00, 22.59, 14.06 (OCH<sub>2</sub>C<sub>6</sub>H<sub>13</sub>), 20.95, 20.81, 20.74 (3xCOCH<sub>3</sub>); HR-MS calcd. for C<sub>19</sub>H<sub>32</sub>O<sub>8</sub> [M+Na]<sup>+</sup>, 411.1995, found 411.1992.

***n*-Heptyl 6-deoxy- $\alpha$ -D-mannopyranoside (11)**

To a solution of **16** (80 mg, 0.206 mmol) in methanol (2 mL) was added 0.5 M CH<sub>3</sub>ONa/MeOH (41  $\mu$ L) at rt. The reaction mixture was stirred at rt overnight, then neutralized with Amberlyst 15. The reaction mixture was filtered and the residue was washed thoroughly with MeOH. The filtrate was concentrated and the residue was purified with flash chromatography on silica gel (PE:EA 1:1-1:2) to give compound **11** (36 mg, 66%) as a colorless oil.  $[\alpha]_D^{20} = +50.40$  ( $c = 0.33$ , MeOH); <sup>1</sup>H NMR (500 MHz, CD<sub>3</sub>OD):  $\delta$  4.65 (s, 1H, H-1), 3.78 (m, 1H, H-2), 3.69-3.62 (m, 2H, H-OCH<sub>2</sub>C<sub>6</sub>H<sub>13</sub>, H-3), 3.57 (m, 1H, H-5), 3.42-3.35 (m, 2H, H-OCH<sub>2</sub>C<sub>6</sub>H<sub>13</sub>, H-4), 1.58 (m, 2H, OCH<sub>2</sub>CH<sub>2</sub>C<sub>5</sub>H<sub>11</sub>), 1.32 (m, 8H, OC<sub>2</sub>H<sub>4</sub>C<sub>4</sub>H<sub>8</sub>CH<sub>3</sub>), 1.26 (d,  $J = 6.0$  Hz, 3H, H-6), 0.91 (t,  $J = 7.0$  Hz, 3H, OC<sub>2</sub>H<sub>4</sub>C<sub>4</sub>H<sub>8</sub>CH<sub>3</sub>); <sup>13</sup>C NMR (125 MHz, CD<sub>3</sub>OD):  $\delta$  101.61 (C-1), 73.96 (C-4), 72.44 (C-3), 72.34 (C-2), 69.70 (C-5), 68.52 (OCH<sub>2</sub>C<sub>6</sub>H<sub>13</sub>), 32.97, 30.64, 30.20, 27.30, 23.68, 14.43 (OCH<sub>2</sub>C<sub>6</sub>H<sub>13</sub>), 18.00 (C-6); HR-MS Calcd for C<sub>13</sub>H<sub>26</sub>O<sub>5</sub> [M+Na]<sup>+</sup>, 285.1675, found 285.1678.

## References

1. Rabbani, S., Jiang, X., Schwardt, O., & Ernst, B. (2010). Expression of the carbohydrate recognition domain of FimH and development of a competitive binding assay. *Anal. Biochem.*, 407(2): 188-195.
2. Wang, Z.-X. (1995). An exact mathematical expression for describing competitive binding of two different ligands to a protein molecule. *FEBS Lett.*, 360(2): 111-114.
3. Kleeb, S., *et al.* (2015). FimH antagonists: bioisosteres to improve the in vitro and in vivo PK/PD profile. *J. Med. Chem.*, 58(5): 2221-2239.
4. Kabsch, W. (2010). Integration, scaling, space-group assignment and post-refinement. *Acta Crystallogr. D Biol. Crystallogr.*, 66(Pt 2): 133-144.
5. Kabsch, W. (2010). Xds. *Acta Crystallogr. D Biol. Crystallogr.*, 66(Pt 2): 125-132.
6. McCoy, A.J. (2007). Solving structures of protein complexes by molecular replacement with Phaser. *Acta Crystallogr. Sect. D.-Biol. Crystallogr.*, 63(1): 32-41.
7. Emsley, P. & Cowtan, K. (2004). Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.*, 60(Pt 12 Pt 1): 2126-2132.
8. Adams, P.D., *et al.* (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.*, 66(Pt 2): 213-221.
9. Blanc, E., *et al.* (2004). Refinement of severely incomplete structures with maximum likelihood in BUSTER-TNT. *Acta Crystallogr. Sect. D.-Biol. Crystallogr.*, 60(12): 2210-2221.
10. van Aalten, D.M., *et al.* (1996). PRODRG, a program for generating molecular topologies and unique molecular descriptors from coordinates of small molecules. *J. Comput. Aided Mol. Des.*, 10(3): 255-262.
11. Chen, V.B., *et al.* (2010). MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D Biol. Crystallogr.*, 66(Pt 1): 12-21.
12. Schrödinger Suite 2014-4. MacroModel version 10.6; Schrödinger, LLC, New York, NY, 2014.
13. Frisch, M., *et al.* (2009). Gaussian 09, revision A. 02; Gaussian, Inc. Wallingford, CT, 19: 227-238.
14. Desmond Molecular Dynamics System. version 4.4; D. E. Shaw Research, New York, NY, 2014.
15. Schrödinger Suite 2015-4: Desmond Molecular Dynamics System. Maestro-Desmond Interoperability Tools Version 4.4 ed; Schrödinger, LLC, New York, NY, 2015.
16. Guo, Z., *et al.* (2010). Probing the alpha-helical structural stability of stapled p53 peptides: molecular dynamics simulations and analysis. *Chem. Biol. Drug Des.*, 75(4): 348-359.
17. Shivakumar, D., *et al.* (2010). Prediction of Absolute Solvation Free Energies using Molecular Dynamics Free Energy Perturbation and the OPLS Force Field. *J. Chem. Theory Comput.*, 6(5): 1509-1519.
18. Bowers, K.J., Dror, R.O., & Shaw, D.E. (2006). The midpoint method for parallelization of particle simulations. *J. Chem. Phys.*, 124(18): 184109.
19. Schrödinger Suite 2014-4. Prime version 3.8; Schrödinger, LLC, New York, NY, 2014.
20. Jacobson, M.P., Friesner, R.A., Xiang, Z., & Honig, B. (2002). On the role of the crystal environment in determining protein side-chain conformations. *J. Mol. Biol.*, 320(3): 597-608.
21. Jacobson, M.P., *et al.* (2004). A hierarchical approach to all-atom protein loop prediction. *Proteins*, 55(2): 351-367.
22. Vranken, W.F., *et al.* (2005). The CCPN data model for NMR spectroscopy: development of a software pipeline. *Proteins*, 59(4): 687-696.
23. Fiege, B., *et al.* (2015). The Tyrosine Gate of the Bacterial Lectin FimH: A Conformational Analysis by NMR Spectroscopy and X-ray Crystallography. *ChemBioChem*, 16(8): 1235-1246.
24. Pellecchia, M., Sebbel, P., Hermanns, U., Wuthrich, K., & Glockshuber, R. (1999). Pilus chaperone FimC-adhesin FimH interactions mapped by TROSY-NMR. *Nat. Struct. Biol.*, 6(4): 336-339.
25. Edelhoch, H. (1967). Spectroscopic determination of tryptophan and tyrosine in proteins. *Biochemistry*, 6(7): 1948-1954.
26. AFFINImeter. KinITC AFFINImeter kinetics version 01-07-2015; Software for Science Developments, Santiago de Compostela, Spain, 2015.
27. Burnouf, D., *et al.* (2012). kinITC: a new method for obtaining joint thermodynamic and kinetic data by isothermal titration calorimetry. *J. Am. Chem. Soc.*, 134(1): 559-565.
28. Du, Y., Zhang, M., & Kong, F. (2001). Efficient and practical syntheses of three pentasaccharides core structures corresponding to N-glycans. *Tetrahedron*, 57(9): 1757-1763.
29. Ding, X., Wang, W., & Kong, F. (1997). Detritylation of mono- and di-saccharide derivatives using ferric chloride hydrate. *Carbohydr. Res.*, 303(4): 445-448.



## Manuscript 2

### The price of flexibility – a case study on septanoses as pyranose mimetics

Christoph P. Sager,<sup>†#</sup> Brigitte Fiege,<sup>†#</sup> Pascal Zihlmann,<sup>†#</sup> Raghu Vannam,<sup>†#</sup>  
Said Rabbani,<sup>†</sup> Roman P. Jakob,<sup>§</sup> Roland C. Preston,<sup>†</sup> Adam Zalewski,<sup>†</sup>  
Timm Maier,<sup>§</sup> Beat Ernst,<sup>†\*</sup> Mark W. Pecuh<sup>¶</sup>

<sup>#</sup> These authors contributed equally to the project

<sup>¶</sup> University of Connecticut, Department of Chemistry,  
55 N. Eagleville Road U3060, Storrs, CT, 06279 USA

<sup>†</sup> University of Basel, Institute of Molecular Pharmacy,  
Klingelbergstr. 50, 4056 Basel, Switzerland

<sup>§</sup> University of Basel, Institute of Structural Biology,  
Klingelbergstr. 70, 4056 Basel, Switzerland

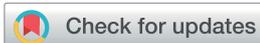
\*Corresponding author.

Tel.: 0041 (0)61 267 15 51; Fax: 0041 (0)61 267 15 52;  
E-mail: beat.ernst@unibas.ch

#### Contributions of Pascal Zihlmann:

- Aiding in manuscript preparation
- Thermodynamic profiling by ITC and data analysis





Cite this: DOI: 10.1039/c7sc04289b

## The price of flexibility – a case study on septanoses as pyranose mimetics†

Christoph P. Sager,<sup>‡a</sup> Brigitte Fiege,<sup>‡a</sup> Pascal Zihlmann,<sup>‡a</sup> Raghu Vannam,<sup>‡b</sup> Said Rabbani,<sup>a</sup> Roman P. Jakob,<sup>‡c</sup> Roland C. Preston,<sup>a</sup> Adam Zalewski,<sup>‡a</sup> Timm Maier,<sup>‡c</sup> Mark W. Pecuh<sup>‡\*b</sup> and Beat Ernst<sup>‡\*a</sup>

Seven-membered ring mimetics of mannose were studied as ligands for the mannose-specific bacterial lectin FimH, which plays an essential role in the first step of urinary tract infections (UTI). A competitive binding assay and isothermal titration calorimetry (ITC) experiments indicated an approximately ten-fold lower affinity for the seven-membered ring mannose mimetic 2-*O*-*n*-heptyl-1,6-anhydro-*D*-glycero-*D*-galactitol (**7**) compared to *n*-heptyl  $\alpha$ -*D*-mannopyranoside (**2**), resulting exclusively from a loss of conformational entropy. Investigations by solution NMR, X-ray crystallography, and molecular modeling revealed that **7** establishes a superimposable H-bond network compared to mannoside **2**, but at the price of a high entropic penalty due to the loss of its pronounced conformational flexibility. These results underscore the importance of having access to the complete thermodynamic profile of a molecular interaction to “rescue” ligands from entropic penalties with an otherwise perfect fit to the protein binding site.

Received 2nd October 2017  
Accepted 7th November 2017

DOI: 10.1039/c7sc04289b

rsc.li/chemical-science

## Introduction

The flexibility of a molecule defines its intrinsic conformational entropy. Rigid molecules adopt a highly populated conformational ground state with no or only a few alternative conformations. In contrast, flexible molecules populate multiple conformers with low barriers for interconversion between them. Flexibility has implications on intermolecular interactions in different manners (Fig. 1).<sup>1</sup> Emil Fischer's lock-and-key model (A) assumes two rigid partners and therefore requires precise complementary shapes to allow tight binding.<sup>2–6</sup> However, in most cases, either the ligand (B)<sup>7–12</sup> or the receptor (C)<sup>13,14</sup> exhibits flexibility while its binding partner is rather rigid. In these cases, complex formation requires an induced fit<sup>5</sup> or a conformational selection mechanism.<sup>15–18</sup> Finally, considerable reorganizational energy is required for interactions in which both ligand and receptor are flexible (D).<sup>19</sup> A detailed knowledge of the thermodynamic profile allows for categorization of a binding event with respect to these four scenarios and

affects the design of ligands. The enthalpic and entropic contributions to these intermolecular associations, fundamental for characterizing the thermodynamic profile of an

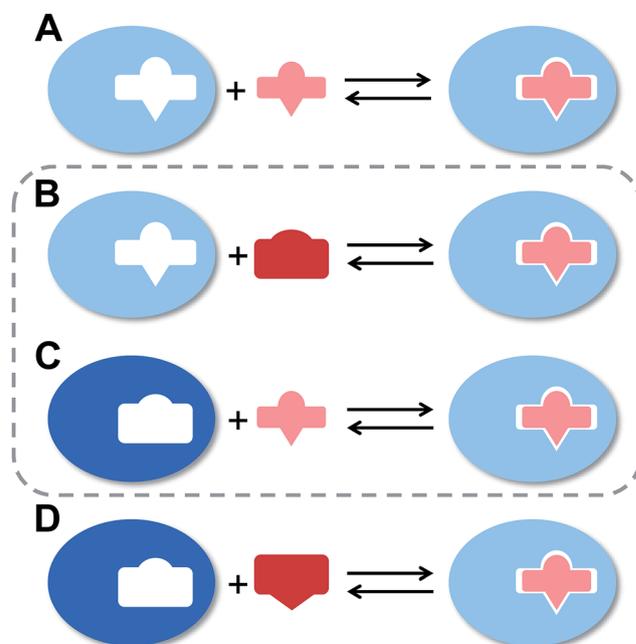


Fig. 1 The four types of receptor–ligand interactions. (A) Rigid receptor, rigid ligand. (B) Rigid receptor, flexible ligand. (C) Flexible receptor, rigid ligand. (D) Flexible receptor, flexible ligand.

<sup>a</sup>University of Basel, Institute of Molecular Pharmacy, Pharmazentrum of the University of Basel, Klingelbergstrasse 50, 4056, Basel, Switzerland. E-mail: beat.ernst@unibas.ch

<sup>b</sup>Department of Chemistry, University of Connecticut, 55 N. Eagleville Road U3060, Storrs, CT, 06279 USA. E-mail: mark.pecuh@uconn.edu

<sup>c</sup>University of Basel, Biozentrum: Focal Area Structural Biology, Klingelbergstrasse 70, 4056 Basel, Switzerland

† Electronic supplementary information (ESI) available. See DOI: 10.1039/c7sc04289b

‡ These authors contributed equally to this work.



association, can be determined by isothermal titration calorimetry (ITC).

In the present study, an example of a receptor–ligand interaction according to type B (Fig. 1) was studied in detail. The receptor is the lectin FimH, located at the tip of bacterial pili. It mediates the attachment of uropathogenic *E. coli* (UPEC) to urothelial cells of the mammalian host and therefore represents the first step of the process leading to urinary tract infection (UTI).<sup>20–22</sup> This adhesion impedes the clearance of bacteria from the urinary tract and is therefore essential for the infection of urothelial cells.<sup>23–25</sup> Since the physiological ligand of FimH is the highly mannosylated glycoprotein uroplakin 1A located on urothelial cells,<sup>20–22</sup> a broad variety of mannose-based FimH antagonists have been developed for blocking this interaction and thereby the bacterial infection of bladder cells.<sup>26–34</sup> Their successful application in a UTI mouse model clearly supports their therapeutic potential for treatment and prevention of UTI.<sup>26,31,35</sup>

To date, only the aglycone of mannosidic FimH antagonists has been successfully varied,<sup>26–37</sup> whereas the replacement of mannose by other hexoses, *e.g.* D-glucose or D-galactose, led to a severe reduction or even loss of affinity.<sup>38</sup> We were interested to know whether septanose derivatives, *i.e.* seven-membered ring homologs of pyranoses,<sup>39</sup> could be utilized as ring-expanded FimH antagonists. Our interest was motivated by previous studies where the jackbean lectin concanavalin A (ConA) bound methyl septanosides in its carbohydrate binding pocket, albeit with reduced affinity relative to the pyranose ligand.<sup>40,41</sup> ConA naturally recognizes the  $\alpha$ -1-3,  $\alpha$ -1-6-dimannosylmannoside<sup>42</sup> as well as methyl  $\alpha$ -pyranosides of mannose and glucose.<sup>43,44</sup> In contrast to the natural selectivity for  $\alpha$ -configured saccharides, methyl  $\beta$ -septanosides were bound by ConA whereas  $\alpha$ -septanosides were not.

Here we report results from an investigation of septanoses binding to the isolated lectin domain of FimH (FimH<sub>LD</sub>), which is locked in the high-affinity state.<sup>45,46</sup> Flexibility was fundamentally important for binding of the best seven-membered ring ligand, 2-*O*-*n*-heptyl-1,6-anhydro-D-*glycero*-D-galactitol (7) (informally referred to as 2-*O*-*n*-heptyl-1-deoxy-septanose), as FimH<sub>LD</sub> selected one conformation among those on its shallow conformational energy surface.

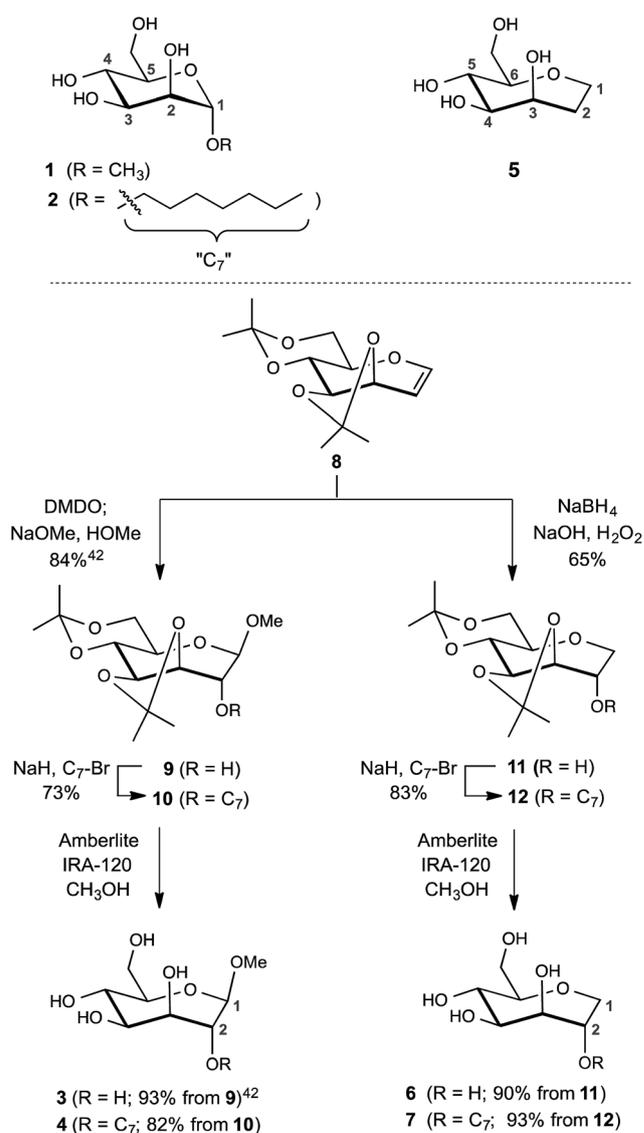
## Results

### Synthesis of septanose ligands

The design of the septanose ligands was guided by the structure of the mannose-binding pocket of FimH. It consists of a deep pocket lined with polar amino acid residues forming an extended H-bond network with the natural ligand. The entrance to the mannose-binding pocket, the so-called “tyrosine gate”, is furnished with two tyrosines and an isoleucine and interacts with aliphatic or aromatic aglycones of FimH antagonists.<sup>38</sup> For mimicking mannose-based FimH antagonists, methyl  $\beta$ -D-*glycero*-D-galacto-septanosides 3 and 4 along with the related 1,6-anhydro-D-*glycero*-D-galactitols 5–7 were synthesized (3 and 4 will be referred to as methyl septanosides, and 5–7 as 1-deoxy-septanoses). For all these seven-membered ring analogs of  $\alpha$ -D-mannose, the configuration at C3 to C6 (septanose numbering) is identical to the configuration at C2 to C5 in D-

mannose. In addition, the *n*-heptyl substituent on the C2 hydroxyl group of septanose 4 and 7 was expected to mimic the *n*-heptyl aglycone at the anomeric center (C1) of  $\alpha$ -D-mannopyranoside 2 (Scheme 1).

For the syntheses of 3 & 4 and 6 & 7, acetonide protected oxepine 8 was utilized as a key starting material.<sup>47</sup> The previously described diastereoselective epoxidation and methanolysis of 8 provided methyl  $\beta$ -D-*glycero*-D-galacto-septanoside 9 (84%).<sup>47,48</sup> Alkylation of its C2 hydroxyl group gave 2-*O*-*n*-heptyl septanoside 10 in 73% yield. By removal of the acetonide group on 9 and 10, methyl septanosides 3 and 4 were obtained in 93% and 82% yields, respectively. The preparation of 6 and 7 relied on the regio- and diastereoselective hydroboration–oxidation of 8 to give 11 (65%).<sup>49</sup> Deprotection of 1-deoxy-septanose 11 then yielded 6 (90%). Alternatively, 7 was obtained by alkylation of 11 at C2 to form compound 12 followed by deprotection (77% over two steps).



Scheme 1 Syntheses of FimH antagonists with a septanose core as replacement for  $\alpha$ -D-mannopyranoside 1 and 2. Methyl  $\beta$ -septanoside 3 and 1,2-di-deoxy-septanose 5 have been previously described.<sup>40</sup>



**Table 1** Affinity ( $IC_{50}$  &  $K_D$ ) and thermodynamic data for the interaction of mannosides 1 & 2, septanosides 3 & 4 and 1-deoxy-septanosides 5–7 with FimH<sub>LD</sub>

Compound	$IC_{50}^a$ [ $\mu$ M]	$K_D^b$ [ $\mu$ M]	$\Delta C_{obs}^b$ [ $kJ\ mol^{-1}$ ]	$\Delta H_{obs}^b$ [ $kJ\ mol^{-1}$ ]	$-T\Delta S_{obs}^b$ [ $kJ\ mol^{-1}$ ]	$N^b$
1 (methyl $\alpha$ -D-mannoside)	$5.6 \pm 0.5$	n.d.				
2 ( <i>n</i> -heptyl $\alpha$ -D-mannoside)	$0.064 \pm 0.02$	0.029	-43.0	-50.3	7.3	1.00
3	$31.6 \pm 4.2$	n.d.				
4	$2.83 \pm 0.4$	2.20	-32.3	-27.9	-4.4	0.98
5	$86.8 \pm 7.9$	n.d.				
6	$421 \pm 30.5$	n.d.				
7	$1.37 \pm 0.3$	0.26	-37.5	-49.4	11.8	1.00

<sup>a</sup>  $IC_{50}$  values were measured in a competitive binding assay.<sup>36</sup> <sup>b</sup>  $K_D$  values and thermodynamic parameters were determined by ITC experiments at a temperature of 25 °C.

### Competitive binding assay

Compounds 1–7, along with other analogs (see ESI†), were evaluated for their binding to FimH<sub>LD</sub>. Compounds were first screened in a competitive binding assay,<sup>36</sup> where a polymeric oligomannose–biotin conjugate coupled *via* streptavidin to horseradish peroxidase is added simultaneously with a serial dilution of the test ligand to wells coated with FimH<sub>LD</sub>. After incubation, the peroxidase substrate 3,3',5,5'-tetramethylbenzidine is added and the binding is quantified by colorimetric detection. For the compounds 1–7, the competitive binding assay revealed  $IC_{50}$  values ranging from 64 nM to ~400  $\mu$ M (Table 1). The addition of an *n*-heptyl substituent at C2 of 3 ( $\rightarrow$  4) and 6 ( $\rightarrow$  7) led to significantly improved inhibition. Septanoside 4 exhibits an 11-fold increase in affinity over 3, and 1-deoxy-septanose 7 a 307-fold improvement over 6. Thus, the *n*-heptyl substituent in 4 and 7 likely mimics the aglycone of the corresponding pyranoside 2.

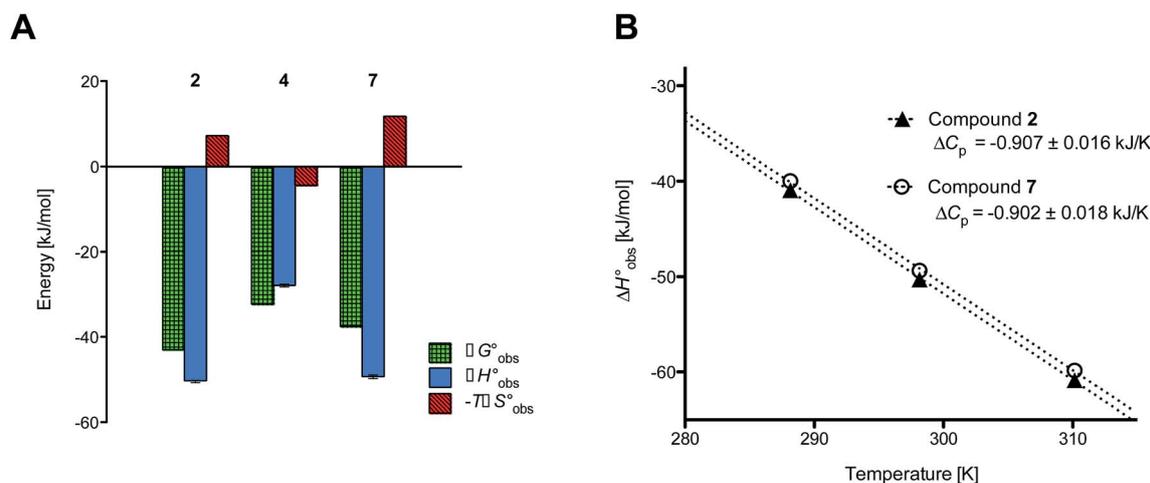
### ITC experiments with compounds 2, 4 and 7

Isothermal titration calorimetry (ITC) provided a deeper insight into the nature of the intermolecular interaction of FimH<sub>LD</sub> with 2-*n*-heptyl-septanoside 4 and 2-*n*-heptyl-1-deoxy-septanose 7 in comparison with *n*-heptyl  $\alpha$ -D-mannopyranoside (2). All

three test compounds contain an *n*-heptyl unit – either as aglycone ( $\rightarrow$  2) or as substituent in the 2-position ( $\rightarrow$  4 and 7, respectively). The *n*-heptyl group is essential to establish the important hydrophobic contacts within the tyrosine gate. The  $K_D$  values obtained from ITC experiments were in relative agreement with the  $IC_{50}$  values from the competitive binding assay, *i.e.* in both assays 2 was the strongest binder followed by 7 and then 4 (Table 1).

ITC data (Table 1 and Fig. 2A) revealed that the almost 100-fold loss in affinity of compound 4 compared to compound 2 is due to a diminished enthalpic contribution ( $\Delta H_{obs}^b$ ) that is partly compensated by a more favorable entropy term ( $T\Delta S_{obs}^b$ ). The less beneficial enthalpy term for septanoside 4 likely results from an increased desolvation penalty due to its increased polar surface area (108.6 Å<sup>2</sup> for 4 compared to 99.4 Å<sup>2</sup> for 2)<sup>50</sup> as well as from looser interactions with the FimH<sub>LD</sub>. As a consequence of a looser fit, the flexibility of septanoside 4 is less compromised, explaining the substantially improved entropy term by 11.7 kJ mol<sup>-1</sup> (Table 1).

In contrast, the binding enthalpies of 2 and 7 are almost identical, so that the affinity loss of 1-deoxy-septanose 7 in comparison to mannoside 2 is exclusively due to a significant entropic penalty ( $T\Delta S_{obs}^b$  term). According to eqn (1), the entropy  $\Delta S^b$  consists of desolvation entropy  $\Delta S_{solv}^b$ ,



**Fig. 2** (A) Thermodynamic fingerprints of compounds 2, 4, and 7 binding to FimH<sub>LD</sub>. (B) Heat capacity ( $\Delta C_p$ ) of compounds 2 and 7 binding to FimH<sub>LD</sub> from ITC experiments performed at three different temperatures (details in ESI†).



Table 2 Dissected entropic contributions of mannoside 2 and 1-deoxy-septanose 7 binding to FimH<sub>LD</sub> at 298.15 K

Contributor	Mannoside 2 [kJ mol <sup>-1</sup> ]	1-Deoxy-septanose 7 [kJ mol <sup>-1</sup> ]
$-T\Delta S_{\text{solv}}^{\circ}$	-69.1	-68.8
$-T\Delta S_{\text{conf}}^{\circ}$	66.4	70.6
$-T\Delta S_{\text{mix}}^{\circ}$	10.0	10.0
$-T\Delta S_{\text{obs}}^{\circ}$	7.3	11.8

conformational entropy  $\Delta S_{\text{conf}}^{\circ}$ , and a term  $\Delta S_{\text{mix}}^{\circ}$  reflecting the loss in translational and rotational degrees of freedom of both binding partners upon complex formation. Because  $\Delta S_{\text{mix}}^{\circ}$  for the two antagonists is equal (10 kJ mol<sup>-1</sup>, eqn (2), where  $R$  is the universal gas constant and 55.6 the molarity of water in [M]),<sup>51,52</sup> the entropy penalty originates exclusively from a loss of conformational flexibility ( $\Delta S_{\text{conf}}^{\circ}$ ) – either caused by the ligand and/or the receptor.

$$\Delta S_{\text{obs}}^{\circ} = \Delta S_{\text{solv}}^{\circ} + \Delta S_{\text{conf}}^{\circ} + \Delta S_{\text{mix}}^{\circ} \quad (1)$$

$$\Delta S_{\text{mix}}^{\circ} = R \ln\left(\frac{1}{55.6}\right) \quad (2)$$

To quantify the contributions of  $\Delta S_{\text{solv}}^{\circ}$  and  $\Delta S_{\text{conf}}^{\circ}$ , ITC experiments were conducted at three different temperatures to determine the heat capacity ( $\Delta C_p$ ) of mannoside 2 and 1-deoxy-septanose 7 (Fig. 2B and Table 2).  $\Delta C_p$  (eqn (3)), reporting on the entropy contributions from desolvation of ligand and protein, is equal for both complex formations (Fig. 2B). Since at 385 K,  $\Delta S_{\text{solv}}^{\circ}$  reaches zero, *i.e.* the hydration shell no longer exists,  $\Delta S_{\text{solv}, 298 \text{ K}}^{\circ}$  can be calculated according to eqn (4).<sup>52-54</sup>

$$\Delta C_p = \left(\frac{\partial \Delta H_{\text{obs}}^{\circ}}{\partial T}\right) \quad (3)$$

$$\Delta S_{\text{solv}, 298.15 \text{ K}}^{\circ} = \Delta C_p \ln\left(\frac{298.15 \text{ K}}{385 \text{ K}}\right) \quad (4)$$

The identical  $\Delta C_p$  values ( $0.902 \text{ kJ K}^{-1} \pm 0.018$  and  $0.907 \text{ kJ K}^{-1} \pm 0.016$ , respectively) obtained for the interaction of FimH<sub>LD</sub> with 1-deoxy-septanose 7 and mannoside 2 are a clear indication that both complex formations release comparable numbers of water molecules leading to a comparable gain of desolvation entropy  $\Delta S_{\text{solv}}^{\circ}$ . Consequently, the entropic penalty for 7 originates from a loss of conformational freedom of either the ligand and/or the protein upon binding (Table 2).

To comprehensively interpret the thermodynamics of binding, additional information regarding solution and bound conformations of mannoside 2 and 1-deoxy-septanose 7 are indispensable. The conformational flexibility of septanoses *vs.* pyranoses as well as the *n*-heptyl aglycone deserves special attention in this regard. Furthermore, the degree of conformational flexibility of FimH<sub>LD</sub> has to be addressed. Finally, structural analysis by X-ray crystallography to elucidate the hydrogen

bond network established by the two ligands is essential for a final assessment of the two binding processes.

### *In silico* comparison of the solution conformation of mannoside 2 and 1-deoxy-septanose 7

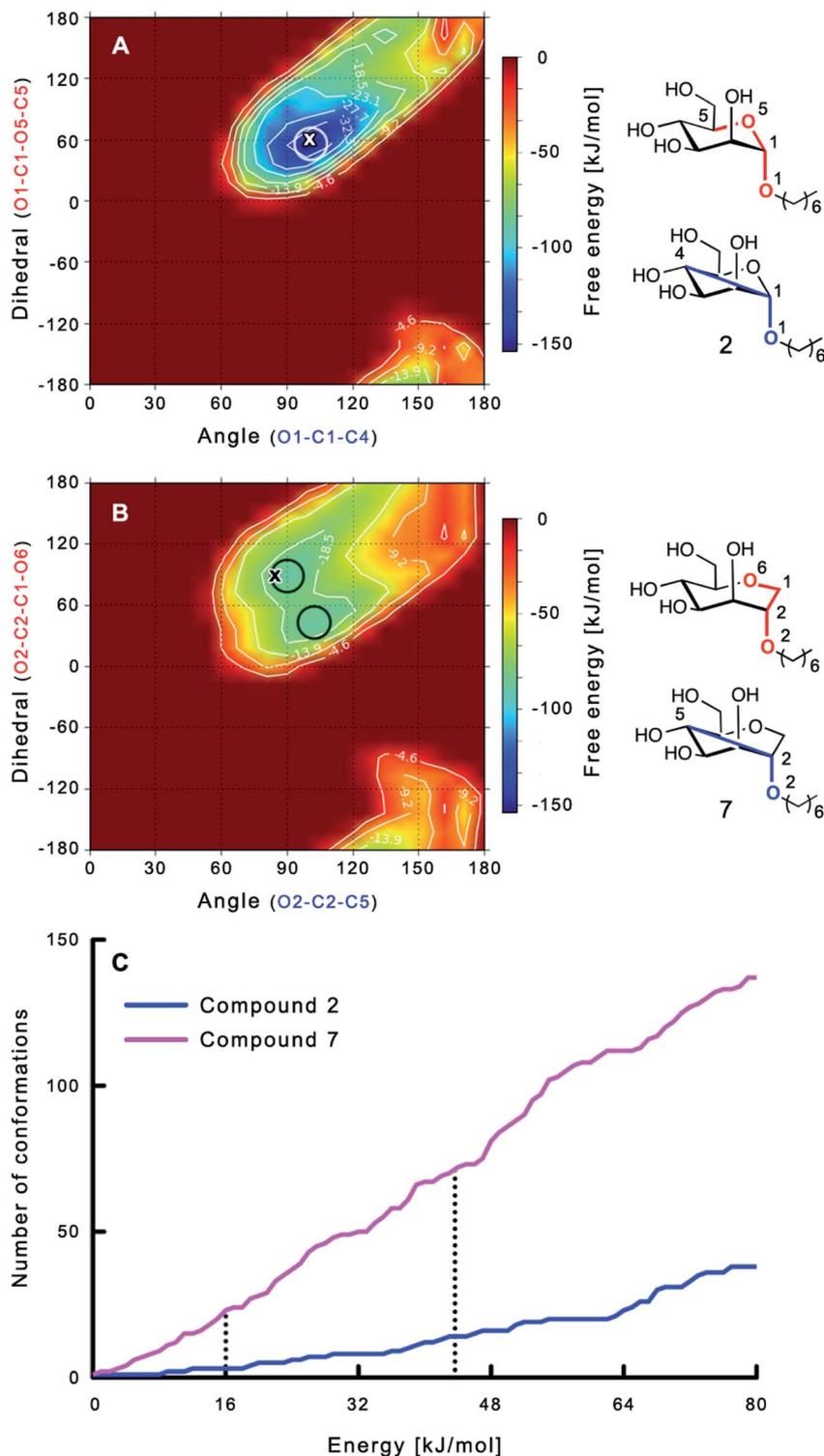
The large conformational entropy penalty for 1-deoxy-septanose 7 in comparison to mannoside 2 could be related to the extent of flexibility of their respective solution conformations. Therefore, the dynamic behavior in solution of each was studied by monitoring two collected variables using metadynamics simulations. The collected variables were defined as the angle between O1–C1–C4 and the dihedral torsion of O1–C1–O5–C5 for mannoside 2 and the corresponding angle O2–C2–C5 and dihedral torsion O2–C2–C1–O6 for compound 7 (for numbering see Fig. 3). Metadynamics simulations revealed an energy landscape for mannoside 2 with one distinct minimum of  $-154.8 \text{ kJ mol}^{-1}$  of  $99^\circ$  for the angle O1–C1–C4 and  $54^\circ$  for the dihedral angle (indicated by a circle in Fig. 3A), whereas the energy surface for 1-deoxy-septanose 7 was much shallower, showing two minima of  $-90.4 \text{ kJ mol}^{-1}$  at  $99^\circ$  and  $36^\circ$ , and  $-89.5 \text{ kJ mol}^{-1}$  at  $90^\circ$  and  $90^\circ$ , respectively (indicated by circles in Fig. 3B). Thus, the metadynamic simulations clearly indicate that 1-deoxy-septanose 7 exhibits a much larger ring flexibility in comparison to mannoside 2. In addition, the minima observed in the metadynamics simulations coincide closely with the bound conformation of the crystal structure. However, compound 7 shows an additional minimum that does not overlap with the bound conformation and therefore adds to the entropic loss.

To further quantify the flexibility issue, all ring conformations were normalized to their respective global minimum. Cumulative frequency analysis revealed that compound 7 exhibits 23 accessible conformations within  $16 \text{ kJ mol}^{-1}$  corresponding to the energy range of a hydrogen bond, whereas compound 2 has only access to 3 conformations. When a  $44 \text{ kJ mol}^{-1}$  energy range is taken into consideration (equal to the highest barrier in a cyclohexane ring-flip), compound 7 has access to 72 conformations in contrast to the rather rigid mannoside 2 with only 14 accessible conformations (Fig. 3C). In summary, the much higher number of conformations available to septanose derivative 7 results in a higher loss of conformational flexibility upon complexation to FimH and accounts for the difference in entropies of binding to FimH<sub>LD</sub> for 7 compared to 2.

### *Ab initio* calculations involving interactions of the *n*-heptyl group

The preference of 1-deoxy-septanose 7 to position its *n*-heptyl tail in an axial orientation was demonstrated qualitatively by solution NMR experiments in the absence of protein. Karplus analysis of coupling constants  $^3J_{2,3}$  and  $^3J_{3,4}$  (septanose numbering) of 7 obtained from a 1D <sup>1</sup>H NMR spectrum at 900 MHz indicated torsion angles in agreement with those expected for the axial conformation (see ESI† for details). To assess whether the tyrosine gate affects the *n*-heptyl tail in 2 and 7 identically, non-covalent interaction energies were calculated based on the available crystal structures (*cf.* Fig. 5) with





**Fig. 3** (A, B) Energy surface diagrams of the solution conformation of 2 (A) and 7 (B) obtained from metadynamics analysis. Circles represent the distinct minima observed in metadynamics simulations and crosses indicate the bound conformations in the crystal structures [PDB code 4BUQ (A) & 5CGB (B)]. Energies are given in  $\text{kJ mol}^{-1}$  and are color-coded from bordeaux-red ( $0 \text{ kJ mol}^{-1}$ ) to dark-blue ( $-150 \text{ kJ mol}^{-1}$ ). Vertical axes describe the observed dihedral angle (red), horizontal axes the observed angle (blue), both indicated in the structures to the right. (C) Cumulative frequency distribution analysis of all conformations (within  $80 \text{ kJ mol}^{-1}$  from the global minimum). Dashed lines indicate examples that occur at similar energies as denoted in the text.



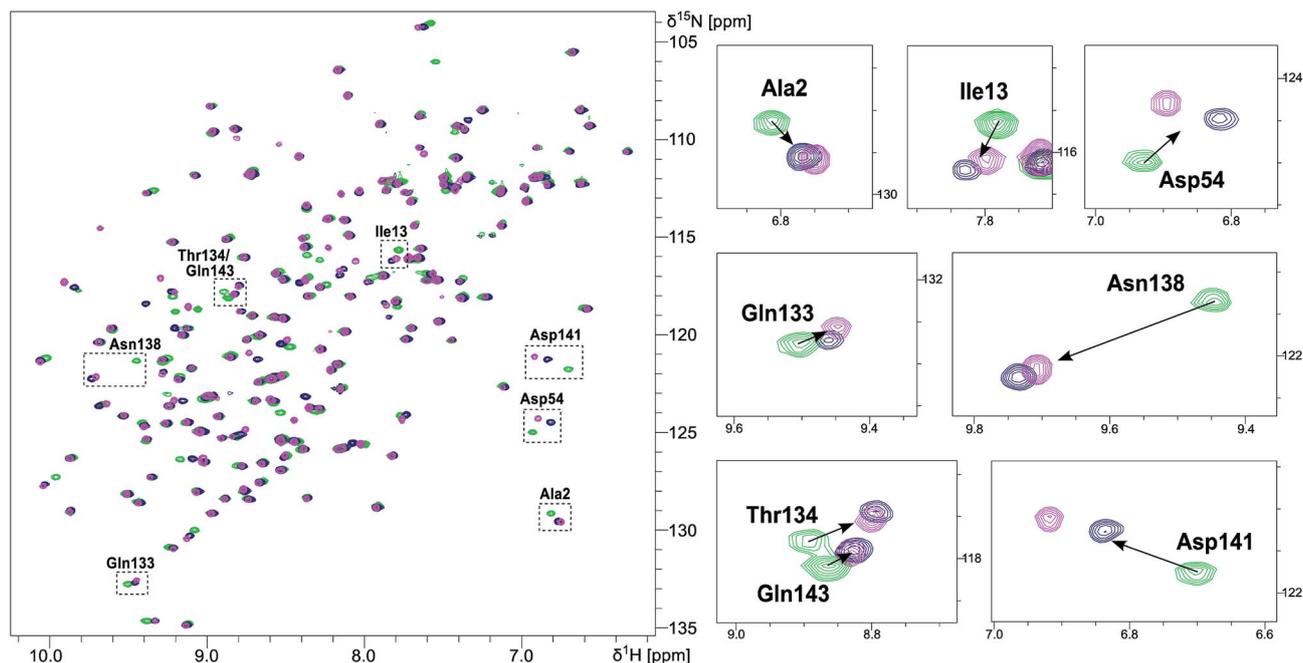


Fig. 4 NMR chemical shift perturbation experiments of FimH<sub>LD</sub> with mannoside 2 and 1-deoxy-septanose 7. <sup>1</sup>H,<sup>15</sup>N-HSQC spectra of FimH<sub>LD</sub> in the absence of antagonist (green) and in the presence of an excess of 2 (blue) and 7 (magenta). Important residues in the binding sites are shown in detail on the panel at the right.

Jaguar.<sup>55–57</sup> For the proper quantification of the non-covalent interactions, we performed *ab initio* DFT calculations using the B3LYP-MM functional in combination with the cc-pVDZ++ basis-set in the gas phase. The calculations led to comparable interaction energies of  $-18.5 \text{ kJ mol}^{-1}$  and  $-19.2 \text{ kJ mol}^{-1}$  for the *n*-heptyl tails of 2 and 7 with the tyrosine-gate. We therefore assume that their contributions to enthalpy as well as entropy are of comparable size.

#### Conformation of FimH<sub>LD</sub> upon binding of ligand 2 and 7 assessed by NMR CSP experiments

NMR chemical shift perturbation (CSP) experiments of the FimH<sub>LD</sub> backbone amide resonances were used to assess the

conformation of the protein backbone.<sup>58</sup> Thus, <sup>1</sup>H,<sup>15</sup>N-HSQC experiments of uniformly <sup>15</sup>N-labeled FimH<sub>LD</sub> (see ESI†) were conducted in the absence and presence of an excess of each ligand. The resonance assignment of FimH<sub>LD</sub> was available from previous studies.<sup>58,59</sup> For both antagonists 2 and 7, the set of protein residues of the mannose binding pocket (*e.g.* Ala2, Asp54 and Gln133) and a nearby loop (*e.g.* Thr134, Asn138, Asp141, Gln143) exhibited highly similar CSP effects (Fig. 4; see also Fig. S5 in ESI†). The absence of line broadening effects demonstrates that the protein is rather rigid in complex with both antagonists. Furthermore, the almost identical signals for amino acids remote from the mannose-binding pocket indicate that FimH<sub>LD</sub> does not undergo global conformational adaptations

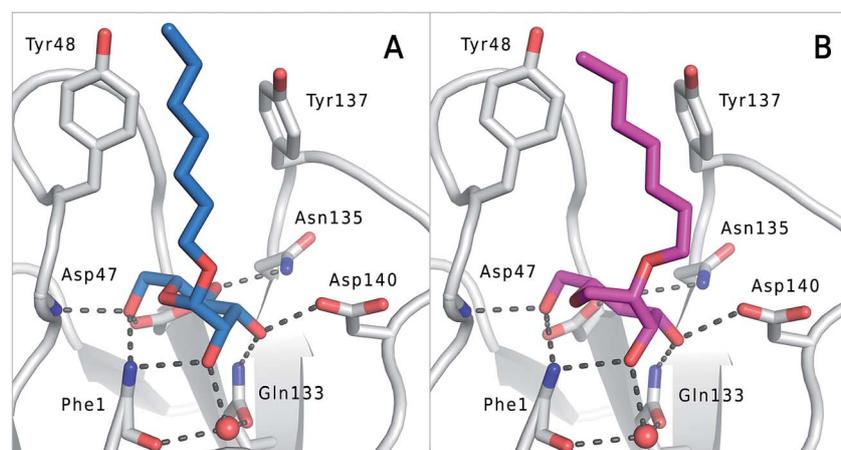


Fig. 5 (A) Co-crystal structure of FimH<sub>LD</sub> with mannoside 2 (PDB code 4BUQ) and (B) co-crystal structure of FimH<sub>LD</sub> with 1-deoxy-septanose 7 (PDB code 5CGB); the intermolecular hydrogen bond network is depicted by dashed lines.



upon ligand binding. Thus, we can assume that the protein conformation upon binding mannoside 2 and 1-deoxy-septanose 7 is highly similar, consistent with the comparable thermodynamic outcomes of the interactions.

### Crystal structure of FimH<sub>LD</sub> in complex with mannoside 2 and 1-deoxy-septanose 7

The previously reported co-crystal structure of mannoside 2–FimH<sub>LD</sub><sup>60</sup> (PDB code 4BUQ, Fig. 5A) and the new high-resolution X-ray structure of FimH<sub>LD</sub> co-crystallized with 7 (PDB code 5CGB; Fig. 5B and Table S3 in ESI†) exhibit a surprisingly high structural similarity with a backbone RMSD of 0.2 Å. Noteworthy, the well-defined, tight FimH mannoside-binding pocket installs identical hydrogen bond networks with mannoside 2 and 1-deoxy-septanose 7 (Fig. 5). For both ligands, nine hydrogen bonds with the side-chains of the residues Asn46, Asp54, Gln133, Asn135 and Asp140, and to the backbone of Phe1 and Asp47 are formed (for statistics on diffraction data and structure refinement of ligand 7–FimH<sub>LD</sub> complex see ESI†).

The dynamic stability of the hydrogen bond networks was assessed by MD simulations starting from the conformations of compound 2 and 7 from the crystal structures in Fig. 5. The simulations were run for the duration of 4.8 ns, followed by an analysis of the hydrogen bond network, where the occupancy of a given hydrogen bond interaction was monitored. The crucial hydrogen bonds with Phe1, Asp54 and Gln133 were retained with comparable occupancies by both compounds 2 and 7 (see Table S3 in ESI†). Therefore, in accordance with ITC experiments, enthalpic contributions from ligand–protein interactions can be predicted to be of similar size.

## Discussion

The bacterial lectin FimH mediates the adhesion of uropathogenic *E. coli* to the bladder epithelium of its host. FimH<sub>LD</sub> consists of a mannoside-binding pocket, which is reached through a hydrophobic cleft equipped with two tyrosines and an isoleucine (tyrosine gate). Its high affinity interaction with mannoside 2 is achieved by an extended hydrogen bond network established between the mannoside moiety and the carbohydrate recognition domain of the lectin and by hydrophobic  $\sigma$ – $\pi$  interactions of the *n*-heptyl aglycone with the tyrosine gate.<sup>38,58,61</sup> Replacement of the mannoside moiety by other hexoses, *e.g.* D-glucose or D-galactose, leads to a substantial reduction or even loss of binding affinity.<sup>38</sup>

Seven-membered ring analogs designed to mimic  $\alpha$ -D-mannopyranosides in general exhibit reduced affinities for FimH<sub>LD</sub>. From a series of septanosides and analogs, methyl septanoside 4 and 1-deoxy-septanose 7 were identified as the best representative binders. The *n*-heptyl substituent on their 2-position enables  $\sigma$ – $\pi$  stacking interaction with the tyrosine gate in analogy to mannoside 2.<sup>58,61</sup> Nevertheless, 4 (2.2  $\mu$ M) and 7 (260 nM) showed a 75- and 9-fold loss of inhibitory potency compared to mannoside 2 (29 nM).

To explain the large difference in affinity, a comparison of the thermodynamic fingerprints of their interactions with

FimH<sub>LD</sub> turned out to be instructive. The ITC profile of *n*-heptyl mannoside (2) served as reference. Binding of 2, like many other mannoside-based antagonists, is largely enthalpy-driven with a major contribution by the nine hydrogen bonds formed between the mannoside moiety and the lectin.<sup>29,35,38,58</sup> The severe drop in affinity of methyl septanoside 4 compared to mannoside 2 is predominantly due to a substantially smaller enthalpy term ( $\Delta\Delta H_{\text{obs}}^{\circ} = 22.4 \text{ kJ mol}^{-1}$ ), accompanied by a pronounced improvement of the entropy term ( $-T\Delta\Delta S_{\text{obs}}^{\circ} = 11.7 \text{ kJ mol}^{-1}$ ). One possible explanation for the severe enthalpy penalty is a steric clash induced by the methoxy group in the 1-position, disordering the essential hydrogen bond network. A less well-defined hydrogen bond network on the other hand would be consistent with an entropy gain of 4 in comparison to mannoside 2 because of the increased conformational flexibility. Interestingly, the almost identical enthalpy terms for mannoside 2 and 1-deoxy-septanose 7 ( $\Delta H_{\text{obs}}^{\circ} \sim -50 \text{ kJ mol}^{-1}$ ) suggest that FimH adopts very similar conformations enabling the formation of all critical hydrogen bonds to both ligands, a hypothesis supported by solution NMR and X-ray co-crystallography.

The origin for the reduced affinity of 7 compared to 2 is a significantly higher entropic penalty. ITC measurements at different temperatures revealed almost identical heat capacities  $\Delta C_p$  for both mannoside 2 and 1-deoxy-septanose 7, indicating comparable desolvation entropies ( $\Delta S_{\text{sol}}^{\circ}$ ). Since  $\Delta S_{\text{mix}}^{\circ}$  reflecting the loss in translational and rotational degrees of freedom upon complex formation can be assumed to be identical, the increased entropic penalty for 7 can be attributed solely to a reduced conformational freedom of the ligand and/or the protein. With NMR chemical shift perturbation experiments it was shown that 1-deoxy-septanose 7 exhibited comparable chemical shift changes and line widths for amino acid residues of the FimH binding pocket as mannoside 2. Thus, both ligands form comparable complexes with FimH<sub>LD</sub> and therefore exhibit comparable  $\Delta S_{\text{conf}}^{\circ}$  terms. Quantum-mechanical calculations revealed almost identical contributions of the aglycones of 2 and 7 to the interaction energy, suggesting that the loss of conformational entropy is solely due to the higher flexibility of the seven-membered ring of 7 compared to the six-membered ring of 2 in solution. Upon binding to FimH<sub>LD</sub>, a specific conformation of the septanose ring is required, leading to a significant rigidification of the compound relative to its unbound state. Metadynamics simulations, revealing a flat energy profile for 1-deoxy-septanose 7 with various ring conformations separated by relatively low energetic barriers, support this assumption.

The bound conformation of 7 can therefore be regarded as the result of a conformational selection of the ligand. Whereas conformational selection in terms of protein dynamics has been widely described and quantified by the use of NMR dynamics experiments of the protein,<sup>7,15–18</sup> a quantitative description of the impact of ligand flexibility on the thermodynamics of binding is still rare, likely due to the numerous contributing factors, which could be separated and analyzed in this specific example.



## Conclusions

ITC results, NMR CSP experiments, computational considerations, and X-ray crystallographic data support the conclusion that 2-*n*-heptyl 1-deoxy-septanose 7 adopts a conformation enabling the formation of a H-bond network with FimH<sub>LD</sub> identical to the corresponding mannoside 2. Furthermore, the loss of affinity of 7 is almost exclusively due to a loss in conformational entropy  $\Delta S_{\text{conf}}$ . *In silico*-analysis and solution NMR data ultimately showed that, although the main solution conformation of the seven-membered ring is similar to the bound conformation, the pronounced flexibility of the ring in solution causes significant entropic penalties upon complex formation.

In view of this result, the evaluation of glycomimetics should always include thermodynamic profiling because the quantification of enthalpic and entropic terms can be highly informative for the design of improved antagonists. In the case of FimH antagonists, 1-deoxy-septanose 7 provides an excellent mannoside mimic, *i.e.* can establish an identical H-bond network as the parent compound. However, it suffers from a substantial entropic penalty due to rigidification of the inherent ring flexibility upon complexation. This study excellently exemplifies that flexibility can be an important element in the design of small molecule ligands of target proteins.<sup>62,63</sup>

## Conflicts of interest

There are no conflicts to declare.

## Acknowledgements

We thank Prof. Stephan Grzesiek at the Biozentrum, University of Basel for the access to 600 and 900 MHz NMR spectrometers. The financial support by the Swiss National Science Foundation (CS & PZ: Swiss National Science Foundation (SNF) grant 200020\_129935; RP: SNF grant 200020\_146202; TM: SNF grant R'EQIP 145023) is gratefully acknowledged. B. F. thanks the German Academic Exchange Service (DAAD) for a stipendship. Finally, support for this work was provided to MWP by the National Science Foundation (CAREER CHE-0546311).

## Notes and references

- 1 S. Moghaddam, Y. Inoue and M. K. Gilson, *J. Am. Chem. Soc.*, 2009, **131**, 4012.
- 2 E. Fischer, *Ber. Dtsch. Chem. Ges.*, 1894, **27**, 2985.
- 3 F. W. Lichtenthaler, *Angew. Chem., Int. Ed.*, 1995, **33**, 2364.
- 4 R. U. Lemieux and U. Spohr, *Adv. Carbohydr. Chem. Biochem.*, 1994, **50**, 1.
- 5 D. E. Koshland, *Angew. Chem., Int. Ed.*, 1995, **33**, 2375.
- 6 P. W. Snyder, J. Mecinovic, D. T. Moustakas, S. W. Thomas III, M. Harder, E. T. Mack, M. R. Lockett, A. Heroux, W. Sherman and G. M. Whitesides, *Proc. Natl. Acad. Sci. U. S. A.*, 2011, **108**, 17889.
- 7 C. Diehl, O. Engstrom, T. Delaine, M. Hakansson, S. Genheden, K. Modig, H. Leffler, U. Ryde, U. J. Nilsson and M. Akke, *J. Am. Chem. Soc.*, 2010, **132**, 14577.
- 8 F. Marcelo, Y. He, S. A. Yuzwa, L. Nieto, J. Jiménez-Barbero, M. Sollogoub, D. J. Vocadlo, G. D. Davies and Y. Bleriot, *J. Am. Chem. Soc.*, 2009, **131**, 5390.
- 9 A. R. Patel, G. Ball, L. Hunter and F. Liu, *Org. Biomol. Chem.*, 2013, **11**, 3781.
- 10 D. D. Staas, K. L. Savage, V. L. Sherman, H. L. Shimp, T. A. Lyle, L. O. Tran, C. M. Wiscount, D. R. McMasters, P. E. Sanderson, P. D. Williams, B. J. Lucas Jr, J. A. Krueger, S. D. Lewis, R. B. White, S. Yu, B. K. Wong, C. J. Kochansky, M. R. Anari, Y. Yan and J. P. Vacca, *Bioorg. Med. Chem.*, 2006, **14**, 6900.
- 11 Y. Endo, S. Takehana, M. Ohno, P. E. Driedger, S. Stabel, M. Y. Mizutani, N. Tomioka, A. Itai and K. Shudo, *J. Med. Chem.*, 1998, **41**, 1476.
- 12 Y. Endo, M. Ohno, M. Hirano, A. Itai and K. Shudo, *J. Am. Chem. Soc.*, 1996, **118**, 1841.
- 13 K. Gunasekaran and R. Nussinov, *J. Mol. Biol.*, 2007, **365**, 257.
- 14 D. Rauh, G. Klebe and M. T. Stubbs, *J. Mol. Biol.*, 2004, **335**, 1325.
- 15 D. D. Boehr, R. Nussinov and P. E. Wright, *Nat. Chem. Biol.*, 2009, **5**, 789.
- 16 K. Okazaki and S. Takada, *Proc. Natl. Acad. Sci. U. S. A.*, 2008, **105**, 11182.
- 17 I. V. Nesmelova, E. Ermakova, V. A. Daragan, M. Pang, M. Menendez, L. Lagartera, D. Solis, L. G. Baum and K. H. Mayo, *J. Mol. Biol.*, 2010, **397**, 1209.
- 18 T. R. Weikl and F. Paul, *Protein Sci.*, 2014, **23**, 1508.
- 19 C. H. Reynolds and M. K. Holloway, *ACS Med. Chem. Lett.*, 2011, **2**, 433.
- 20 X. R. Wu, T. T. Sun and J. J. Medina, *Proc. Natl. Acad. Sci. U. S. A.*, 1996, **93**, 9630.
- 21 G. Zhou, W. J. Mo, P. Sebbel, G. Min, T. A. Neubert, R. Glockshuber, X. R. Wu, T. T. Sun and X. P. Kong, *J. Cell Sci.*, 2001, **114**, 4095.
- 22 B. Xie, G. Zhou, S. Y. Chan, E. Shapiro, X. P. Kong, X. R. Wu, T. T. Sun and C. E. Costello, *J. Biol. Chem.*, 2006, **281**, 14644.
- 23 A. Ronald, *Am. J. Med.*, 2002, **113**, 14.
- 24 D. Abgottspon, G. Rölli, L. Hosch, A. Steinhuber, X. Jiang, O. Schwardt, B. Cutting, M. Smieško, U. Jenal, B. Ernst and A. Trampuz, *J. Microbiol. Methods*, 2010, **82**, 249.
- 25 M. Scharenberg, D. Abgottspon, E. Cicek, X. Jiang, O. Schwardt, S. Rabbani and B. Ernst, *Assay Drug Dev. Technol.*, 2011, **9**, 455.
- 26 T. Klein, D. Abgottspon, M. Wittwer, S. Rabbani, J. Herold, X. Jiang, S. Kleeb, C. Luthi, M. Scharenberg, J. Bezençon, E. Gubler, L. Pang, M. Smieško, B. Cutting, O. Schwardt and B. Ernst, *J. Med. Chem.*, 2010, **53**, 8627.
- 27 O. Schwardt, S. Rabbani, M. Hartmann, D. Abgottspon, M. Wittwer, S. Kleeb, A. Zalewski, M. Smieško, B. Cutting and B. Ernst, *Bioorg. Med. Chem.*, 2011, **19**, 6454.
- 28 X. Jiang, D. Abgottspon, S. Kleeb, S. Rabbani, M. Scharenberg, M. Wittwer, M. Haug, O. Schwardt and B. Ernst, *J. Med. Chem.*, 2012, **55**, 4700.



- 29 L. Pang, S. Kleeb, K. Lemme, S. Rabbani, M. Scharenberg, A. Zalewski, F. Schadler, O. Schwardt and B. Ernst, *ChemMedChem*, 2012, 7, 1404.
- 30 M. Aronson, O. Medalia, L. Schori, D. Mirelman, N. Sharon and I. Ofek, *J. Infect. Dis.*, 1979, 139, 329.
- 31 C. K. Cusumano, J. S. Pinkner, Z. Han, S. E. Greene, B. A. Ford, J. R. Crowley, J. P. Henderson, J. W. Janetka and S. J. Hultgren, *Sci. Transl. Med.*, 2011, 3, 109ra115.
- 32 D. Abgottspon and B. Ernst, *Chimia*, 2012, 66, 166.
- 33 M. Totsika, M. Kostakioti, T. J. Hannan, M. Upton, S. A. Beatson, J. W. Janetka, S. J. Hultgren and M. A. Schembri, *J. Infect. Dis.*, 2013, 208, 921.
- 34 J. Bouckaert, Z. Li, C. Xavier, M. Almant, V. Caveliers, T. Lahoutte, S. D. Weeks, J. Kovensky and S. G. Gouin, *Chem.–Eur. J.*, 2013, 19, 7847.
- 35 S. Kleeb, L. Pang, K. Mayer, D. Eriş, A. Sigl, R. C. Preston, P. Zihlmann, T. Sharpe, R. P. Jakob, D. Abgottspon, A. S. Hutter, M. Scharenberg, X. Jiang, G. Navarra, S. Rabbani, M. Smieško, N. Lüdin, J. Bezençon, O. Schwardt, T. Maier and B. Ernst, *J. Med. Chem.*, 2015, 58, 2221.
- 36 S. Rabbani, X. Jiang, O. Schwardt and B. Ernst, *Anal. Biochem.*, 2010, 407, 188.
- 37 Z. Han, J. S. Pinkner, B. Ford, R. Obermann, W. Nolan, S. A. Wildman, D. Hobbs, T. Ellenberger, C. K. Cusumano, S. J. Hultgren and J. W. Janetka, *J. Med. Chem.*, 2010, 53, 4779.
- 38 J. Bouckaert, J. Berglund, M. Schembri, E. De Genst, L. Cools, M. Wuhrer, C.-S. Hung, J. Pinkner, R. Slättegård, A. Zavialov, D. Choudhury, S. Langermann, S. J. Hultgren, L. Wyns, P. Klemm, S. Oscarson, S. D. Knight and H. De Greve, *Mol. Microbiol.*, 2005, 55, 441.
- 39 J. Saha and M. W. Pecuh, *Adv. Carbohydr. Chem. Biochem.*, 2011, 66, 121.
- 40 M. R. Duff Jr, W. S. Fyvie, S. D. Markad, A. E. Frankel, C. V. Kumar, J. A. Gascón and M. W. Pecuh, *Org. Biomol. Chem.*, 2011, 9, 154.
- 41 S. Castro, M. Duff, N. L. Snyder, M. Morton, C. V. Kumar and M. W. Pecuh, *Org. Biomol. Chem.*, 2005, 3, 3869.
- 42 D. Gupta, T. K. Dam, S. Oscarson and C. F. Brewer, *J. Biol. Chem.*, 1997, 272, 6388.
- 43 M. C. Chervenak and E. J. Toone, *Biochemistry*, 1995, 34, 5685.
- 44 F. P. Schwarz, K. D. Puri, R. G. Bhat and A. Surolia, *J. Biol. Chem.*, 1993, 268, 7668.
- 45 D. Eriş, R. C. Preston, M. Scharenberg, F. Hulliger, D. Abgottspon, L. Pang, X. Jiang, O. Schwardt and B. Ernst, *ChemBioChem*, 2016, 17, 1012.
- 46 M. M. Sauer, R. P. Jakob, J. Eras, S. Baday, D. Eriş, G. Navarra, S. Bernèche, B. Ernst, T. Maier and R. Glockshuber, *Nat. Commun.*, 2016, 7, 10738.
- 47 S. D. Markad, S. Xia, N. L. Snyder, B. Surana, M. D. Morton, C. M. Hadad and M. W. Pecuh, *J. Org. Chem.*, 2008, 73, 6341.
- 48 M. P. DeMatteo, N. L. Snyder, M. Morton, D. M. Baldisseri, C. M. Hadad and M. W. Pecuh, *J. Org. Chem.*, 2005, 70, 24.
- 49 R. Murali and M. Nagarajan, *Carbohydr. Res.*, 1996, 280, 351.
- 50 P. Ertl, B. Rohde and P. Selzer, *J. Med. Chem.*, 2000, 43, 3714.
- 51 B. M. Baker and K. P. Murphy, *J. Mol. Biol.*, 1997, 268, 557.
- 52 K. P. Murphy, D. Xie, K. S. Thompson, L. M. Amzel and E. Freire, *Proteins*, 1994, 18, 63.
- 53 K. P. Murphy, P. L. Privalov and S. J. Gill, *Science*, 1990, 247, 559.
- 54 R. L. Baldwin, *Proc. Natl. Acad. Sci. U. S. A.*, 1986, 83, 8069.
- 55 *Schrödinger Suite 2015-4, Jaguar, Version 9.0*, Schrödinger, LLC, New York, NY, 2015.
- 56 S. T. Schneebeli, A. D. Bochevarov and R. A. Friesner, *J. Chem. Theory Comput.*, 2011, 7, 658.
- 57 A. D. Bochevarov, E. Harder, T. F. Hughes, J. R. Greenwood, D. A. Braden, D. M. Philipp, D. Rinaldo, M. D. Halls, J. Zhang and R. A. Friesner, *Int. J. Quantum Chem.*, 2013, 113, 2110.
- 58 B. Fiege, S. Rabbani, R. C. Preston, R. P. Jakob, P. Zihlmann, O. Schwardt, X. Jiang, T. Maier and B. Ernst, *ChemBioChem*, 2015, 16, 1235.
- 59 S. Vanwetswinkel, A. N. Volkov, Y. G. J. Sterckx, A. Garcia-Pino, L. Buts, W. F. Vranken, J. Bouckaert, R. Roy, L. Wyns and N. A. J. van Nuland, *J. Med. Chem.*, 2014, 57, 1416.
- 60 G. Roos, A. Wellens, M. Touaibia, N. Yamakawa, P. Geerlings, R. Roy, L. Wyns and J. Bouckaert, *ACS Med. Chem. Lett.*, 2013, 4, 1085.
- 61 A. Wellens, M. Lahmann, M. Touaibia, J. Vaucher, S. Oscarson, R. Roy, H. Remaut and J. Bouckaert, *Biochemistry*, 2012, 51, 4790.
- 62 S. Chung, J. B. Parker, M. Bianchet, L. M. Amzel and J. T. Stivers, *Nat. Chem. Biol.*, 2009, 5, 407.
- 63 S. A. Moura-Tamames, M. J. Ramos and P. A. Fernandes, *J. Mol. Graphics Modell.*, 2009, 27, 908.





### Manuscript 3

## The Tyrosine Gate of the Bacterial Lectin FimH: A Conformational Analysis by NMR Spectroscopy and X-ray Crystallography

Brigitte Fiege,<sup>#</sup> Said Rabbani,<sup>#</sup> Roland C. Preston,<sup>#</sup> Roman P. Jakob,<sup>§</sup>  
Pascal Zihlmann,<sup>#</sup> Oliver Schwardt,<sup>#</sup> Xiaohua Jiang,<sup>#</sup>  
Timm Meier<sup>§</sup> and Beat Ernst<sup>#\*</sup>

<sup>#</sup> University of Basel, Institute of Molecular Pharmacy,  
Klingelbergstr. 50, 4056 Basel, Switzerland

<sup>§</sup> University of Basel, Institute of Structural Biology,  
Klingelbergstr. 70, 4056 Basel, Switzerland

\*Corresponding author.

Tel.: 0041 (0)61 267 15 51; Fax: 0041 (0)61 267 15 52;  
E-mail: beat.ernst@unibas.ch

Contribution of Pascal Zihlmann:

- Thermodynamic profiling by ITC and data analysis



VIP

# The Tyrosine Gate of the Bacterial Lectin FimH: A Conformational Analysis by NMR Spectroscopy and X-ray Crystallography

Brigitte Fiege,<sup>[a]</sup> Said Rabbani,<sup>[a]</sup> Roland C. Preston,<sup>[a]</sup> Roman P. Jakob,<sup>[b]</sup> Pascal Zihlmann,<sup>[a]</sup> Oliver Schwardt,<sup>[a]</sup> Xiaohua Jiang,<sup>[a]</sup> Timm Maier,<sup>\*,[b]</sup> and Beat Ernst<sup>\*,[a]</sup>

Urinary tract infections caused by uropathogenic *E. coli* are among the most prevalent infectious diseases. The mannose-specific lectin FimH mediates the adhesion of the bacteria to the urothelium, thus enabling host cell invasion and recurrent infections. An attractive alternative to antibiotic treatment is the development of FimH antagonists that mimic the physiological ligand. A large variety of candidate drugs have been developed and characterized by means of in vitro studies and animal models. Here we present the X-ray co-crystal structures of FimH with members of four antagonist classes. In three of these cases no structural data had previously been available.

We used NMR spectroscopy to characterize FimH–antagonist interactions further by chemical shift perturbation. The analysis allowed a clear determination of the conformation of the tyrosine gate motif that is crucial for the interaction with aglycone moieties and was not obvious from X-ray structural data alone. Finally, ITC experiments provided insight into the thermodynamics of antagonist binding. In conjunction with the structural information from X-ray and NMR experiments the results provide a mechanism for the often-observed enthalpy–entropy compensation of FimH antagonists that plays a role in fine-tuning of the interaction.

## Introduction

Urinary tract infection (UTI) is one of the most frequent infectious diseases, affecting millions of people every year.<sup>[1]</sup> Women have a 50% risk of experiencing at least one symptomatic UTI during their lifetime. The large majority of UTIs are caused by uropathogenic *Escherichia coli* (UPEC) that are able to invade the urothelial cells in the bladder, form biofilms, and cause recurrent infections.<sup>[2]</sup> To date, UTIs are mainly treated with antibiotics, thus inducing antimicrobial resistance, a serious threat to patients worldwide.<sup>[3]</sup> Therefore, new treatment strategies are urgently needed.

Adherence of UPEC to the urothelial surface is mediated through the mannose-specific lectin FimH, located at the tip of bacterial type 1 pili.<sup>[4]</sup> FimH therefore represents a major virulence factor of UPEC. It consists of two immunoglobulin-like domains: the N-terminal lectin domain, or carbohydrate recognition domain (CRD), and, connected by a short linker, the C-terminal pilin domain. The pilin domain anchors the adhesin to the pilus and regulates the switch between the two conforma-

tional states—with either high or low affinity for mannoses—of the CRD.<sup>[5]</sup> Most in vitro interaction studies with FimH antagonists have been performed only with the lectin domain FimH-CRD that is trapped in the high-affinity state.<sup>[6]</sup>

The natural ligand for FimH is the mannosylated glycoprotein uroplakin Ia present on urothelial cells.<sup>[7]</sup> Mannose-based FimH antagonists compete with this interaction and prevent bacterial adhesion and hence infection. A substantial advantage of this anti-adhesion therapy over antibiotics treatment is the reduced risk of resistance development, because no direct selection pressure is imposed on the pathogen.<sup>[8]</sup> The first successful demonstration of the anti-adhesion strategy was of the protective effect of methyl  $\alpha$ -D-mannoside in a UTI mouse model.<sup>[9]</sup> Since then, significant progress has been made in the optimization of FimH antagonists, yielding improved affinities in the low nanomolar range.<sup>[10]</sup>

In part, these modifications are the result of rational drug design based on X-ray crystal structures of FimH bound to various  $\alpha$ -D-mannosides.<sup>[11]</sup> The mannose moiety is in each case tightly bound and involved in an extended hydrogen bond network. A promising prospect for optimizing binding is offered by a rim of hydrophobic residues lining the entrance to the mannose pocket. This so-called tyrosine gate, consisting of Tyr48 and Tyr137, with Ile52 positioned in-between, is involved in interactions with oligomannosides, as revealed by the X-ray co-crystal structure of FimH-CRD with oligomannose-3.<sup>[12]</sup> For drug design, the tyrosine gate has been exploited for hydrophobic stacking interactions with suitable aglycones. In particular, the finding that mannosides with hydrophobic aryl and

[a] Dr. B. Fiege, Dr. S. Rabbani, Dr. R. C. Preston, P. Zihlmann, Dr. O. Schwardt, Dr. X. Jiang, Prof. B. Ernst  
Institute of Molecular Pharmacy, University of Basel  
Klingelbergstrasse 50, 4056 Basel (Switzerland)  
E-mail: beat.ernst@unibas.ch

[b] Dr. R. P. Jakob, Prof. T. Maier  
Structural Biology, Biozentrum, University of Basel  
Klingelbergstrasse 70, 4056 Basel (Switzerland)  
E-mail: timm.maier@unibas.ch

Supporting information for this article is available on the WWW under <http://dx.doi.org/10.1002/cbic.201402714>.

alkyl aglycones show stronger affinities has led to many promising drug candidates. Through binding studies with a series of alkyl mannosides, *n*-heptyl  $\alpha$ -D-mannoside (**1**, Table 1, below) was identified as the most efficient non-aromatic FimH binder.<sup>[11a]</sup> For aryl and heteroaryl derivatives a large variety of antagonists differing in the number of aromatic rings and substituents and the type of linkers exist. Chemically easily accessible aromatic derivatives such as 4-nitrophenyl and 4-methylumbelliferyl mannosides were soon discovered.<sup>[11a,13]</sup> These were followed by derivatives of squaric acid,<sup>[14]</sup> biphenyls and other diaryls,<sup>[15]</sup> triazoles,<sup>[16]</sup> and indolines,<sup>[17]</sup> among others.

We routinely test new candidates for *in vitro* binding to FimH-CRD<sup>[15b,16–18]</sup> and to UPEC,<sup>[19]</sup> as well as for efficacy in a UTI mouse model.<sup>[20]</sup> A special focus is placed on the PK/PD (pharmacokinetic/pharmacodynamic) properties for oral bioavailability.<sup>[21]</sup> The ideal FimH antagonist for UTI treatment needs to be balanced between a reasonable solubility for effective dosage and a certain degree of lipophilicity for efficient membrane permeation during oral absorption. Despite the availability of structural information, rational design of FimH antagonists has not always led to the expected improvement in binding affinities. Specifically, thermodynamic profiles from isothermal titration calorimetry (ITC) have frequently revealed enthalpy–entropy compensations,<sup>[15d,21]</sup> the reasons for which are not yet fully understood.

In this study, we analyzed the binding of a representative set of FimH antagonists to FimH-CRD by means of a combination of high-resolution X-ray crystal structures and NMR chemical shift perturbation (CSP) experiments. NMR spectroscopy allows the study of protein–carbohydrate interactions in solution, with the native dynamic behavior of the protein thus being maintained.<sup>[22]</sup> In contrast, X-ray crystallography provides structural information from a “static” crystal with limited information on flexibility. The two methods are therefore highly complementary.

The potential of NMR spectroscopy to contribute to the drug design process has been widely acknowledged.<sup>[23]</sup> The chemical shifts of protein resonances are highly sensitive to the chemical environment, for example to hydrogen bonds and aromatic ring currents.<sup>[24]</sup> CSP effects hence identify residues in direct proximity to a bound ligand and, in addition, indicate conformational changes in allosteric sites. We used <sup>1</sup>H,<sup>15</sup>N HSQC experiments with <sup>15</sup>N-enriched FimH-CRD in the presence of various antagonists to monitor the CSP effects on the backbone amide groups. Previous NMR studies of binding of methyl  $\alpha$ -D-mannoside<sup>[6]</sup> and *n*-heptyl  $\alpha$ -D-mannoside<sup>[25]</sup> to FimH-CRD only noted a general match of the CSP maps and the ligand binding interface identified from X-ray co-crystal structures. With our analysis we demonstrate that NMR and X-ray structural data can be combined to characterize FimH–antagonist complex formation in unprecedented detail. In addition to the structural information, ITC experiments were performed to quantify the thermodynamics of antagonist binding to FimH-CRD.

## Results and Discussion

### FimH antagonist classes studied in this work

In this study we examined FimH antagonists composed of an  $\alpha$ -D-mannosyl residue linked to an aliphatic or aromatic aglycone (Table 1). *n*-Heptyl  $\alpha$ -D-mannoside (**1**)<sup>[11a]</sup> is the only non-aromatic antagonist with an affinity (low nanomolar range)

**Table 1.** FimH antagonists analyzed for binding to FimH-CRD by X-ray crystallography (PDB ID and reference for reported structures given) and/or CSP NMR experiments (indicated with +).

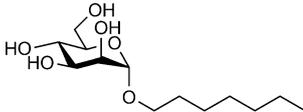
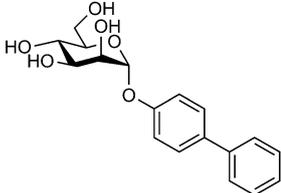
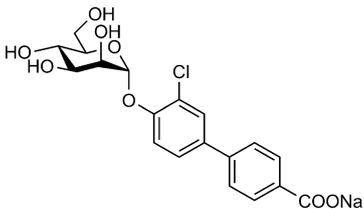
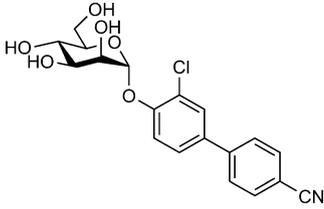
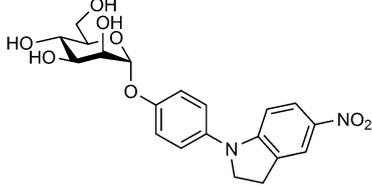
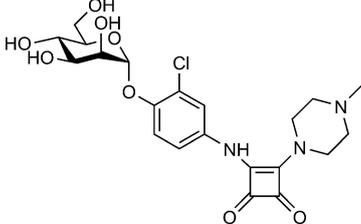
Compd	Structure	X-ray	NMR
1		4BUQ <sup>[27]</sup>	+
2		4X50	+
3		n.d.	+
4		4CST <sup>[21]</sup>	+
5		4X5Q	+
6		4X5R	+

Table 1. (Continued)			
Compd	Structure	X-ray	NMR
7		4X5P	+
8		n.d.	+
9		1UWF, 1TR7 <sup>[11a]</sup>	+
10		n.d.	+
11		n.d.	+

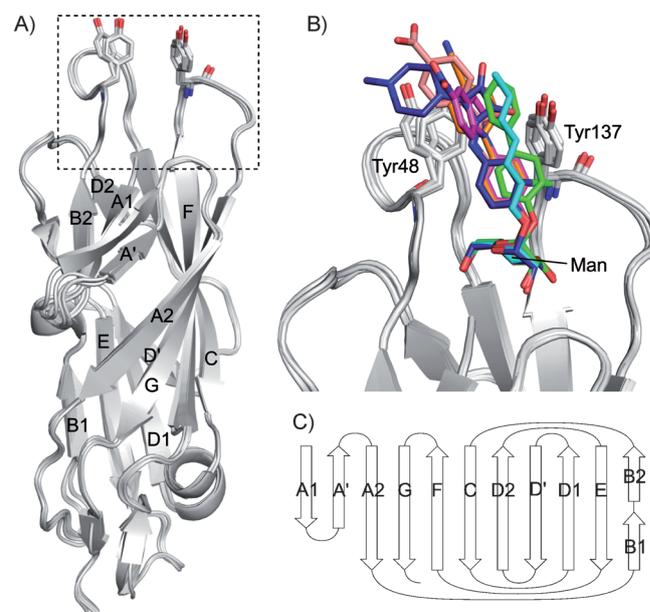
similar to those of the aromatic antagonists **2** to **7** (Table 1 and Table 2, below). Although its pharmacokinetic properties render **1** ineffective for oral administration, it is often used as a reference compound in screening studies. Biphenyl mannoside **2**<sup>[15a]</sup> was rationally designed to form aromatic stacking interactions with the side chains of Tyr48 and Tyr137 in the tyrosine gate. Later on, the aryl moieties were derivatized to optimize this interaction, as reported by our group and others.<sup>[15]</sup> An impressive example of this is represented by biphenyl mannoside **4**, with an *ortho*-chloro substituent on the inner aromatic ring and a cyano group in the *para* position of the outer ring as a bioorthogonal replacement for the carboxylate group present in antagonist **3**.<sup>[21]</sup> Both ring substituents of **4** reduce the electron density in the aglycone and thereby enhance the  $\pi$ - $\pi$  interaction, thus leading to a more than tenfold affinity improvement relative to the unsubstituted antagonist **2** (Table 2, below). Besides the biphenyls **2** to **4**, three other compound classes with multiple aromatic rings in nonplanar arrangements have been tested: **5** features a 5-nitroindolinyl moiety *N*-linked to an inner phenyl ring,<sup>[17]</sup> **6** extends an *ortho*-chlorophenyl system with squaric acid and *N*-methyl-piperazine,<sup>[14b]</sup> and in **7** (for synthesis see the Supporting Information) an amide bond is inserted between two phenyl rings with *ortho*-chlorine and *para*-carboxylate substituents.

In addition to high-affinity antagonists, we also studied methyl  $\alpha$ -D-mannoside (**8**), *n*-butyl  $\alpha$ -D-mannoside (**9**), and 1,5-anhydromannitol (**10**). Finally, the 4-deoxy-4-fluoro derivate **11** (for synthesis see the Supporting Information) exhibits a dramatically reduced affinity relative to **1**, emphasizing the impor-

ance of the hydrogen bond network in the mannoside binding pocket.

### X-ray crystal structures of FimH with antagonists

We obtained high-resolution X-ray structures of FimH-CRD co-crystallized with antagonists **2** and **5–7**, with resolutions ranging from 1.00 to 2.00 Å (Table 1 and Table S1 in the Supporting Information), and compared them with the reported structures with antagonists **1** and **4**.<sup>[21]</sup> FimH-CRD consists of eleven mainly antiparallel  $\beta$ -strands and two short helices (Figure 1).



**Figure 1.** Superposition of X-ray co-crystal structures of FimH-CRD with six antagonists: A) only protein is shown, with Tyr48 and Tyr137 as gray sticks, and B) with antagonists shown as colored sticks (1 cyan, 2 green, 4 orange, 5 magenta, 6 blue, and 7 salmon). C)  $\beta$ -sheet topology diagram of FimH-CRD.<sup>[11a,29]</sup>

Superposition of the six protein structures reveals a high structural similarity with a backbone RMSD of around 0.5 Å between any two structures. In all six structures the ligand is bound within the mannoside binding pocket. The mannosyl moieties form hydrogen bonds to the side chains of residues Asp54, Gln133, Asn135, and Asp140 and to the backbone atoms of Phe1 and Asp47 (Figure S1). Tyr48 is involved in stacking interactions with the outer aromatic ring of the aglycones or, in the case of **1**, with the alkyl chain of the ligand (Figure 1). Tyr137 stabilizes the loop containing Tyr48 through a hydrophobic interaction with Ile52. In the cases of **6** and **7**, the side chain hydroxy group of Tyr137 can form a hydrogen bond with the oxygen atoms of the squaric acid and the amide bond, respectively. Moreover, the backbone amide groups of Tyr137, Asn138, and Asp140 and the side chains of Asn135, Asn138, and Asp140 form a complex hydrogen bond network within the loop and with the 3- and 4-positions of the mannosyl residues of the ligands.

It is noteworthy that in all six structures the side chain of tyrosine gate residue Tyr137 has a single orientation whereas Tyr48 shows two slightly different orientations. In all cases, the side chain positions correspond to the *closed conformation* of the tyrosine gate as described previously.<sup>[11a,15a,d,26]</sup> The *closed conformation* was also observed in co-crystal structures of FimH-CRD with *n*-butyl  $\alpha$ -D-mannoside (**9**) (PDB ID: 1UWF)<sup>[11a]</sup> or with a biphenyl mannoside with a methylcarboxylate group in the *meta*-position of the outer aromatic ring (PDB ID: 3MCY).<sup>[15a]</sup>

A significantly different orientation of the tyrosine gate, termed the *open conformation*, was observed in the crystal structure of the full-length FimH in complex with  $\alpha$ -D-mannose stabilized by the chaperone FimC (PDB ID: 1KLF),<sup>[11b]</sup> and of FimH-CRD with ethane-1,2-diol in the mannoside binding site (PDB ID: 4AUU).<sup>[26]</sup> The latter is considered a "pseudo-apo" structure, although the bound ethane-1,2-diol could still have an influence on the binding site. Whereas the Tyr137 side chain in these structures remains in the orientation seen in the *closed conformation*, the Tyr48 side chain is rotated towards the mannoside pocket by about 3.5 Å. This *open conformation* of FimH-CRD was also observed in the co-crystal structure with the native ligand epitope oligomannose-3, Man $\alpha$ 1,3(Man $\alpha$ 1,6)-Man $\beta$ 1,4GlcNAc $\beta$ 1,4GlcNAc (PDB ID: 2VCO).<sup>[12]</sup> Here, the  $\alpha$ 1,3-linked mannoside at the nonreducing end is recognized in the orientation typical for all mannosides. This brings the first GlcNAc residue into a position in which it would clash with Tyr48 in the *closed conformation*. In view of the limited conformational freedom around the glycosidic linkages, it can be speculated that oligomannose-3 might only be able to bind to FimH in the *open conformation*.

In computational studies, FimH antagonists with flexible aglycones have been docked to FimH in both conformations of the tyrosine gate.<sup>[11a,15d]</sup> As an experimental verification, *n*-butyl  $\alpha$ -D-mannoside (**9**)<sup>[11a]</sup> and several monoaryl antagonists<sup>[27]</sup> have been co-crystallized with FimH-CRD in both the open and the closed forms. Wellens et al. proposed that the *open conformation* represents the minimum-energy conformer of FimH and that the *closed conformation* is only stabilized by favorable interactions with hydrophobic aglycones.<sup>[26]</sup> This would correspond to a conformational selection process in which the ligand binds to a conformer from a preexisting set of substates.<sup>[28]</sup> Nevertheless, an induced-fit mechanism also seems possible.

To conclude, Tyr48 can be regarded as a key element in ligand interaction, and knowledge of the Tyr48 side chain conformation is critical to the discussion. Unfortunately, in many FimH-mannoside co-crystal structures the binding pocket residues or the ligand are involved in crystal lattice contacts potentially affecting the binding pocket geometry.<sup>[26–27]</sup> Therefore, more experiments are needed to study the structure and dynamics of FimH-antagonist complex formation in solution.

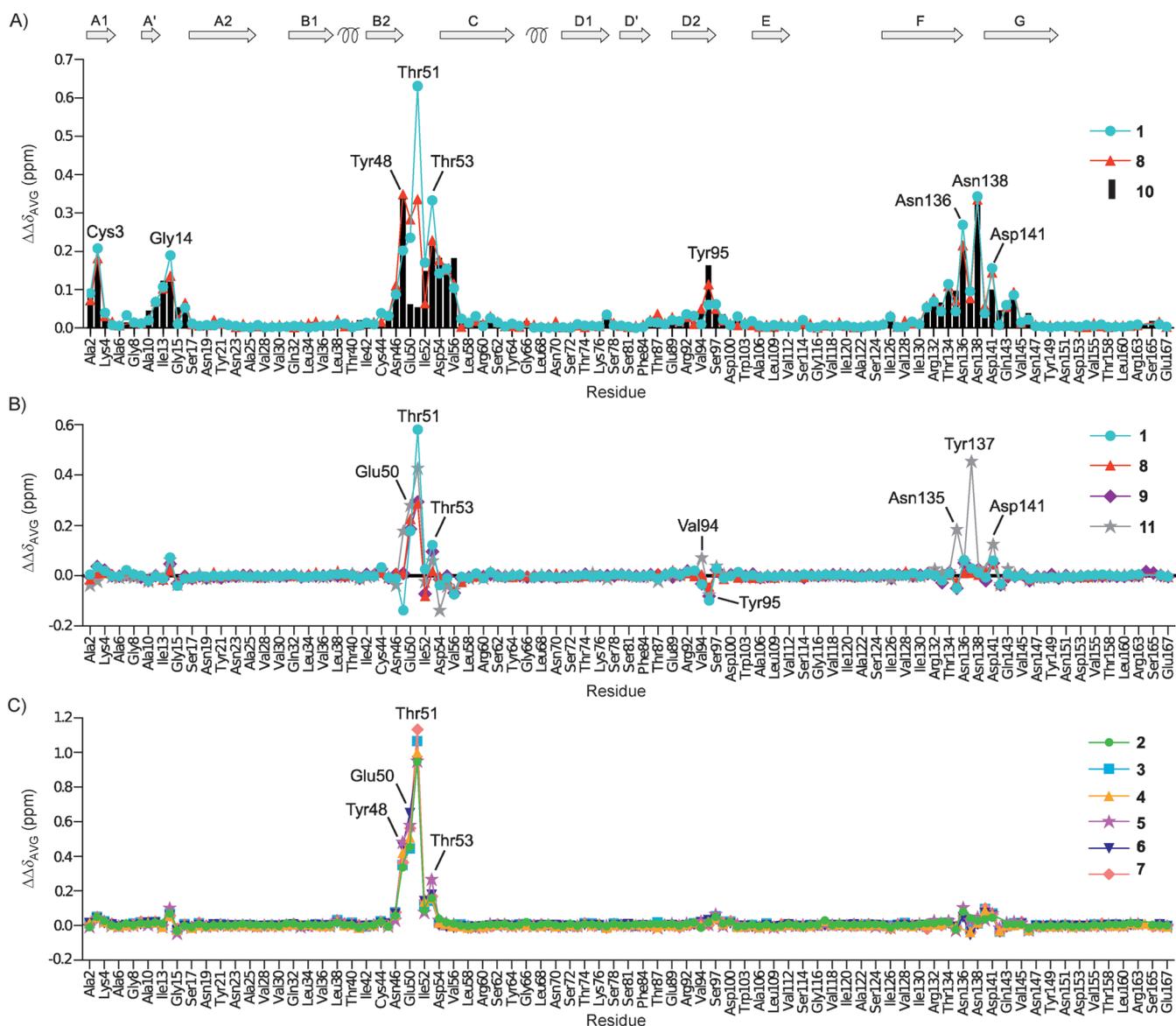
### NMR experiments with FimH-CRD

We performed NMR experiments to obtain structural information on FimH-antagonist complexes in solution. This allows full

retention of the molecule dynamics and avoids potential distortion from crystal packing. The binding of antagonists to <sup>15</sup>N-labeled FimH-CRD (173 residues) was monitored by CSP analysis of <sup>1</sup>H,<sup>15</sup>N HSQC fingerprint spectra representing the backbone amide groups. For efficient affinity purification, the protein contained a C-terminal His<sub>6</sub>-tag, which does not influence mannoside binding.<sup>[18]</sup> A complete backbone resonance assignment was performed by measuring a set of triple-resonance NMR spectra with a <sup>13</sup>C,<sup>15</sup>N-labeled FimH-CRD sample. The high chemical shift dispersion of the amide resonances allowed sequential backbone assignment mainly on the basis of HNCACB and CBCACONH spectra.<sup>[30]</sup> A total of 152 (94.4%) out of 161 assignable residues were assigned.<sup>[31]</sup> The His<sub>6</sub>-tag remained unassigned. Three residues were absent in the <sup>1</sup>H,<sup>15</sup>N HSQC spectrum: the backbone and side chain signals of Asn96 and Arg98, which are located in a loop close to the binding pocket, as well as Gly79, which is located in a solvent-exposed loop about 20 Å away from the binding pocket. Carbon and proton signals of the side chains of these three residues were observed as (*i*−1) correlations of the succeeding residues, thus suggesting that fast exchange with the solvent and not (solely) conformational exchange was responsible for the absence of signals. The assignment matches that from a recently published study of FimH-CRD (BMRB entry: 19066)<sup>[25]</sup> that lacked a His<sub>6</sub>-tag. Those authors were able to observe and assign Gly79, Asn96, and Arg98 amide signals. This is likely due to different measuring conditions, in particular a lower pH of 6.0,<sup>[25]</sup> in comparison with pH 6.8 (for assignment) and pH 7.0 (for CSP experiments) in this study.

**Chemical shift perturbation experiments:** We subjected <sup>15</sup>N-labeled FimH-CRD to CSP experiments with antagonists **1** to **11** (Table 1). Methyl  $\alpha$ -D-mannoside (**8**) and *n*-heptyl  $\alpha$ -D-mannoside (**1**) have been analyzed in similar studies before.<sup>[6,25]</sup> With all compounds, separate signals for the bound and free forms were observed in the <sup>1</sup>H,<sup>15</sup>N HSQC spectra at 500 MHz, thus indicating slow exchange on the NMR timescale. The observation is in accordance with nanomolar to low micromolar affinities (Table 2) and with kinetic data obtained from surface plasmon resonance experiments.<sup>[32]</sup> Interestingly, even antagonist **11**, which binds to FimH-CRD with a greatly diminished affinity of 83  $\mu$ M (Table 2), was found to be in slow or intermediate exchange. Slow exchange kinetics might suggest rearrangements of conformation during complex formation.

For the non-overlapping signals in the <sup>1</sup>H,<sup>15</sup>N HSQC spectra, combined chemical shift changes ( $\Delta\delta_{\text{AVG}}$ ) with respect to the protein signals in the absence of ligand were determined (Figure 2A). 1,5-Anhydromannitol (**10**) is the smallest structural motif that still shows specific binding to the mannoside pocket of FimH-CRD. Its CSP effects were therefore used to calculate differential CSP effects ( $\Delta\Delta\delta_{\text{AVG}}$ ) for the other ligands that should mainly reflect the influence of the aglycone moieties (Figure 2B and C). In the presence of antagonists **1** to **11**, FimH-CRD displayed chemical shift changes for residues near the known mannoside binding site. Only minor CSP effects of below 0.1 ppm were observed for residues distal from this site. Because proton chemical shifts are highly sensitive to changes in hydrogen bond length,<sup>[24]</sup> global conformational changes of

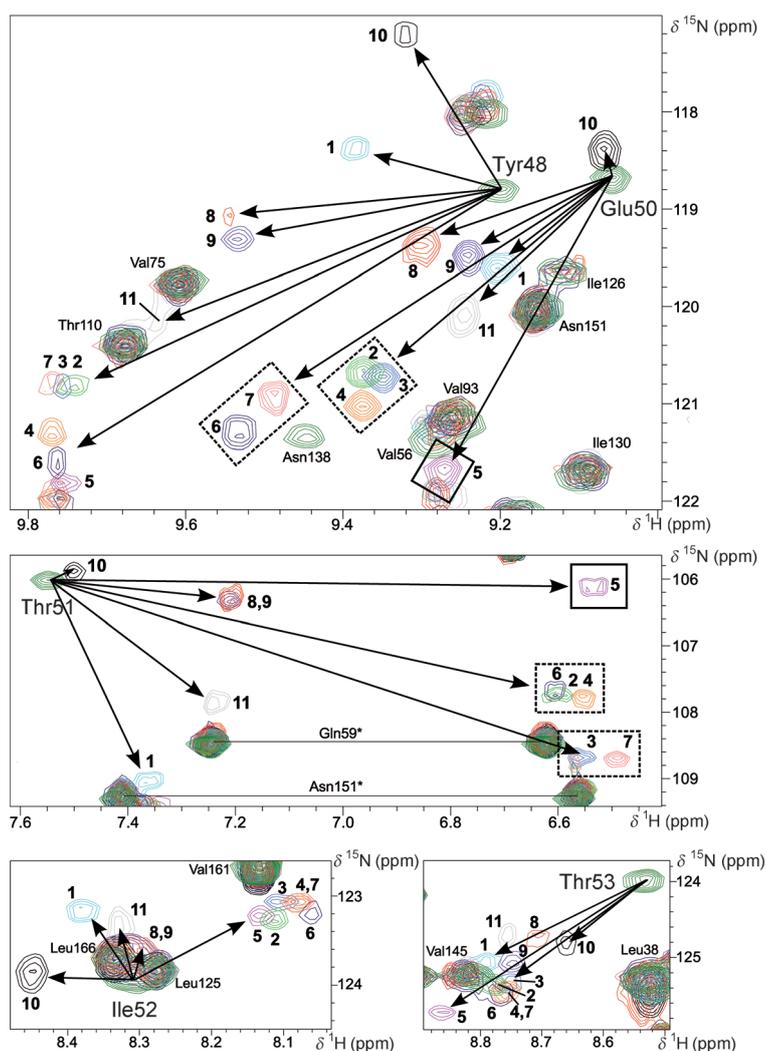


**Figure 2.** Chemical shift changes of FimH-CRD backbone amide signals upon addition of FimH antagonists. A) Absolute combined chemical shift changes ( $\Delta\delta_{\text{AVG}}$ ) of **1**, **8**, and **10**. B), C) Differential CSP effects ( $\Delta\Delta\delta_{\text{AVG}}$ ) of B) aliphatic, and C) aromatic ligands relative to 1,5-anhydromannitol (**10**). Secondary structure elements are schematized on top. Residues missing from the chart are Phe1 (N terminus), 11 proline residues (residues 12, 26, 49, 83, 85, 91, 102, 104, 111, 157, and 162), Gly79, Asn96 and Arg98 (not observed), Asp47, Tyr108 and Phe142 (overlap), Ser139 (exchange broadening), as well as the C-terminal His<sub>6</sub>-tag (residues His168 to His173).

FimH-CRD in response to antagonist binding can be excluded, in view of the fact that the protein is mainly composed of anti-parallel  $\beta$ -sheets that form numerous hydrogen bonds.<sup>[29]</sup> The observation is in agreement with the high structural similarity of the X-ray co-crystal structures discussed above. In presence of the “core” motif **10**, significant  $\Delta\delta_{\text{AVG}}$  values, above 0.1 ppm, were observed for (in order of decreasing  $\Delta\delta_{\text{AVG}}$ ) Tyr48, Asn138, Thr53, Asn136, Asp54, Val56, Cys3, Tyr55, Asp47, Tyr95, Ile52, Ile13, Gly14, Thr134, and Asn46. Most of these residues directly constitute the binding pocket, in particular residues from the loop regions between  $\beta$ -sheets A' and A2 (Ile11 to Gly16), B2 and C (Asn46 to Val56), and F and G (Arg132 to Phe144). Of these, Tyr48, Ile52, and Tyr137 form the tyrosine gate that is known to interact with the aglycone moieties of

FimH antagonists. In addition, CSP effects were observed for residues Ala2 and Cys3, located at the bottom of the binding pocket, and for Tyr95, which forms a hydrogen bond with the pocket residue Asp54. For most residues, only very small  $\Delta\delta_{\text{AVG}}$  values were observed. Larger deviations were observed for loop residues around Tyr137 in the presence of **11** (see later discussion) and for loop residues around Tyr48.

*CSP data report on the Tyr48 side chain conformation:* The strongest and most heterogeneous chemical shift changes of FimH-CRD with different antagonists were observed for the loop residues Tyr48, Glu50, Thr51, Ile52, and Thr53 (Figures 2 and 3). In the X-ray structures of FimH-CRD co-crystallized with **1**, **2**, and **4** to **7**, the Tyr48 side chain directly interacts with the aglycone moieties and adopts an orientation significantly dif-



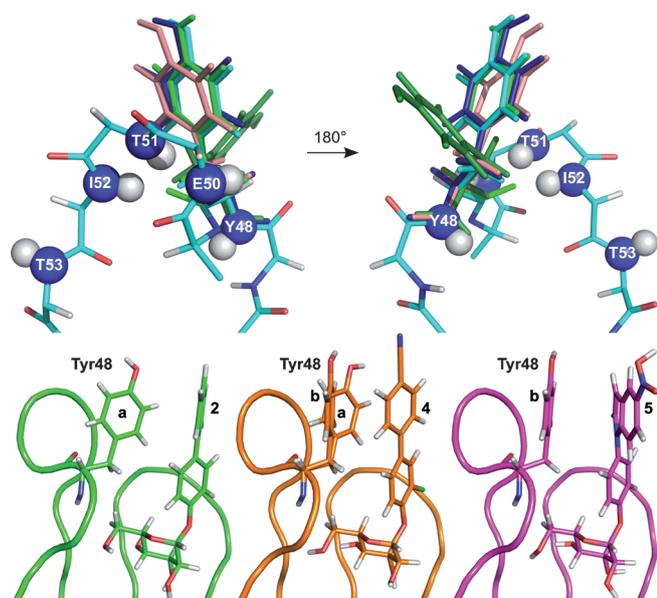
**Figure 3.** Chemical shift changes of the binding loop residues Tyr48, Glu50, Thr51, Ile52, and Thr53 in the presence of antagonists 1 to 11. Spectra in absence of ligand are colored in green. Signals corresponding to Tyr48 conformations a and b in Figure 4 are marked by dashed and solid line boxes, respectively.

ferent from that observed in the “pseudo-apo” X-ray structure<sup>[26]</sup> and in a recently solved NMR solution structure<sup>[25]</sup> (Figure S2). Apart from Tyr48, the conformation of the loop residues is remarkably conserved in all FimH-CRD crystal structures. This includes hydrogen bonds between Tyr48 H<sup>N</sup> and Asn46 O<sup>δ1</sup>, between Glu50 H<sup>N</sup> and its side chain oxygen (stabilized by interaction with the guanidinium group of Arg98, Figure S3), and between Thr51 H<sup>N</sup> and Tyr48 O. We hypothesized that the observed backbone CSP effects of Tyr48 to Thr53 mainly reflect the change in conformation of the Tyr48 side chain rather than direct ligand effects. In general, aromatic rings dramatically influence the chemical shifts of nearby nuclei through the local magnetic fields induced by their delocalized  $\pi$  electrons. For protons located on the outside of the ring, the local magnetic field is parallel to the external field, causing a downfield shift (higher ppm values). In contrast, protons located inside or above the ring experience an opposing field and hence an upfield shift. For Tyr48 and Glu50, we ob-

served strong downfield shifts of both proton and <sup>15</sup>N chemical shifts in the presence of antagonists, whereas Thr51 experienced a strong proton upfield shift of up to 1 ppm and above (Figure 3). The shifts were larger for aromatic aglycones than for aliphatic ones. In the *open* conformation seen in the “pseudo-apo” crystal structure<sup>[26]</sup> the Tyr48 ring is distant from the loop residues. In the *closed* conformation, it is rotated towards the loop such that Glu50 H<sup>N</sup> comes closer and within 30° of the ring plane, where it experiences a strong deshielding aromatic ring current. In contrast, Thr51 H<sup>N</sup> points almost directly into the ring in a T-shaped orientation in which a strong shielding field causes upfield shifts.<sup>[33]</sup> Previous studies had demonstrated that T-shaped N–H $\cdots\pi$  interactions can significantly contribute to ligand binding affinities.<sup>[34]</sup> Enhancement of the aromatic ring current effect through  $\pi$ – $\pi$  stacking provides the explanation for the generally larger CSP effects observed for aromatic antagonists than for their aliphatic counterparts. Importantly, Glu50 and Thr51 show almost no shifts in the presence of 10; this strongly indicates that FimH-CRD remains in the unperturbed *open* conformation.

A more detailed analysis of the CSP effects with aromatic antagonists allows even further differentiation of the Tyr48 orientation in the *closed* conformation. The signals of Glu50 and Thr51 are grouped together in the presence of compounds 2, 3, 4, 6, and 7 (Figure 3, dashed line boxes). With antagonist 5, slightly different shifts were observed, thus suggesting a different orientation of Tyr48. Aliphatic compounds 1 and 8–11 cannot be directly compared to the aromatic antagonists, because the aromatic ring current effect of Tyr48 as the main source of the shifts is drastically different. In good agreement with the NMR results, the high-resolution X-ray structure of FimH-CRD in complex with 5 shows a distinct orientation of the Tyr48 ring, which is tilted by about 40° in relation to the co-crystal structures with compounds 1, 2, 6, and 7 (Figure 4, bottom). This tilt is likely the result of the unique geometry of the nonplanar indolinyphenyl moiety of 5 and the propensity of the system to optimize the stacking interaction. Interestingly, in the co-crystal structure with biphenyl derivative 4 the electron density clearly allows positioning of the Tyr48 ring in both conformations (Figure 4). We cannot find any confirmation for this in the NMR experiments, because only a single set of bound signals is observed with 3 and the CSP effects are very similar to those with compounds 2 and 6 with a single Tyr48 orientation in the crystal structure. Nevertheless, it is conceivable that Tyr48 could retain a certain degree of flexibility in the bound state (see later discussion on ITC data). NMR chemical shifts of the binding loop residues thereby would show an average of rapidly exchanging Tyr48 orientations in the complex.

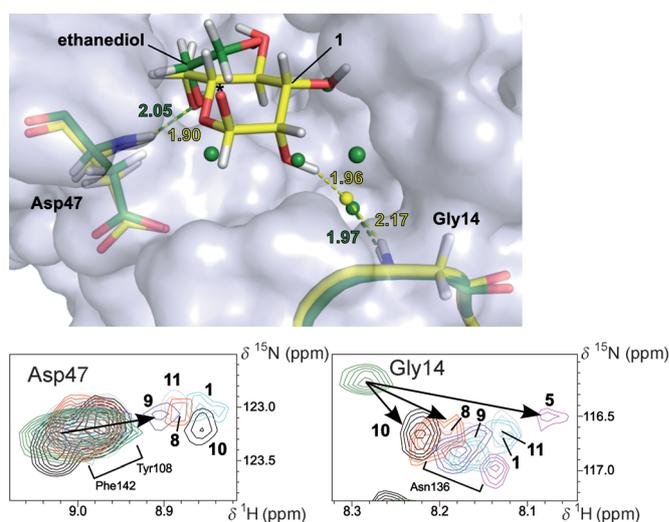
*Water coordination in the FimH binding pocket:* Some CSP effects are expected to be due to direct interactions of the pro-



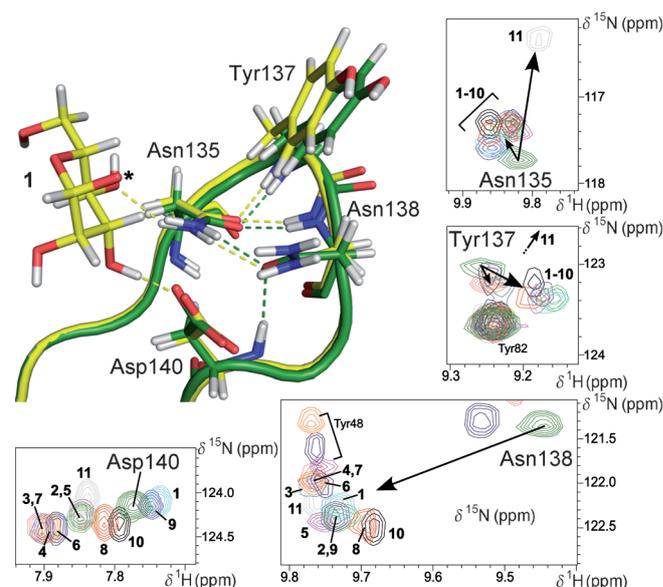
**Figure 4.** Orientation of the Tyr48 side chain in FimH-CRD crystal structures. Top: Loop residues Tyr48 to Thr53 in the co-crystal structure with **1** (backbone in cyan, amides shown as spheres); Tyr48 side chain in the “pseudo-apo” structure<sup>[26]</sup> (green) and with **1** (cyan), **2** (pale green), **6** (dark blue), and **7** (salmon). Bottom: Different Tyr48 orientations (a and b) with **2**, **4** and **5**.

tein with the mannosyl moiety. X-ray structural data implicate the formation of a direct hydrogen bond between Asp47 H<sup>N</sup> and Man OH6 and of a water-mediated hydrogen bond between Gly14 H<sup>N</sup> and Man OH2 (Figure 5). Against the expectation of a <sup>1</sup>H downfield shift from deshielding upon hydrogen bond formation, both amide protons showed upfield shifts (Figure 5). This instead points to weakening of an existing hydrogen bond. Although other more complex effects cannot be excluded, the results indicate that Gly14 and Asp47 coordinate water in the absence of ligand, and that upon mannoside binding this is replaced by polar groups of the ligand in an effectively weaker hydrogen bond. The higher degree of freedom of water molecules might indeed allow hydrogen bonds with more ideal geometries than in the ligands. In the X-ray co-crystal structures with antagonists **1**, **2**, and **4** to **7** the distances between the bridging water and Gly14 H<sup>N</sup> are 0.1 to 0.2 Å less than in the “pseudo-apo” crystal structure, in agreement with our hypothesis (Figure 5). Finally, molecular dynamics simulations also indicated the presence of a structural water molecule close to Gly14.<sup>[35]</sup>

**Hydrogen bond network in the binding loop with Tyr137:** The loop residues Asn135 to Asp141 display similar CSP effects in the presence of all tested antagonists except **11** (Figure 2). X-ray co-crystal structures revealed a complex interresidue hydrogen bond network between backbone and side chain atoms within this loop (Figure 6). Asn138 H<sup>N</sup>, for example, forms a hydrogen bond to the side chain oxygen of Asn135. The strong deshielding of Asn138 H<sup>N</sup> in the presence of ligand suggests a strengthening of this hydrogen bond (Figure 6). The shift could also be the result of a change in the conformation of the aromatic ring of Tyr137. However, Tyr137 H<sup>N</sup> displayed no or only very small shifts, thus suggesting that its aromatic ring



**Figure 5.** Water coordination in the FimH binding pocket. Top: X-ray structures of “pseudo-apo” FimH-CRD (with ethane-1,2-diol in the binding pocket, green sticks and transparent surface)<sup>[26]</sup> and in complex with **1** (yellow, asterisk indicates attachment point of aglycone); water molecules are shown as spheres, and hydrogen bond lengths are given in Å. Bottom: Chemical shift changes of Asp47 (only assigned for nonaromatic antagonists) and Gly14 (only subset of spectra shown).

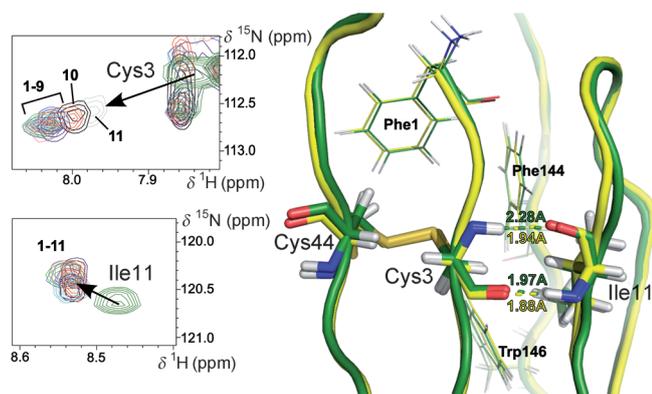


**Figure 6.** Hydrogen bond network in the binding loop with Tyr137. Top left: X-ray crystal structures of “pseudo-apo” FimH-CRD (green)<sup>[26]</sup> and in complex with **1** (yellow, attachment point of aglycone marked with an asterisk). Observed hydrogen bonds: Asn135 H<sup>δ22</sup>...Asn138 O<sup>δ1</sup>, Tyr137 H<sup>N</sup>...Asn135 O<sup>δ1</sup>, Asn138 H<sup>N</sup>...Asn135 O<sup>δ1</sup>, and Asp140 H<sup>N</sup>...Asn138 O<sup>δ1</sup>. With ligand: Man O-H3...Asp140 O<sup>δ1</sup> and Asn135 H<sup>δ21</sup>...Man O4. Bottom and right: Chemical shift changes of Asn138 and Asp140 with antagonists **1** to **11**.

remains largely unchanged, in agreement with very similar orientations in the X-ray structures. Furthermore, deshielding of Asp140 H<sup>N</sup> with all ligands except **1** and **9** indicates a slight shortening of the hydrogen bond to the side chain oxygen of Asn138 (Figure 6).

In summary, mannoside coordination seems to induce subtle but specific conformational changes in the hydrogen bond network within the binding loop. The differences in the CSP effects of Asn135, Tyr137, and Asp141 with the 4-deoxy-4-fluoro derivative **11** are hence caused by a different electronic environment of the fluorine atom that also imposes a different loop arrangement through disruption of hydrogen bonds. The X-ray structural data are not entirely conclusive on the effect of mannoside binding to the loop conformation. In all six co-crystal structures discussed above, the loop residues are well ordered, but hydrogen bonds in the loop are identical within error to those in the “pseudo-apo” structure. Apparently, very small conformational changes, notably in a solvent-exposed loop with low-populated fluctuating states, cannot be captured by X-ray crystallography.

**Conformational changes in the pocket zipper motif:** The CSP effects for Cys3 are unlikely to be caused by direct ligand effects, in view of its remoteness ( $>8.5$  Å) from the bound mannose. In FimH-CRD, Cys3 is part of a highly structured and stable hydrogen bond network to neighboring  $\beta$ -strands further stabilized by a disulfide bond to Cys44 (Figure 7). This so-



**Figure 7.** Conformational changes in the pocket zipper motif. Left: Chemical shift changes of Cys3 and Ile11 in the presence of **1** to **11**. Right: X-ray crystal structures of “pseudo-apo” FimH-CRD (green)<sup>[26]</sup> and in complex with **1** (yellow); hydrogen bond lengths are given in Å, and aromatic residues in the vicinity are shown as lines.

called “pocket zipper” motif is one of four regions that are suspected to play a key role in the conformation transition between the low- and the high-affinity states in the full-length FimH.<sup>[6,36]</sup> Upon transition to the high-affinity state, the “pocket zipper” becomes tightly hydrogen bonded, causing a constriction of the binding pocket.<sup>[6]</sup> The isolated lectin domain FimH-CRD is believed to be locked in the high-affinity state. In the CSP experiments with FimH antagonists, we observed relatively uniform downfield shifts for Cys3  $\text{H}^{\text{N}}$ , which forms a hydrogen bond to the backbone oxygen of Ile11 (Figure 7). A similar, albeit smaller, effect was seen for Ile11  $\text{H}^{\text{N}}$ , which forms an adjacent hydrogen bond to the Cys3 backbone oxygen. This suggests a strengthening of the hydrogen bonds, although a contribution of nearby aromatic rings, in particular of Phe144 (ca. 3 Å from Cys3  $\text{H}^{\text{N}}$ ) also seems possible. In the six FimH-CRD

crystal structures discussed above, the hydrogen bond between Cys3  $\text{H}^{\text{N}}$  and Ile11 O is 0.2 to 0.3 Å shorter than in the “pseudo-apo” structures, in agreement with the CSP effects (Figure 7). Such a difference is considered to be significant at the given resolution.<sup>[37]</sup> A strengthening of the “pocket zipper” upon ligand binding could point to a residual conformation transition of FimH-CRD. However, no CSP effects were observed in the other allosteric regions, thus confirming that the isolated lectin domain is not undergoing significant conformation adaptations characteristic of the full-length protein. Corresponding NMR solution studies of full-length FimH would be highly desirable but have so far been hampered by limited protein yield and stability.

### Thermodynamics data from ITC experiments

FimH antagonists **1** to **11** have been subjected to thermodynamic profiling by ITC analysis with FimH-CRD (Table 2 and Table S2). To obtain reliable thermodynamics data, the high affinities of some of the compounds required the establishment of a competitive ITC assay<sup>[38]</sup> (see the Supporting Information).

**Table 2.** Thermodynamics of binding of FimH antagonists analyzed by ITC.

Compd	$K_{\text{D}}$ [nM]	$\Delta G^{\circ}_{\text{obs}}$ [kJ mol <sup>-1</sup> ]	$\Delta H^{\circ}_{\text{obs}}$ [kJ mol <sup>-1</sup> ]	$-T\Delta S^{\circ}_{\text{obs}}$ [kJ mol <sup>-1</sup> ]
<b>1</b> <sup>[a]</sup>	28.9	-43.0	-50.3	7.3
<b>2</b> <sup>[a]</sup>	17.7	-44.2	-45.0	0.8
<b>3</b> <sup>[b]</sup>	3.5	-48.3	-56.2	8.0
<b>4</b> <sup>[b]</sup>	1.3	-50.7	-60.9	10.1
<b>5</b> <sup>[b]</sup>	1.0	-51.3	-62.1	10.8
<b>6</b> <sup>[a]</sup>	14.0	-44.8	-63.6	18.8
<b>7</b> <sup>[a]</sup>	6.2	-46.8	-71.8	25.0
<b>8</b> <sup>[a]</sup>	1222	-33.8	-37.2	3.5
<b>9</b> <sup>[a]</sup>	144	-39.1	-39.1	0.1
<b>10</b> <sup>[a]</sup>	1125	-34.0	-42.9	9.0
<b>11</b> <sup>[a]</sup>	89 990	-23.1	-21.1	-2.0

[a] Direct ITC assay format. [b] Competitive ITC assay format with a medium-affinity ligand.

The results indicate that all tested antagonists bind in an enthalpy-driven manner with mostly unfavorable entropic contributions, in accordance with previous studies.<sup>[15d,21,26]</sup> The enthalpy term  $\Delta H^{\circ}_{\text{obs}} = -42.9$  kJ mol<sup>-1</sup> of 1,5-anhydromannitol (**10**) includes the binding energies from van der Waals contacts and ten specific hydrogen bonds with the protein.<sup>[11]</sup> These favorable terms are partly compensated by enthalpic penalties from desolvation of ligand and protein. The enthalpic cost for desolvation of a single hydroxy group had been estimated at 29 kJ mol<sup>-1</sup>.<sup>[39]</sup> On the protein side, X-ray and NMR data and molecular dynamics simulations found evidence of highly structured hydrogen-bonded water in the binding pocket, with this being replaced by the hydroxy groups of the ligand. The classic hydrophobic effect predicts a strong entropic gain for the release of bound water upon complex formation.<sup>[40]</sup> However, other studies also suggested an enthalpy-driven “non-classical” hydrophobic effect that is usually explained by water

being forced into an enthalpically unfavorable configuration in the binding pocket.<sup>[41]</sup> The entropy term for **10** suggests that any entropic gain from the release of water into the bulk is overcompensated by unfavorable contributions such as the loss of rotational and translational entropy of the ligand (estimated at about 25 kJ mol<sup>-1</sup> in aqueous solutions),<sup>[42]</sup> as well as by conformational restriction of ligand and protein in the complex. Relevant for the latter is the formation of several hydrogen bonds that reduce the flexibility of the ligand's hydroxy groups and the protein residues involved. Quantification of the individual entropic contributions is cumbersome and in fact still represents a major obstacle in current research.<sup>[41,43]</sup>

The thermodynamic data for methyl  $\alpha$ -D-mannoside (**8**) reveal a less favorable enthalpy ( $\Delta\Delta H^\circ_{\text{obs}}$ : 5.7 kJ mol<sup>-1</sup>) and a smaller entropy penalty ( $-T\Delta\Delta S^\circ_{\text{obs}}$ : -5.5 kJ mol<sup>-1</sup>) than in the case of **10**. The additional methoxy group of **8** would be expected to increase the binding enthalpy through van der Waals interactions with the protein (Ile13, Asp47, and Tyr48 within 4 Å). The effective enthalpy loss is due to an additional desolvation penalty for the methoxy group, whereas the entropy gain reflects the release of additional water molecules into the bulk. From NMR CSP experiments we also expect a change in the Tyr48 side chain conformation for all mannosides with methyl or larger aglycones. The implications for the thermodynamics depend on whether the complex formation follows an induced-fit mechanism or conformational selection or a mixture of both.<sup>[28]</sup> In a recent solution NMR study of FimH-CRD, <sup>15</sup>N relaxation and CPMG relaxation dispersion experiments investigating the backbone amides did not provide any indication of multiple conformations of Tyr48 in the apo protein, neither of a significant change in the backbone flexibility in the presence of **1**.<sup>[25]</sup> The data therefore rather argue for an induced-fit mechanism of the tyrosine gate instead of conformational selection. Nevertheless, conformational equilibria of Tyr48 in the apo state might exist on a timescale not easily accessible by NMR experiments ( $\mu$ s to ms). Hence, the exact binding mechanism remains to be elucidated. Elongation of the aglycone from methyl (compound **8**) to butyl (compound **9**) and heptyl (compound **1**) led to significant improvement in the binding enthalpy, due to the enhanced C-H $\cdots\pi$  interaction between the elongated alkyl chain and the Tyr48 ring. An expected improvement in the entropy term from the release of more water into the bulk is observed for **9** but not for **1**, which even shows a loss of entropy relative to **8**. We postulated that the stacking interaction of the heptyl chain of **1** with Tyr48 leads to rigidification of both interaction partners and hence to significant entropic costs. For compound **11**, the 4-deoxy-4-fluoro analogue of **1**, a significantly smaller enthalpy term was observed, as would be expected from the disruption of the hydrogen bonds in position 4. The CSP experiments also indicated structural changes in the binding loop containing Tyr137 that can translate into enthalpic costs. Interestingly, the entropy is improved by more than 9 kJ mol<sup>-1</sup> relative to **1**, due to the greater flexibility of the protein binding pocket and the ligand.

For the aromatic antagonists **2** to **7**, strongly favorable enthalpy terms ranging from -45.0 to -71.8 kJ mol<sup>-1</sup> were ob-

tained. Again, an increase in enthalpy is always accompanied by an increase in entropic costs, as becomes obvious in the enthalpy-entropy compensation plot (Figure S4). X-ray structure data indicate that the outer aromatic rings of antagonists **2** and **4-7** make parallel-displaced  $\pi$ - $\pi$  interactions with the Tyr48 ring in the *closed* conformation, in analogy to the C-H $\cdots\pi$  interaction of the heptyl chain of **1**. However, a direct comparison of the thermodynamic data for aliphatic and aromatic antagonists is difficult in view of the large structural difference of the aglycones and because of the many contributing factors. The unsubstituted biphenyl mannoside **2** displayed the smallest enthalpy term and likewise the smallest entropic penalty of all aromatic antagonists. For the substituted biphenyls **3** and **4** we observed significant increases in the enthalpy, by 11.2 and 15.9 kJ mol<sup>-1</sup>, respectively, relative to **2**. Firstly, the *ortho*-chloro substituent is positioned to enhance the binding enthalpy through favorable van der Waals interactions with the protein in a small pocket formed by Ile52, Tyr137, and Asn138.<sup>[15d]</sup> Secondly, the electron-withdrawing character of the substituents allows stronger  $\pi$ - $\pi$  interaction of the aglycones with Tyr48. As already suggested for the comparison of **1** with **8**, a stronger stacking interaction accounts for a reduced flexibility both in the Tyr48 side chain and in the aglycone, leading to compensating entropic penalties. For antagonist **5**, the enthalpy is slightly better than for the best biphenyl compound **4**. The entropy compensation is only partial, resulting in the strongest dissociation constant of all tested antagonists. Compounds **6** and **7** display the most favorable enthalpies but also the highest entropic penalties. Compound **6** could suffer from entropic costs for rigidification of additional rotatable bonds in the aglycone moiety. Compound **7** is structurally very similar to biphenyl **3** apart from insertion of an amide bond linkage. The dramatically enhanced enthalpy ( $\Delta\Delta H^\circ_{\text{obs}}$ : -15.6 kJ mol<sup>-1</sup>) and entropic cost ( $-T\Delta\Delta S^\circ_{\text{obs}}$ : 17.0 kJ mol<sup>-1</sup>) of **7** relative to **3** can only be explained in terms of a very efficient stacking interaction with Tyr48 and subsequent loss of conformational entropy in the complex. A comparison of all thermodynamic data demonstrates that the antagonist with the highest enthalpic contribution, presumably from the strongest interaction with Tyr48, is not necessarily the strongest binder. Antagonist **5** exhibits an optimal thermodynamic profile, although its enthalpy term is nearly 10 kJ mol<sup>-1</sup> weaker than that of **7**. From the NMR and X-ray data we deduced a distinct orientation of Tyr48 in the presence of **5**, likely as a result of the unique geometry of its indolylphenyl aglycone. We assume that this complex conformation is optimal in terms of interaction efficiency and minimized compensating entropic penalties.

## Conclusion

We have combined NMR chemical shift perturbation (CSP) experiments with high-resolution X-ray structures to elucidate the interaction of the FimH lectin domain with antagonists. One advantage of NMR spectroscopy is its ability to identify even subtle conformational changes of the protein in solution. This is particularly helpful for the determination of ligand bind-

ing modes in cases in which either crystallization was unsuccessful or crystal packing effects distorted the ligand binding site. We demonstrated that the CSP effects of FimH-CRD can be used as an indicator for the conformation of the tyrosine gate motif in the binding pocket. The *open* conformation was observed in the apo protein and with 1,5-anhydromannitol (**10**), whereas antagonists with aliphatic or aromatic aglycones are bound to FimH in the *closed* conformation of the tyrosine gate. The CSP data additionally revealed slight differences in the Tyr48 conformation as a result of different aglycone geometries. Furthermore, the NMR results indicate the presence of highly structured water in the binding pocket of FimH-CRD. In combination with information from crystal structures and molecular modeling, CSP experiments might therefore help to analyze water coordination as an important contributor to the thermodynamics of ligand binding.

We also performed ITC experiments to access full thermodynamic profiles of the antagonists. The results suggest enthalpy–entropy compensation in which enthalpically favorable stacking of hydrophobic aglycones with the Tyr48 side chain leads to conformational restriction of both the protein and the ligand and hence to unfavorable entropy. NMR relaxation experiments with FimH-CRD did not show any changes in the backbone flexibility upon binding of antagonists.<sup>[25]</sup> However, we speculate that the flexibility of the protein side chains, in particular of Tyr48, plays a pivotal role in the modulation of FimH–antagonist binding. These effects can be dissected by NMR relaxation experiments for side chain dynamics, and incorporation of the results into rational design might lead to further improved FimH antagonists urgently needed for UTI therapy.

## Experimental Section

**Protein preparation:** FimH-CRD from the *E. coli* K-12 strain was expressed with a C-terminal thrombin cleavage site and a His<sub>6</sub>-tag (FimH-CRD-Th-His<sub>6</sub>, 173 residues) by a previously published protocol.<sup>[18]</sup> The clone containing the FimH-CRD construct was expressed in protease-deficient *E. coli* HM125 at 30 °C and 180 rpm in M9 minimal medium supplemented with MgSO<sub>4</sub> (2 mM), CaCl<sub>2</sub> (0.1 mM), glucose (2 g L<sup>-1</sup>), and ampicillin (100 μg mL<sup>-1</sup>). The protein expression was induced by addition of IPTG (1 mM) at an OD<sub>600</sub> of 0.8. The cells were further cultivated for 16 h and harvested by centrifugation for 20 min at 2000g and 4 °C. The pellet was resuspended in lysis buffer containing Tris (pH 7.4, 50 mM), NaCl (150 mM), EDTA (5 mM), and polymyxin B sulfate (1 mg mL<sup>-1</sup>). The supernatant containing the periplasmic extract was dialyzed against sodium phosphate buffer and purified on Ni-NTA columns. For crystallization, the protein without tag was prepared as described in ref. [44] and dialyzed against HEPES buffer (pH 7.4, 20 mM). For production of <sup>15</sup>N- and <sup>13</sup>C,<sup>15</sup>N-labeled FimH-CRD-Th-His<sub>6</sub> for NMR experiments, *E. coli* HM125 was cultivated in M9 minimal medium containing <sup>15</sup>NH<sub>4</sub>Cl (1 g L<sup>-1</sup>) or <sup>5</sup>NH<sub>4</sub>Cl (1 g L<sup>-1</sup>) and <sup>13</sup>C-glucose (2 g L<sup>-1</sup>) (Sigma–Aldrich) as the only sources of nitrogen and carbon, respectively. The labeled proteins were purified as described above and dialyzed against phosphate buffer (pH 6.8, 25 mM) in the case of <sup>13</sup>C,<sup>15</sup>N-FimH-CRD-Th-His<sub>6</sub> or phosphate buffer (pH 7.0, 20 mM) in that of <sup>15</sup>N-FimH-CRD-Th-His<sub>6</sub>. The purities of the proteins were verified by nonreducing SDS-PAGE analysis,

and the concentrations were determined by UV absorption (NanoDrop ND-1000, Thermo Scientific). The molecular weights of <sup>15</sup>N-FimH-CRD-Th-His<sub>6</sub> (18860.2 Da) and <sup>13</sup>C,<sup>15</sup>N-FimH-CRD-Th-His<sub>6</sub> (19687.0 Da), determined by mass spectrometry, demonstrated isotope incorporation of >99.9%.

**NMR spectroscopy:** NMR assignment experiments were performed at 298 K with a Bruker Avance III 700 MHz NMR spectrometer equipped with a 5 mm QCI-P cryogenic probe. A set of triple-resonance NMR spectra of a sample of <sup>13</sup>C,<sup>15</sup>N-FimH-CRD (600 μM) in phosphate buffer (pH 6.8, 25 mM) and H<sub>2</sub>O/D<sub>2</sub>O (90%/10%) in a 5 mm Shigemi Tube (Shigemi Inc., USA) was recorded for backbone assignment: HNCA, HN(CO)CA, HNCO, HN(CA)CO, HNCACB, and CBCA(CO)NH. Protein stability during data acquisition was tested by regular inspection of 1D <sup>1</sup>H and <sup>1</sup>H,<sup>15</sup>N HSQC NMR spectra. Spectra were acquired and processed with Topspin 3.2 (Bruker BioSpin, Switzerland). CcpNmr Analysis (versions 2.2 and 2.3) was used for NMR resonance assignment.<sup>[45]</sup> The backbone assignment of FimH-CRD has been deposited in the Biological Magnetic Resonance Data Bank (<http://www.bmrb.wisc.edu>, accession code 26541).

Chemical shift perturbation (CSP) experiments with FimH antagonists were performed with <sup>15</sup>N-FimH-CRD (100 to 200 μM) in non-deuterated phosphate buffer (pH 7.0, 20 mM). FimH antagonists were dissolved in D<sub>2</sub>O at 2.5 to 100 mM stock concentrations. Biphenyl compound **4** was dissolved at 6.3 mM in D<sub>2</sub>O with [D<sub>6</sub>]DMSO (40%), due to poor solubility. The antagonists were added at slight molar excesses to ensure complete saturation of the protein. NMR samples were prepared in 3 mm NMR tubes (Hilgenberg, Germany) with 5% D<sub>2</sub>O and [D<sub>4</sub>]TSP [3-(trimethylsilyl)-2,2',3,3'-tetradeuteropropionic acid, Armar Chemicals, Switzerland] (0.1 mM) added as an internal reference. <sup>1</sup>H,<sup>15</sup>N HSQC spectra were measured with a Bruker Avance III 500 MHz spectrometer equipped with a BBO double resonance probe at 298 K. Spectra were acquired and processed with Topspin 2.1 and analyzed with CcpNmr Analysis (versions 2.2 and 2.3).<sup>[45]</sup> All antagonists bound in the slow exchange regime, and amide signals of the bound protein were assigned from chemical shift proximity. Combined chemical shift changes of FimH-CRD signals were calculated as weighted averages of <sup>1</sup>H and <sup>15</sup>N chemical shift changes according to Equation (1):<sup>[46]</sup>

$$\Delta\delta_{AV} = \sqrt{(\Delta\delta^1H)^2 + (0.2\Delta\delta^{15}N)^2} \quad (1)$$

**Co-crystallization of FimH antagonists:** Crystallization of all FimH-CRD/ligand complexes was performed by sitting-drop vapor diffusion. For crystallization trials of ligands **5–7**, FimH-CRD (residues 1–158) was used at a final concentration of 18 mg mL<sup>-1</sup> (ca. 1 mM) with a fivefold molar excess of ligand (5 mM) in HEPES buffer (pH 7.4, 20 mM). For crystallization trials with ligand **2**, an FimH-CRD solution (18 mg mL<sup>-1</sup>) was diluted to 10 mg mL<sup>-1</sup> with a saturated solution of ligand **2** in HEPES buffer (pH 7.4, 20 mM). Crystals with ligand **5** were grown as previously described.<sup>[21]</sup> Co-crystals with ligand **6** were grown in the Proplex HT-96 screen (Molecular Dimensions, USA) with (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub> (1.5 M) and HEPES (pH 7.0, 0.1 M). After four weeks equilibration at 20 °C, crystals appeared after subsequent equilibration at 4 °C within a few days. Crystals with ligand **7** were grown in PEG 4000 (20%) and NaH<sub>2</sub>PO<sub>4</sub> (pH 4.7, 0.2 M) at 20 °C. After two days equilibration, streak seeding with FimH-CRD/**5** co-crystals yielded crystals within 24 h. Crystals with ligand **2** grew within a few days with (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub> (0.2 M), PEG 8000 (19%), and HEPES (pH 7.4, 0.1 M) at 20 °C. All crystals were flash-cooled to 100 K with perfluoropolyether cryo oil (Hampton Re-

search, USA). Data were collected with synchrotron radiation at the PXI (ligand 2) or PXIII (ligands 5–7) beamlines of the Swiss Light Source (Paul Scherrer Institute, Switzerland).

**Structure determination and refining:** Data were indexed, integrated, and scaled with XDS<sup>[47]</sup> or iMOSFLM.<sup>[48]</sup> Structures were solved by molecular replacement with PHASER<sup>[49]</sup> with use of the FimH-CRD-*n*-butyl  $\alpha$ -D-mannopyranoside complex (PDB ID: 1UWF<sup>[11a]</sup>) as search model. The structures were built by use of the COOT software<sup>[50]</sup> and periodically refined with the PHENIX software.<sup>[51]</sup> Geometric restraints for the ligands were generated with PRODRG<sup>[52]</sup> or Grade Version 1.1.1.<sup>[53]</sup> Molprobit<sup>[54]</sup> was used to validate the available atomic coordinates and to add protons to PDB files without hydrogens for distance calculation. The structures are deposited in the Protein Data Bank with PDB codes: 4X50 (2), 4X5Q (5), 4X5R (6), and 4X5P (7).

**Isothermal titration calorimetry (ITC):** All ITC experiments were performed with FimH-CRD-Th-His<sub>6</sub> and use of a MicroCal VP-ITC instrument (Malvern Instruments Ltd, Worcestershire, UK) with a sample cell volume of 1.4523 mL. Measurements were performed at 25 °C with a stirring speed of 307 rpm and 10  $\mu\text{cal s}^{-1}$  reference power. The protein was dialyzed against assay buffer [HEPES (10 mM), NaCl (150 mM), pH 7.4]. Injections of ligand solution (3–12  $\mu\text{L}$ ) were added at 10 min intervals to a sample cell containing protein (5–50  $\mu\text{M}$ ). The  $c$  values [ $c = M_i(0)K_D^{-1}$ , where  $M_i(0)$  is the initial protein concentration] were in a reliable range between 5 and 1000 for compounds 1, 2, and 6–10. For compounds 3, 4, and 5, the  $c$  values of the direct titrations were above 1000, so additional competitive ITC experiments were performed.<sup>[38]</sup> The ligands were titrated into protein preincubated with an eight- to ninefold excess of compound 12 (structure and synthesis in the Supporting Information), resulting in sigmoidal titration curves. For the low-affinity compound 11 the  $c$  value was only 0.5, but reliable thermodynamics data could be extracted by fixing the stoichiometry to 1.0. Baseline correction and peak integration was performed with Origin 7 software (OriginLab, USA). Baseline subtraction and curve fitting with the three variables  $N$  (concentration correction factor),  $K_D$  (dissociation constant), and  $\Delta H^\circ$  (change in enthalpy) were performed with SEDPHAT version 10.40 (National Institutes of Health).<sup>[55]</sup> Global fitting to obtain  $K_D$  values was performed for the competitive titrations of compound 12 with 3, 4, and 5 and for the direct titration of 12. The  $\Delta H^\circ$  and  $N$  values were then obtained by fitting of the direct titrations of 3, 4, and 5. For directly titrated compounds 1, 2, and 6–11, all three variables were determined from a global analysis. The 95% confidence intervals of  $K_D$  and  $\Delta H^\circ$  were calculated with the 1-dimensional error surface projection. The  $\Delta G^\circ$  and  $-\Delta S^\circ$  (change in entropy) values were calculated from Eq. (2)

$$\Delta G^\circ = \Delta H^\circ - T\Delta S^\circ = -RT \ln K_A \quad (2)$$

with  $T$  being the absolute temperature and  $R$  the universal gas constant (8.314 J mol<sup>-1</sup> K<sup>-1</sup>).

## Acknowledgements

We sincerely thank Dr. Helene Kovacs of Bruker BioSpin (Fällanden, Switzerland) for acquisition of NMR assignment experiments. B.F. thanks the German Academic Exchange Service (DAAD) for a stipend.

**Keywords:** FimH • ITC • NMR spectroscopy • urinary tract infections • X-ray diffraction

- [1] B. Foxman, *Am. J. Med.* **2002**, *113 Suppl. 1A*, 5S–13S.
- [2] T. J. Wiles, R. R. Kulesus, M. A. Mulvey, *Exp. Mol. Pathol.* **2008**, *85*, 11–19.
- [3] a) G. G. Zhanel, T. L. Hisanaga, N. M. Laing, M. R. DeCorby, K. A. Nichol, B. Weshnoweski, J. Johnson, A. Noreddin, D. E. Low, J. A. Karlowsky, N. Group, D. J. Hoban, *Int. J. Antimicrob. Agents* **2006**, *27*, 468–475; b) G. C. Schito, K. G. Naber, H. Botto, J. Palou, T. Mazzei, L. Gualco, A. Marchese, *Int. J. Antimicrob. Agents* **2009**, *34*, 407–413.
- [4] a) N. Sharon, *FEBS Lett.* **1987**, *217*, 145–157; b) E. V. Sokurenko, V. Chesnokova, R. J. Doyle, D. L. Hasty, *J. Biol. Chem.* **1997**, *272*, 17880–17886.
- [5] G. Waksman, S. J. Hultgren, *Nat. Rev. Microbiol.* **2009**, *7*, 765–774.
- [6] I. Le Trong, P. Aprikian, B. A. Kidd, M. Forero-Shelton, V. Tchesnokova, P. Rajagopal, V. Rodriguez, G. Interlandi, R. Klevit, V. Vogel, R. E. Stenkamp, E. V. Sokurenko, W. E. Thomas, *Cell* **2010**, *141*, 645–655.
- [7] G. Zhou, W.-J. Mo, P. Sebbel, G. Min, T. A. Neubert, R. Glockshuber, X.-R. Wu, T.-T. Sun, X.-P. Kong, *J. Cell Sci.* **2001**, *114*, 4095–4103.
- [8] a) I. Ofek, D. L. Hasty, N. Sharon, *FEMS Immunol. Med. Microbiol.* **2003**, *38*, 181–191; b) R. J. Pieters, *Med. Res. Rev.* **2007**, *27*, 796–816.
- [9] M. Aronson, O. Medalia, L. Schori, D. Mirelman, N. Sharon, I. Ofek, *J. Infect. Dis.* **1979**, *139*, 329–332.
- [10] M. Hartmann, T. K. Lindhorst, *Eur. J. Org. Chem.* **2011**, 3583–3609.
- [11] a) J. Bouckaert, J. Berglund, M. Schembri, E. De Genst, L. Cools, M. Wuhler, C. S. Hung, J. Pinkner, R. Slattegard, A. Zavalov, D. Choudhury, S. Langermann, S. J. Hultgren, L. Wyns, P. Klemm, S. Oscarson, S. D. Knight, H. De Greve, *Mol. Microbiol.* **2005**, *55*, 441–455; b) C. S. Hung, J. Bouckaert, D. Hung, J. Pinkner, C. Widberg, A. DeFusco, C. G. Auguste, R. Strouse, S. Langermann, G. Waksman, S. J. Hultgren, *Mol. Microbiol.* **2002**, *44*, 903–915.
- [12] A. Wellens, C. Garofalo, H. Nguyen, N. Van Gerven, R. Slattegard, J. P. Hernalsteens, L. Wyns, S. Oscarson, H. De Greve, S. Hultgren, J. Bouckaert, *PLoS One* **2008**, *3*, e2040.
- [13] a) N. Firon, S. Ashkenazi, D. Mirelman, I. Ofek, N. Sharon, *Infect. Immun.* **1987**, *55*, 472–476; b) T. K. Lindhorst, C. Kieburg, U. Krallmann-Wenzel, *Glycoconjugate J.* **1998**, *15*, 605–613.
- [14] a) O. Sperling, A. Fuchs, T. K. Lindhorst, *Org. Biomol. Chem.* **2006**, *4*, 3913–3922; b) M. Scharenberg, O. Schwardt, S. Rabbani, B. Ernst, *J. Med. Chem.* **2012**, *55*, 9810–9816.
- [15] a) Z. Han, J. S. Pinkner, B. Ford, R. Obermann, W. Nolan, S. A. Wildman, D. Hobbs, T. Ellenberger, C. K. Cusumano, S. J. Hultgren, J. W. Janetka, *J. Med. Chem.* **2010**, *53*, 4779–4792; b) T. Klein, D. Abgottspon, M. Wittwer, S. Rabbani, J. Herold, X. Jiang, S. Kleeb, C. Luthi, M. Scharenberg, J. Bezencon, E. Gubler, L. Pang, M. Smiesko, B. Cutting, O. Schwardt, B. Ernst, *J. Med. Chem.* **2010**, *53*, 8627–8641; c) Z. Han, J. S. Pinkner, B. Ford, E. Chorell, J. M. Crowley, C. K. Cusumano, S. Campbell, J. P. Henderson, S. J. Hultgren, J. W. Janetka, *J. Med. Chem.* **2012**, *55*, 3945–3959; d) L. Pang, S. Kleeb, K. Lemme, S. Rabbani, M. Scharenberg, A. Zalewski, F. Schadler, O. Schwardt, B. Ernst, *ChemMedChem* **2012**, *7*, 1404–1422.
- [16] O. Schwardt, S. Rabbani, M. Hartmann, D. Abgottspon, M. Wittwer, S. Kleeb, A. Zalewski, M. Smiesko, B. Cutting, B. Ernst, *Bioorg. Med. Chem.* **2011**, *19*, 6454–6473.
- [17] X. Jiang, D. Abgottspon, S. Kleeb, S. Rabbani, M. Scharenberg, M. Wittwer, M. Haug, O. Schwardt, B. Ernst, *J. Med. Chem.* **2012**, *55*, 4700–4713.
- [18] S. Rabbani, X. Jiang, O. Schwardt, B. Ernst, *Anal. Biochem.* **2010**, *407*, 188–195.
- [19] a) D. Abgottspon, G. Rolli, L. Hosch, A. Steinhuber, X. Jiang, O. Schwardt, B. Cutting, M. Smiesko, U. Jenal, B. Ernst, A. Trampuz, *J. Microbiol. Methods* **2010**, *82*, 249–255; b) M. Scharenberg, D. Abgottspon, E. Cicek, X. Jiang, O. Schwardt, S. Rabbani, B. Ernst, *Assay Drug Dev. Technol.* **2011**, *9*, 455–464.
- [20] D. Abgottspon, B. Ernst, *Chimia* **2012**, *66*, 166–169.
- [21] S. Kleeb, L. Pang, K. Mayer, D. Eris, A. Sigl, R. C. Preston, P. Zihlmann, T. Sharpe, R. P. Jakob, D. Abgottspon, A. S. Hutter, M. Scharenberg, X. Jiang, G. Navarra, S. Rabbani, M. Smiesko, N. Ludin, J. Bezencon, O. Schwardt, T. Maier, B. Ernst, *J. Med. Chem.* **2015**, *58*, 2221–2239.

- [22] a) A. Cléry, M. Schubert, F. H. Allain, *Chimia* **2012**, *66*, 741–746; b) M. del Carmen Fernandez-Alonso, D. Diaz, M. A. Berbis, F. Marcelo, J. Canada, J. Jimenez-Barbero, *Curr. Protein Pept. Sci.* **2012**, *13*, 816–830.
- [23] a) S. B. Shuker, P. J. Hajduk, R. P. Meadows, S. W. Fesik, *Science* **1996**, *274*, 1531–1534; b) L. Skjærven, L. Codutti, A. Angelini, M. Grimaldi, D. Latek, P. Monecke, M. K. Dreyer, T. Carlomagno, *J. Am. Chem. Soc.* **2013**, *135*, 5819–5827.
- [24] a) G. Wagner, A. Pardi, K. Wuthrich, *J. Am. Chem. Soc.* **1983**, *105*, 5948–5949; b) M. P. Williamson, *Prog. Nucl. Magn. Reson. Spectrosc.* **2013**, *73*, 1–16.
- [25] S. Vanwetswinkel, A. N. Volkov, Y. G. Sterckx, A. Garcia-Pino, L. Buts, W. F. Vranken, J. Bouckaert, R. Roy, L. Wyns, N. A. van Nuland, *J. Med. Chem.* **2014**, *57*, 1416–1427.
- [26] A. Wellens, M. Lahmann, M. Touaibia, J. Vaucher, S. Oscarson, R. Roy, H. Remaut, J. Bouckaert, *Biochemistry* **2012**, *51*, 4790–4799.
- [27] G. Roos, A. Wellens, M. Touaibia, N. Yamakawa, P. Geerlings, R. Roy, L. Wyns, J. Bouckaert, *ACS Med. Chem. Lett.* **2013**, *4*, 1085–1090.
- [28] D. D. Boehr, R. Nussinov, P. E. Wright, *Nat. Chem. Biol.* **2009**, *5*, 789–796.
- [29] D. Choudhury, A. Thompson, V. Stojanoff, S. Langermann, J. Pinkner, S. J. Hultgren, S. D. Knight, *Science* **1999**, *285*, 1061–1066.
- [30] M. Sattler, J. Schleucher, C. Griesinger, *Prog. Nucl. Magn. Reson. Spectrosc.* **1999**, *34*, 93–158.
- [31] The backbone resonance assignment was deposited in the Biological Magnetic Resonance Database (BMRB) entry: 26541.
- [32] M. Scharenberg, X. Jiang, L. Pang, G. Navarra, S. Rabbani, F. Binder, O. Schwardt, B. Ernst, *ChemMedChem* **2014**, *9*, 78–83.
- [33] M. J. Cloninger, H. W. Whitlock, *J. Org. Chem.* **1998**, *63*, 6153–6159.
- [34] E. A. Meyer, R. K. Castellano, F. Diederich, *Angew. Chem. Int. Ed.* **2003**, *42*, 1210–1250; *Angew. Chem.* **2003**, *115*, 1244–1287.
- [35] A. Zalewski, PhD thesis, University of Basel (Basel), **2013**.
- [36] V. B. Rodriguez, B. A. Kidd, G. Interlandi, V. Tchesnokova, E. V. Sokurenko, W. E. Thomas, *J. Biol. Chem.* **2013**, *288*, 24128–24139.
- [37] D. W. Cruickshank, *Acta Crystallogr. Sect. D Biol. Crystallogr.* **1999**, *55*, 583–601.
- [38] B. W. Sigurskjold, *Anal. Biochem.* **2000**, *277*, 260–266.
- [39] a) L. Frick, C. Yang, V. E. Marquez, R. Wolfenden, *Biochemistry* **1989**, *28*, 9423–9430; b) W. M. Kati, R. Wolfenden, *Biochemistry* **1989**, *28*, 7919–7927.
- [40] R. U. Lemieux, *Acc. Chem. Res.* **1996**, *29*, 373–380.
- [41] a) J. M. Myslinski, J. E. DeLorbe, J. H. Clements, S. F. Martin, *J. Am. Chem. Soc.* **2011**, *133*, 18518–18521; b) K. L. Portman, J. Long, S. Carr, L. Briand, D. J. Winzor, M. S. Searle, D. J. Scott, *Biochemistry* **2014**, *53*, 2371–2379; c) P. W. Snyder, J. Mecinovic, D. T. Moustakas, S. W. Thomas 3rd, M. Harder, E. T. Mack, M. R. Lockett, A. Heroux, W. Sherman, G. M. Whitesides, *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 17889–17894.
- [42] a) J. J. Lundquist, E. J. Toone, *Chem. Rev.* **2002**, *102*, 555–578; b) W. B. Turnbull, B. L. Precious, S. W. Homans, *J. Am. Chem. Soc.* **2004**, *126*, 1047–1054.
- [43] M. Ambrosi, N. R. Cameron, B. G. Davis, *Org. Biomol. Chem.* **2005**, *3*, 1593–1608.
- [44] M. Vetsch, P. Sebbel, R. Glockshuber, *J. Mol. Biol.* **2002**, *322*, 827–840.
- [45] W. F. Vranken, W. Boucher, T. J. Stevens, R. H. Fogh, A. Pajon, M. Llinas, E. L. Ulrich, J. L. Markley, J. Ionides, E. D. Laue, *Proteins Struct. Funct. Bioinf.* **2005**, *59*, 687–696.
- [46] M. Pellecchia, P. Sebbel, U. Hermanns, K. Wuthrich, R. Glockshuber, *Nat. Struct. Biol.* **1999**, *6*, 336–339.
- [47] a) W. Kabsch, *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2010**, *66*, 133–144; b) W. Kabsch, *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2010**, *66*, 125–132.
- [48] T. G. Batty, L. Kontogiannis, O. Johnson, H. R. Powell, A. G. Leslie, *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2011**, *67*, 271–281.
- [49] A. J. McCoy, *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2007**, *63*, 32–41.
- [50] P. Emsley, K. Cowtan, *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2004**, *60*, 2126–2132.
- [51] P. D. Adams, P. V. Afonine, G. Bunkoczi, V. B. Chen, I. W. Davis, N. Echols, J. J. Headd, L. W. Hung, G. J. Kapral, R. W. Grosse-Kunstleve, A. J. McCoy, N. W. Moriarty, R. Oeffner, R. J. Read, D. C. Richardson, J. S. Richardson, T. C. Terwilliger, P. H. Zwart, *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2010**, *66*, 213–221.
- [52] D. M. van Aalten, R. Bywater, J. B. Findlay, M. Hendlich, R. W. Hooft, G. Vriend, *J. Comput. Aided Mol. Des.* **1996**, *10*, 255–262.
- [53] O. S. Smart, T. O. Womack, A. Sharff, C. Flensburg, P. Keller, W. Paciorek, C. Vonrhein, G. Bricogne, Grade Version 1.1.1, Copyright by Global Phasing Limited, **2011**.
- [54] V. B. Chen, W. B. Arendall III, 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.
- [55] J. C. Houtman, P. H. Brown, B. Bowden, H. Yamaguchi, E. Appella, L. E. Samelson, P. Schuck, *Protein Sci.* **2007**, *16*, 30–42.

Manuscript received: December 23, 2014

Final article published: May 4, 2015

## Manuscript 4

### 2-C-Branched mannosides as a novel family of FimH antagonists – Synthesis and biological evaluation

Wojciech Schönemann, Marcel Lindegger, Said Rabbani, Pascal Zihlmann,  
Oliver Schwardt and Beat Ernst<sup>†#\*</sup>

University of Basel, Institute of Molecular Pharmacy,  
Klingelbergstr. 50, 4056 Basel, Switzerland

\*Corresponding author.

Tel.: 0041 (0)61 267 15 51; Fax: 0041 (0)61 267 15 52;

E-mail: beat.ernst@unibas.ch

Contribution of Pascal Zihlmann:

- Thermodynamic profiling by ITC

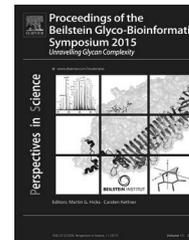




Available online at [www.sciencedirect.com](http://www.sciencedirect.com)

ScienceDirect

journal homepage: [www.elsevier.com/pisc](http://www.elsevier.com/pisc)



## 2-C-Branched mannosides as a novel family of FimH antagonists—Synthesis and biological evaluation<sup>☆</sup>



Wojciech Schönemann, Marcel Lindegger, Said Rabbani, Pascal Zihlmann, Oliver Schwardt, Beat Ernst\*

Institute of Molecular Pharmacy, Pharmacenter, University of Basel, Klingelbergstrasse 50, CH-4056 Basel, Switzerland

Received 5 July 2016; accepted 9 September 2016  
Available online 20 October 2016

### KEYWORDS

Urinary tract infections;  
Uropathogenic *Escherichia coli*;  
FimH antagonists;  
Carbohydrate-recognition domain;  
2-C-Branched carbohydrates

**Summary** Urinary tract infections (UTIs), which are among the most prevalent bacterial infections worldwide, are mainly attributed to uropathogenic *Escherichia coli* (UPEC). Because of frequent antibiotic treatment, antimicrobial resistance constitutes an increasing therapeutic problem. Antagonists of the mannose-specific bacterial lectin FimH, a key protein mediating the adhesion of UPEC to human bladder cells, would offer an alternative anti-adhesive treatment strategy. In general, FimH antagonists consist of a mannose moiety and a wide range of lipophilic aglycones. Modifications of the mannose core led to a distinct drop in affinity. A visual inspection of the crystal structure of FimH revealed a previously unexplored cavity surrounded by Ile13, Phe142 and Asp140, which could be reached by functional groups in the equatorial 2-position of the mannose. Here, we describe the synthesis of 2-C-branched mannosides and evaluation of their pharmacodynamic properties. ITC experiments with the selected antagonists revealed a drastic enthalpy loss for all 2-C-branched antagonists, which, however, is partially compensated by an entropy gain. This supports the hypothesis that the target cavity is too small to accommodate 2-C-substituents.

© 2016 Published by Elsevier GmbH. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

**Abbreviations:** UPEC, uropathogenic *Escherichia coli*; UTI, urinary tract infection; CRD, carbohydrate-recognition domain; IC<sub>50</sub>, half maximal inhibitory concentration; ITC, isothermal titration calorimetry; K<sub>D</sub>, dissociation constant.

<sup>☆</sup> This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. This article is part of a special issue entitled Proceedings of the Beilstein Glyco-Bioinformatics Symposium 2015 with copyright © 2017 Beilstein-Institut. Published by Elsevier GmbH. All rights reserved.

\* Corresponding author. Fax: +41 61 267 15 52.

E-mail address: [beat.ernst@unibas.ch](mailto:beat.ernst@unibas.ch) (B. Ernst).

<http://dx.doi.org/10.1016/j.pisc.2016.10.002>

2213-0209/© 2016 Published by Elsevier GmbH. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

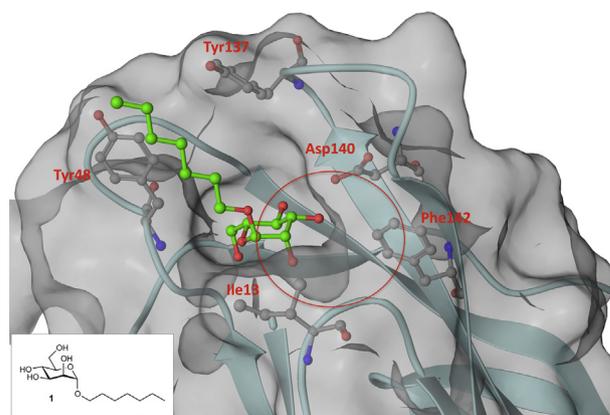
## Introduction

Urinary tract infections (UTIs) are among the most prevalent bacterial infections affecting millions of people (Foxman et al., 2000). They are mainly associated with uropathogenic *Escherichia coli* (UPEC) (Roland, 2002). Currently, the first-line treatment involves antibiotics (Hooton et al., 2004; Fihn, 2003) which can induce resistance, especially when frequently applied (Sanchez et al., 2012). Therefore, novel and efficient non-antibiotic approaches are urgently needed.

In the first step of the infection cycle, UPEC attach to urothelial cells of the host by means of the bacterial adhesin called FimH, which is located at the tip of the approximately 300 bacterial type 1 pili (Mulvey et al., 2000; Schilling et al., 2001). This allows UPEC to evade elimination from the host organism by the bulk flow of the urine. FimH is composed of a lectin domain (FimH<sub>LD</sub>) containing a carbohydrate recognition domain (CRD) and a pilin domain (FimH<sub>PD</sub>) regulating the switch between the high and low affinity states of the CRD (Le Trong et al., 2010).

More than thirty years ago, Firon et al. (1982, 1983, 1987) reported on aryl  $\alpha$ -D-mannosides abolishing FimH-mediated aggregation of UPEC with mannan-containing yeast cells (*Saccharomyces cerevisiae*) in *in vitro* assays. Over the course of the last few years, a range of highly potent monovalent antagonists consisting of a mannose moiety and a lipophilic aglycone was reported (Bouckaert et al., 2005; Sperling et al., 2006; Han et al., 2010; Klein et al., 2010; Cusumano et al., 2011; Han et al., 2012; Pang et al., 2012; Jiang et al., 2012; Schwardt et al., 2011; Kleeb et al., 2015; Brument et al., 2013; Jarvis et al., 2016; Chalopin et al., 2016). The various aglycones provide hydrophobic contacts or  $\pi$ - $\pi$  stacking interactions to amino acids forming the entrance to the mannose binding pocket. This entrance called 'tyrosine gate' is composed of two tyrosines and one isoleucine. However, the pharmacokinetic properties, e.g., solubility and/or permeability, of most of the reported FimH antagonists are not suitable for an oral application. For physicochemical and pharmacokinetic reasons, the numerous reported multivalent FimH antagonists (Lindhorst et al., 1998; Nagahori et al., 2002; Appeldoorn et al., 2005; Patel and Lindhorst, 2006; Touaibia et al., 2007; Durka et al., 2011; Bouckaert et al., 2013) are rather suited for the therapy of *E. coli* induced colitis ulcerosa, a form of inflammatory bowel disease (Barnich et al., 2007; Carvalho et al., 2009).

When interacting with FimH, the mannose moiety establishes a perfect hydrogen bond network (Hung et al., 2002). Since every hydroxyl group of mannose is part of this network, the removal/replacement of individual various hydroxyl groups or the replacement of the whole mannose moiety by other hexoses (e.g., glucose, galactose, fructose) resulted in a significant loss of affinity (Bouckaert et al., 2005; Han et al., 2010; Old, 1972; Fiege et al., 2015). Moreover, recently reported 1-C-branched mannose derivatives bearing additional equatorial groups at the anomeric carbon also showed reduced activity compared to methyl  $\alpha$ -D-mannoside (Gloe et al., 2015). In contrast, when the anomeric oxygen was replaced by carbon or nitrogen,



**Figure 1** The crystal structure of FimH<sub>LD</sub> co-crystallized with *n*-heptyl  $\alpha$ -D-mannopyranoside (**1**, PDB ID: 4BUQ) (Fiege et al., 2015). A mainly hydrophobic cavity formed by Ile13, Phe142 and Asp140 is located next to the entrance of the mannose binding site and can be reached by equatorial substituents in the 2-position of the mannose moiety.

nanomolar affinity could still be reached (Schwardt et al., 2011; Brument et al., 2013; Chalopin et al., 2016).

A visual inspection of the crystal structure of FimH<sub>LD</sub> co-crystallized with *n*-heptyl  $\alpha$ -D-mannoside (**1**, PDB ID: 4BUQ) (Fiege et al., 2015) revealed a previously unexplored hydrophobic cavity formed by Ile13, Phe142 and Asp140, which is located close to the entrance to the mannose-binding pocket (Fig. 1). By extending the 2-position of the mannose moiety with equatorial substituents ( $\rightarrow$  derivatives **2a–k**, Fig. 2), an interaction with the hydrophobic cavity should become possible.

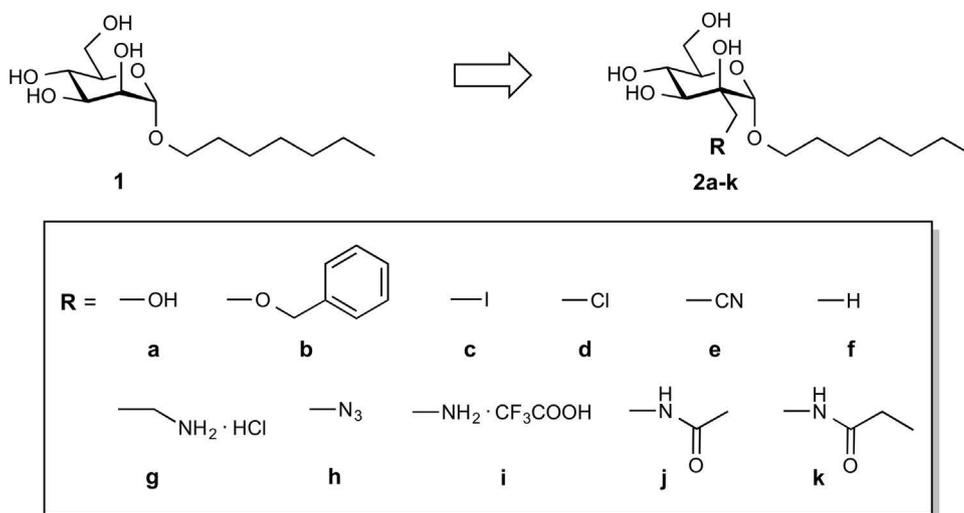
An adaption of the synthetic pathway of previously reported 2-C-branched mannose derivatives, in which the 2-position is modified at an early stage, lead to rather laborious approaches (Mitchell et al., 2007). We therefore planned a more convergent synthesis with a more flexible introduction of aglycones as well as equatorial substituents in the 2-position.

## Result and discussion

The synthetic route to 2-C-branched FimH antagonists fulfils two requirements: The facile introduction of various aglycones as well as various equatorial C-substituents in the 2-C-position of the mannose moiety.

## Synthesis

The synthesis of the 2-C-branched mannoside donor **5** is depicted in Scheme 1. The 2-C-modified D-mannofuranose **3** was synthesized according to a literature procedure starting from commercially available D-mannose (Witczak et al., 1984). Selective benzylation of the hydroxymethyl group using dibutyltin oxide (Malleron and David, 1998) followed by cleavage of the acetonides under acidic conditions yielded the 2-C-branched D-mannopyranose **4** (Waschke et al., 2011). For its perbenzylation with benzoyl chloride in presence of a catalytic amount of 4-dimethylamino-pyridine (DMAP) in dry pyridine, elevated temperature



**Figure 2** Modifications of the mannose moiety of *n*-heptyl  $\alpha$ -D-mannopyranoside (**1**) by equatorial substituents in the 2-position.

(110–120 °C) had to be applied. Subsequently, the glycosyl donor **5** was obtained by reaction with thiophenol using  $\text{BF}_3 \cdot \text{Et}_2\text{O}$  as a promoter. To couple donor **5** with 1-heptanol, different promoters were tested. Whereas with NIS/TMSOTf or NIS/TfOH donor **5** was only partially consumed after 24 h, it reacted within minutes in the presence of commercially available *p*-nitrobenzenesulfonyl chloride (*p*- $\text{NO}_2\text{PhSCL}$ ) accompanied by silver triflate ( $\text{AgOTf}$ ) (Crich et al., 2008). Apart from the desired  $\alpha$ -anomeric mannoside **6** (32%), the 2-OH deprotected  $\alpha$ -anomer **7** (32%) and the 2-OH deprotected  $\beta$ -anomer **8** (5%) were obtained as well. As silver triflate-mediated glycosylation has been reported to lead to partial transesterification affecting acetyl groups at the 2-O-position, low stereoselectivity of the glycosylation reaction was not unexpected (Ziegler et al., 1990; Nukada et al., 1999; Murakami et al., 2007).

To functionalize the equatorial substituent in the 2-C-position, **6** was debenzylated by catalytic hydrogenolysis to afford the primary alcohol **9**. However, attempts to mesylate its primary hydroxyl group failed. Since we attributed the low reactivity of the hydroxyl group in **9** to steric hindrance, we switched to mannoside **7** with an unprotected axial hydroxyl group in the 2-position.

Indeed, after hydrogenolysis of **7** ( $\rightarrow$  **11**), we were able to selectively mesylate the primary hydroxyl groups ( $\rightarrow$  **12**). However, displacement of the mesylate by fluoride using KF in aprotic solvent in presence of crown ether at elevated temperature afforded epoxide **14** instead of the desired fluoride **13**. Under these reaction conditions, the strongly basic fluoride is obviously deprotonating the axial hydroxyl group followed by conversion of mesylate **12** into the epoxide **14** by an intramolecular  $\text{S}_{\text{N}}2$  mechanism. With an excess of LiCl, epoxide **14** could be opened, leading to the chloride **15** in 41% yield. By acting as Lewis acid, the large excess of lithium ions can facilitate opening of the epoxide. Apart from the epoxide route, substituents can be introduced directly by nucleophilic substitution (see Scheme 2). However, prevalence of one mechanism over the other may depend on the nucleophile, i.e., its nucleophilicity and basicity as well as temperature and concentration.

The synthesis of a series of 2-C-branched FimH antagonists is depicted in Scheme 2. Debenzylation of **6** under Zemplén conditions ( $\rightarrow$  **2b**) followed by  $\text{Pd}(\text{OH})_2$ -catalyzed hydrogenolysis afforded test compound **2a**. The configuration at the anomeric carbon of deprotected derivative **2a** ( $^1J_{\text{H,C}} = 169 \text{ Hz}$ ) was unambiguously confirmed by the  $^{13}\text{C}$ – $^1\text{H}$  coupling constant of the anomeric nuclei using uncoupled  $^{13}\text{C}$  NMR. In general, the coupling constant for the equatorial anomeric proton amounts to  $\sim 170 \text{ Hz}$ , while a value of  $\sim 160 \text{ Hz}$  is indicative for an anomeric proton in axial orientation (Bubb, 2003).

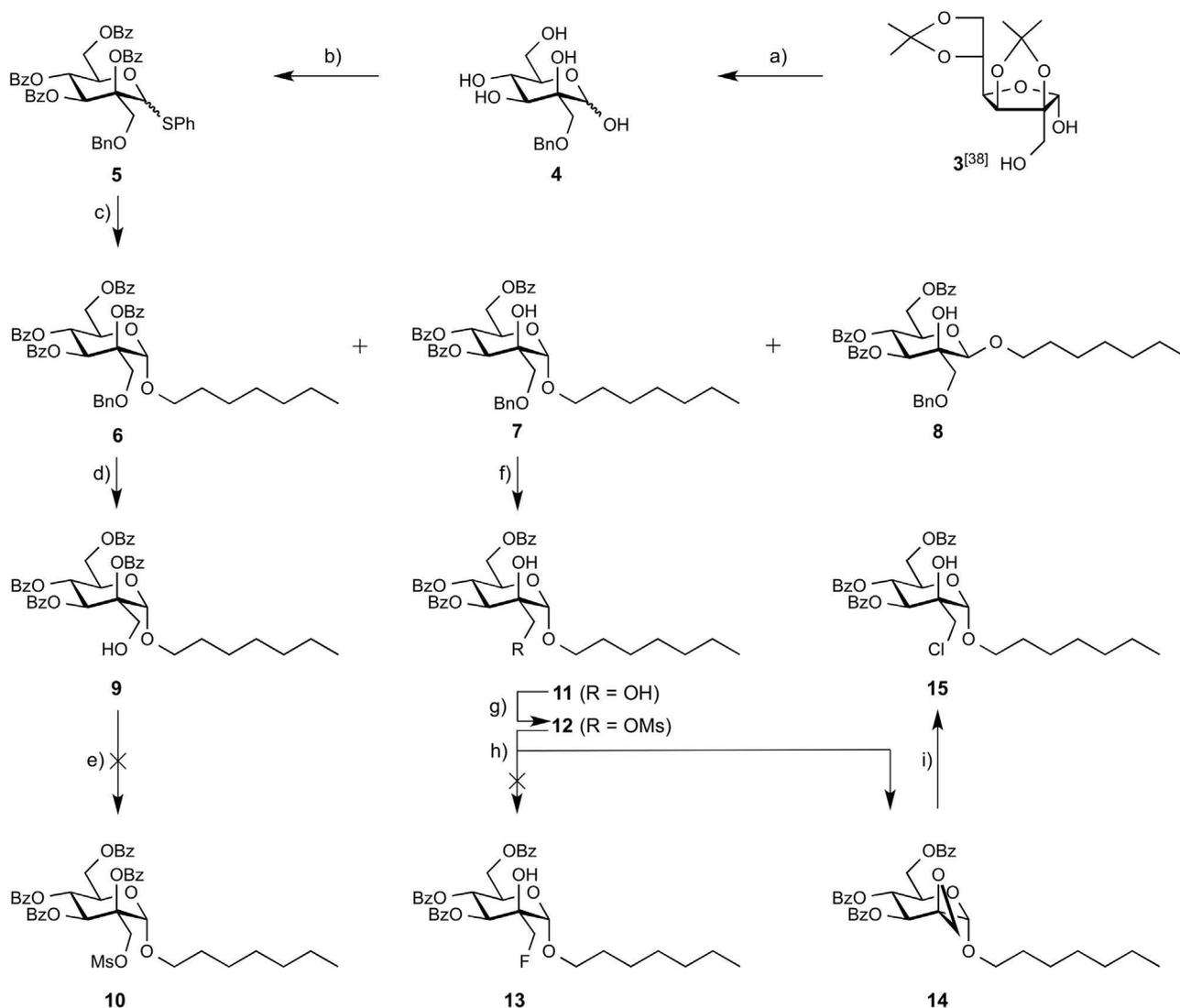
Upon mesylation of **11**, chloride was introduced using LiCl followed by deprotection of the intermediate with sodium methoxide to afford derivative **2d**. Direct introduction of chloride starting from mesylate was faster and gave a higher yield than the already discussed opening of epoxide **14**. Since only traces of epoxide were observed by TLC control,  $\text{S}_{\text{N}}2$  reaction seems to be the prevailing mechanism in this particular case. However, more basic nucleophiles may lead to different results.

Using an identical synthetic approach, iodide **2c**, cyanide **2e** and azide **2h** were obtained with NaI, KCN and  $\text{NaN}_3$  as nucleophiles. Hydrogenation of iodide **2c**, cyanide **2e** and azide **2h** in presence of  $\text{Pd}(\text{OH})_2$  on carbon yielded the methyl derivative **2f** and amine derivatives **2g** and **2i**, respectively. In addition, when **2h** was hydrogenated and subsequently acylated with acetyl chloride or propionyl chloride followed by deacetylation under Zemplén conditions, amides **2j** and **2k** were obtained.

To evaluate the  $\beta$ -anomeric derivative as well in the biological assay,  $\beta$ -mannoside **8** was debenzoylated ( $\rightarrow$  **16b**, Scheme 3) followed by hydrogenolysis to yield **16a** ( $^1J_{\text{H,C}} = 159 \text{ Hz}$ ).

### Affinity and thermodynamic profile

The affinities of the 2-C-branched mannosides were determined in a cell-free competitive binding assay (Table 1) (Rabbani et al., 2010). The assay uses FimH<sub>LD</sub>-Th-His<sub>6</sub> (Th: thrombin cleavage site) as a target protein and a

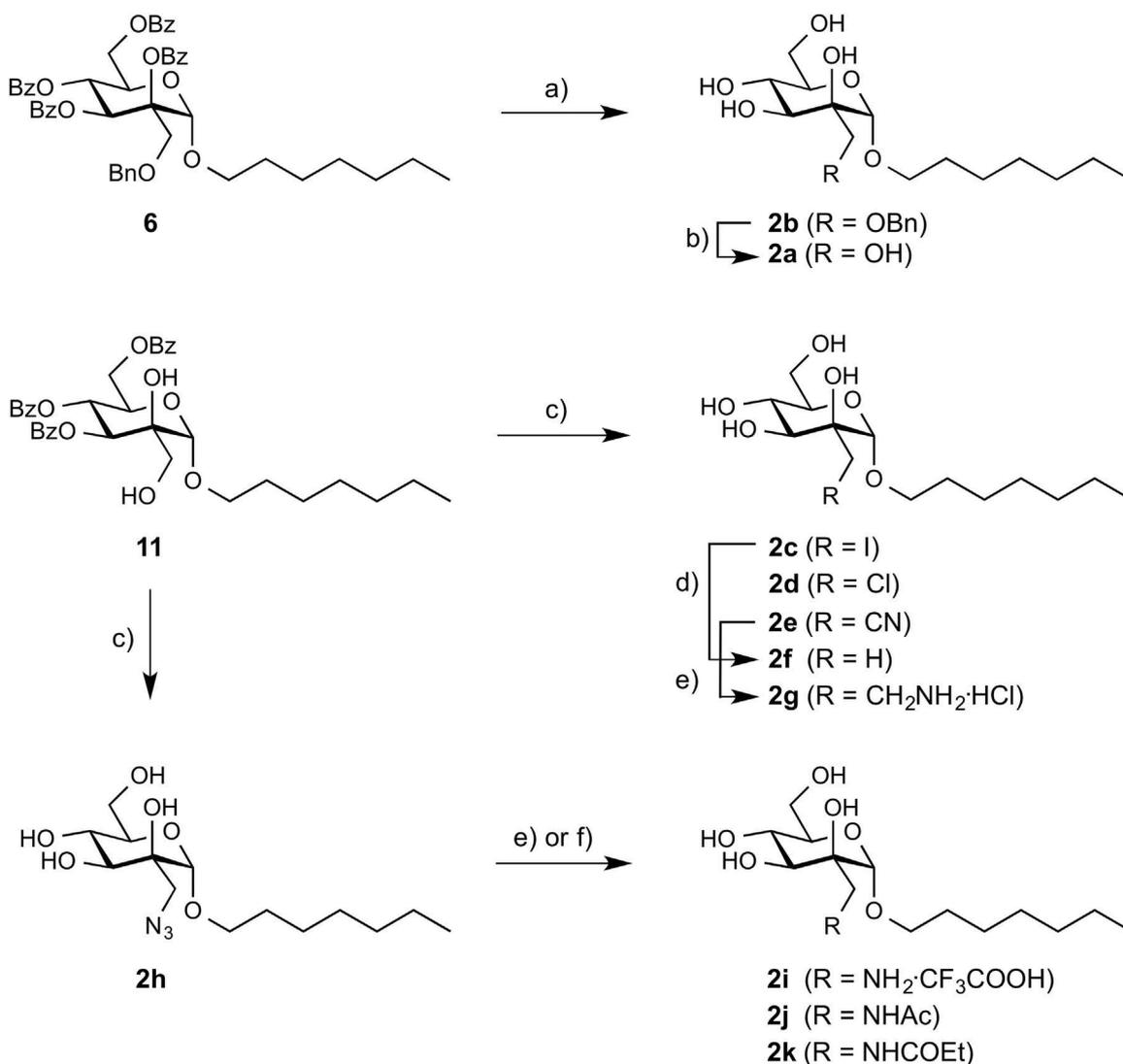


**Scheme 1** (a) i. BnBr, Bu<sub>2</sub>SnO, TBAB, toluene, 120 °C → 80 °C, 24 h; ii. Amberlyst-15 (H<sup>+</sup>), EtOH, H<sub>2</sub>O, 50 °C, 43 h, 90%; (b) i. BzCl, DMAP, pyridine, 120 °C, 48 h; ii. PhSH, BF<sub>3</sub>·Et<sub>2</sub>O, DCM, 0 °C → rt, 24 h, 42%; (c) 1-heptanol, *p*-NO<sub>2</sub>PhSCL, AgOTf, DCM, 4 Å MS, 0 °C → rt, 5 h, 32% for **6**, 32% for **7**, 5% for **8**; (d) Pd(OH)<sub>2</sub>/C, H<sub>2</sub>, EtOH, AcOH, rt, 22 h, 86%; (e) MsCl, TEA, DCM, 0 °C → 65 °C, 7 h, 0%; (f) Pd(OH)<sub>2</sub>/C, H<sub>2</sub>, EtOH, rt, 5 h, 93%; (g) MsCl, TEA, DCM, 0 °C → rt, 6 h, 86%; (h) KF, Kryptofix 2.2.2, DMSO, 100 °C, 2 h, 85%; (i) LiCl, DMF, 90 °C, 7 h, 41%.

biotinylated polyacrylamide glycopolymer as competitor. Conjugation of biotin with streptavidin-horseradish peroxidase allows quantification of the bound polymer and therefore the determination of the IC<sub>50</sub>. The activity of all antagonists was measured twice in duplicates. The antagonist *n*-heptyl α-D-mannopyranoside (**1**) was used as a reference compound and tested in parallel to ensure comparability. The affinities are referred to the activity of **1** as rIC<sub>50</sub>.

In addition to the competitive binding assay, ITC experiments were performed with mannosides **1**, **2a** and **2f** to reveal a thermodynamic fingerprint of mannose-modified FimH antagonists (Table 2). ITC enables direct measurement of the dissociation constant (*K*<sub>D</sub>) and the change in enthalpy ( $\Delta H^\circ$ ), which are further used to calculate the changes in free energy ( $\Delta G^\circ$ ) and entropy ( $\Delta S^\circ$ ) (Chen and Wadsö, 1982; Freire et al., 1990).

Unfortunately, all 2-C modifications proved to be detrimental to the affinity. Already the smallest substituent, a methyl group (→ **2f**), resulted in a 2.8-fold higher IC<sub>50</sub> value. A comparable 4.1-fold drop in activity was observed in ITC. This finding might be explained by an unexpected unfavourable steric clash of Ile13 and/or Phe142 with the methyl group already too big to fit to the targeted cavity. This hypothesis is supported by the considerably improved entropy term ( $-T\Delta\Delta S^\circ -17.2$  kJ/mol) compared to the reference **1**, indicating an increased conformational flexibility of the ligand. The resulting disruption of the hydrogen bond network within the pocket is also reflected by a substantial decline of enthalpy ( $\Delta\Delta H^\circ +20.6$  kJ/mol). A further reduction in affinity to the micromolar level for larger substituents, e.g., iodomethyl and chloromethyl (→ **2c** and **2d**, Table 1) is in full agreement with this argumentation.

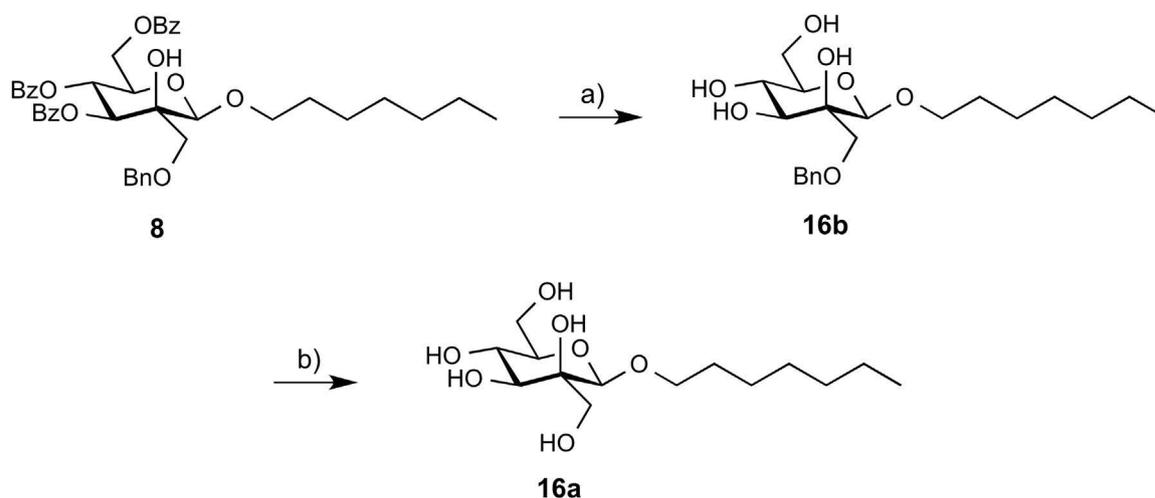


**Scheme 2** (a) MeONa/MeOH, rt, 4 h, 85%; (b) Pd(OH)<sub>2</sub>/C, H<sub>2</sub>, EtOH, rt, 6 h, 75%; (c) i. MsCl, TEA, DCM, 0 °C → rt, 1.5–6 h; ii. NaI, LiCl, KCN or NaN<sub>3</sub>, DMF or DMSO, 70–100 °C, 3–64 h; iii. MeONa/MeOH, rt, 1–6 h, 39% for **2c**, 38% for **2d**, 11% for **2e**, 73% for **2h**; (d) Pd(OH)<sub>2</sub>/C, H<sub>2</sub>, EtOH, TEA, rt, 44 h, 91%; (e) i. Pd(OH)<sub>2</sub>/C, H<sub>2</sub>, MeOH or EtOH, rt, 3–11 h; ii. 0.1% TFA or 0.01 M HCl, H<sub>2</sub>O, MeOH, 76% for **2g**, 86% for **2i**; (f) i. Pd(OH)<sub>2</sub>/C, H<sub>2</sub>, EtOH, rt, 7 h; ii. AcCl or CH<sub>3</sub>CH<sub>2</sub>COCl, pyridine, DCM, rt, 2.5–5.5 h; iii. MeONa/MeOH, rt, 1.5–4 h, 20% for **2j**, 38% for **2k**.

Unexpectedly, the benzyloxymethyl group in **2b**, despite its bulkiness, only slightly reduced the activity compared to the methyl substituent (→**2f**). Furthermore, **2b** performed better compared to the halogens **2c** and **2d**. This may result from a smaller van der Waals radius of oxygen compared to chloride or iodide. Moreover, a solvent exposed phenyl ring can be involved in non-specific hydrophobic interactions with the surface of the protein, attenuating the negative effect of the size of the 2-C-branch.

The antagonists bearing hydrogen bond donating groups, such as a hydroxyl group (→**2a**), an amine (→**2i**) or an amide (→**2j** and **2k**), were among the most active derivatives. The hydroxymethyl group (→**2a**) led to roughly a 5-fold drop in affinity in both, competitive binding assay and ITC. Compared to the methyl group (→**2f**), this substituent was expected to disrupt the hydrogen bond network even further due to its larger size and to cause additional

enthalpy costs due to a desolvation penalty related to the hydroxyl group. However, the enthalpy loss in this case was smaller ( $\Delta\Delta H^\circ_{2a-2f} = 6.1$  kJ/mol) implying additional beneficial interactions formed by **2a**. As a consequence, the entropy gain was limited ( $-T\Delta\Delta S^\circ = -10.4$  kJ/mol) compared to **2f**. However, this beneficial effect was almost compensated by a loss in entropy. Furthermore, the structurally similar aminomethyl derivative **2i** was the most active compound within the series with only 2.1-fold lower affinity compared to reference **1**. Presumably, the hydroxymethyl (→**2a**) and aminomethyl (→**2i**) groups are involved in electrostatic interactions with a hydrogen bond acceptor, i.e., Asp140 or the backbone amide of Ile13. The improved affinity of **2i** may result from the fact that the ammonium group in **2i** can form a slightly stronger interaction (Lopes Jesus and Redinha, 2011). Finally, when we incorporated a longer linker between the nitrogen and the sugar moiety (→**2g**)

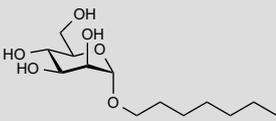
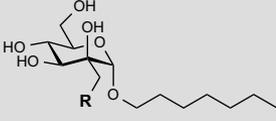
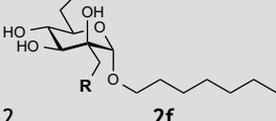


**Scheme 3** (a) MeONa/MeOH, rt, 2 h, 68%; (b) Pd(OH)<sub>2</sub>/C, H<sub>2</sub>, EtOH, rt, 3 h, 93%.

**Table 1** Affinity of FimH antagonists. The rIC<sub>50</sub> values were calculated by dividing the IC<sub>50</sub> of the compound of interest by the IC<sub>50</sub> of the reference compound *n*-heptyl  $\alpha$ -D-mannopyranoside (**1**). rIC<sub>50</sub> values below 1.0 are obtained for antagonists more active than reference compound **1**, whereas rIC<sub>50</sub> above 1.0 are obtained for antagonists less active than reference compound **1**.

Entry	Compound	R	IC <sub>50</sub> [nM]	rIC <sub>50</sub>
1	<b>1</b>		48.9–65.2	1
2	<b>2a</b>	–OH	311.0	4.9
3	<b>2b</b>		340.9	5.4
4	<b>2c</b>	–I	>2000	38.9
5	<b>2d</b>	–Cl	>2000	304
6	<b>2e</b>	–CN	522.3	9.7
7	<b>2f</b>	–H	181.2	2.8
8	<b>2g</b>	–CH <sub>2</sub> NH <sub>2</sub> ·HCl	>2000	37.1
9	<b>2h</b>	–N <sub>3</sub>	437.0	6.9
10	<b>2i</b>	–NH <sub>2</sub> ·CF <sub>3</sub> COOH	135.0	2.1
11	<b>2j</b>		182.5	3.4
12	<b>2k</b>		295.2	5.5
13	<b>16a</b>	–OH	446.4	7.0
14	<b>16b</b>		4400	69.4

**Table 2** Thermodynamic profile of FimH antagonists in binding to FimH<sub>LD</sub>-Th-His<sub>6</sub> at 25 °C and pH 7.4.

Entry	Compound	R	K <sub>D</sub> [nM]	ΔG° [kJ/mol]	ΔH° [kJ/mol]	−TΔS° [kJ/mol]
1			28.9	−43.0	−50.3	7.3
2		−H	118.6	−39.5	−29.7	−9.9
3		−OH	154.0	−38.9	−35.8	−3.1

the affinity was lowered to the micromolar level. Similarly to the aminomethyl derivative **2i**, the amide **2j** had one of the best affinities in the series. However, elongation of the aliphatic chain attached to the amide ( $\rightarrow$ **2k**) again resulted in reduction of potency.

Quite surprisingly, the  $\beta$ -anomeric derivative **16a** (Table 1, entry 13) performed only slightly worse than its  $\alpha$ -anomeric analogue **2a**. However, when a benzyloxymethyl group was introduced ( $\rightarrow$ **16b**, entry 14) the affinity was almost 13-fold lower than for its  $\alpha$ -anomeric counterpart **2b**, according to docking studies (data not shown) due to a steric clash with Ile13.

## Conclusions

A new family of mannose-based FimH antagonists equipped with equatorial substituents at the 2-position of the sugar moiety was designed and synthesized to target a cavity located close to the entrance of FimH-CRD. Only when the axial 2-hydroxyl group was unprotected ( $\rightarrow$ **11**), the otherwise unsuccessful substitution at the equatorial 2-position of the mannose moiety could be performed, leading to the test compounds **2c–k**. In one case, the intermediate epoxide **14** could be isolated, indicating two possible reaction pathways; one via direct S<sub>N</sub>2-substitution and one via the epoxide **14**. The resulting epoxide intermediate could be opened, however, affording only a moderate yield.

The activities of the 2-C-branched FimH antagonists were evaluated in a cell-free competitive binding assay and compared to the reference compound *n*-heptyl  $\alpha$ -D-mannopyranoside (**1**). None of the modifications proved to be advantageous for binding to FimH-CRD. The loss of affinity is probably related to steric hindrance as it was already observed upon introduction of the smallest substituent, a methyl group ( $\rightarrow$ **2f**). With hydrogen bond donating substituents ( $\rightarrow$ **2a**, **2i–k**), affinity could be partially regained. Unexpectedly, the  $\beta$ -anomer **16a** performed only slightly worse than its  $\alpha$ -anomeric counterpart **2a**. However, as already experienced in the  $\alpha$ -series, a larger benzyloxymethyl substituent ( $\rightarrow$ **16b**) severely compromised affinity.

Finally, ITC experiments with the selected antagonists **1**, **2a** and **2f** revealed a drastic enthalpy loss for

the 2-C-branched antagonists, which, however, is partially compensated by an entropy gain. This supports the hypothesis that the target cavity is too small to accommodate 2-C-substituents. However, with larger, hydrogen bond donating substituents the enthalpy loss could be substantially reduced.

## Acknowledgement

Financial support for PZ from the Swiss National Science Foundation (200020\_146202) is gratefully acknowledged.

## Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.pisc.2016.10.002>.

## References

- Appeldoorn, C.C.M., Joosten, J.A.F., el Maate, F.A., Dobrindt, U., Hacker, J., Liskamp, R.M.J., Khan, A.S., Pieters, R.J., 2005. Novel multivalent mannose compounds and their inhibition of the adhesion of type 1 fimbriated uropathogenic *E. coli*. *Tetrahedron: Asymmetry* 16, 361–372. <http://dx.doi.org/10.1016/j.tetasy.2004.11.014>.
- Barnich, N., Carvalho, F.A., Glasser, A.-L., Darcha, C., Jantschkeff, P., Allez, M., Peeters, H., Bommelaer, G., Desreumaux, P., Colombel, J.-F., Darfeuille-Michaud, A., 2007. CEACAM6 acts as a receptor for adherent-invasive *E. coli*, supporting ileal mucosa colonization in Crohn disease. *J. Clin. Invest.* 117, 1566–1574. <http://dx.doi.org/10.1172/jci30504>.
- Bouckaert, J., Berglund, J., Schembri, M., De Genst, E., Cools, L., Wuhrer, M., Hung, C.-S., Pinkner, J., Slättergard, R., Zavialov, A., Choudhury, D., Langermann, S., Hultgren, S.J., Wyns, L., Klemm, P., Oscarson, S., Knight, S.D., Greve, H.D., 2005. Receptor binding studies disclose a novel class of high-affinity inhibitors of the *Escherichia coli* FimH adhesin. *Mol. Microbiol.* 55, 441–455. <http://dx.doi.org/10.1111/j.1365-2958.2004.04415.x>.
- Bouckaert, J., Li, Z., Xavier, C., Almant, M., Caveliers, V., Lahoutte, T., Weeks, S.D., Kovensky, J., Guoin, S.G., 2013. Heptyl  $\alpha$ -D-mannosides grafted on a  $\beta$ -cyclodextrin core to interfere with *Escherichia coli* adhesion: an

- in vivo multivalent effect. *Chem. Eur. J.* 19, 7847–7855, <http://dx.doi.org/10.1002/chem.201204015>.
- Brument, S., Sivignon, A., Dumych, T.I., Moreau, N., Roos, G., Guérardel, Y., Chalopin, T., Deniaud, D., Bilyy, R.O., Darfeuille-Michaud, A., Bouckaert, J., Gouin, S.G., 2013. Thiazolylaminomannosides as potent antiadhesives of type 1 piliated *Escherichia coli* isolated from Crohn's disease patients. *J. Med. Chem.* 56, 5395–5406, <http://dx.doi.org/10.1021/jm400723n>.
- Bubb, W.A., 2003. NMR spectroscopy in the study of carbohydrates: characterizing the structural complexity. *Concepts Magn. Reson.* A 19A, 1–19, <http://dx.doi.org/10.1002/cmr.a.10080>.
- Carvalho, F.A., Barnich, N., Sivignon, A., Darcha, C., Chan, C.H.F., Stanners, C.P., Darfeuille-Michaud, A., 2009. Crohn's disease adherent-invasive *Escherichia coli* colonize and induce strong gut inflammation in transgenic mice expressing human CEACAM. *J. Exp. Med.* 206, 2179–2189, <http://dx.doi.org/10.1084/jem.20090741>.
- Chalopin, T., Alvarez Dorta, D., Sivignon, A., Caudan, M., Dumych, T.I., Bilyy, R.O., Deniaud, D., Barnich, N., Bouckaert, J., Gouin, S.G., 2016. Second generation of thiazolylmannosides, FimH antagonists for *E. coli*-induced Crohn's disease. *Org. Biomol. Chem.* 14, 3913–3925, <http://dx.doi.org/10.1039/c6ob00424e>.
- Chen, A.-t., Wadsö, I., 1982. Simultaneous determination of  $\Delta G$ ,  $\Delta H$  and  $\Delta S$  by an automatic microcalorimetric titration technique. Application to protein ligand binding. *J. Biochem. Biophys. Methods* 6, 307–316, [http://dx.doi.org/10.1016/0165-022X\(82\)90012-4](http://dx.doi.org/10.1016/0165-022X(82)90012-4).
- Crich, D., Cai, F., Yang, F., 2008. A stable, commercially available sulfonyl chloride for the activation of thioglycosides in conjunction with silver trifluoromethanesulfonate. *Carbohydr. Res.* 343, 1858–1862, <http://dx.doi.org/10.1016/j.carres.2008.03.002>.
- Cusumano, C.K., Pinkner, J.S., Han, Z., Greene, S.E., Ford, B.A., Crowley, J.R., Henderson, J.P., Janetka, J.W., Hultgren, S.J., 2011. Treatment and prevention of urinary tract infection with orally active FimH inhibitors. *Sci. Transl. Med.* 3, 109ra115, <http://dx.doi.org/10.1126/scitranslmed.3003021>.
- Durka, M., Buffet, K., Iehl, J., Holler, M., Nierengarten, J.-F., Taganna, J., Bouckaert, J., Vincent, S.P., 2011. The functional valency of dodecamannosylated fullerenes with *Escherichia coli* FimH—towards novel bacterial antiadhesives. *Chem. Commun.* 47, 1321–1323, <http://dx.doi.org/10.1039/c0cc04468g>.
- Fiege, B., Rabbani, S., Preston, R.C., Jakob, R.P., Zihlmann, P., Schwarzt, O., Jiang, X., Maier, T., Ernst, B., 2015. The tyrosine gate of the bacterial lectin FimH: a conformational analysis by NMR spectroscopy and X-ray crystallography. *ChemBioChem* 16, 1235–1246, <http://dx.doi.org/10.1002/cbic.201402714>.
- Fihn, S.D., 2003. Acute uncomplicated urinary tract infection in women. *N. Engl. J. Med.* 349, 259–266, <http://dx.doi.org/10.1056/nejmcp030027>.
- Firon, N., Ofek, I., Sharon, N., 1982. Interaction of mannose-containing oligosaccharides with the fimbrial lectin of *Escherichia coli*. *Biochem. Biophys. Res. Commun.* 105, 1426–1432, [http://dx.doi.org/10.1016/0006-291x\(82\)90947-0](http://dx.doi.org/10.1016/0006-291x(82)90947-0).
- Firon, N., Ofek, I., Sharon, N., 1983. Carbohydrate specificity of the surface lectins of *Escherichia coli*, *Klebsiella pneumoniae*, and *Salmonella typhimurium*. *Carbohydr. Res.* 120, 235–249, [http://dx.doi.org/10.1016/0008-6215\(83\)88019-7](http://dx.doi.org/10.1016/0008-6215(83)88019-7).
- Firon, N., Ashkenazi, S., Mirelman, D., Ofek, I., Sharon, N., 1987. 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.* 55, 472–476.
- Foxman, B., Barlow, R., D'Arcy, H., Gillespie, B., Sobel, J.D., 2000. Urinary tract infection: self reported incidence and associated costs. *Ann. Epidemiol.* 10, 509–515, [http://dx.doi.org/10.1016/s1047-2797\(00\)00072-7](http://dx.doi.org/10.1016/s1047-2797(00)00072-7).
- Freire, E., Mayorga, O.L., Straume, M., 1990. Isothermal titration calorimetry. *Anal. Chem.* 62, 950A–959A, <http://dx.doi.org/10.1021/ac00217a002>.
- Gloe, T.-E., Stamer, I., Hojnik, C., Wrodnigg, T.M., Lindhorst, T.K., 2015. Are D-manno-configured Amadori products ligands of the bacterial lectin FimH? *Beilstein J. Org. Chem.* 11, 1096–1104, <http://dx.doi.org/10.3762/bjoc.11.123>.
- Han, Z., Pinkner, J.S., Ford, B., Obermann, R., Nolan, W., Wildman, S.A., Hobbs, D., Ellenberger, T., Cusumano, C.K., Hultgren, S.J., Janetka, J.W., 2010. Structure-based drug design and optimization of mannoside bacterial FimH antagonists. *J. Med. Chem.* 53, 4779–4792, <http://dx.doi.org/10.1021/jm100438s>.
- Han, Z., Pinkner, J.S., Ford, B., Chorell, E., Crowley, J.M., Cusumano, C.K., Campbell, S., Henderson, J.P., Hultgren, S.J., Janetka, J.W., 2012. Lead optimization studies on FimH antagonists: discovery of potent and orally bioavailable ortho-substituted biphenyl mannosides. *J. Med. Chem.* 55, 3945–3959, <http://dx.doi.org/10.1021/jm300165m>.
- Hooton, T.M., Besser, R., Foxman, B., Fritsche, T.R., Nicolle, L.E., 2004. Acute uncomplicated cystitis in an era of increasing antibiotic resistance: a proposed approach to empirical therapy. *Clin. Infect. Dis.* 39, 75–80, <http://dx.doi.org/10.1086/422145>.
- Hung, C.-S., Bouckaert, J., Hung, D., Pinkner, J., Widberg, C., DeFusco, A., Auguste, C.G., Strouse, R., Langermann, S., Waksman, G., Hultgren, S.J., 2002. Structural basis of tropism of *Escherichia coli* to the bladder during urinary tract infection. *Mol. Microbiol.* 44, 903–915, <http://dx.doi.org/10.1046/j.1365-2958.2002.02915.x>.
- Jarvis, C., Han, Z., Kalas, V., Klein, R., Pinkner, J.S., Ford, B., Binkley, J., Cusumano, C.K., Cusumano, Z., Mydock-McGrane, L., Hultgren, S.J., Janetka, J.W., 2016. Antivirulence isoquinolone mannosides: optimization of the biaryl aglycone for FimH lectin binding affinity and efficacy in the treatment of chronic UTI. *ChemMedChem* 11, 367–373, <http://dx.doi.org/10.1002/cmdc.201600045>.
- Jiang, X., Abgottsson, D., Kleeb, S., Rabbani, S., Scharenberg, M., Wittwer, M., Haug, M., Schwarzt, O., Ernst, B., 2012. Antiadhesion therapy for urinary tract infections—a balanced PK/PD profile proved to be key for success. *J. Med. Chem.* 55, 4700–4713, <http://dx.doi.org/10.1021/jm300192x>.
- Kleeb, S., Pang, L., Mayer, K., Eris, D., Sigl, A., Preston, R.C., Zihlmann, P., Sharpe, T., Jakob, R.P., Abgottsson, D., Hutter, A.S., Scharenberg, M., Jiang, X., Navarra, G., Rabbani, S., Smieško, M., Lüdin, N., Bezençon, J., Schwarzt, O., Maier, T., Ernst, B., 2015. FimH antagonists: bioisosteres to improve the in vitro and in vivo PK/PD profile. *J. Med. Chem.* 58, 2221–2239, <http://dx.doi.org/10.1021/jm501524q>.
- Klein, T., Abgottsson, D., Wittwer, M., Rabbani, S., Herold, J., Jiang, X., Kleeb, S., Lüthi, C., Scharenberg, M., Bezençon, J., Gubler, E., Pang, L., Smieško, M., Cutting, B., Schwarzt, O., Ernst, B., 2010. FimH antagonists for the oral treatment of urinary tract infections: from design and synthesis to in vitro and in vivo evaluation. *J. Med. Chem.* 53, 8627–8641, <http://dx.doi.org/10.1021/jm101011y>.
- Le Trong, I., Aprikian, P., Kidd, B.A., Forero-Shelton, M., Tchenokova, V., Rajagopal, P., Rodriguez, V., Interlandi, G., Klevit, R., Vogel, V., Stenkamp, R.E., Sokurenko, E.V., Thomas, W.E., 2010. Structural basis for mechanical force regulation of the adhesin FimH via finger trap-like  $\beta$  sheet twisting. *Cell* 141, 645–655, <http://dx.doi.org/10.1016/j.cell.2010.03.038>.
- Lindhorst, T.K., Kieburg, C., Krallmann-Wenzel, U., 1998. Inhibition of the type 1 fimbriae-mediated adhesion of *Escherichia coli* to erythrocytes by multiantennary d-mannosyl clusters: the effect of multivalency. *Glycoconj. J.* 15, 605–613, <http://dx.doi.org/10.1023/a:1006920027641>.
- Lopes Jesus, A.J., Redinha, J.S., 2011. Charge-assisted intramolecular hydrogen bonds in disubstituted cyclohex-

- ane derivatives. *J. Phys. Chem. A* 115, 14069–14077, <http://dx.doi.org/10.1021/jp206193a>.
- Malleron, A., David, S., 1998. A preparation of protected 2-deoxy-2-hydroxymethyl-D-mannose and -D-glucose derivatives not involving organometallic reagents. *Carbohydr. Res.* 308, 93–98, [http://dx.doi.org/10.1016/s0008-6215\(98\)00080-9](http://dx.doi.org/10.1016/s0008-6215(98)00080-9).
- Mitchell, D.A., Jones, N.A., Hunter, S.J., Cook, J.M.D., Jenkinson, S.F., Wormald, M.R., Dwek, R.A., Fleet, G.W.J., 2007. Synthesis of 2-C-branched derivatives of D-mannose: 2-C-aminomethyl-D-mannose binds to the human C-type lectin DC-SIGN with affinity greater than an order of magnitude compared to that of D-mannose. *Tetrahedron: Asymmetry* 18, 1502–1510, <http://dx.doi.org/10.1016/j.tetasy.2007.06.003>.
- Mulvey, M.A., Schilling, J.D., Martinez, J.J., Hultgren, S.J., 2000. Bad bugs and beleaguered bladders: interplay between uropathogenic *Escherichia coli* and innate host defenses. *Proc. Natl. Acad. Sci. U. S. A.* 97, 8829–8835, <http://dx.doi.org/10.1073/pnas.97.16.8829>.
- Murakami, T., Hirono, R., Sato, Y., Furusawa, K., 2007. Efficient synthesis of  $\omega$ -mercaptoalkyl 1,2-*trans*-glycosides from sugar peracetates. *Carbohydr. Res.* 342, 1009–1020, <http://dx.doi.org/10.1016/j.carres.2007.02.024>.
- Nagahori, N., Lee, R.T., Nishimura, S.-L., Pagé, D., Roy, R., Lee, Y.C., 2002. Inhibition of adhesion of type 1 fimbriated *Escherichia coli* to highly mannose ligands. *ChemBioChem* 3, 836–844, [http://dx.doi.org/10.1002/1439-7633\(20020902\)3:9<836:aid-cbic836>3.0.co;2-2](http://dx.doi.org/10.1002/1439-7633(20020902)3:9<836:aid-cbic836>3.0.co;2-2).
- Nukada, T., Berces, A., Whitfield, D.M., 1999. Acyl transfer as a problematic side reaction in polymer-supported oligosaccharide synthesis. *J. Org. Chem.* 64, 9030–9045, <http://dx.doi.org/10.1021/jo990712b>.
- Old, D.C., 1972. Inhibition of the interaction between fimbrial haemagglutinins and erythrocytes by D-mannose and other carbohydrates. *J. Gen. Microbiol.* 71, 149–157, <http://dx.doi.org/10.1099/00221287-71-1-149>.
- Pang, L., Kleeb, S., Lemme, K., Rabbani, S., Scharenberg, M., Zalewski, A., Schädler, F., Schwardt, O., Ernst, B., 2012. FimH antagonists: structure-activity and structure-property relationships for biphenyl  $\alpha$ -D-mannopyranosides. *ChemMedChem* 7, 1404–1422, <http://dx.doi.org/10.1002/cmdc.201200125>.
- Patel, A., Lindhorst, T.K., 2006. A modular approach for the synthesis of oligosaccharide mimetics. *Carbohydr. Res.* 341, 1657–1668, <http://dx.doi.org/10.1016/j.carres.2006.01.024>.
- Rabbani, S., Jiang, X., Schwardt, O., Ernst, B., 2010. Expression of the carbohydrate recognition domain of FimH and development of a competitive binding assay. *Anal. Biochem.* 407, 188–195, <http://dx.doi.org/10.1016/j.ab.2010.08.007>.
- Roland, A., 2002. The etiology of urinary tract infection: traditional and emerging pathogens. *Am. J. Med.* 113 (Suppl. 1A), 14S–19S, [http://dx.doi.org/10.1016/s0002-9343\(02\)01055-0](http://dx.doi.org/10.1016/s0002-9343(02)01055-0).
- Sanchez, G.V., Master, R.N., Karlowky, J.A., Bordon, J.M., 2012. *In vitro* antimicrobial resistance of urinary *Escherichia coli* isolates among U.S. outpatients from 2000 to 2010. *Antimicrob. Agents Chemother.* 56, 2181–2183, <http://dx.doi.org/10.1128/aac.06060-11>.
- Schilling, J.D., Mulvey, M.A., Hultgren, S.J., 2001. Structure and function of *Escherichia coli* type 1 pili: new insight into the pathogenesis of urinary tract infections. *J. Infect. Dis.* 183 (Suppl. 1), S36–S40, <http://dx.doi.org/10.1086/318855>.
- Schwardt, O., Rabbani, S., Hartmann, M., Abgottspon, D., Wittwer, M., Kleeb, S., Zalewski, A., Smieško, M., Cutting, B., Ernst, B., 2011. Design, synthesis and biological evaluation of mannosyl triazoles as FimH antagonists. *Bioorg. Med. Chem.* 19, 6454–6473, <http://dx.doi.org/10.1016/j.bmc.2011.08.057>.
- Sperling, O., Fuchs, A., Lindhorst, T.K., 2006. Evaluation of the carbohydrate recognition domain of the bacterial adhesin FimH: design, synthesis and binding properties of mannoside ligands. *Org. Biomol. Chem.* 4, 3913–3922, <http://dx.doi.org/10.1039/b610745a>.
- Touaibia, M., Wellens, A., Shiao, T.C., Wang, Q., Sirois, S., Bouckaert, J., Roy, R., 2007. Mannosylated G(0) dendrimers with nanomolar affinities to *Escherichia coli* FimH. *ChemMedChem* 2, 1190–1201, <http://dx.doi.org/10.1002/cmdc.200700063>.
- Waschke, D., Thimm, J., Thiem, J., 2011. Highly efficient synthesis of ketoheptoses. *Org. Lett.* 13, 3628–3631, <http://dx.doi.org/10.1021/ol2012764>.
- Witczak, Z.J., Whistler, R.L., Daniel, J.R., 1984. Synthesis of 3-C-(hydroxymethyl)erythritol and 3-C-methylerythritol. *Carbohydr. Res.* 133, 235–245, [http://dx.doi.org/10.1016/0008-6215\(84\)85201-5](http://dx.doi.org/10.1016/0008-6215(84)85201-5).
- Ziegler, T., Kováč, P., Gludemans, C.P.J., 1990. Transesterification during glycosylation promoted by silver trifluoromethanesulfonate. *Eur. J. Org. Chem.*, 613–615, <http://dx.doi.org/10.1002/jlac.1990199001115>.



## Manuscript 5

### FimH Antagonists: Bioisosteres To Improve the in Vitro and in Vivo PK/PD Profile

Simon Kleeb,<sup>†#</sup> Lijuan Pang,<sup>†#</sup> Katharina Mayer,<sup>†#</sup> Deniz Eris,<sup>†#</sup> Anja Sigl,<sup>†#</sup>  
Roland C. Preston,<sup>†</sup> Pascal Zihlmann,<sup>†</sup> Timothy Sharpe,<sup>§</sup> Roman P. Jakob,<sup>‡</sup>  
Daniela Abgottspon,<sup>†</sup> Aline S. Hutter,<sup>†</sup> Meike Scharenberg,<sup>†</sup> Xiaohua Jiang,<sup>†</sup>  
Giulio Navarra,<sup>†</sup> Said Rabbani,<sup>†</sup> Martin Smiesko,<sup>†</sup> Nathalie Lüdin,<sup>†</sup> Jacqueline  
Bezençon,<sup>†</sup> Oliver Schwaradt,<sup>†</sup> Timm Maier,<sup>‡</sup> and Beat Ernst<sup>†\*</sup>

<sup>#</sup>These authors contributed equally to the project

<sup>†</sup>University of Basel, Institute of Molecular Pharmacy,  
Klingelbergstr. 50, 4056 Basel, Switzerland

<sup>‡</sup>University of Basel, Institute of Structural Biology,  
Klingelbergstr. 70, 4056 Basel, Switzerland

<sup>§</sup> University of Basel, Biophysics Facility,  
Klingelbergstr. 70, 4056 Basel, Switzerland

\*Corresponding author.

Tel.: 0041 (0)61 267 15 51; Fax: 0041 (0)61 267 15 52;  
E-mail: beat.ernst@unibas.ch

#### Contribution of Pascal Zihlmann:

- Thermodynamic profiling by ITC and data analysis



## FimH Antagonists: Bioisosteres To Improve the in Vitro and in Vivo PK/PD Profile

Simon Kleeb,<sup>†,||</sup> Lijuan Pang,<sup>†,||</sup> Katharina Mayer,<sup>†,||</sup> Deniz Eris,<sup>†,||</sup> Anja Sigl,<sup>†,||</sup> Roland C. Preston,<sup>†</sup> Pascal Zihlmann,<sup>†</sup> Timothy Sharpe,<sup>§</sup> Roman P. Jakob,<sup>‡</sup> Daniela Abgottspon,<sup>†</sup> Aline S. Hutter,<sup>†</sup> Meike Scharenberg,<sup>†</sup> Xiaohua Jiang,<sup>†</sup> Giulio Navarra,<sup>†</sup> Said Rabbani,<sup>†</sup> Martin Smiesko,<sup>†</sup> Nathalie Lüdin,<sup>†</sup> Jacqueline Bezençon,<sup>†</sup> Oliver Schwardt,<sup>†</sup> Timm Maier,<sup>‡</sup> and Beat Ernst<sup>\*,†</sup>

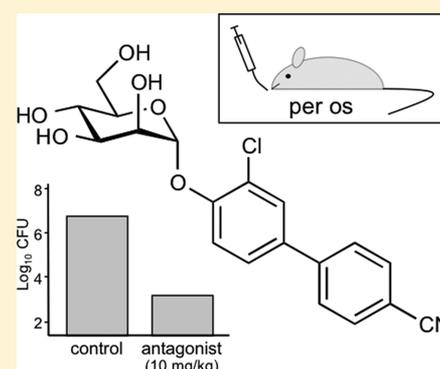
<sup>†</sup>Institute of Molecular Pharmacy, Pharmacenter, University of Basel, Klingelbergstrasse 50, CH-4056 Basel, Switzerland

<sup>‡</sup>Structural Biology, Biocenter, University of Basel, Klingelbergstrasse 70, CH-4056 Basel, Switzerland

<sup>§</sup>Biophysical Facility, Biocenter, University of Basel, Klingelbergstrasse 70, CH-4056 Basel, Switzerland

**S** Supporting Information

**ABSTRACT:** Urinary tract infections (UTIs), predominantly caused by uropathogenic *Escherichia coli* (UPEC), belong to the most prevalent infectious diseases worldwide. The attachment of UPEC to host cells is mediated by FimH, a mannose-binding adhesin at the tip of bacterial type 1 pili. To date, UTIs are mainly treated with antibiotics, leading to the ubiquitous problem of increasing resistance against most of the currently available antimicrobials. Therefore, new treatment strategies are urgently needed. Here, we describe the development of an orally available FimH antagonist. Starting from the carboxylate substituted biphenyl  $\alpha$ -D-mannoside **9**, affinity and the relevant pharmacokinetic parameters (solubility, permeability, renal excretion) were substantially improved by a bioisosteric approach. With 3'-chloro-4'-( $\alpha$ -D-mannopyranosyloxy)biphenyl-4-carbonitrile (**10j**) a FimH antagonist with an optimal in vitro PK/PD profile was identified. Orally applied, **10j** was effective in a mouse model of UTI by reducing the bacterial load in the bladder by about 1000-fold.

**■ INTRODUCTION**

Urinary tract infection (UTI) is one of the most frequent infectious diseases worldwide and affects millions of people every year.<sup>1</sup> In more than 70% of the reported cases, uropathogenic *Escherichia coli* (UPEC) is the causal pathogen.<sup>2</sup> Acute, uncomplicated lower urinary tract infection, commonly referred to as cystitis, requires an antibiotic treatment for symptom relief (i.e., reduction of dysuria, frequent and urgent urination, bacteriuria, pyuria) and for prevention of more devastating or even life threatening complications like pyelonephritis and urosepsis.<sup>3,4</sup> However, the repeated use of antibacterial chemotherapeutics provokes antimicrobial resistance leading to treatment failure.<sup>5</sup> Hence, a new approach for the prevention and treatment of UTI with orally applicable therapeutics is urgently needed.<sup>6</sup>

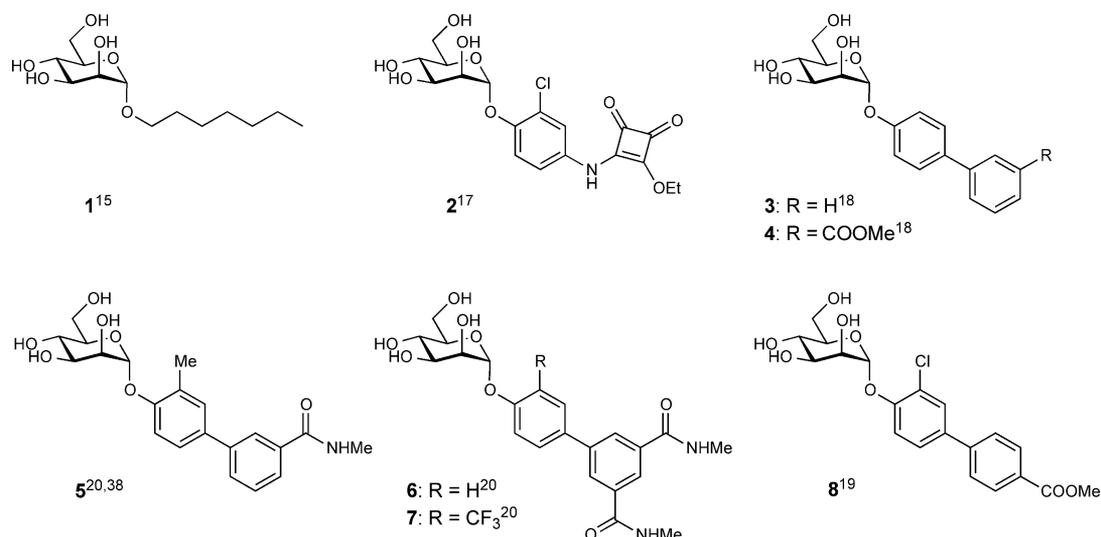
UPEC undergo a well-defined infection cycle within the host.<sup>7</sup> The key step in pathogenesis is bacterial adhesion to the epithelial cells in the lower urinary tract.<sup>8</sup> This interaction prevents UPEC from clearance by the bulk flow of urine and enables the bacteria to colonize the epithelial cells. The adhesion is mediated by the virulence factor FimH located at the tip of bacterial type 1 pili.<sup>9,10</sup> FimH consists of two immunoglobulin-like domains: the N-terminal lectin domain and (connected by a short linker) the C-terminal pilin domain.<sup>11</sup> The lectin domain encloses the carbohydrate recognition domain (CRD) that binds

to the oligomannosides of the glycoprotein uroplakin Ia on the epithelial cell surface.<sup>12</sup> The pilin domain anchors the adhesin to the pilus and regulates the switch between two conformational states of the CRD with high and low affinity for mannoses, respectively.

More than 3 decades ago, Sharon and co-workers described various oligomannosides and aryl  $\alpha$ -D-mannosides as potential antagonists of the FimH-mediated bacterial adhesion.<sup>13,14</sup> However, only weak interactions in the milli- to micromolar range were observed. In recent years, several high-affinity monovalent mannose-based FimH antagonists with various aglycones like *n*-alkyl,<sup>15</sup> phenyl,<sup>16</sup> dioxocyclobutenyl-aminophenyl,<sup>17</sup> umbelliferyl,<sup>16</sup> biphenyl,<sup>18–22</sup> indol(in)-ylphenyl,<sup>23</sup> triazolyl,<sup>24</sup> and thiazolylamino<sup>25</sup> have been reported. In addition, different multivalent presentations of the mannose have been synthesized<sup>26–32</sup> and a heptavalent presentation of *n*-heptyl  $\alpha$ -D-mannoside (**1**) tethered to  $\beta$ -cyclodextrin proved to be highly effective when applied together with the UTI89 bacterial strain through a catheter into the bladder of C3H/HeN mice.<sup>32</sup> Importantly, adverse side effects resulting from nonselective binding of FimH antagonists (they are all  $\alpha$ -D-mannopyrano-

Received: October 3, 2014

Published: February 10, 2015



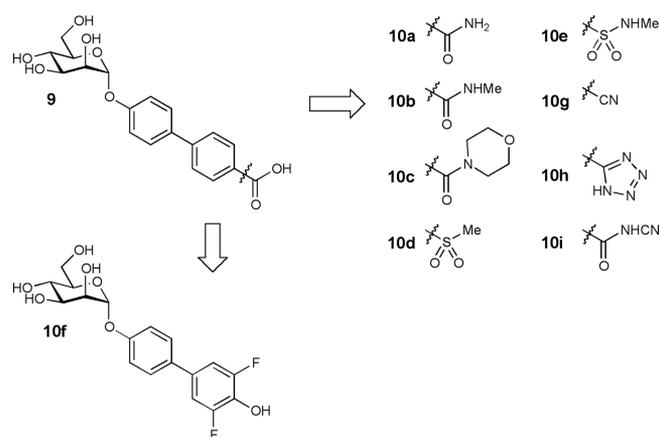
**Figure 1.** Monovalent FimH antagonists 1–4 acting as reference compounds and 5–8 which have been orally explored in in vivo disease models.

sides) to mannose receptors of the human host system have recently been ruled out.<sup>33</sup>

The high affinities of the monovalent  $\alpha$ -D-mannopyranosides are based on optimal interactions with the main structural features of the CRD:<sup>34–37</sup> first, the mannose binding pocket accommodating the mannose moiety by means of an extended hydrogen bond network and, second, the entrance to the binding site composed of three hydrophobic amino acids (Tyr48, Tyr137, and Ile52) and therefore referred to as “tyrosine gate” hosting aliphatic and aromatic aglycones. As an example, *n*-heptyl  $\alpha$ -D-mannopyranoside (**1**) exhibits nanomolar affinity due to hydrophobic contacts of the alkyl aglycone with the hydrophobic residues of the tyrosine gate.<sup>15</sup> Furthermore, aromatic aglycones, such as present in mannosides **2** and **3** (Figure 1), provide strong  $\pi$ – $\pi$  stacking interactions with the tyrosine gate. This interaction is further favored by the addition of an electron withdrawing substituent on the terminal ring of the biaryl portion ( $\rightarrow$ **4**).<sup>18,19</sup>

Recent in vivo PK studies in mice proved the high potential of the biphenyl  $\alpha$ -D-mannosides **5**–**8** for an oral treatment, although high doses ( $\geq 50$  mg/kg) were necessary to achieve the minimal concentrations required for the antiadhesive effect in the urinary bladder.<sup>19–21</sup> Moreover, the therapeutic effect could only be maintained for a few hours, i.e., 4 h for a po (per os) single-dose application of **7** (50 mg/kg), because of rapid elimination by glomerular filtration and low reabsorption from the primary urine in the renal tubules.<sup>20</sup>

To date, the physicochemical properties affecting the rate of renal excretion, i.e., lipophilicity and plasma protein binding (PPB), or metabolic liabilities promoting nonrenal elimination pathways have been barely investigated for FimH antagonists. The goal of the present study was to optimize the biphenyl  $\alpha$ -D-mannoside with respect to oral bioavailability and renal excretion. Starting from antagonist **9**<sup>19</sup> (Figure 2), we synthesized new biphenyl derivatives, characterized their affinity to the CRD, structurally investigated their binding mode, and determined physicochemical and pharmacokinetic parameters predictive for intestinal absorption and renal elimination. Furthermore, we determined in vivo PK (pharmacokinetics) of the most promising new antagonists in a mouse model. After oral administration, the compound with the best PK profile proved effective in reducing the bacterial loads upon bladder infection in a mouse model of UTI.



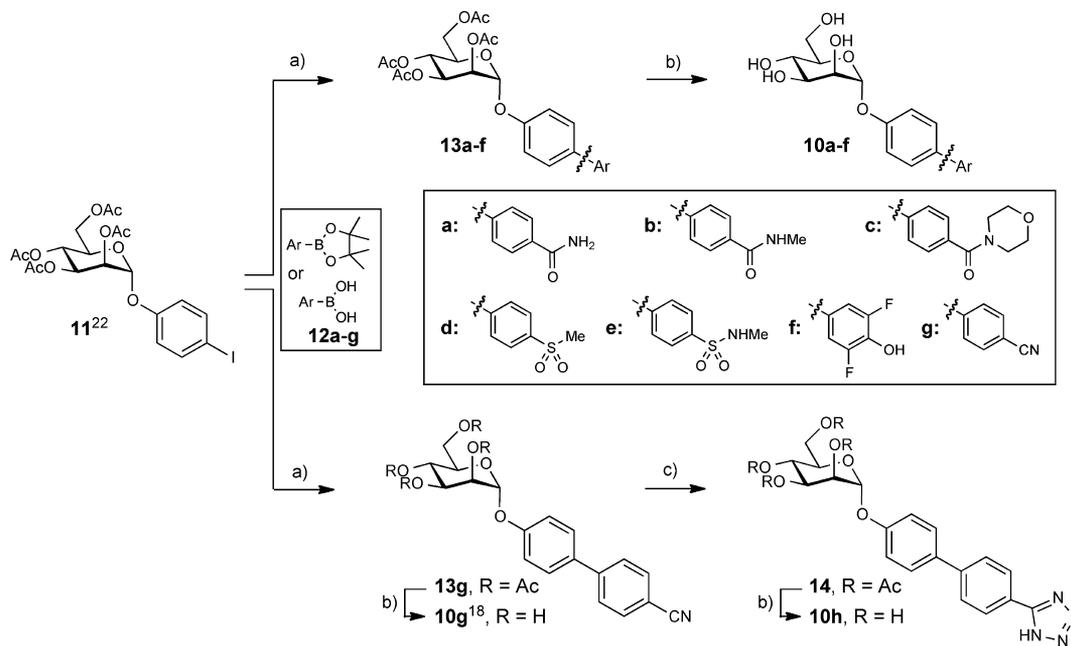
**Figure 2.** Bioisosteric replacement of the carboxylic acid substituent of biphenyl  $\alpha$ -D-mannopyranoside **9**.

## RESULTS AND DISCUSSION

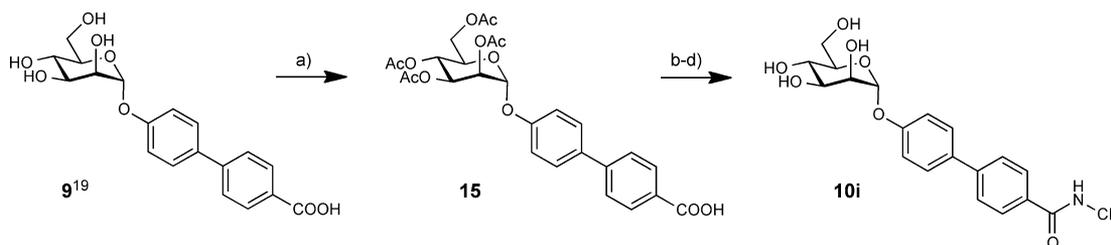
As previously reported, the carboxylate substituent present in the biphenyl mannoside **9** (its electron withdrawing potential being essential for an enhanced drug target interaction) strongly decreases the lipophilicity of the antagonist ( $\log D_{7.4} < -1.5$ <sup>19</sup>) in comparison to the *n*-heptyl ( $\rightarrow$ **1**,  $\log P = 1.7$ <sup>19</sup>) or the unsubstituted biphenyl aglycone ( $\rightarrow$ **3**,  $\log P = 2.1$ <sup>22</sup>). Since low lipophilicity is a major reason for low intestinal absorption and rapid renal excretion of the systemically available antagonist,<sup>19,23</sup> we aspired to improve oral bioavailability as well as renal excretion by replacing the carboxylate in **9** with various bioisosteric groups<sup>39</sup> (Figure 2).

**Synthesis.** Iodide **11** was prepared from peracetylated mannose and 4-iodophenol in the presence of  $\text{BF}_3 \cdot \text{Et}_2\text{O}$ .<sup>22</sup> In a palladium-catalyzed Miyaura–Suzuki coupling<sup>40</sup> with the boronic acid or boronate derivatives **12a–g**, the biphenyl derivatives **13a–g** were obtained in good to excellent yields. Final deprotection yielded the test compounds **10a–g**. When microwave-assisted reaction conditions<sup>41</sup> were utilized, the conversion of aryl nitrile **13g** to tetrazole **14** proceeded rapidly and with good yield. After deprotection of **14** using Zemplén conditions, the test compound **10h** was obtained (Scheme 1).

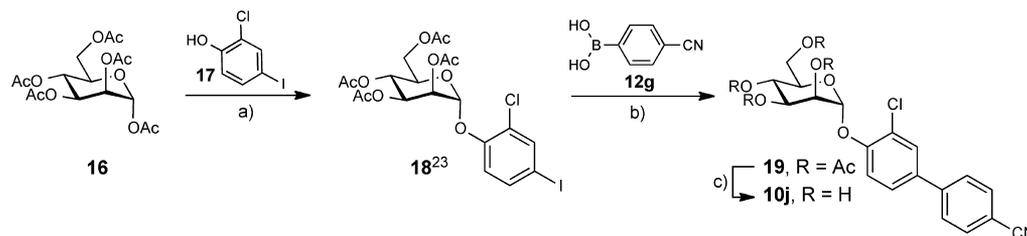
The cyanobenzamide derivative **10i** (Scheme 2) was obtained from **9** by peracetylation ( $\rightarrow$ **15**) followed by conversion of the

Scheme 1<sup>a</sup>

<sup>a</sup>(a) Pd(Cl<sub>2</sub>)dppf·CH<sub>2</sub>Cl<sub>2</sub>, K<sub>3</sub>PO<sub>4</sub>, DMF, 80 °C, 4 h (**13a–g**, 44–99%); (b) NaOMe, MeOH, rt, 4 h (**10a–h**, 29–86%); (c) TMSN<sub>3</sub>, Bu<sub>2</sub>Sn(O), DME, 150 °C, microwave, 10 min (81%).

Scheme 2<sup>a</sup>

<sup>a</sup>(a) (i) Ac<sub>2</sub>O, DMAP, pyridine, 0 °C to rt, overnight; (ii) sat. NaHCO<sub>3</sub> aq, DCM, rt, 2 h (**15**, 53%); (b) 1-chloro-*N,N*,2-trimethyl-1-propenylamine, toluene, 0 °C to rt, 2 h; (c) NaH, NH<sub>2</sub>CN, DMF, 0 °C to rt, overnight; (d) NaOMe, MeOH, rt, 4 h (**10i**, 21% for three steps).

Scheme 3<sup>a</sup>

<sup>a</sup>(a) BF<sub>3</sub>·Et<sub>2</sub>O, CH<sub>2</sub>Cl<sub>2</sub>, 40 °C (76%); (b) Pd(Cl<sub>2</sub>)dppf·CH<sub>2</sub>Cl<sub>2</sub>, K<sub>3</sub>PO<sub>4</sub>, DMF, 80 °C (75%); (c) NaOMe, MeOH, rt, 4 h (48%).

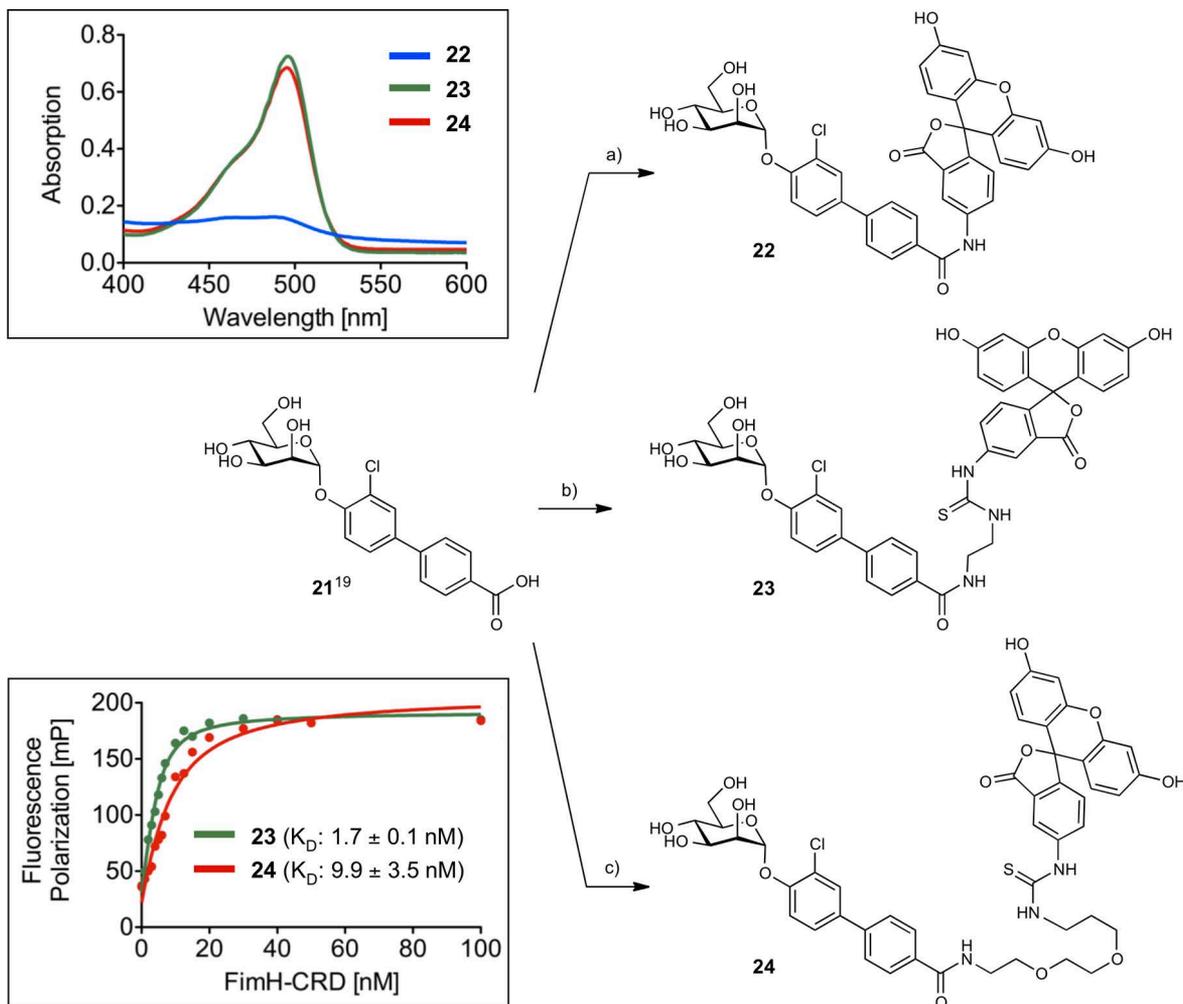
carboxylic acid into its acid chloride with 1-chloro-*N,N*,2-trimethyl-1-propenylamine.<sup>42</sup> Without isolation, the acid chloride was reacted with sodium hydrogen cyanamide in DMF followed by deacetylation under Zemplén conditions to yield the test compound **10i**.

Finally, to further improve the pharmacokinetic properties of mannoside **10g**<sup>18</sup> (see Table 3), a chloride substituent was introduced to the ortho-position of the aromatic ring adjacent to the anomeric oxygen. For its synthesis, peracetylated  $\alpha$ -D-mannose (**16**) was coupled with 2-chloro-4-iodophenol (**17**) using BF<sub>3</sub>·Et<sub>2</sub>O as promoter ( $\rightarrow$ **18**, 76%). After the introduction

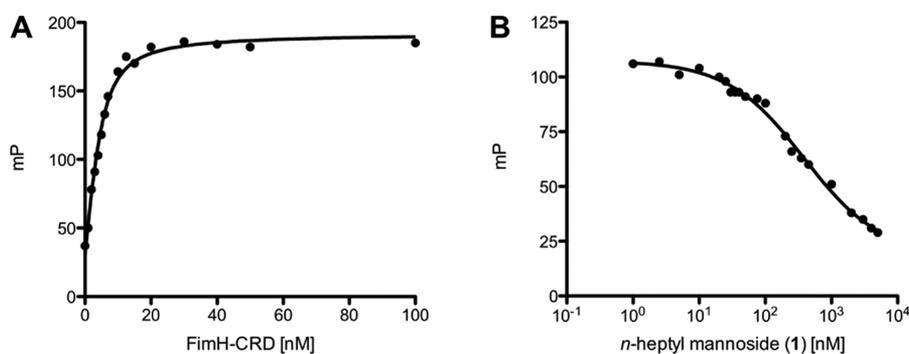
of the second aromatic ring by Miyaura–Suzuki coupling ( $\rightarrow$ **19**, 75%), deprotection yielded mannoside **10j** (Scheme 3).

**Binding Affinity.** The binding affinity of heptyl mannoside **1**, the biphenyl mannosides **3**, **9**, **20**,<sup>18</sup> and the bioisosteres **10a–j** was determined in a competitive fluorescence polarization assay (FP assay) and with isothermal titration calorimetry (ITC). A protein construct consisting of the CRD with a C-terminal His-tag with a thrombin cleavage site (FimH-CRD-Th-His<sub>6</sub>) was used for all experiments.<sup>43</sup>

**Competitive Fluorescence Polarization Assay.** For the rapid evaluation of binding affinity, we established a competitive

Scheme 4<sup>a</sup>

<sup>a</sup>(a) 1-[1-(Cyano-2-ethoxy-2-oxoethylideneaminoxy)dimethylaminomorpholinomethylene]methanaminium hexafluorophosphate (COMU),  $\text{NEt}_3$ , fluoresceinamine, DMF, rt, 7 h (**22**, 19%); b) (i) DIC, NHS, *N*-Boc-ethylenediamine, DMF, rt, 12 h; (ii) TFA, DCM, rt, 10 min (68% over two steps), (iii) fluorescein isothiocyanate (FITC),  $\text{NEt}_3$ , DMF, rt, 3 h (**23**, 48%); (c) (i) DIC, NHS, *N*-Boc-PEG2- $\text{NH}_2$ , DMF, rt, 14 h; (ii) TFA, DCM, rt, 30 min (62% over two steps); (iii) FITC, DMF, rt (**24**, 65%).



**Figure 3.** (A) Direct binding curve of the labeled competitor **23** obtained by adding a linear dilution of FimH-CRD (0–100 nM) and a constant concentration of competitor **23** (5 nM). The  $K_D$  was determined by fitting the experimental data to a single-site binding fit that accounts for ligand depletion. In three FP based direct binding experiments the  $K_D$  of competitor **23** was determined to be 1.7 nM. (B) Inhibition curve of *n*-heptyl mannoside (**1**) from the competitive FP assay. The  $\text{IC}_{50}$  value was determined by nonlinear least-squares fitting to a standard four-parameter equation. A modified Cheng–Prusoff equation<sup>45</sup> was used to calculate the corresponding  $K_D$  value ( $K_D = 28.3$  nM).

binding assay based on fluorescence polarization (FP). Similar formats have been applied before for the detection of carbohydrate–lectin interactions.<sup>18,44</sup> In this assay, the antagonist of interest displaces a fluorescently labeled competitor from

the binding site, thereby causing a reduction in fluorescence polarization.<sup>45</sup> To identify the optimal competitor, fluorescein isothiocyanate (FITC) was connected to the FimH ligand **21** by three linkers of different lengths ( $\rightarrow$ **22–24**, Scheme 4). For

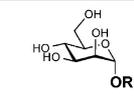
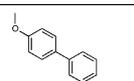
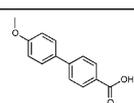
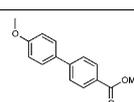
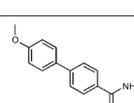
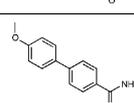
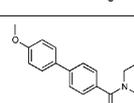
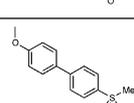
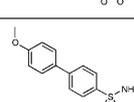
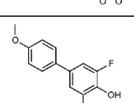
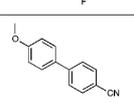
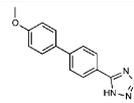
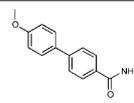
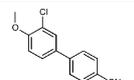
optimal sensitivity and signal-to-noise ratio, three main parameters need to be considered: (i) the affinity of the competitor should not be impaired by the fluorescent label; (ii) the conformational flexibility of the label upon binding of the competitor to the CRD should be low; (iii) the fluorescence properties of the label should not be affected by the connected ligand.<sup>46–48</sup> A change in fluorescence properties was observed for reporter ligand **22** in which the label was linked to the biphenyl aglycone by an amide bond. The absorption spectrum revealed a lack of the characteristic fluorescein absorption peak at 494 nm (Scheme 4), likely due to an extension of the conjugated system to the biphenyl moiety of the ligand. The elongated saturated spacer groups in competitors **23** and **24** ensured that the expected spectral properties of the dye were retained (Scheme 4).

For the determination of their binding affinity, fixed concentrations of the reporter ligands **23** and **24** were incubated for 24 h with a linear dilution of the FimH-CRD (0–100 nM). FP was measured using a plate reader, with polarized excitation at 485 nm and emission at 528 nm measured through appropriately oriented polarizers. Fitting the single-site binding function of Cooper<sup>49</sup> to the observed FP data resulted for compound **23** in a dissociation constant ( $K_D = 1.7$  nM, Figure 3A) similar to that of the unlabeled parent compound **21**,<sup>19</sup> whereas **24** showed a 5-fold lower affinity (9.9 nM) (Scheme 4). Therefore, the reporter ligand **23** fulfills all characteristics as an optimal competitor and was used for the FP assay.

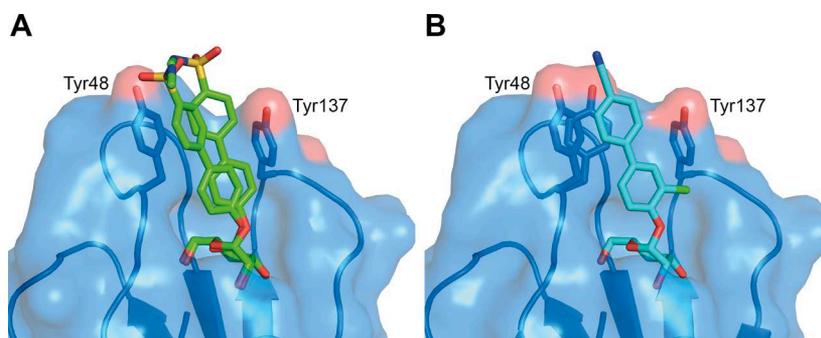
For the test compounds **1**, **3**, **9**, **20**, and **10a–j**, a 24 h incubation time was applied before FP was measured because of the long residence time of FimH antagonists ( $t_{1/2} > 3.5$  h, Figure 3B<sup>50</sup>). The 24 h incubation period was empirically determined to be necessary to reach equilibrium between reporter ligand and compound of interest.  $IC_{50}$  values were obtained by nonlinear least-squares regression (standard four-parameter dose–response curve) and converted to  $K_D$  values using a modified Cheng–Prusoff equation.<sup>45</sup> This equation accounts for the ligand depletion effect in competitive titrations involving high-affinity interaction partners present in similar concentrations. Under these conditions, the free concentration of an interacting species cannot be assumed to equal the total concentration.

The  $K_D$  values determined for the test compounds **1**, **3**, **9**, **20**, and **10a–j** are summarized in Table 1. Against our expectations, the biphenyl mannosides **3** and **9** exhibit similar affinities (Table 1), despite the presence of an electron withdrawing carboxylate substituent in antagonist **9**. According to the crystal structure of FimH cocrystallized with the sulfonamide derivative **10e** (Figure 4A), the outer aromatic ring of the biphenyl aglycone forms  $\pi$ – $\pi$  interactions with the electron rich Tyr48, which is part of the tyrosine gate of FimH.<sup>15</sup> A reduction of electron density of the aglycone by the electron withdrawing carboxylate was expected to enforce these  $\pi$ – $\pi$  stacking interactions and lead to improved affinity. However, this beneficial effect might be compensated by an entropic penalty originating from the improved  $\pi$ – $\pi$  stacking to Tyr48 that might lead to the reduced flexibility of both protein and antagonist. Furthermore, a beneficial enthalpy effect might be partially compensated by an enthalpy penalty originating from the desolvation of the charged carboxylate in **9**<sup>51</sup> (see also Experimental Section). Although this substituent is solvent exposed, at least a partial desolvation may be necessary upon antagonist binding. To prove this assumption, we replaced the carboxylate by the corresponding methyl ester ( $\rightarrow$ **20**)<sup>18</sup> in order to reduce the desolvation penalty and, as predicted by the Hammett constant  $\sigma_p$ ,<sup>52</sup> to further improve the  $\pi$ – $\pi$  stacking.

**Table 1.** Affinities ( $K_D$ ) of FimH Antagonists to FimH-CRD-Th-His<sub>6</sub><sup>b</sup>

Entry	Compd		Affinity $K_D$ [nM]
1	<b>1</b>		28.3 ± 5.0
2	<b>3</b>		15.1 ± 2.2
3	<b>9</b>		17.9 ± 1.5
4	<b>20</b>		3.6 ± 0.9
5	<b>10a</b>		2.8 ± 0.3
6	<b>10b</b>		2.9 ± 0.5
7	<b>10c</b>		3.0 ± 0.1
8	<b>10d</b>		1.7 ± 0.2
9	<b>10e</b>		2.7 ± 0.4
10	<b>10f</b>		3.7 ± 0.2
11	<b>10g</b>		2.0 ± 0.6
12	<b>10h</b>		5.7 ± 0.1
13	<b>10i</b>		8.4 ± 0.3
14	<b>10j</b>		< 1 <sup>a</sup> )

<sup>a</sup>The  $K_D$  value of **10j** was approximated to be in the subnanomolar range. The  $IC_{50}$  value obtained in the competitive FP assay was equal to the lowest value that can be resolved by the assay, indicating stoichiometric titration of **10j** due to its high affinity. Consequently, its  $K_D$  must be below the  $K_D$  of competitor **23**. <sup>b</sup>Dissociation constants ( $K_D$ ) were determined in a competitive fluorescence polarization assay.



**Figure 4.** Ligand binding poses determined by X-ray cocrystallization with compounds **10e** resolved to 1.07 Å (A) and **10j** resolved to 1.10 Å (B). The electron density surrounding the aglycone of **10e** indicates flexibility of the aglycone and was modeled in two poses. Both compounds bind in a similar pose with a well-defined hydrogen network surrounding the mannose moiety and  $\pi$ - $\pi$  stacking interactions between the second aromatic ring and Tyr48 side chain (A). In contrast, in the FimH-CRD/**10j** structure the amino acid side chain of Y48 can be modeled in two distinct rotamers, suggesting flexibility also of the receptor (B).

Indeed, a 6-fold improvement in affinity was achieved. However, since the methyl ester undergoes rapid enzyme-mediated hydrolysis *in vivo*,<sup>19</sup> it will not be available at the place of action in the urinary bladder. The methyl ester was therefore replaced by metabolically stable bioisosteres<sup>39</sup> exhibiting comparable electron withdrawing properties<sup>52</sup> (Table 1, entries 5–13). The most potent derivatives **10d**, **10e**, and **10g** showed affinities in the low nanomolar range.

As previously reported,<sup>22</sup> a chloro substituent in the ortho-position of the aromatic ring adjacent to the anomeric oxygen is favorable for affinity and improves the physicochemical properties relevant for oral bioavailability. Indeed, the corresponding antagonist **10j** was the most potent compound tested in this study.

**Isothermal Titration Calorimetry (ITC).** To further confirm our hypothesis regarding  $\pi$ - $\pi$  stacking and desolvation, we performed ITC experiments with the reference compound **1**, the unsubstituted biphenyl mannoside **3**, the carboxylic acid **9**, and the bioisosteres **10b–e,g,j** (Table 2). ITC allows the simultaneous determination of the stoichiometry ( $N$ ), the change in enthalpy ( $\Delta H$ ) and the dissociation constant ( $K_D$ ) for ligand–protein binding.<sup>53,54</sup> The reliable determination of these three parameters requires well-defined sigmoidal titration curves characterized by the dimensionless Wiseman parameter  $c$  ( $c = Mt(0)K_D^{-1}$ , where  $Mt(0)$  is the initial macromolecule concentration).<sup>55</sup> To be sure that data can be fitted with confidence, the  $c$ -value should be between 1 and 1000 (ideally between 5 and 500),<sup>56</sup> which could be achieved for the antagonists **3** and **9**. For titrations involving low micromolar  $Mt(0)$  and interactions in the low nanomolar or picomolar range, as suggested for the bioisosteres **10b–j**,  $c$ -values above 1000 were expected. Since these conditions lead to steep titration curves that do not allow the determination of the curve slope representing  $1/K_D$ , we applied an alternative, competitive format referred to as displacement assay.<sup>57,58</sup> First, FimH-CRD-Th-His<sub>6</sub> was preincubated with the low affinity antagonist *n*-heptyl 2-deoxy- $\alpha$ -D-mannopyranoside (**25**, for synthesis see Supporting Information). The high-affinity bioisosteres of interest were titrated into the protein–ligand complex giving well-defined sigmoidal titration curves.

The resulting  $K_D$  values (Table 2) correspond well with the data obtained from the FP assay (Table 1). A comparison of the thermodynamic fingerprints of antagonists **3** and **9** reveals that the more favorable enthalpic contribution resulting from facilitated  $\pi$ - $\pi$  stacking leads to a net enthalpy gain ( $\Delta\Delta H =$

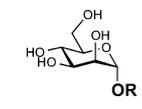
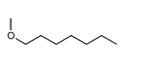
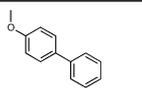
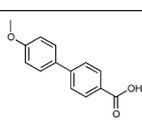
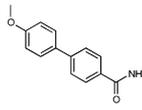
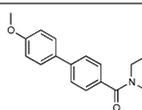
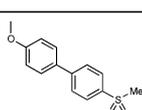
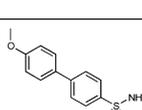
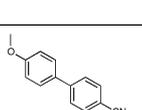
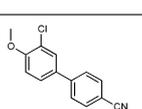
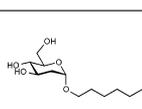
$-3.7$  kJ/mol). However, an even greater increase in enthalpy is likely countered by the enthalpy costs for desolvation of the electron withdrawing carboxylate.

The gain in enthalpy is in turn compensated by an unfavorable entropy ( $-T\Delta\Delta S = 3.2$  kJ/mol) as a result of the reduced flexibility of both the antagonist and the Tyr48 side chain caused by the improved interaction. This is not entirely outweighed by the beneficial entropy contribution related to the partial desolvation of the carboxylate and the related release of water into the bulk. Added together, the enthalpy and entropy contributions of antagonists **3** and **9** result in similar affinities ( $K_D$  of 17.7 and 15.0 nM, respectively).

In contrast, the replacement of the carboxylate group by various neutral bioisosteres (entries 4–7) reduces the enthalpy costs for desolvation (see calculated free energies of desolvation, Experimental Section) and therefore leads to a markedly improved enthalpy ( $\Delta\Delta H$  from  $-3.5$  to  $-5.8$  kJ/mol). As a result, an up to 5-fold improvement of the  $K_D$  values was achieved. Finally, with a cyano substituent (entries 8 and 9), the enthalpy term was further improved ( $\Delta\Delta H = -3.7$  kJ/mol) because of a reduced desolvation penalty and improved  $\pi$ - $\pi$  stacking interactions. However, this beneficial component is again partially compensated by a decrease in entropy. This can be attributed, first, to the loss of flexibility of the tightly bound ligand (Figure 4B) and, second, to the smaller surface area of the cyano substituent compared to amide, sulfonamide, and sulfone, which results in a smaller number of water molecules being released to bulk upon binding.

**X-ray Crystallography.** To determine the binding poses of the bioisosteres, we cocrystallized the compounds **10e** and **10j** with the FimH-CRD (Figure 4). Atomic resolution crystal structures were obtained at 1.07 Å (**10e**) and 1.10 Å (**10j**). As observed in previous mannoside cocrystal structures,<sup>15,18,36</sup> the mannose moiety forms an extensive hydrogen bond network to the well-defined binding site with all of its hydroxyl groups. The biphenyl aglycone is located between the tyrosine gate residues (Tyr48/Tyr137). The  $\pi$ - $\pi$  stacking of the second aromatic ring of the aglycone to the side chain of Tyr48 contributes most to the interaction energy of the aglycone moiety. Interactions to the Tyr137 side chain on the other hand are only limited. Whereas a previously published crystal structure of a biphenyl mannoside in complex with FimH-CRD suffers from crystal contacts of binding site residues (Tyr48 side chain to backbone oxygen of Val27) possibly causing the distortion of the binding site,<sup>18</sup> the binding sites of our structures are mostly solvent exposed. This

Table 2. Thermodynamic Parameters from ITC for Selected FimH Antagonists Binding to FimH-CRD-Th-His<sub>6</sub><sup>a</sup>

Entry	Compd		$K_D$ <sup>[a]</sup> [nM]	$\Delta G$ [kJ/mol]	$\Delta H$ <sup>[a]</sup> [kJ/mol]	$-T\Delta S$ [kJ/mol]	$n$	Type of measurement
1	<b>1</b> <sup>[b,c]</sup>		28.9 (25.8 – 32.3)	-43.0	-50.3 (-50.2 – -50.7)	7.3	1.00	direct
2	<b>3</b> <sup>[b]</sup>		17.7 (14.1 – 22.3)	-44.2	-45.0 (-44.5 – -45.6)	0.8	1.07	direct
3	<b>9</b>		15.0 (13.4 – 16.7)	-44.7	-48.7 (-48.4 – -49.0)	4.0	1.05	direct
4	<b>10b</b>		4.3 (3.2 – 5.6)	-47.8	-54.5 (-54.1 – -54.9)	6.7	1.02	competitive vs. <b>25</b>
5	<b>10c</b>		5.0 (3.8 – 6.6)	-47.4	-54.5 (-54.1 – -54.8)	7.1	0.97	competitive vs. <b>25</b>
6	<b>10d</b>		3.0 (2.1 – 4.2)	-48.7	-52.3 (-51.5 – -53.1)	3.6	0.99	competitive vs. <b>25</b>
7	<b>10e</b>		3.5 (2.9 – 4.3)	-48.2	-52.2 (-51.6 – -52.8)	3.9	1.06	competitive vs. <b>25</b>
8	<b>10g</b>		2.8 (2.3 – 3.3)	-48.8	-58.2 (-57.8 – -58.6)	9.4	1.00	competitive vs. <b>25</b>
9	<b>10j</b>		1.3 (1.1 – 1.6)	-50.7	-60.9 (-60.4 – -61.4)	10.1	1.01	competitive vs. <b>25</b>
10	<b>25</b>		9'386 (8'555 – 10'287)	-28.7	-19.5 (-19.1 – -20.0)	-9.1	1.00	direct

<sup>a</sup>95% confidence interval from fitting in parentheses. <sup>b</sup>Global fit including two direct titration measurements. <sup>c</sup>ITC data were previously published with an  $n$ -value of 0.82.<sup>37</sup> <sup>d</sup> $n$ , stoichiometric correction factor.

revealed the flexibility of the aglycone in the FimH-CRD/**10e** structure, since the electron density toward the solvent-exposed sulfonamide indicates that there is not one single orientation. Therefore, the aglycone was modeled in two distinct poses. In contrast, in the FimH-CRD/**10j** structure the amino acid side chain of Y48 can be modeled in two distinct rotamers, suggesting flexibility also of the receptor.

**Physicochemical Properties and in Vitro Pharmacokinetics.** Intestinal absorption and renal excretion are prerequisites for a successful oral treatment of UTI with FimH

antagonists. Furthermore, reabsorption of antagonist from the renal ultrafiltrate is desirable for maintaining the minimal antiadhesive concentration in the target organ, namely, the bladder, over an extended period of time. To estimate the influence of the bioisostere approach on oral bioavailability and the rate of renal excretion, we determined lipophilicity by means of the octanol–water distribution coefficient ( $\log D_{7.4}$ ),<sup>59</sup> aqueous solubility, and membrane permeability in the artificial membrane permeability assay (PAMPA)<sup>60</sup> and the colorectal adenocarcinoma (Caco-2) cell monolayer model.<sup>61</sup>

Table 3. Physicochemical and in Vitro Pharmacokinetic Parameters<sup>h</sup>

compd	pK <sub>a</sub> <sup>a</sup>	log D <sub>7.4</sub> <sup>b</sup>	solubility [μg/mL]/pH <sup>c</sup>	PAMPA log P <sub>e</sub> <sup>d</sup> [cm/s]/pH <sup>e</sup>	Caco-2 P <sub>app</sub> [10 <sup>-6</sup> cm/s] <sup>e</sup>		PPB f <sub>b</sub> <sup>f</sup> [%] <sup>g</sup>	metabolic stability t <sub>1/2</sub> <sup>g</sup> [min] <sup>g</sup>
					a → b	b → a		
1		1.65	>3000	-4.89	7.0 ± 0.6	9.4 ± 0.2	81	13
3		2.1 ± 0.1	21 ± 1/7.4	-4.7 ± 0.1/7.4	10.0 ± 0.9	19.0 ± 1.2	93 ± 1	nd
20		2.14	33.8/6.51	-4.7	4.23	nd	93	1.0
9	3.88	<-1.5	>3000/6.61	no permeation	nd	nd	73	>60
10a		0.5 ± 0.1	12 ± 1/7.4	-6.8 ± 0.3/7.4	0.12 ± 0.01	0.61 ± 0.03	nd	nd
10b		0.8 ± 0.0	122 ± 13/7.4	-9.2 ± 1.4/7.4	1.10 ± 0.82	0.87 ± 0.15	nd	nd
10c		0.2 ± 0.1	>250/7.4	-7.8 ± 0.3/7.4	0.18 ± 0.07	1.30 ± 0.03	48 ± 2	>60
10d		0.4 ± 0.0	246 ± 17/7.4	-7.2 ± 0.0/7.4	0.36 ± 0.01	1.76 ± 0.12	99 ± 1	>60
10e		0.7 ± 0.1	>250/7.4	-8.6 ± 0.2/7.4	0.28 ± 0.23	1.82 ± 0.14	>99	>60
10f	6.5	1.1 ± 0.0	>150/3.0	-7.7 ± 0.8/5.0	0.40 ± 0.02	1.90 ± 0.17	nd	nd
			>150/7.4	-8.8 ± 0.1/7.4				
10g		1.4 ± 0.0	186 ± 4/7.6	-5.7 ± 0.0/7.4	2.0 ± 0.1	13.2 ± 2.1	99 ± 0	>60
10h	3.7	-1.4 ± 0.1	11 ± 0/3.0	-9.3 ± 1.4/5.0	0.17 ± 0.00	0.22 ± 0.01	nd	nd
			273 ± 2/7.4	-8.8 ± 1.4/7.4				
10i	2.5	-1.1 ± 0.1	>150/3.0	-6.8 ± 0.2/5.0	0.22 ± 0.14	0.29 ± 0.03	nd	nd
			>150/7.4	-7.0 ± 0.1/7.4				
10j		2.1 ± 0.0	192 ± 5/7.4	-5.2 ± 0.0/7.4	2.2 ± 0.4	22.1 ± 1.5	89 ± 1	>60

<sup>a</sup>pK<sub>a</sub> values were determined by NMR spectroscopy. <sup>b</sup>Octanol–water distribution coefficients (log D<sub>7.4</sub>) were determined by a miniaturized shake-flask procedure at pH 7.4. Values represent the mean ± SD of sextuplicate measurements.<sup>59</sup> <sup>c</sup>Kinetic solubility was measured in a 96-well format using the μSOL Explorer solubility analyzer at the indicated pH in triplicate. <sup>d</sup>P<sub>e</sub> = effective permeability. Passive permeation through an artificial membrane was determined by the parallel artificial membrane permeation assay (PAMPA). Values represent the mean ± SD of quadruplicate measurements performed at the indicated pH.<sup>60</sup> <sup>e</sup>P<sub>app</sub> = apparent permeability. Permeation through a Caco-2 cell monolayer was assessed in the absorptive (a → b) and secretory (b → a) directions in triplicate.<sup>61</sup> <sup>f</sup>Plasma protein binding (PPB) was determined by equilibrium dialysis in triplicate.<sup>62</sup> <sup>g</sup>Metabolic stability was determined by incubating the compounds (2 μM) with pooled rat liver microsomes (RLM, 0.5 mg/mL) in the presence of NADPH (1 mM, compounds 1, 9, 10c–e, g, j) or without NADPH (compound 20).<sup>63</sup> <sup>h</sup>nd = not determined.

**Oral Bioavailability.** Oral bioavailability of a compound relies on solubility, permeation through the membranes lining the intestine, and stability against first pass metabolism.<sup>64,65</sup> As discussed by Lipinski<sup>66</sup> and Curatolo,<sup>67</sup> dose and permeability define the minimum aqueous solubility required for oral administration. Thus, a dose of 1 mg/kg of a moderately permeable compound requires a solubility of at least 52 μg/mL. Whereas sufficient aqueous solubility (>3000 μg/mL) was reported for *n*-heptyl α-mannopyranoside (1),<sup>19</sup> the unsubstituted biphenyl α-D-mannopyranoside 3 and the antagonists bearing a methylcarboxylate, carboxamide, or tetrazole substituent (compounds 20, 10a, and 10h) were found to be scarcely soluble.<sup>22</sup> As proposed by Ishikawa,<sup>68</sup> a possible reason is the apolar and planar aglycone. By contrast, the polar carboxylic acid moiety present in antagonist 9 or the substituents in the bioisosteres 10b–j enhance solubility to 122–273 μg/mL, a level sufficient for in vivo PK studies. For in vivo disease studies, however, dosages of up to 10 mg/kg were foreseen (see below), requiring a solubility of 520 μg/mL.<sup>66,67</sup> For this reason, surfactant Tween 80 (1%) had to be added.

Furthermore, permeability data derived from PAMPA<sup>69</sup> and the Caco-2 model<sup>70</sup> suggest moderate to high permeation of the moderately lipophilic antagonists 1, 3, and 20 (log D<sub>7.4</sub> > 1.6) through the intestinal membranes. The bioisosteres 10a–f, h, i, although slightly more permeable than the strongly hydrophilic carboxylic acid derivative 9, show only low values of permeability compared to *n*-heptyl α-D-mannopyranoside (1) or the unsubstituted biphenyl mannoside 3. However, the *p*-cyanobiphenyl derivatives 10g and 10j display elevated log D<sub>7.4</sub> and effective permeability (log P<sub>e</sub>) in the range for successful intestinal absorption. Regarding both sufficient aqueous solubility and elevated membrane permeability, the *p*-cyano substituted bioisosteres 10g and 10j are thus the most promising

candidates for oral absorption. Moreover, combining the bioisosteric replacement with the addition of a chloro substituent in the ortho-position of the aromatic ring adjacent to the anomeric oxygen (→10j)<sup>22</sup> resulted in the most advantageous physicochemical profile for oral bioavailability.

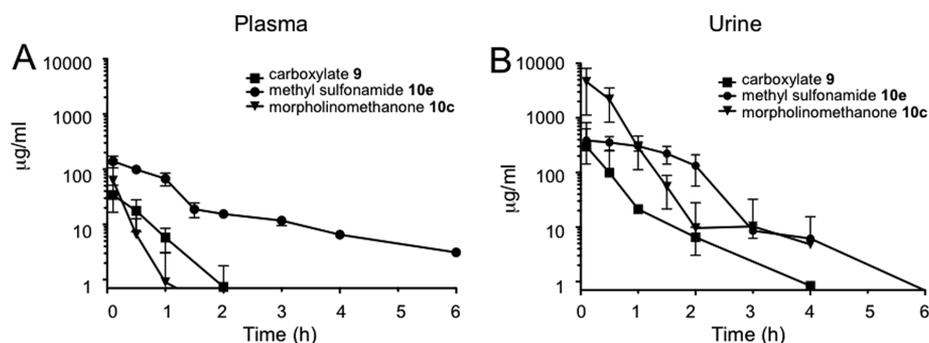
**Renal Excretion.** The rate of renal excretion depends on the rate of glomerular filtration and the propensity to tubular secretion and reabsorption of an antagonist.<sup>71</sup> Only the fraction that is not bound to plasma proteins is expected to enter the glomerular filtrate.<sup>72</sup> Plasma protein binding (PPB) data indicating the fraction bound (f<sub>b</sub>) are listed in Table 2.<sup>62</sup> The biphenyls 9 and 10c were identified as moderate binders to plasma proteins (f<sub>b</sub> ≤ 65%), which suggests a low impact of PPB on antagonist filtration. The f<sub>b</sub> values of the antagonists 1, 3, 20, and 10j were between 80% and 93%, whereas the bioisosteres 10d, e, g showed particularly high protein binding (f<sub>b</sub> ≥ 99%) implying slow compound entry into the primary urine. However, the kinetic aspects of PPB, that is, association and dissociation rate constants, remain to be determined to quantify precisely the influence of PPB on filtration.<sup>73</sup>

Furthermore, log D<sub>7.4</sub> was identified as key determinant of tubular reabsorption.<sup>74–76</sup> Accordingly, lipophilic compounds are predominantly reabsorbed from the renal filtrate. Given that renal clearance is the major route of elimination, this will result in a slow but steady excretion into the bladder. In contrast, hydrophilic compounds are poorly reabsorbed and thus quickly renally eliminated, which leads to high initial compound levels in the urine but narrows the time range where the minimal antiadhesive concentration is maintained. Consequently, low log D<sub>7.4</sub> as shown for the antagonists 9, 10h, and 10i implies low tubular reabsorption and rapid elimination of the filtered molecules by the urine. Otherwise, log D<sub>7.4</sub> between 0.2 and 0.7, such as determined for the bioisosteres 10a–e, suggests

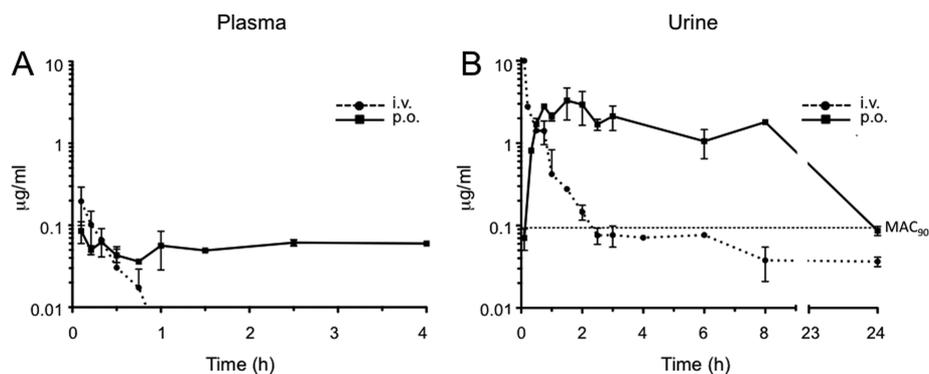
**Table 4. Pharmacokinetic Parameters Determined after a Single iv Application of Compounds 9, 10c, 10e, and 10j in Female C3H/HeN Mice<sup>a</sup>**

compd	plasma						urine, $C_{max}$ ( $\mu\text{g/mL}$ )
	$C_0$ ( $\mu\text{g/mL}$ )	dose (mg/kg)	$V_z$ (mL)	$t_{1/2}$ (h)	$AUC_{0-\infty}$ ( $\mu\text{g}\cdot\text{h/mL}$ )	$CL_{tot}$ (mL/h)	
9	40	50	25.2	0.33	23.5	53.1	300
10c	109.7	50	28.3	0.4	25.3	49.4	4611
10e	151.6	50	19.5	1.9	175.1	7.1	387
10j	0.36	0.625	52.8	0.17	0.07	218	10

<sup>a</sup>Values were calculated using PKSolver.<sup>78</sup>  $C_0$ , initial concentration;  $V_z$ , volume of distribution in terminal phase; AUC, area under the curve;  $CL_{tot}$ , total clearance;  $C_{max}$ , maximal concentration.



**Figure 5.** Antagonist concentrations in (A) plasma and (B) urine after a single iv application of 9, 10c, and 10e (50 mg/kg).



**Figure 6.** Antagonist concentrations in (A) plasma and (B) urine after a single iv and po application of compound 10j (iv, 0.625 mg/kg; po, 1.25 mg/kg).  $MAC_{90}$  is the minimal antiadhesive concentration to inhibit 90% adhesion (0.094  $\mu\text{g/mL}$ ).

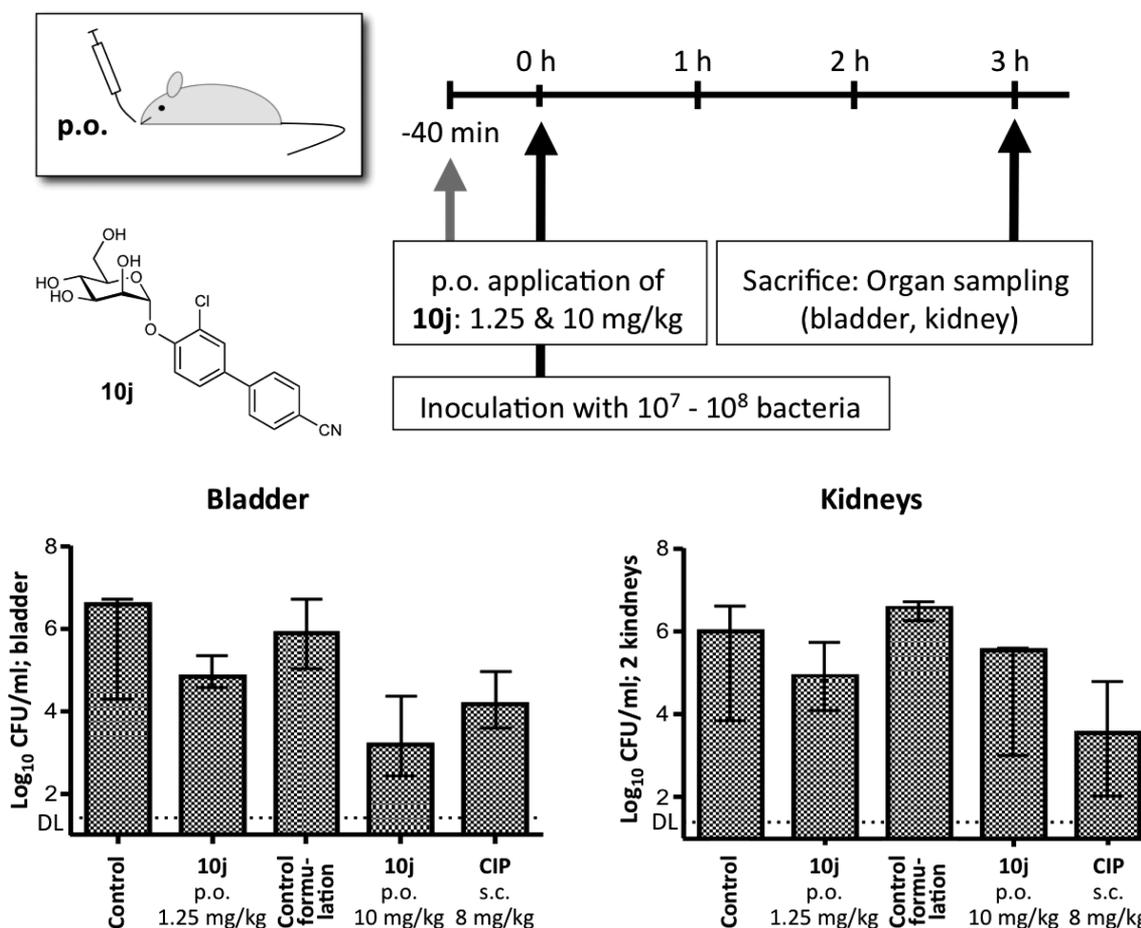
increasing propensity to tubular reuptake, whereas  $\log D_{7,4} > 1$  as shown for heptyl mannoside **1** and the biphenyl mannosides **3**, **20**, **10g**, **10f**, and **10j** is optimal for tubular reabsorption from the glomerular filtrate and thus for slow renal clearance.

**Metabolic Stability.** Increasing lipophilicity is usually paralleled by increasing susceptibility to metabolism.<sup>77</sup> Liabilities toward metabolic clearance pathways that prevent the intact antagonist from reaching the target in the bladder were therefore of interest. To assess their propensity to cytochrome P450 (CYP450) mediated metabolism, heptyl mannoside **1**, the carboxylic acid derivative **9**, and the bioisosteres **10c–e,g,j** were incubated with rat liver microsomes (RLM, 0.5 mg/mL) in the presence of the cofactor  $\beta$ -nicotinamide adenine dinucleotide phosphate (NADPH).<sup>63</sup> To confirm the high propensity of the methyl ester present in antagonist **20** to carboxylesterase (CES) mediated hydrolysis, this antagonist was incubated with RLM only. The profiles of unchanged compound versus time revealed high susceptibility of heptyl mannoside **1** to CYP450-mediated metabolism ( $t_{1/2} = 13$  min) and rapid hydrolysis of the ester **20** by the hepatic CES ( $t_{1/2} = 1.0$  min). Otherwise, the bioisosteres **10c–e,g,j** were stable against enzyme-mediated bioconversion

( $t_{1/2} > 60$  min), suggesting lower propensity to metabolic, nonrenal elimination pathways.

Considering PPB, lipophilicity, and metabolic stability data, we therefore expected (i) a steady release of compounds **10d,e,g,j** into the bladder because of high PPB decelerating glomerular filtration (**10d,e,g**) and/or high  $\log D_{7,4}$  supporting tubular reabsorption (**10g,j**), (ii) a fast excretion of antagonists **9** and **10c** via the urine due to low PPB and low  $\log D_{7,4}$ , and (iii) a rapid clearance of heptyl mannoside **1** from the body by renal and metabolic pathways. Compounds featuring high propensity to renal excretion as major route of elimination (**10c**, **10e** and **10j**) were selected for in vivo PK studies in a mouse model.

**Pharmacokinetic Studies in C3H/HeN Mice.** This first part of our study explored the predicted effects of lipophilicity, PPB, and metabolic stability on antagonist disposition and elimination upon a single dose iv application (50 mg/kg) of compounds **10c** and **10e**. The PK parameters of these applications and those of the previously published carboxylate **9** are summarized in Table 4. The table also contains the results of the iv administration of compound **10j** (0.625 mg/kg).



**Figure 7.** Preventive efficacy of 10j in the UTI mouse model 3 h after infection. The bars depict the median bacterial load with the interquartile range in the different study groups. Shown are the results of the control group (PBS), control group formulation (5% DMSO in PBS containing 1% Tween 80), and the intervention groups with the preventive applications of either 1.25 or 10 mg/kg 10j po or 8 mg/kg CIP sc (representing the murine dose equivalent to a human standard dose).<sup>81</sup> DL, detection limit. CFU, colony forming units.

In contrast to the fast plasma clearance of antagonists **9** and **10c** (Figure 5A), the methylsulfonamide bioisostere **10e** attained higher initial concentration in plasma ( $C_0$ ) and lower total clearance ( $CL_{tot}$ ). Therefore, it could be detected until 6 h after application, resulting in markedly higher plasma AUC. The observed high  $C_0$  of compound **10e** may be attributed to a small volume of distribution ( $V_z$ ) resulting from the high PPB ( $f_b \geq 99\%$ ).<sup>72</sup> In urine (Figure 5B), the carboxylic acid **9** and the morpholinomethanone **10c** displayed high levels immediately following administration and a rapid concentration decrease within the first 2 h, reflecting the rapid elimination from plasma. Fast renal excretion as major route of elimination can be rationalized by the physicochemical properties of the antagonists **9** and **10c**, that is, moderate PPB and  $\log D_{7.4}$ , as well as high metabolic stability. Otherwise, the methylsulfonamide bioisostere **10e** showed sustained compound levels in urine over a period of 2 h and subsequent slow decrease until 6 h after administration. This sustained renal excretion is a result of the interplay of the antagonist's elevated PPB and  $\log D_{7.4}$ .

In a second study, the *p*-cyano bioisostere **10j**, characterized by a high oral absorption potential, was administered as a single dose iv (0.625 mg/kg) and po (1.25 mg/kg). The plasma concentration curve upon iv dosing displays a steep decline within the first hour after application, while the po curve shows a prolonged period where absorption and elimination are in equilibrium (Figure 6A). The urine concentration profiles

(Figure 6B) parallel the plasma curves obtained by the two modes of application; i.e., high plasma clearance upon iv bolus injection led to high initial antagonist levels in urine and a rapid concentration decline. By contrast, sustained plasma concentrations upon po administration resulted in prolonged urine levels.

As a result, urine concentrations exceed the minimum level required for the antiadhesive effect as estimated from the *in vitro* cell infection model<sup>79</sup> (minimal antiadhesion concentration,<sup>23</sup>  $MAC_{90} = 0.094 \mu\text{g/mL}$ ) for more than 8 h upon oral single-dose administration (Figure 6B).

**Infection Study in C3H/HeN Mice.** In a preventive study, six mice were inoculated with UT189 following an oral application of **10j** (1.25 mg/kg) 40 min prior to infection. Three hours after inoculation, the animals were sacrificed and bladder and kidneys were removed. Organs were homogenized and analyzed for bacterial counts. The effect of the FimH antagonist was compared to a 8 mg/kg dose of ciprofloxacin (CIP), applied subcutaneously (sc) 10 min before infection. CIP is used as standard antibiotic therapy in humans for the treatment of UTI.<sup>80</sup> In mice, the dose of 8 mg/kg sc was shown to mimic the standard human dose regarding peak levels and the  $AUC_{24}$  in serum.<sup>81</sup> The median reductions in bacterial counts in mice treated with **10j** and CIP compared to the control group 3 h after infection are displayed in Figure 7.

The median value in the untreated control group showed bacterial counts of 6.6 log<sub>10</sub> colony forming units (CFU) in the bladder and 6 log<sub>10</sub> CFU in the kidneys. After oral application of 1.25 mg/kg **10j**, bacterial loads in the bladder decreased by 1.78 log<sub>10</sub> CFU and 1.07 log<sub>10</sub> CFU in the kidneys. The lower reduction in the kidneys is most likely due to the differing adhesion mechanisms between bladder and kidneys (type 1 pili vs P-pili), which is not targeted by **10j**.<sup>82</sup> With CIP (8 mg/kg sc) a substantial reduction in both bladder and kidneys (median reductions of 2.44 log<sub>10</sub> and 2.47 log<sub>10</sub>, respectively) was observed. Despite the low oral dose of **10j** (1.25 mg/kg), the approximately 100-fold reduction of CFU in the bladder promised an even higher effect upon dose increase to 10 mg/kg. Since the solubility of **10j** for this increased dose is too low (192 µg/mL), we used 5% DMSO and surfactant Tween 80 (1%) as solubilizer. To effectively compare the effect of a higher dose of **10j**, a control group receiving the formulation only (5% DMSO in PBS containing 1% Tween 80, termed control group formulation) was tested in parallel. When 10 mg/kg **10j** was applied, bacterial loads in the bladder decreased by 2.68 log<sub>10</sub> CFU/mL compared to the control group formulation, clearly exceeding the effect of CIP with a reduction of 2.44 log<sub>10</sub> CFU/mL. However, only a moderate reduction of 1.04 log<sub>10</sub> CFU was achieved in the kidneys.

## SUMMARY AND CONCLUSION

Recently, numerous monovalent alkyl and aryl  $\alpha$ -D-mannopyranosides have been described as potent FimH antagonists. However, most of them suffer from insufficient pharmacokinetic properties, i.e., modest bioavailability and short duration of the therapeutic effect in the bladder, their site of action. As a consequence, high doses at short intervals are required to achieve antiadhesive effects over an extended period of time. Therefore, the goal of the present study was an appropriate optimization of the pharmacokinetic profile of biphenyl  $\alpha$ -D-mannopyranosides while keeping their high affinity to the CRD of FimH. The starting point was the biphenylcarboxylate **9** where the critical carboxylate was replaced by bioisosteres.<sup>39,83</sup>

With a series of bioisosteres, a 3- to 5-fold improvement of affinity was achieved compared to **9**. Although binding necessitates only partial desolvation of the carboxylate and its bioisosteric replacements, a reduction of the enthalpy penalty for desolvation<sup>51</sup> was identified as the source of the improved affinity exhibited by the bioisosteres. Thermodynamic evaluation of antagonists **10b–e** revealed almost identical enthalpy contribution to binding. However, for antagonists with the *p*-cyano substituent (**10g** and **10j**) an enhancement of up to  $-8.7$  kJ/mol was observed, indicating a reduced desolvation penalty and an improved stacking as derived from the crystal structure of **10j** cocrystallized with the CRD of FimH (Figure 4B). On the other hand, higher affinity originating from a reduction of conformational flexibility of ligand and protein resulted in a concomitant entropy penalty of up to 6.5 kJ/mol.

In addition to the improved pharmacodynamics, the relevant pharmacokinetic parameters (solubility, permeability, renal excretion) were substantially improved. With 3'-chloro-4'-( $\alpha$ -D-mannopyranosyloxy)biphenyl-4-carbonitrile (**10j**), a FimH antagonist with an optimal in vitro PK/PD profile was identified. The *p*-cyano substituent conferred lipophilicity and high binding to plasma proteins, which slowed the rate of renal excretion. Despite higher lipophilicity, antagonist **10j** was unsusceptible to CYP450-mediated metabolism and therefore predominantly eliminated via the renal pathway. In vivo experiments confirmed

the excellent PK profile of **10j** with steady renal excretion for more than 8 h after oral application (1.25 mg/kg), suggesting a long-lasting antiadhesive effect. Finally, the preventive oral application of **10j** (10 mg/kg) reduced the bacterial load in the bladder by almost 1000-fold 3 h after infection. Although the first 3 h of the infection do not represent the complete infection cycle, they represent the time span of bacteria adhering and invading urothelial cells.<sup>84,85</sup> Nevertheless, the effect of FimH antagonist **10j** within a longer infection time and at higher dosing will be the subject of future investigations.

## EXPERIMENTAL SECTION

**Synthesis.** The synthesis of compounds **10a–d**, **10f**, **10g**, **10i**, **13a–d**, **13f**, **13g**, **15**, **18**, and **25**, including compound characterization data, can be found in the Supporting Information.

**General Methods.** NMR spectra were recorded on a Bruker Avance DMX-500 (500.1 MHz) spectrometer. Assignment of <sup>1</sup>H and <sup>13</sup>C NMR spectra was achieved using 2D methods (COSY, HSQC, HMBC). Chemical shifts are expressed in ppm using residual CHCl<sub>3</sub>, CHD<sub>2</sub>OD, or HDO as references. Optical rotations were measured using PerkinElmer polarimeter 341. Electron spray ionization mass spectra were obtained on a Waters micromass ZQ. The LC/HRMS analyses were carried out using a Agilent 1100 LC equipped with a photodiode array detector and a Micromass QTOF I equipped with a 4 GHz digital time converter. Microwave-assisted reactions were carried out with a CEM Discover and Explorer. Reactions were monitored by TLC using glass plates coated with silica gel 60 F<sub>254</sub> (Merck) and visualized by using UV light and/or by charring with a molybdate solution (a 0.02 M solution of ammonium cerium sulfate hydrate and ammonium molybdate tetrahydrate in aqueous 10% H<sub>2</sub>SO<sub>4</sub>). MPLC separations were carried out on a CombiFlash Companion or Rf (Teledyne Isco) equipped with RediSep normal-phase or RP-18 reversed-phase flash columns. LC–MS separations were done on a Waters system equipped with sample manager 2767, pump 2525, PDA 2525, and Micromass ZQ. All compounds used for biological assays are at least of 95% purity based on HPLC analytical results. Commercially available reagents were purchased from Fluka, Aldrich, Alfa Aesar, or abcr GmbH & Co. KG (Germany). Solvents were purchased from Sigma-Aldrich or Acros and were dried prior to use where indicated. Methanol (MeOH) was dried by refluxing with sodium methoxide and distilled immediately before use. Dimethoxyethane (DME) was dried by filtration over Al<sub>2</sub>O<sub>3</sub> (Fluka, type 5016 A basic).

**4'-(2,3,4,6-Tetra-O-acetyl- $\alpha$ -D-mannopyranosyloxy)-N-methylbiphenyl-4-sulfonamide (13e).** A Schlenk tube was charged with aryl iodide **11**<sup>22</sup> (116 mg, 0.21 mmol), 4-(N-methylsulfamoyl)phenylboronic acid (**12e**, 50 mg, 0.23 mmol), Pd(dppf)Cl<sub>2</sub>·CH<sub>2</sub>Cl<sub>2</sub> (5 mg, 0.006 mmol), K<sub>3</sub>PO<sub>4</sub> (67 mg, 0.32 mmol), and a stirring bar. The tube was closed with a rubber septum and was evacuated and flushed with argon. This procedure was repeated once, and then anhydrous DMF (1 mL) was added under a stream of argon. The mixture was degassed in an ultrasonic bath and flushed with argon for 5 min and then stirred at 80 °C overnight. The reaction mixture was cooled to rt, diluted with EtOAc (50 mL), and washed with water (50 mL) and brine (50 mL). The organic layer was dried over Na<sub>2</sub>SO<sub>4</sub> and concentrated in vacuo. The residue was purified by MPLC on silica gel (petroleum ether/EtOAc) to afford **13e** (105 mg, 84%) as a white solid. [ $\alpha$ ]<sub>D</sub><sup>20</sup> +56.4 (c 0.50, MeOH). <sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>):  $\delta$  = 7.92–7.90 (m, 2H, Ar–H), 7.70–7.68 (m, 2H, Ar–H), 7.57–7.55 (m, 2H, Ar–H), 7.21–7.19 (m, 2H, Ar–H), 5.60–5.57 (m, 2H, H-1, H-3), 5.48 (dd, *J* = 1.8, 3.4 Hz, 1H, H-2), 5.40 (t, *J* = 10.0 Hz, 1H, H-4), 4.38 (dd, *J* = 5.4, 10.8 Hz, 1H, NH), 4.30 (dd, *J* = 4.9, 12.3 Hz, 1H, H-6a), 4.13–4.08 (m, 2H, H-5, H-6b), 2.72 (d, *J* = 5.4 Hz, 3H, NCH<sub>3</sub>), 2.22, 2.07, 2.05, 2.04 (4 s, 12H, 4 COCH<sub>3</sub>). <sup>13</sup>C NMR (126 MHz, CDCl<sub>3</sub>):  $\delta$  = 170.55, 170.06, 170.03, 169.75 (4 CO), 155.97, 144.81, 137.16, 134.09, 128.62, 127.85, 127.39, 117.01 (Ar–C), 95.78 (C-1), 69.34 (C-5), 69.31 (C-2), 68.81 (C-3), 65.86 (C-4), 62.07 (C-6), 29.44 (NHCH<sub>3</sub>), 20.92, 20.74, 20.72 (4C, 4 COCH<sub>3</sub>). ESI-MS *m/z*, calcd for C<sub>27</sub>H<sub>31</sub>NNaO<sub>12</sub>S [M + Na]<sup>+</sup>: 616.1. Found: 616.1.

**4'-( $\alpha$ -D-Mannopyranosyloxy)-N-methylbiphenyl-4-sulfonamide (10e).** To a solution of **13e** (40 mg, 0.07 mmol) in dry MeOH (5 mL) was added freshly prepared 1 M NaOMe/MeOH (0.1 equiv) under argon. The mixture was stirred at rt until the reaction was complete (monitored by TLC), then neutralized with Amberlyst-15 ( $H^+$ ) ion-exchange resin, filtered, and concentrated in vacuo. The residue was purified by MPLC on silica gel (DCM/MeOH, 10:1 to 7:1) to afford **10e** (22 mg, 76%) as white solid.  $[\alpha]_D^{20} +105.7$  (c 0.30, MeOH).  $^1H$  NMR (500 MHz,  $CD_3OD$ ):  $\delta = 7.90$ – $7.88$  (m, 2H, Ar–H),  $7.80$ – $7.79$  (m, 2H, Ar–H),  $7.66$ – $7.64$  (m, 2H, Ar–H),  $7.26$ – $7.25$  (m, 2H, Ar–H),  $5.58$  (d,  $J = 1.7$  Hz, 1H, H-1),  $4.06$  (dd,  $J = 1.8, 3.3$  Hz, 1H, H-2),  $3.96$  (dd,  $J = 3.4, 9.5$  Hz, 1H, H-3),  $3.79$ – $3.74$  (m, 3H, H-4, H-6a, H-6b),  $3.63$  (ddd,  $J = 2.5, 5.2, 9.7$  Hz, 1H, H-5),  $2.57$  (s, 3H,  $NHCH_3$ ).  $^{13}C$  NMR (126 MHz,  $CD_3OD$ ):  $\delta = 158.34, 146.13, 138.67, 134.55, 129.53, 128.82, 128.21, 118.29$  (Ar–C),  $100.09$  (C-1),  $75.53$  (C-5),  $72.42$  (C-3),  $71.96$  (C-2),  $68.32$  (C-4),  $62.68$  (C-6),  $29.31$  ( $NHCH_3$ ). HRMS  $m/z$ , calcd for  $C_{19}H_{23}NNaO_8S [M + Na]^+$ : 448.1037. Found: 448.1038.

**5-(4'-(2,3,4,6-Tetra-O-acetyl- $\alpha$ -D-mannopyranosyloxy)biphenyl-4-yl)-1H-tetrazole (14).** A Schlenk tube was charged with **13g** (30 mg, 0.06 mmol), trimethylsilyl azide (16  $\mu$ L, 0.12 mmol), dibutyltin oxide (2 mg, 0.006 mmol), DME (1 mL), and a stirring bar. The mixture was heated to 150 °C for 10 min by microwave irradiation. The reaction mixture was cooled to rt and then concentrated in vacuo. The residue was purified by MPLC on silica gel (DCM/MeOH, 9:1 to 8:1) to afford **14** (26 mg, 81%) as a colorless oil.  $[\alpha]_D^{20} +56.1$  (c 0.3, MeOH).  $^1H$  NMR (500 MHz,  $CDCl_3$ ):  $\delta = 8.25$ – $8.15$  (m, 2H, Ar–H),  $7.75$ – $7.65$  (m, 2H, Ar–H),  $7.60$ – $7.55$  (m, 2H, Ar–H),  $7.20$ – $7.17$  (m, 2H, Ar–H),  $5.64$ – $5.55$  (m, 2H, H-1, H-3),  $5.49$  (dd,  $J = 1.7, 3.3$  Hz, 1H, H-2),  $5.40$  (t,  $J = 10.1$  Hz, 1H, H-4),  $4.31$  (dd,  $J = 5.3, 12.4$  Hz, 1H, H-6a),  $4.17$ – $4.06$  (m, 2H, H-5, H-6b),  $2.22, 2.07, 2.06, 2.05$  (4 s, 12H, 4  $COCH_3$ ).  $^{13}C$  NMR (126 MHz,  $CDCl_3$ ):  $\delta = 170.67, 170.14, 170.11, 169.81$  (4 CO),  $155.61, 128.36, 127.84, 127.49, 116.93$  (Ar–C),  $95.78$  (C-1),  $69.36$  (C-5),  $69.26$  (C-2),  $68.90$  (C-3),  $65.89$  (C-4),  $62.12$  (C-6),  $20.92, 20.76, 20.73$  (4  $COCH_3$ ). ESI-MS  $m/z$ , calcd for  $C_{27}H_{28}N_4NaO_{10} [M + Na]^+$ : 591.2. Found: 591.1.

**5-(4'-( $\alpha$ -D-Mannopyranosyloxy)biphenyl-4-yl)-1H-tetrazole (10h).** Prepared according to the procedure described for **10e** from **14** (26 mg, 0.03 mmol). Yield: 18 mg (quant) as a white solid.  $[\alpha]_D^{20} +112.1$  (c 0.1, MeOH/ $H_2O$ , 2:1).  $^1H$  NMR (500 MHz,  $CD_3OD$ ):  $\delta = 7.98$ – $7.96$  (m, 2H, Ar–H),  $7.72$ – $7.71$  (m, 2H, Ar–H),  $7.58$ – $7.54$  (m, 2H, Ar–H),  $7.16$ – $7.13$  (m, 2H, Ar–H),  $5.46$  (d,  $J = 1.7$  Hz, 1H, H-1),  $3.94$  (dd,  $J = 1.9, 3.5$  Hz, 1H, H-2),  $3.83$  (dd,  $J = 3.4, 9.5$  Hz, 1H, H-3),  $3.68$ – $3.61$  (m, 3H, H-4, H-6a, H-6b),  $3.52$  (ddd,  $J = 2.5, 5.4, 9.7$  Hz, 1H, H-5).  $^{13}C$  NMR (126 MHz,  $CD_3OD$ ):  $\delta = 158.19, 145.07, 134.97, 129.29, 128.74, 128.55, 118.26$  (Ar–C),  $100.13$  (C-1),  $75.52$  (C-5),  $72.42$  (C-3),  $71.98$  (C-2),  $68.33$  (C-4),  $62.69$  (C-6). HRMS  $m/z$ , calcd for  $C_{19}H_{21}N_4O_6 [M + H]^+$ : 401.1456. Found: 401.1450.

**4'-(2,3,4,6-Tetra-O-acetyl- $\alpha$ -D-mannopyranosyloxy)-3'-chlorobiphenyl-4-carbonitrile (19).** Prepared according to the procedure described for **13e** from aryl iodide **18<sup>23</sup>** (79 mg, 0.135 mmol), **12g** (22 mg, 0.15 mmol), Pd(dppf)Cl<sub>2</sub>·CH<sub>2</sub>Cl<sub>2</sub> (3.3 mg, 4  $\mu$ mol), and K<sub>3</sub>PO<sub>4</sub> (57 mg, 0.27 mmol). Yield: 57 mg (75%) as a white solid.  $[\alpha]_D^{20} +77.7$  (c 0.5,  $CHCl_3$ ).  $^1H$  NMR (500 MHz,  $CDCl_3$ ):  $\delta = 7.72$  (d,  $J = 8.3$  Hz, 2H, Ar–H),  $7.63$  (m, 3H, Ar–H),  $7.43$  (dd,  $J = 2.2, 8.6$  Hz, 1H, Ar–H),  $7.27$  (d,  $J = 8.6$  Hz, 1H, Ar–H),  $5.64$ – $5.59$  (m, 2H, H-1, H-2),  $5.54$  (dd,  $J = 1.9, 3.2$  Hz, 1H, H-3),  $5.41$  (t,  $J = 10.1$  Hz, 1H, H-4),  $4.28$  (dd,  $J = 5.2, 12.3$  Hz, 1H, H-6a),  $4.17$  (ddd,  $J = 2.1, 5.1, 10.0$  Hz, 1H, H-5),  $4.10$  (dd,  $J = 2.2, 12.3$  Hz, 1H, H-6b),  $2.21$  (s, 3H,  $COCH_3$ ),  $2.12$ – $2.00$  (m, 9H, 3  $COCH_3$ ).  $^{13}C$  NMR (126 MHz,  $CDCl_3$ ):  $\delta = 170.54, 170.08, 169.90, 169.84, 143.00, 151.67, 143.61, 135.29, 132.87, 129.41, 127.53, 126.60, 125.20, 118.79, 117.36, 111.47$  (Ar–C, CN),  $96.72$  (C-1),  $70.00$  (C-5),  $69.39$  (C-3),  $68.82$  (C-2),  $65.86$  (C-4),  $62.16$  (C-6),  $20.98, 20.81, 20.79, 20.78$  (4  $COCH_3$ ). ESI-MS  $m/z$ , calcd for  $C_{27}H_{26}ClNNaO_{10} [M + Na]^+$ : 582.1. Found: 582.1.

**3'-Chloro-4'-( $\alpha$ -D-mannopyranosyloxy)biphenyl-4-carbonitrile (10j).** Prepared according to the procedure described for **10e** from **19** (36 mg, 0.06 mmol). Yield: 12 mg (48%) as a white solid.  $[\alpha]_D^{20} +109.4$  (c 0.23, MeOH).  $^1H$  NMR (500 MHz,  $CD_3OD$ ):  $\delta = 7.80$ – $7.72$  (m, 5H, Ar–H),  $7.59$  (dd,  $J = 2.2, 8.6$  Hz, 1H, Ar–H),  $7.48$  (d,  $J = 8.7$  Hz, 1H, Ar–H),  $5.62$  (d,  $J = 1.4$  Hz, 1H, H-1),  $4.12$  (dd,  $J = 1.8, 3.3$  Hz, 1H,

H-2),  $4.00$  (dd,  $J = 3.4, 9.5$  Hz, 1H, H-3),  $3.83$ – $3.68$  (m, 3H, H-4, H-6a, H-6b),  $3.63$  (ddd,  $J = 2.3, 5.4, 9.6$  Hz, 1H, H-5).  $^{13}C$  NMR (126 MHz,  $CD_3OD$ ):  $\delta = 153.65, 145.15, 135.42, 133.86, 129.82, 128.53, 127.87, 125.47, 119.70, 118.59$  (Ar–C),  $111.97$  (CN),  $100.66$  (C-1),  $76.05$  (C-5),  $72.39$  (C-3),  $71.80$  (C-2),  $68.20$  (C-4),  $62.65$  (C-6). IR (KBr),  $\nu = 3400$  (OH),  $2227$  ( $C\equiv N$ ),  $1606, 1487$  (Ar–C=C)  $cm^{-1}$ . HRMS  $m/z$ , calcd for  $C_{19}H_{18}ClNNaO_6 [M + Na]^+$ : 414.0715. Found: 414.0721.

**3'-Chloro-N-(3',6'-dihydroxy-3-oxo-3H-spiro[isobenzofuran-1,9'-xanthen]-5-yl)-4'-( $\alpha$ -D-mannopyranosyloxy)biphenyl-4-carboxamide (22).** Compound **21** (10.0 mg, 0.024 mmol), fluoresceinamine isomer I (12.7 mg, 0.037 mmol), and COMU (20.9 mg, 0.049 mmol) were dissolved in dry DMF (1 mL). Then  $NEt_3$  (10  $\mu$ L, 0.073 mmol) was added and the mixture was stirred at rt for 7 h. 1 N HCl in DMF was added until acid reaction on pH paper and the mixture was concentrated. The residue was dissolved in DCM/MeOH (3:1) and loaded onto a silica gel column. The complex mixture of compounds was only partially resolved. The fractions containing the product were collected, concentrated, and purified by preparative HPLC (gradient  $H_2O$ /MeCN, +0.2%  $HCO_2H$ ) to afford compound **22** (5 mg, 19%).  $[\alpha]_D^{20} +21.1$  (c 0.10, MeOH).  $^1H$  NMR (500 MHz,  $CD_3OD$ ):  $\delta = 8.26$  (d,  $J = 8.4$  Hz, 2H, Ar–H),  $7.88$ – $7.74$  (m, 3H, Ar–H),  $7.66$  (dd,  $J = 2.2, 8.6$  Hz, 1H, Ar–H),  $7.51$  (d,  $J = 8.7$  Hz, 1H, Ar–H),  $7.29$  (dd,  $J = 1.9, 5.3$  Hz, 2H, Ar–H),  $7.19$  (dd,  $J = 2.1, 8.3$  Hz, 1H, Ar–H),  $7.08$ – $6.99$  (m, 2H, Ar–H),  $6.95$  (d,  $J = 8.7$  Hz, 1H, Ar–H),  $6.72$  (dd,  $J = 5.5, 10.6$  Hz, 2H, Ar–H),  $6.61$  (dd,  $J = 2.3, 8.7$  Hz, 1H, Ar–H),  $5.65$  (s, 1H, H-1),  $4.15$  (dd,  $J = 1.8, 3.2$  Hz, H-2),  $4.03$  (dd,  $J = 3.4, 9.5$  Hz, H-3),  $3.87$ – $3.72$  (m, 3H, H-4, H-6a, H-6b),  $3.65$  (m, 1H, H-5).  $^{13}C$  NMR (126 MHz,  $CD_3OD$ ):  $\delta = 137.50, 136.01, 131.90, 130.24, 130.20, 129.87, 129.24, 128.03, 127.91, 125.79, 125.46, 124.73, 118.99, 118.76, 118.65$  (Ar–C),  $100.73$  (C-1),  $76.06$  (C-5),  $72.42$  (C-3),  $71.85$  (C-2),  $68.24$  (C-4),  $62.69$  (C-2). ESI-MS  $m/z$ , calcd for  $C_{39}H_{31}ClNO_{12} [M + H]^+$ : 740.2. Found: 740.2.

**3'-Chloro-N-(2-(3-(3',6'-dihydroxy-3-oxo-3H-spiro[isobenzofuran-1,9'-xanthen]-5-yl)thioureido)ethyl)-4'-( $\alpha$ -D-mannopyranosyloxy)biphenyl-4-carboxamide (23).** To a stirred solution of compound **21** (25 mg, 0.061 mmol) in dry DMF (1 mL), NHS (21 mg, 0.183 mmol) was added, followed by DIC (9.2 mg, 0.073 mmol). The mixture was stirred at rt for 2 h. Then *N*-Boc-ethylendiamine (10.7 mg, 0.067 mmol) was added and the reaction was stirred for 10 h. It was then cooled down to 0 °C, diluted with water, and concentrated. Chromatography on silica gel (DCM/MeOH) yielded 23 mg (0.042 mmol, 68%) of *tert*-butyl (3'-chloro-4'-( $\alpha$ -D-mannopyranosyloxy)biphenyl-4-yl-carboxamido)ethyl)carbamate. This product was dissolved in DCM (3 mL), and TFA (1 mL) was added. The solid dissolved during addition of TFA. After 10 min the reaction was complete. The mixture was evaporated, and excess TFA was removed in high vacuum. The intermediate *N*-(2-aminoethyl)-3'-chloro-4'-( $\alpha$ -D-mannopyranosyloxy)biphenyl-4-carboxamide TFA salt (23 mg, 0.042 mmol, quant) was used directly in the next step. It was dissolved in dry DMF (0.5 mL), and  $NEt_3$  (12.8 mg, 0.127 mmol) was added. The mixture was cooled to 0 °C. Then FITC (14.8 mg, 0.038 mmol) was added and the mixture was stirred for 3 h in the dark. The mixture was then coevaporated with water, taken up in MeOH/10% aq acetic acid and evaporated. Chromatography on silica gel (DCM/MeOH) yielded compound **23**, contaminated with triethylammonium acetate. The compound was then redissolved in MeOH, and 0.5 N HCl in MeOH was added. The mixture was evaporated and chromatographed on silica gel to yield pure **23** (15 mg, 47%).  $[\alpha]_D^{20} +12.1$  (c 0.30, MeOH).  $^1H$  NMR (500 MHz,  $CD_3OD$ ):  $\delta = 8.12$  (s, 1H),  $7.92$  (d,  $J = 8.3$  Hz, 2H, Ar–H),  $7.70$  (dd,  $J = 5.0, 13.1$  Hz, 2H, Ar–H),  $7.64$  (d,  $J = 8.3$  Hz, 2H, Ar–H),  $7.54$  (dd,  $J = 2.2, 8.6$  Hz, 1H, Ar–H),  $7.46$  (d,  $J = 8.7$  Hz, 1H, Ar–H),  $7.09$  (d,  $J = 8.2$  Hz, 1H, Ar–H),  $6.74$  (s, 2H),  $6.69$  (d,  $J = 1.4$  Hz, 2H, Ar–H),  $6.55$  (d,  $J = 8.4$  Hz, 2H, Ar–H),  $5.63$  (d,  $J = 1.3$  Hz, H-1),  $4.15$  (dd,  $J = 1.8, 3.1$  Hz, H-2),  $4.03$  (dd,  $J = 3.4, 9.5$  Hz, H-3),  $3.94$  (s, 2H,  $CH_2$ ),  $3.86$ – $3.64$  (m, 6H, H-4, H-5, H-6,  $CH_2$ ).  $^{13}C$  NMR (126 MHz,  $CD_3OD$ ):  $\delta = 153.21, 143.84, 136.41, 129.66, 129.18, 127.76, 127.70, 125.37, 118.64, 103.62$  (Ar–C),  $100.75$  (C-1),  $76.00$  (C-5),  $72.41$  (C-3),  $71.86$  (C-2),  $68.24$  (C-4),  $62.69$  (C-6),  $40.76$  ( $CH_2$ ). ESI-MS  $m/z$ , calcd for  $C_{42}H_{37}ClN_3O_{12}S [M + H]^+$ : 842.2. Found: 842.2.

**3'-Chloro-N-(2-(2-(2-(3-(3',6'-dihydroxy-3-oxo-3H-spiro[isobenzofuran-1,9'-xanthen]-5-yl)thioureido)ethoxy)ethoxy)ethyl)-4'-( $\alpha$ -D-mannopyranosyloxy)biphenyl-4-carboxamide (24).** Compound 21 (280 mg, 0.68 mmol) was dissolved in dry DMF (5 mL) under argon. Then NHS (235 mg, 2.04 mmol) was added, followed by DIC (0.12 mL, 0.78 mmol) and the mixture was stirred at rt for 4 h. Then Boc-PEG2-NH<sub>2</sub> (186 mg, 0.75 mmol) was added, and the mixture was stirred at rt under argon for 10 h. It was then slowly diluted with water and concentrated. The residue was purified by chromatography on silica gel (DCM/MeOH) to give *tert*-butyl (2-(2-(2-(3'-chloro-4'-( $\alpha$ -D-mannopyranosyloxy)biphenyl-4-ylcarboxamido)ethoxy)ethoxy)ethyl)-carbamate (300 mg, 0.468 mmol, 69%). Then the carbamate was suspended in DCM (3 mL), and TFA (1 mL) was added dropwise at rt. After 30 min, the solvents were evaporated and the crude mixture was dissolved in CHCl<sub>3</sub>/MeOH (6:4, +0.5% conc NH<sub>4</sub>OH) and transferred to a silica gel column, eluting with the same solvent mixture, to yield *N*-(2-(2-(2-aminoethoxy)ethoxy)ethyl)-3'-chloro-4'-( $\alpha$ -D-mannopyranosyloxy)biphenyl-4-carboxamide (228 mg, 90%). A fraction of the amine (10 mg, 0.018 mmol) was dissolved in dry DMF (0.5 mL) and cooled to 0 °C. FITC (6.5 mg, 0.017 mmol) was added, and the mixture was stirred for 1 h. The mixture was concentrated and the residue was purified by chromatography on silica (DCM/MeOH) to yield 24 (10 mg, 65%). <sup>1</sup>H NMR (500 MHz, CD<sub>3</sub>OD):  $\delta$  = 8.21 (d, *J* = 1.4 Hz, 1H, Ar-H), 7.88 (d, *J* = 8.3 Hz, 2H, Ar-H), 7.68 (d, *J* = 2.2 Hz, 2H, Ar-H), 7.63 (d, *J* = 8.3 Hz, 2H, Ar-H), 7.53 (dd, *J* = 2.2, 8.6 Hz, 1H, Ar-H), 7.43 (d, *J* = 8.7 Hz, 1H, Ar-H), 7.09 (d, *J* = 8.2 Hz, 1H, Ar-H), 6.68 (d, *J* = 2.3 Hz, 2H, Ar-H), 6.65 (dd, *J* = 2.6, 8.6 Hz, 2H, Ar-H), 6.53 (dd, *J* = 1.6, 8.7 Hz, 2H, Ar-H), 5.61 (d, *J* = 1.3 Hz, 1H, H-1), 4.14 (dd, *J* = 1.8, 3.2, Hz, 1H, H-2), 4.03 (dd, *J* = 3.4, 9.5 Hz, 1H, H-3), 3.93–3.53 (m, 16H), 3.37 (s, 2H, NCH<sub>2</sub>), 1.30 (s, 2H, CH<sub>2</sub>). <sup>13</sup>C NMR (126 MHz, CD<sub>3</sub>OD):  $\delta$  = 170.01 (CO), 153.17, 143.72, 136.37, 134.37, 130.39, 129.69, 129.04, 127.78, 127.73, 125.35, 118.60, 103.60 (Ar-C), 100.72 (C-1), 75.97 (C-5), 72.41 (C-3), 71.86, 71.40, 70.59 (5C, C-2, OCH<sub>2</sub>), 68.23 (C-4), 62.64 (C-6), 49.88, 45.49, 40.97 (CH<sub>2</sub>). ESI-MS *m/z*, calcd for C<sub>46</sub>H<sub>45</sub>ClN<sub>3</sub>O<sub>14</sub>S [M + H]<sup>+</sup>: 930.2. Found: 930.4.

**Competitive Fluorescence Polarization Assay.** *Expression and Purification of CRD of FimH.* A recombinant protein consisting of the CRD of FimH linked to a 6His-tag via a thrombin cleavage site (FimH-CRD-Th-His<sub>6</sub>) was expressed in *E. coli* strain HM125 and purified by affinity chromatography as previously described.<sup>43</sup>

*K<sub>D</sub> Determination of FITC-Labeled Ligands.* The functionalized ligands (23, 24) were prepared as a 10 mM stock solution in pure DMSO (Sigma-Aldrich, Buchs, Switzerland). All further dilutions of compounds and FimH-CRD-Th-His<sub>6</sub> protein were prepared in assay buffer (20 mM HEPES, 150 mM NaCl, 50  $\mu$ g/mL BSA, pH 7.4). BSA was added to the assay buffer to prevent nonspecific binding of protein to the plastic surface. Binding isotherms for the fluorescent ligands were obtained in direct binding studies by adding a constant concentration of ligand (final concentration 5 nM) and a linear dilution of protein (final concentration 0–100 nM) to a final volume of 200  $\mu$ L in 96-well, black, flat bottom NBS plates (Corning Inc., Corning, NY, USA). After incubation of the plate for 24 h at rt with gentle shaking, the fluorescence polarization was measured with the Synergy HI hybrid multimode microplate reader (BioTek Instruments Inc., Winooski, VT, USA) with polarized excitation at 485 nm and emission measured at 528 nm through polarizing filters parallel and perpendicularly oriented to the incident polarized light. *K<sub>D</sub>* values were determined by plotting the FP readout as a function of the protein concentration and applying the following single-site binding equation (eq 1) that accounts for ligand depletion:

$$S_{\text{obs}} = S_{\text{F}} + (S_{\text{B}} - S_{\text{F}}) \times \left( \frac{C_{\text{p}} + C_{\text{L}} + K_{\text{D}} - \sqrt{(C_{\text{p}} + C_{\text{L}} + K_{\text{D}})^2 - 4C_{\text{p}}C_{\text{L}}}}{2C_{\text{L}}} \right) \quad (1)$$

where *S<sub>obs</sub>* is the observed signal from the ligand, *S<sub>F</sub>* is the signal from free ligand, *S<sub>B</sub>* is the signal from bound ligand, *C<sub>p</sub>* is the total concentration of protein, and *C<sub>L</sub>* is the total concentration of ligand.<sup>49</sup>

*K<sub>D</sub> Determination of FimH Antagonists.* The fluorescently labeled ligand 23 was used for the competitive fluorescence polarization assay. A linear dilution of nonlabeled FimH antagonist with final concentrations ranging from 0 to 10  $\mu$ M was titrated into 96-well, black, flat-bottom NBS plates (Corning Inc.) to a final volume of 200  $\mu$ L containing a constant concentration of protein (final concentration 25 nM) and FITC-labeled ligand which was fixed at a higher concentration in competitive binding assays than in direct binding experiments to obtain higher fluorescence intensities (final concentration 20 nM). Prior to measuring the fluorescence polarization, the plates were incubated on a shaker for 24 h at rt until the reaction reached equilibrium. The IC<sub>50</sub> value was determined with Prism (GraphPad Software Inc., La Jolla, CA, USA) by applying a standard four-parameter IC<sub>50</sub> function. The obtained IC<sub>50</sub> values were converted into their corresponding *K<sub>D</sub>* values using the derivation of the Cheng–Prusoff equation.<sup>45</sup> This variation of the Cheng–Prusoff equation is applied to competition assays with tight-binding inhibitors and includes terms to correct for ligand depletion effects. However, the *K<sub>D</sub>* for antagonists having a higher affinity toward FimH than the labeled ligand could not be accurately determined.<sup>45</sup>

**Isothermal Titration Calorimetry (ITC).** All ITC experiments were performed with the FimH-CRD-Th-His<sub>6</sub> protein using a VP-ITC instrument from MicroCal, Inc. (Malvern Instruments, Worcestershire, U.K.) with a sample cell volume of 1.4523 mL. The measurements were performed with 0–5% DMSO at 25 °C, a stirring speed of 307 rpm, and 10  $\mu$ cal s<sup>-1</sup> reference power. The protein samples were dialyzed in assay buffer prior to all experiments. Because of the high protein consumption of ITC, only the experiments for the reference compounds (1, 3, and 25) were measured in duplicates. Compounds 1, 3, 9, and 25 were measured in a direct fashion by titration of ligand (100–2,000  $\mu$ M) into protein (8.6–55  $\mu$ M) with injections of 3–8  $\mu$ L at intervals of 10 min to ensure nonoverlapping peaks. The quantity *c* = Mt(0) *K<sub>D</sub>*<sup>-1</sup>, where Mt(0) is the initial macromolecule concentration, is of importance in titration microcalorimetry. The *c*-values of the direct titrations were below 1000 and thus within the reliable range. For the compounds 10b–e, 10g, and 10j additional competitive ITC experiments were performed because of their high affinity resulting in *c*-values above 1000 for direct titrations. These ligands (600  $\mu$ M) were titrated into protein (30  $\mu$ M), which was preincubated with compound 25 (300  $\mu$ M) resulting in sigmoidal titration curves. Because of slow reaction kinetics, titration intervals of 20 min were used.

Baseline correction and peak integration were performed using the Origin 7 software (OriginLab, Northampton, MA, USA). An initial 2  $\mu$ L injection was excluded from data analysis. Baseline subtraction and curve-fitting with the three variables *N* (concentration correction factor), *K<sub>D</sub>* (dissociation constant), and  $\Delta H^\circ$  (change in enthalpy) were performed with the SEDPHAT software, version 10.40 (National Institutes of Health).<sup>86</sup> A global fitting analysis was performed for the competition titration (10b–e, 10g, or 10j competing for the protein binding site with compound 25) and the direct titration of the competitor (compound 25 binding to protein) to fit for *K<sub>D</sub>*,  $\Delta H^\circ$  and *N* were fitted from direct titrations of 10b–e, 10g, or 10j into protein. For the compounds 3, 9, and 25 binding to protein all variables could be determined from a global analysis of the direct titration.

The thermodynamic parameters were calculated with the following equation (eq 2):

$$\Delta G^\circ = \Delta H^\circ - T\Delta S^\circ = RT \ln K_{\text{D}} = -RT \ln K_{\text{A}} \quad (2)$$

where  $\Delta G^\circ$ ,  $\Delta H^\circ$ , and  $\Delta S^\circ$  are the changes in free energy, enthalpy, and entropy of binding, respectively, *T* is the absolute temperature, and *R* is the universal gas constant (8.314 J mol<sup>-1</sup> K<sup>-1</sup>). The 95% confidence intervals of the measurements were calculated for the two variables *K<sub>D</sub>* and  $\Delta H^\circ$  with the one-dimensional error surface projection within the SEDPHAT software.

*Calculation of the Free Energy of Desolvation.* The three-dimensional representation for each of the aglycons (4-methoxybiphenyl scaffold, Figure 8) was built in the Maestro<sup>87</sup> modeling environment, and the global minimum conformation was identified by performing 500 iterations of the mixed torsional/low-mode conformational sampling in combination with the OPLS-2005 force-field and the implicit solvent model (water) as implemented in the MacroModel

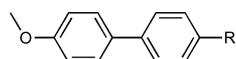


Figure 8. 4-Methoxybiphenyl scaffold of aglycons.

9.9.<sup>88</sup> The global minimum structures were used as input for the AMSOL 7.1 program<sup>89</sup> to obtain the free energy of desolvation  $\Delta G_{\text{des}}$  (Table 5) with the SM5.4A solvation model<sup>90</sup> and the AM1<sup>91</sup> level of theory (used keywords “AM1 SM5.4A SOLVNT=WATER TRUES”).

Table 5. Aqueous Free Energy of Desolvation

R	$\Delta G_{\text{des}}$ [kJ/mol]
neutral	
H	15.6
CONHCH <sub>3</sub>	39.9
COOCH <sub>3</sub>	23.0
SO <sub>2</sub> NHCH <sub>3</sub>	65.5
SO <sub>2</sub> CH <sub>3</sub>	56.4
4-morpholineamide	45.3
CN	22.0
deprotonated	
COO <sup>-</sup>	298.2
SO <sub>2</sub> -N <sup>-</sup> -Me	342.0

**Determination of the MAC<sub>90</sub> by Flow Cytometry.** The MAC<sub>90</sub> was determined in principle as in the previously published flow cytometry assay<sup>79</sup> but with some modifications. The human epithelial bladder carcinoma cell line 5637 (DSMZ, Braunschweig, Germany) was grown in RPMI 1640 medium, supplemented with 10% fetal calf serum (FCS), 100 U/mL penicillin, and 100  $\mu$ g/mL streptomycin at 37 °C, 5% CO<sub>2</sub>. All solutions were purchased from Invitrogen (Basel, Switzerland). The cells were subcultured 1:6 twice per week [using trypsin/EDTA (Sigma-Aldrich) for the detachment]. Two days before infection,  $1.8 \times 10^5$  cells were seeded in each well of a 24-well plate in RPMI 1640 containing 10% FCS without antibiotics. The cell density was approximately  $(3-5) \times 10^5$  cells/well at the assay day.

For infection, the GFP-expressing clinical *E. coli* isolate UTI89<sup>92</sup> (UTI89 wt) and the GFP-expressing FimA-H knockout strain UTI89  $\Delta$ fimA-H were used (strains were provided by Prof. Urs Jenal, Biocenter, University of Basel, Switzerland).<sup>79</sup> Bacteria were cultivated at 37 °C in 10 mL Luria–Bertani (LB) broth (Becton, Dickinson and Company) overnight, harvested by centrifugation (3800 rpm, 10 min), and washed three times in phosphate buffered saline (PBS, Sigma-Aldrich), and a bacterial solution of OD<sub>600</sub> of 0.75 in RPMI + 10% FCS was prepared. For the determination of the MAC<sub>90</sub> value, the IC<sub>90</sub>, linear dilutions of the FimH antagonist were prepared in 5% DMSO and PBS. Bacteria and antagonists were preincubated for 10 min at 37 °C, before cells were infected with either only 200  $\mu$ L of bacterial solution of UTI89 or UTI89  $\Delta$ fimA-H (positive and negative controls), or 225  $\mu$ L of the preincubated bacteria–antagonist mixture. Infection lasted for 1.5 h. During this time infected cells were incubated at 37 °C. Then, cells were washed with PBS and detached from wells by the addition of 150  $\mu$ L of trypsin and incubation at 37 °C for 10 min, before flushing from wells PBS containing 2% FCS and transferred to tubes. To dilute the trypsin, cells were centrifuged at 13 000 rpm, 1 min, 600  $\mu$ L of the supernatant was discarded, and the pellet was resuspended in the remaining 300  $\mu$ L of PBS containing 2% FCS. Samples were stored on ice until measurement. Before analysis with the flow cytometer (Becton Dickinson, FACSCanto II), the samples were gently mixed and filtered using a 35  $\mu$ m nylon mesh (Corning Life Sciences) to prevent cellular aggregation. Cells were gated with linear scaling for side scatter (SSC) and forward scatter (FSC) and GFP intensity of live cells was evaluated. IC<sub>90</sub> values were determined by plotting the concentration of the antagonist in a logarithmic mode versus the mean fluorescence intensity (MFI) of living cells and by fitting a dose–response curve (variable slope, four parameters) with the Prism software (GraphPad Prism).

**X-ray Analysis of the Antagonists 10e and 10j Cocrystallized with FimH-CRD.** *FimH-CRD/10e Cocrystallization.* Initial FimH-CRD (18 mg/mL in 20 mM HEPES, pH 7.4) crystals were obtained in complex with 4-(5-nitroindolin-1-yl)phenyl  $\alpha$ -D-mannopyranoside (5 mM).<sup>23</sup> Crystals were grown in sitting-drop vapor diffusion at 20 °C with 200 nL of protein–antagonist mixture together with 200 nL of precipitant solution in well D3 (0.2 M sodium phosphate monobasic monohydrate, 20% w/v PEG 3,350) of the PEG/Ion HT screen (Hampton Research, CA, USA). Cubic crystals appeared within 1 week, which served as cross-seeding crystals. A solution of FimH-CRD (20 mg/mL) and 10e (5 mM) was mixed with 0.2 M sodium phosphate monobasic monohydrate, 20% w/v PEG 400 with 0.5  $\mu$ L of each solution. Streak-seeding was performed after 1 day of incubation. Cubic FimH-CRD/10e crystals formed within 24 h. Crystals were flash cooled to 100 K with perfluoropolyether cryo oil (Hampton Research, CA, USA) as cryoprotectant. Data were collected with synchrotron radiation ( $\lambda = 0.99999$  Å) at the PXIII beamline, Swiss Light Source, Switzerland.

*FimH-CRD/10j Cocrystallization.* Cocrystals were initially grown in sitting-drop vapor diffusion at 20 °C with 0.5  $\mu$ L of a mixture of FimH-CRD (20 mg/mL) and 10j (5 mM) together with 0.5  $\mu$ L of 0.1 M HEPES, pH 7.5, 2 M ammonium sulfate. Platelike crystals formed within 2 weeks and were used as seeds for subsequent crystallization. Diffraction quality crystals were grown by streak-seeding in 0.5  $\mu$ L of FimH-CRD (10 mg/mL) with 10j (2.5 mM) and 0.5  $\mu$ L of 0.1 M HEPES, pH 7.5, 1.25 M ammonium sulfate. The drops were covered with perfluoropolyether cryo oil prior to flash cooling to 100 K. Data were collected with synchrotron radiation ( $\lambda = 1.00003$  Å) at the PXIII beamline, Swiss Light Source, Switzerland.

*Structure Determination and Refinement.* Data were indexed and integrated with the XDS package<sup>93</sup> for the FimH-CRD/10e cocrystal structure, and with mosflm<sup>94</sup> for the FimH-CRD/10j cocrystal structure (Table 6). Scaling was performed with XDS and SCALA included in the CCP4 suite, respectively.<sup>95</sup> Structures were solved by molecular

Table 6. Data Collection and Refinement Statistics for FimH-CRD/10e and FimH-CRD/10j Cocrystals

	FimH-CRD/10e	FimH-CRD/10j
PDB code	4CSS	4CST
space group	<i>P</i> 2 <sub>1</sub> 2 <sub>1</sub>	<i>P</i> 2 <sub>1</sub> 2 <sub>1</sub>
no. of molecules in the asymmetric unit	1	1
Cell Dimensions		
<i>a</i> , <i>b</i> , <i>c</i> (Å)	48.38, 56.23, 61.59	48.84, 55.89, 61.00
$\alpha$ , $\beta$ , $\gamma$ (deg)	90, 90, 90	90, 90, 90
Data Collection		
beamline	Swiss Light Source PXIII	Swiss Light Source PXIII
resolution range (Å) <sup>a</sup>	30.0–1.07 (1.13–1.07)	23.5–1.10 (1.12–1.10)
unique observations <sup>a</sup>	72000 (9354)	66470 (2500)
average multiplicity <sup>a</sup>	10.9 (3.7)	5.4 (2.4)
completeness (%)	96.1 (78.0)	97.2 (76.5)
<i>R</i> <sub>merge</sub> <sup>a</sup>	0.056 (0.57)	0.051 (0.305)
mean <i>I</i> / $\sigma$ ( <i>I</i> ) <sup>a</sup>	21.5 (2.22)	15.5 (2.9)
Refinement		
resolution range (Å)	15.7–1.07	23.5–1.10
<i>R</i> , <i>R</i> <sub>free</sub>	11.2, 13.2	11.4, 13.0
rms deviation from ideal bond length (Å)	0.010	0.010
rms deviation from ideal bond angle (deg)	1.170	1.420

<sup>a</sup>Values in parentheses are for highest-resolution shell.

replacement with PHASER<sup>96</sup> using the FimH-CRD-butyl  $\alpha$ -D-mannopyranoside complex (PDB code 1UWF) as search model. The structures were iteratively built using the COOT software<sup>97</sup> and refined with the PHENIX software.<sup>98</sup> Geometric restraints for **10e** and **10j** were generated with PRODRG.<sup>99</sup> The models were validated using molprobit.<sup>100</sup> Residues 113–115 were not modeled in the **10e** structure because of disorder. Furthermore, the ligand was modeled in two possible conformations. For both ligands, electron density is reduced on the second aromatic ring because of flexibility of the ligand.

#### Physicochemical and in Vitro Pharmacokinetic Studies.

**Materials.** Dimethyl sulfoxide (DMSO), 1-propanol, 1-octanol, Dulbecco's modified Eagle medium (DMEM)–high glucose, L-glutamine solution, penicillin–streptomycin solution, Dulbecco's phosphate buffered saline (DPBS), trypsin–EDTA solution, magnesium chloride hexahydrate, and reduced nicotinamide adenine dinucleotide phosphate (NADPH) were purchased from Sigma-Aldrich. MEM nonessential amino acid (MEM-NEAA) solution, fetal bovine serum (FBS), and DMEM without sodium pyruvate and phenol red were bought from Invitrogen (Carlsbad, CA, USA). PRISMA HT universal buffer, GIT-0 Lipid Solution, and Acceptor Sink Buffer were ordered from pIon (Woburn, MA, USA). Human plasma was bought from Biopredic (Rennes, France), and acetonitrile (MeCN) and methanol (MeOH) were from Acros Organics (Geel, Belgium). Pooled male rat liver microsomes were purchased from BD Bioscience (Franklin Lakes, NJ, USA). Tris(hydroxymethyl)aminomethane (TRIS) was obtained from AppliChem (Darmstadt, Germany). The Caco-2 cells were kindly provided by Prof. G. Imanidis, FHNW, Muttetz, and originated from the American Type Culture Collection (Rockville, MD, USA).

$pK_a$ . The  $pK_a$  values were determined as described elsewhere.<sup>101</sup> In brief, the pH of a sample solution was gradually changed and the chemical shift of protons adjacent to ionizable centers was monitored by <sup>1</sup>H nuclear magnetic resonance (NMR) spectroscopy. The shift was plotted against the pH of the respective sample, and the  $pK_a$  was read out from the inflection point of the resulting sigmoidal curve.

$\log D_{7.4}$ . The in silico prediction tool ALOGPS<sup>102</sup> was used to estimate  $\log P$  values of the compounds. Depending on these values, the compounds were classified into three categories: hydrophilic compounds ( $\log P$  below zero), moderately lipophilic compounds ( $\log P$  between zero and one), and lipophilic compounds ( $\log P$  above one). For each category, two different ratios (volume of 1-octanol to volume of buffer) were defined as experimental parameters (Table 7).

**Table 7. Compound Classification Based on Estimated  $\log P$  Values**

compd type	$\log P$	ratio (1-octanol/buffer)
hydrophilic	<0	30:140, 40:130
moderately lipophilic	0–1	70:110, 110:70
lipophilic	>1	3:180, 4:180

Equal amounts of phosphate buffer (0.1 M, pH 7.4) and 1-octanol were mixed and shaken vigorously for 5 min to saturate the phases. The mixture was left until separation of the two phases occurred, and the buffer was retrieved. Stock solutions of the test compounds were diluted with buffer to a concentration of 1  $\mu$ M. For each compound, six determinations, that is, three determinations per 1-octanol/buffer ratio, were performed in different wells of a 96-well plate. The respective volumes of buffer containing analyte (1  $\mu$ M) were pipetted to the wells and covered by saturated 1-octanol according to the chosen volume ratio. The plate was sealed with aluminum foil, shaken (1350 rpm, 25 °C, 2 h) on a Heidolph Titramax 1000 plate-shaker (Heidolph Instruments GmbH & Co. KG, Schwabach, Germany), and centrifuged (2000 rpm, 25 °C, 5 min, 5804 R Eppendorf centrifuge, Hamburg, Germany). The aqueous phase was transferred to a 96-well plate for analysis by LC–MS.

The  $\log D_{7.4}$  coefficient was calculated from the 1-octanol/buffer ratio ( $o/b$ ), the initial concentration of the analyte in buffer (1  $\mu$ M), and the concentration of the analyte in buffer ( $c_B$ ) with eq 3:

$$\log D_{7.4} = \log \left( \frac{1 \mu\text{M} - c_B}{c_B} \frac{1}{o/b} \right) \quad (3)$$

**Aqueous Solubility.** Solubility was determined in a 96-well format using the  $\mu$ SOL Explorer solubility analyzer (pIon, version 3.4.0.5). For each compound, measurements were performed at pH 3.0 and 7.4 in triplicate. For this purpose, six wells of a deep well plate, that is, three wells per pH value, were filled with 300  $\mu$ L of PRISMA HT universal buffer, adjusted to pH 3.0 or 7.4 by adding the requested amount of NaOH (0.5 M). Aliquots (3  $\mu$ L) of a compound stock solution (10–50 mM in DMSO) were added and thoroughly mixed. The final sample concentration was 0.1–0.5 mM, and the residual DMSO concentration was 1.0% (v/v) in the buffer solutions. After 15 h, the solutions were filtered (0.2  $\mu$ m 96-well filter plates) using a vacuum to collect manifold (Whatman Ltd., Maidstone, U.K.) to remove the precipitates. Equal amounts of filtrate and 1-propanol were mixed and transferred to a 96-well plate for UV/vis detection (190–500 nm, SpectraMax 190). The amount of material dissolved was calculated by comparison with UV/vis spectra obtained from reference samples, which were prepared by dissolving compound stock solution in a 1:1 mixture of buffer and 1-propanol (final concentrations 0.017–0.083 mM).

**Parallel Artificial Membrane Permeation Assay (PAMPA).** Effective permeability ( $\log P_e$ ) was determined in a 96-well format with the PAMPA.<sup>60</sup> For each compound, measurements were performed at pH 5.0 and 7.4 in quadruplicates. Eight wells of a deep well plate, that is, four wells per pH value, were filled with 650  $\mu$ L of PRISMA HT universal buffer adjusted to pH 5.0 or 7.4 by adding the requested amount of NaOH (0.5 M). Samples (150  $\mu$ L) were withdrawn from each well to determine the blank spectra by UV/vis spectroscopy (190–500 nm, SpectraMax 190). Then analyte dissolved in DMSO was added to the remaining buffer to yield 50  $\mu$ M solutions. To exclude precipitation, the optical density was measured at 650 nm, with 0.01 being the threshold value. Solutions exceeding this threshold were filtered. Afterward, samples (150  $\mu$ L) were withdrawn to determine the reference spectra. Further 200  $\mu$ L was transferred to each well of the donor plate of the PAMPA sandwich (pIon, P/N 110163). The filter membranes at the bottom of the acceptor plate were infused with 5  $\mu$ L of GIT-0 lipid solution, and 200  $\mu$ L of Acceptor Sink Buffer was filled into each acceptor well. The sandwich was assembled, placed in the GutBox, and left undisturbed for 16 h. Then it was disassembled and samples (150  $\mu$ L) were transferred from each donor and acceptor well to UV plates for determination of the UV/vis spectra. Effective permeability ( $\log P_e$ ) was calculated from the compound flux deduced from the spectra, the filter area, and the initial sample concentration in the donor well with the aid of the PAMPA Explorer software (pIon, version 3.5).

**Colorectal Adenocarcinoma (Caco-2) Cell Permeation Assay.** Caco-2 cells were cultivated in tissue culture flasks (BD Biosciences) with DMEM high glucose medium, containing L-glutamine (2 mM), nonessential amino acids (0.1 mM), penicillin (100 U/mL), streptomycin (100  $\mu$ g/mL), and fetal bovine serum (10%). The cells were kept at 37 °C in humidified air containing 5% CO<sub>2</sub>, and the medium was changed every second day. When approximately 90% confluence was reached, the cells were split in a 1:10 ratio and distributed to new tissue culture flasks. At passage numbers between 60 and 65, they were seeded at a density of  $5.3 \times 10^5$  cells per well to Transwell six-well plates (Corning Inc.) with 2.5 mL of culture medium in the basolateral and 1.8 mL in the apical compartment. The medium was renewed on alternate days. Permeation experiments were performed between days 19 and 21 after seeding. Prior to the experiment, the integrity of the Caco-2 monolayers was evaluated by measuring the transepithelial electrical resistance (TEER) with an Endohm tissue resistance instrument (World Precision Instruments Inc., Sarasota, FL, USA). Only wells with TEER values higher than 250  $\Omega$  cm<sup>2</sup> were used. Experiments were performed in the apical-to-basolateral (absorptive) and basolateral-to-apical (secretory) directions in triplicate. Transport medium (DMEM without sodium pyruvate and phenol red) was withdrawn from the donor compartments of three wells and replaced by the same volume of compound stock solution (10 mM in DMSO) to reach an initial sample concentration of 62.5  $\mu$ M. The Transwell plate was then shaken (600 rpm, 37 °C) on a Heidolph

Titramax 1000 plate-shaker. Samples (40  $\mu\text{L}$ ) were withdrawn from the donor and acceptor compartments 30 min after initiation of the experiment, and the compound concentrations were determined by LC–MS (see below). Apparent permeability ( $P_{\text{app}}$ ) was calculated according to eq 4:

$$P_{\text{app}} = \frac{dQ}{dt} \frac{1}{Ac_0} \quad (4)$$

where  $dQ/dt$  is the compound flux ( $\text{mol s}^{-1}$ ),  $A$  is the surface area of the monolayer ( $\text{cm}^2$ ), and  $c_0$  is the initial concentration in the donor compartment ( $\text{mol cm}^{-3}$ ).<sup>60</sup> After the experiment, TEER values were assessed again for each well and results from wells with values below 250  $\Omega \text{ cm}^2$  were discarded.

**Plasma Protein Binding (PPB).** PPB was determined in a 96-well format using a high throughput dialysis block (HTD96b; HTDialysis LCC, Gales Ferry, CT, USA). For each compound, measurements were performed in triplicate. Dialysis membranes (MWCO 12–14 K; HTDialysis LCC) were hydrated according to the instructions of the manufacturer and placed into the dialysis block. Human plasma was centrifuged (5800 rpm, 5  $^{\circ}\text{C}$ , 10 min), the pH of the supernatant (without floating plasma lipids) was adjusted to 7.4 by adding the requested amount of HCl (4 M), and analyte was added to yield a final concentration of 10  $\mu\text{M}$ . Equal volumes (150  $\mu\text{L}$ ) of plasma containing the analyte or TRIS–HCl buffer (0.1 M, pH 7.4) were transferred to the compartments separated by the dialysis membrane. The block was covered with a sealing film and left undisturbed (5 h, 37  $^{\circ}\text{C}$ ). Afterward, samples (90  $\mu\text{L}$ ) were withdrawn from the buffer compartments and diluted with plasma (10  $\mu\text{L}$ ). From the plasma compartments, samples (10  $\mu\text{L}$ ) were withdrawn and diluted with TRIS–HCl buffer (90  $\mu\text{L}$ ). The solutions were further diluted with ice-cooled MeCN (300  $\mu\text{L}$ ) to precipitate the proteins and centrifuged (3600 rpm, 4  $^{\circ}\text{C}$ , 10 min). The supernatants (50  $\mu\text{L}$ ) were retrieved, and the analyte concentrations were determined by LC–MS (see below). The fraction bound ( $f_b$ ) was calculated as follows (eq 5):

$$f_b = 1 - \frac{c_b}{c_p} \quad (5)$$

where  $c_b$  is the concentration of the analyte withdrawn from the buffer compartment before dilution and  $c_p$  is the concentration in the plasma compartment. The values were accepted if the recovery of analyte was between 80% and 120% of the initial amount.

**Cytochrome P450 Mediated Metabolism.** Incubations consisted of pooled male rat liver microsomes (0.5 mg microsomal protein/mL), test compound (2  $\mu\text{M}$ ),  $\text{MgCl}_2$  (2 mM), and NADPH (1 mM) in a total volume of 300  $\mu\text{L}$  TRIS–HCl buffer (0.1 M, pH 7.4) and were performed in a 96-well plate on a Thermomixer Comfort (Eppendorf). Compounds and microsomes were preincubated (37  $^{\circ}\text{C}$ , 700 rpm, 10 min) before NADPH was added. Samples (50  $\mu\text{L}$ ) at  $t = 0$  min and after an incubation time of 5, 10, 20, and 30 min were quenched with 150  $\mu\text{L}$  of ice-cooled MeOH, centrifuged (3600 rpm, 4  $^{\circ}\text{C}$ , 10 min), and 80  $\mu\text{L}$  of supernatant was transferred to a 96-well plate for LC–MS analysis (see below). The metabolic half-life ( $t_{1/2}$ ) was calculated from the slope of the linear regression from the log percentage remaining compound versus incubation time relationship. Control experiments without NADPH were performed in parallel.

**LC–MS Measurements.** Analyses were performed using an 1100/1200 series HPLC system coupled to a 6410 triple quadrupole mass detector (Agilent Technologies, Inc., Santa Clara, CA, USA) equipped with electrospray ionization. The system was controlled with the Agilent MassHunter Workstation Data Acquisition software (version B.01.04). The column used was an Atlantis T3 C18 column (2.1 mm  $\times$  50 mm) with a 3  $\mu\text{m}$  particle size (Waters Corp., Milford, MA, USA). The mobile phase consisted of eluent A ( $\text{H}_2\text{O}$  containing 0.1% formic acid (for 10a–f, h, i), or 10 mM ammonium acetate, pH 5.0 in 95:5,  $\text{H}_2\text{O}/\text{MeCN}$  (for 10g, j)) and eluent B (MeCN containing 0.1% formic acid). The flow rate was maintained at 0.6 mL/min. The gradient was ramped from 95% A/5% B to 5% A/95% B over 1 min and then held at 5% A/95% B for 0.1 min. The system was then brought back to 95% A/5% B, resulting in a total duration of 4 min. MS parameters such as fragmentor voltage,

collision energy, polarity were optimized individually for each analyte, and the molecular ion was followed for each compound in the multiple reaction monitoring mode. The concentrations of the analytes were quantified by the Agilent Mass Hunter Quantitative Analysis software (version B.01.04).

**In Vivo Studies. Animals.** Female C3H/HeN mice weighing between 19 and 25 g were obtained from Charles River Laboratories (Sulzfeld, Germany) or Harlan (Venray, The Netherlands) and were housed three or four per cage. The mice were kept under specific pathogen-free conditions in the Animal House of the Department of Biomedicine, University Hospital of Basel, and animal experimentation guidelines according to the regulations of the Swiss veterinary law were followed. After 7 days of acclimatization, 9- to 10-week-old mice were used for the studies. Animals had free access to chow and water at any time and were kept in a 12 h/12 h light/dark cycle. For administration volumes and sampling the good practice guidelines were followed.<sup>103</sup>

**Pharmacokinetic Studies.** The single-dose studies for the first experiment set were performed by intravenous application of FimH antagonists at a dosage of 50 mg/kg body weight, followed by plasma and urine sampling. Antagonists were diluted in PBS (Sigma-Aldrich) for injection into the tail vein. Blood and urine samples (10  $\mu\text{L}$ ) were taken at 6 and 30 min and at 1, 2, 4, 6, and 8 h after injection. For the PK studies with 10j, the antagonist was dissolved in PBS with 5% DMSO (Sigma-Aldrich) and injected into the tail vein (0.625 mg/kg) or given orally (1.25 mg/kg) using a gavage (syringes from BD Micro Fine, U-100 Insuline, 30 G with BD Microlance 3, 25 G needles, Becton Dickinson and Soft-Ject, 1 mL syringes from Henke Sass Wolf; gavage from Fine Science Tools). Blood and urine were sampled (10  $\mu\text{L}$ ) after 7, 13, 20, 30, 45 min and after 1, 1.5, 2, 2.5, 3, 4, 6, 8, and 24 h. Both blood and urine samples were directly diluted after sampling with MeOH (Acros Organics) to precipitate the proteins and centrifuged for 11 min at 13 000 rpm. The supernatants were transferred to a 96-well plate (Agilent Technologies, 0.5 mL, polypropylene), and the analyte concentrations were determined by LC–MS (see above).

**Infection Study.** For all infection studies, the drinking water of the mice was replaced by water containing 5% glucose (monohydrate from AppliChem, BioChemica), 3 days before the start of the experiment. 10j was dosed at 1.25 mg/kg (in 5% DMSO and PBS) and 10 mg/kg (in 5% DMSO in PBS containing 1% Tween 80, all purchased from Sigma-Aldrich) and applied orally via gavage to six and four mice, respectively, as described in the section Pharmacokinetic Studies, 40 min prior to infection. Ciprofloxacin (Ciproxin solution, 2 mg/mL, Bayer) was dosed with 8 mg/kg, which would correspond to a human dose of 500 mg,<sup>81</sup> subcutaneously 10 min prior to infection with UT189 to 4 mice. The values for the control group (PBS, po) resulted from the infection of 11 mice. Four mice were orally treated with the formulation vehicle for 10j (5% DMSO in PBS containing 1% Tween 80) and termed controls formulation. Before infection, remaining urine in the bladder was expelled by gentle pressure on the abdomen. Mice were anesthetized in 2.5 vol % isoflurane/oxygen mixture (Attane, Minrad Inc., USA) and placed on their back. Infection was performed transurethrally using a polyethylene catheter (Intramedic polyethylene tubing, inner diameter 0.28 mm, outer diameter 0.61 mm, Becton Dickinson), on a syringe (Hamilton Gastight Syringe 50  $\mu\text{L}$ , removable 30G needle, BGB Analytik AG, Switzerland). After gentle insertion of the catheter into the bladder, 50  $\mu\text{L}$  of bacterial suspension of UT189 ( $5.5 \times 10^9$  to  $2.25 \times 10^{10}$  CFU/mL) was slowly injected. This corresponded to approximately  $10^7$ – $10^8$  CFU per mouse. Mice were killed by  $\text{CO}_2$  3 h after inoculation, and bladder and kidneys were aseptically removed. Organs were homogenized in 1 mL of PBS using a tissue lyser (Retsch, Germany). Serial dilutions of bladder and kidneys were plated on Levine Eosin Methylene Blue Agar plates (Becton Dickinson), and CFUs were counted after overnight incubation at 37  $^{\circ}\text{C}$ .

## ■ ASSOCIATED CONTENT

### Supporting Information

HPLC data and chromatograms of target compounds and  $^1\text{H}$  NMR spectra of the synthetic compounds. This material is available free of charge via the Internet at <http://pubs.acs.org>.

## ■ AUTHOR INFORMATION

## Corresponding Author

\*Phone: +41 61 267 15 51. Fax: +41 61 267 15 52. E-mail: beat.ernst@unibas.ch.

## Author Contributions

§S.K., L.P., K.M., D.E., and A.S. contributed equally to the project.

## Notes

The authors declare no competing financial interest.

## ■ ACKNOWLEDGMENTS

The authors thank Prof. Dr. med. Radek Skoda, Department of Biomedicine, University Hospital Basel, Switzerland, for giving us access to the animal facility. The financial support by the Swiss National Science Foundation (SNF Interdisciplinary Grant K-32K1-120904) is gratefully acknowledged.

## ■ ABBREVIATIONS USED

$\Delta H$ , change in enthalpy;  $\Delta S$ , change in entropy; AUC, area under the curve; BSA, bovine serum albumin;  $C_{max}$ , maximal concentration; Caco-2, colorectal adenocarcinoma; CFU, colony forming unit;  $CL_{tot}$ , total clearance; CRD, carbohydrate recognition domain;  $C_0$ , initial concentration; DL, detection limit; FITC, fluorescein isothiocyanate; FP, fluorescence polarization; ITC, isothermal titration calorimetry; iv, intravenous;  $K_D$ , dissociation constant;  $MAC_{90}$ , minimal antiadhesion concentration to inhibit 90% adhesion; PAMPA, parallel artificial membrane permeation assay;  $P_{app}$ , apparent permeability; PD, pharmacodynamics;  $P_e$ , effective permeability; PK, pharmacokinetics; po, per os; sc, subcutaneous; UPEC, uropathogenic *Escherichia coli*; UTI, urinary tract infection;  $V_z$ , volume of distribution in terminal phase

## ■ REFERENCES

(1) Foxman, B.; Barlow, R.; D'Arcy, H.; Gillespie, B.; Sobel, J. D. Urinary tract infection: self-reported incidence and associated costs. *Ann. Epidemiol.* **2000**, *10*, 509–515.

(2) Ronald, A. The etiology of urinary tract infection: traditional and emerging pathogens. *Am. J. Med.* **2002**, *113* (Suppl. 1A), 14S–19S.

(3) Fihn, S. D. Acute uncomplicated urinary tract infection in women. *N. Engl. J. Med.* **2003**, *349*, 259–266.

(4) Hooton, T. M.; Besser, R.; Foxman, B.; Fritsche, T. R.; Nicolle, L. E. Acute uncomplicated cystitis in an era of increasing antibiotic resistance: a proposed approach to empirical therapy. *Clin. Infect. Dis.* **2004**, *39*, 75–80.

(5) Sanchez, G. V.; Master, R. N.; Karlowsky, J. A.; Bordon, J. M. In vitro antimicrobial resistance of urinary *Escherichia coli* isolates among U.S. outpatients from 2000 to 2010. *Antimicrob. Agents Chemother.* **2012**, *56*, 2181–2183.

(6) Clatworthy, A. E.; Pierson, E.; Hung, D. T. Targeting virulence: a new paradigm for antimicrobial therapy. *Nat. Chem. Biol.* **2007**, *3*, 541–548.

(7) Mulvey, M. A.; Schilling, J. D.; Martinez, J. J.; Hultgren, S. J. Bad bugs and beleaguered bladders: interplay between uropathogenic *Escherichia coli* and innate host defenses. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 8829–8835.

(8) Schilling, J. D.; Mulvey, M. A.; Hultgren, S. J. Structure and function of *Escherichia coli* type 1 pili: new insight into the pathogenesis of urinary tract infections. *J. Infect. Dis.* **2001**, *183* (Suppl. 1), S36–S40.

(9) Wiles, T. J.; Kulesus, R. R.; Mulvey, M. A. Origins and virulence mechanisms of uropathogenic *Escherichia coli*. *Exp. Mol. Pathol.* **2008**, *85*, 11–19.

(10) Capitani, G.; Eidam, O.; Glockshuber, R.; Grütter, M. G. Structural and functional insights into the assembly of type 1 pili from *Escherichia coli*. *Microbes Infect.* **2006**, *8*, 2284–2290.

(11) Le Trong, I.; Aprikian, P.; Kidd, B. A.; Forero-Shelton, M.; Tchesnokova, V.; Rajagopal, P.; Rodriguez, V.; Interlandi, G.; Klevit, R.; Vogel, V.; Stenkamp, R. E.; Sokurenko, E. V.; Thomas, W. E. Structural basis for mechanical force regulation of the adhesin FimH via finger trap-like  $\beta$  sheet twisting. *Cell* **2010**, *141*, 645–655.

(12) Sharon, N. Carbohydrates as future anti-adhesion drugs for infectious diseases. *Biochim. Biophys. Acta* **2006**, *1760*, 527–537.

(13) Firon, N.; Itzhak, O.; Sharon, N. Interaction of mannose-containing oligosaccharides with the fimbrial lectin of *Escherichia coli*. *Biochem. Biophys. Res. Commun.* **1982**, *105*, 1426–1432.

(14) Firon, N.; Ofek, I.; Sharon, N. Carbohydrate specificity of the surface lectins of *Escherichia coli*, *Klebsiella pneumoniae*, and *Salmonella typhimurium*. *Carbohydr. Res.* **1983**, *120*, 235–249.

(15) Bouckaert, J.; Berglund, J.; Schembri, M.; De Genst, E.; Cools, L.; Wuhler, M.; Hung, C.-S.; Pinkner, J.; Slättegård, R.; Zavalov, A.; Choudhury, D.; Langermann, S.; Hultgren, S. J.; Wyns, L.; Klemm, P.; Oscarson, S.; Knight, S. D.; De Greve, H. Receptor binding studies disclose a novel class of high-affinity inhibitors of the *Escherichia coli* FimH adhesin. *Mol. Microbiol.* **2005**, *55*, 441–455.

(16) 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.

(17) Sperling, O.; Fuchs, A.; Lindhorst, T. K. Evaluation of the carbohydrate recognition domain of the bacterial adhesin FimH. Design, synthesis and binding properties of mannoside ligands. *Org. Biomol. Chem.* **2006**, *4*, 3913–3922.

(18) Han, Z.; Pinkner, J. S.; Ford, B.; Obermann, R.; Nolan, W.; Wildman, S. A.; Hobbs, D.; Ellenberger, T.; Cusumano, C. K.; Hultgren, S. J.; Janetka, J. W. Structure-based drug design and optimization of mannoside bacterial FimH antagonists. *J. Med. Chem.* **2010**, *53*, 4779–4792.

(19) Klein, T.; Abgottspon, D.; Wittwer, M.; Rabbani, S.; Herold, J.; Jiang, X.; Kleeb, S.; Lüthi, C.; Scharenberg, M.; Bezençon, J.; Gubler, E.; Pang, L.; Smiesko, M.; Cutting, B.; Schwardt, O.; Ernst, B. FimH antagonists for the oral treatment of urinary tract infections: from design and synthesis to in vitro and in vivo evaluation. *J. Med. Chem.* **2010**, *53*, 8627–8641.

(20) Cusumano, C. K.; Pinkner, J. S.; Han, Z.; Greene, S. E.; Ford, B. A.; Crowley, J. R.; Henderson, J. P.; Janetka, J. W.; Hultgren, S. J. Treatment and prevention of urinary tract infection with orally active FimH inhibitors. *Sci. Transl. Med.* **2011**, *3*, 109ra115.

(21) Han, Z.; Pinkner, J. S.; Ford, B.; Chorell, E.; Crowley, J. M.; Cusumano, C. K.; Campbell, S.; Henderson, J. P.; Hultgren, S. J.; Janetka, J. W. Lead optimization studies on FimH antagonists: discovery of potent and orally bioavailable ortho-substituted biphenyl mannosides. *J. Med. Chem.* **2012**, *55*, 3945–3959.

(22) Pang, L.; Kleeb, S.; Lemme, K.; Rabbani, S.; Scharenberg, M.; Zalewski, A.; Schädler, F.; Schwardt, O.; Ernst, B. FimH antagonists: structure–activity and structure–property relationships for biphenyl  $\alpha$ -D-mannopyranosides. *ChemMedChem* **2012**, *7*, 1404–1422.

(23) Jiang, X.; Abgottspon, D.; Kleeb, S.; Rabbani, S.; Scharenberg, M.; Wittwer, M.; Haug, M.; Schwardt, O.; Ernst, B. Anti-adhesion therapy for urinary tract infections—a balanced PK/PD profile proved to be key for success. *J. Med. Chem.* **2012**, *55*, 4700–4713.

(24) Schwardt, O.; Rabbani, S.; Hartmann, M.; Abgottspon, D.; Wittwer, M.; Kleeb, S.; Zalewski, A.; Smiesko, M.; Cutting, B.; Ernst, B. Design, synthesis and biological evaluation of mannosyl triazoles as FimH antagonists. *Bioorg. Med. Chem.* **2011**, *19*, 6454–6473.

(25) Brument, S.; Sivignon, A.; Dumych, T. I.; Moreau, N.; Roos, G.; Guérardel, Y.; Chalopin, T.; Deniaud, D.; Bilyy, R. O.; Darfeuille-Michaud, A.; Bouckaert, J.; Gouin, S. G. Thiazolylaminomannosides as potent antiadhesives of type 1 piliated *Escherichia coli* isolated from Crohn's disease patients. *J. Med. Chem.* **2013**, *56*, 5395–5406.

(26) Lindhorst, T. K.; Kieburg, C.; Krallmann-Wenzel, U. Inhibition of the type 1 fimbriae-mediated adhesion of *Escherichia coli* to erythrocytes by multiantennary D-mannosyl clusters: the effect of multivalency. *Glycoconjugate J.* **1998**, *15*, 605–613.

- (27) Nagahori, N.; Lee, R. T.; Nishimura, S.-L.; Pagé, S.; Roy, R.; Lee, Y. C. Inhibition of adhesion of type 1 fimbriated *Escherichia coli* to highly mannose ligands. *ChemBioChem* **2002**, *3*, 836–844.
- (28) Appeldoorn, C. C. M.; Joosten, J. A. F.; Maate, F. A.; Dobrindt, U.; Hacker, J.; Liskamp, R. M. J.; Khan, A. S.; Pieters, R. J. Novel multivalent mannose compounds and their inhibition of the adhesion of type 1 fimbriated uropathogenic *E. coli*. *Tetrahedron: Asymmetry* **2005**, *16*, 361–372.
- (29) Patel, A.; Lindhorst, T. K. A modular approach for the synthesis of oligosaccharide mimetics. *Carbohydr. Res.* **2006**, *341*, 1657–1668.
- (30) Touaibia, M.; Wellens, A.; Shiao, T. C.; Wang, Q.; Sirois, S.; Bouckaert, J.; Roy, R. Mannosylated G(0) dendrimers with nanomolar affinities to *Escherichia coli* FimH. *ChemMedChem* **2007**, *2*, 1190–1201.
- (31) Durka, M.; Buffet, K.; Iehl, J.; Holler, M.; Nierengarten, J.-F.; Taganna, J.; Bouckaert, J.; Vincent, S. P. The functional valency of dodecamannosylated fullerenes with *Escherichia coli* FimH—towards novel bacterial antiadhesives. *Chem. Commun.* **2011**, *47*, 1321–1323.
- (32) Bouckaert, J.; Li, Z.; Xavier, C.; Almant, M.; Caveliers, V.; Lahoutte, T.; Weeks, S. D.; Kovensky, J.; Gouin, S. G. Heptyl  $\alpha$ -D-mannosides grafted on a  $\beta$ -cyclodextrin core to interfere with *Escherichia coli* adhesion: an in vivo multivalent effect. *Chem.—Eur. J.* **2013**, *19*, 7847–7855.
- (33) Scharenberg, M.; Schwardt, O.; Rabbani, S.; Ernst, B. Target selectivity of FimH antagonists. *J. Med. Chem.* **2012**, *55*, 9810–9816.
- (34) Choudhury, D.; Thompson, A.; Stojanoff, V.; Langermann, S.; Pinkner, J.; Hultgren, S. J.; Knight, S. D. X-ray structure of the FimC-FimH chaperone-adhesin complex from uropathogenic *Escherichia coli*. *Science* **1999**, *285*, 1061–1066.
- (35) Hung, C.-S.; Bouckaert, J.; Hung, D.; Pinkner, J.; Widberg, C.; DeFusco, A.; Auguste, C. G.; Strouse, R.; Langermann, S.; Waksman, G.; Hultgren, S. J. Structural basis of tropism of *Escherichia coli* to the bladder drug in urinary tract infection. *Mol. Microbiol.* **2002**, *44*, 903–915.
- (36) Wellens, A.; Garofalo, C.; Nguyen, H.; Van Gerven, N.; Slättegård, R.; Henalsteens, J.-P.; Wyns, L.; Oscarson, S.; De Greve, H.; Hultgren, S. J.; Bouckaert, J. Intervening with urinary tract infections using anti-adhesives based on the crystal structure of the FimH-oligomannose-3 complex. *PLoS One* **2008**, *3*, e2040.
- (37) Wellens, A.; Lahmann, M.; Touaibia, M.; Vaucher, J.; Oscarson, S.; Roy, R.; Remaut, H.; Bouckaert, J. The tyrosine gate as a potential entropic lever in the receptor-binding site of the bacterial adhesin FimH. *Biochemistry* **2012**, *51*, 4790–4799.
- (38) Totsika, M.; Kostakioti, M.; Hannan, T. J.; Upton, M.; Beatson, S. A.; Janetka, J. W.; Hultgren, S. J.; Schembri, M. A. A FimH inhibitor prevents acute bladder infection and treats chronic cystitis caused by multidrug-resistant uropathogenic *Escherichia coli* ST131. *J. Infect. Dis.* **2013**, *208*, 921–928.
- (39) Meanwell, M. A. Synopsis of some recent tactical application of bioisosteres in drug design. *J. Med. Chem.* **2011**, *54*, 2529–2591.
- (40) Prieto, M.; Zurita, E.; Rosa, E.; Luño, L.; Lloyd-Williams, P.; Giralt, E. Arylboronic acids and arylpinacolboronate esters in Suzuki coupling reactions involving indoles. Partner role swapping and heterocycle protection. *J. Org. Chem.* **2004**, *69*, 6812–6820.
- (41) Schulz, M. J.; Coats, S. J.; Hlasta, D. J. Microwave-assisted preparation of aryltetrazoleboronate esters. *Org. Lett.* **2004**, *6*, 3265–3268.
- (42) Devos, A.; Remion, J.; Frisque-Hesbain, A. M.; Colens, A.; Ghosez, L. Synthesis of acyl halides under very mild conditions. *J. Chem. Soc., Chem. Commun.* **1979**, 1180–1181.
- (43) Rabbani, S.; Jiang, X.; Schwardt, O.; Ernst, B. Expression of the carbohydrate recognition domain of FimH and development of a competitive binding assay. *Anal. Biochem.* **2010**, *407*, 188–195.
- (44) Waetherman, R. V.; Kiessling, L. L. Fluorescence anisotropy assay reveals affinities of C- and O-glycosides for concanavalin A. *J. Org. Chem.* **1996**, *61*, 534–538.
- (45) Cer, R. Z.; Mudunuri, U.; Stephens, R.; Lebeda, F. J. IC50-to-Ki: a web-based tool for converting IC50 to Ki values for inhibitors of enzyme activity and ligand binding. *Nucleic Acids Res.* **2009**, *37*, W441–W445.
- (46) Lynch, B. A.; Loiacono, K. A.; Tiong, C. L.; Adams, S. E.; MacNeil, I. A. A fluorescence polarization based Src-SH2 binding assay. *Anal. Biochem.* **1997**, *247*, 77–82.
- (47) Wu, P.; Brasseur, M.; Schindler, U. A high-throughput STAT binding assay using fluorescence polarization. *Anal. Biochem.* **1997**, *249*, 29–36.
- (48) Huang, X. Fluorescence polarization competition assay: the range of resolvable inhibitor potency is limited by the affinity of the fluorescent ligand. *J. Biomol. Screening* **2003**, *8*, 34–38.
- (49) Cooper, A. *Biophysical Chemistry*, 2nd ed.; RSC Publishing: Cambridge, U.K., 2011; pp 122–123.
- (50) Scharenberg, M.; Jiang, X.; Pang, L.; Navarra, G.; Rabbani, S.; Binder, F.; Schwardt, O.; Ernst, B. Kinetic properties of carbohydrate-lectin interactions: FimH antagonists. *ChemMedChem* **2014**, *9*, 78–83.
- (51) Cabani, S.; Gianni, P.; Mollica, V.; Lepori, L. Group contribution to the thermodynamic properties of non-ionic solutes in dilute aqueous solution. *J. Solution Chem.* **1981**, *10*, 563–595.
- (52) Hansch, C.; Leo, A.; Taft, R. W. A survey of Hammett substituent constants and resonance and field parameters. *Chem. Rev.* **1991**, *91*, 165–195.
- (53) Chen, A.; Wadso, I. Simultaneous determination of delta G, delta H and delta S by an automatic microcalorimetric titration technique: application to protein ligand binding. *J. Biochem Biophys Methods* **1982**, *6*, 307–316.
- (54) Freire, E.; Mayorga, O. L.; Straume, M. Isothermal titration calorimetry. *Anal. Chem.* **1990**, *62*, 950A–959A.
- (55) Wiseman, T.; Williston, S.; Brandts, J. F.; Lin, L.-N. Rapid measurement of binding constants and heats of binding using a new titration calorimeter. *Anal. Biochem.* **1989**, *179*, 131–137.
- (56) Turnbull, W. B.; Daranas, A. H. On the value of c: can low affinity systems be studied by isothermal titration calorimetry? *J. Am. Chem. Soc.* **2003**, *125*, 14859–14866.
- (57) Sigurskjold, B. W. Exact analysis of competition ligand binding by displacement isothermal titration calorimetry. *Anal. Biochem.* **2000**, *277*, 260–266.
- (58) Velazquez-Campoy, A.; Freire, E. Isothermal titration calorimetry to determine association constants for high-affinity ligands. *Nat. Protoc.* **2006**, *1*, 186–191.
- (59) Dearden, J. C.; Bresnen, G. M. The measurement of partition coefficients. *QSAR Comb. Sci.* **1988**, *7*, 133–144.
- (60) Kansy, M.; Senner, F.; Gubernator, K. Physicochemical high throughput screening: parallel artificial membrane permeation assay in the description of passive absorption processes. *J. Med. Chem.* **1998**, *41*, 1007–1010.
- (61) Hubatsch, I.; Ragnarsson, E. G. E.; Artursson, P. Determination of drug permeability and prediction of drug absorption in Caco-2 monolayers. *Nat. Protoc.* **2007**, *2*, 2111–2119.
- (62) Banker, M. J.; Clark, T. H.; Williams, J. A. Development and validation of a 96-well equilibrium dialysis apparatus for measuring plasma protein binding. *J. Pharm. Sci.* **2003**, *92*, 967–974.
- (63) Obach, R. S. Prediction of human clearance of twenty-nine drugs from hepatic microsomal intrinsic clearance data: an examination of in vitro half-life approach and nonspecific binding to microsomes. *Drug Metab. Dispos.* **1999**, *27*, 1350–1359.
- (64) Chaturvedi, P. R.; Decker, C. J.; Odinecs, A. Prediction of pharmacokinetic properties using experimental approaches during early drug discovery. *Curr. Opin. Chem. Biol.* **2001**, *5*, 452–463.
- (65) Di, L.; Kerns, E. H. Profiling drug-like properties in discovery research. *Curr. Opin. Chem. Biol.* **2003**, *7*, 402–408.
- (66) Lipinski, C. A. Drug-like properties and the causes of poor solubility and poor permeability. *J. Pharmacol. Toxicol. Methods* **2000**, *44*, 235–249.
- (67) Curatolo, W. Physical chemical properties of oral drug candidates in the discovery and exploratory development settings. *Pharm. Sci. Technol. Today* **1998**, *1*, 387–393.
- (68) Ishikawa, M.; Hashimoto, Y. Improvement in aqueous solubility in small molecule drug discovery programs by disruption of molecular planarity and symmetry. *J. Med. Chem.* **2011**, *54*, 1539–1554.

- (69) Avdeef, A.; Bendels, S.; Di, L.; Faller, B.; Kansy, M.; Sugano, K.; Yamauchi, Y. PAMPA – critical factors for better predictions of absorption. *J. Pharm. Sci.* **2007**, *96*, 2893–2909.
- (70) Artursson, P.; Karlsson, J. Correlation between oral drug absorption in humans and apparent drug permeability coefficients in human intestinal epithelial (Caco-2) cells. *Biochem. Biophys. Res. Commun.* **1991**, *175*, 880–885.
- (71) Feng, B.; LaPerle, J. L.; Chang, G.; Varma, M. V. S. Renal clearance in drug discovery and development: molecular descriptors, drug transporters and disease state. *Expert Opin. Drug. Metab. Toxicol.* **2010**, *6*, 939–952.
- (72) Schmidt, S.; Gonzalez, D.; Derendorf, H. Significance of protein binding in pharmacokinetics and pharmacodynamics. *J. Pharm. Sci.* **2010**, *99*, 1107–1122.
- (73) Weisiger, R. A. Dissociation from albumin: A potentially rate-limiting step in the clearance of substances by the liver. *Proc. Natl. Acad. Sci. U.S.A.* **1985**, *82*, 1563–1567.
- (74) Smith, D. A.; Jones, B. C.; Walker, D. K. Design of drugs involving the concepts and theories of drug metabolism and pharmacokinetics. *Med. Res. Rev.* **1996**, *16*, 243–266.
- (75) Van de Waterbeemd, H.; Smith, D. A.; Beaumont, K.; Walker, D. K. Property-based design: optimization of drug absorption and pharmacokinetics. *J. Med. Chem.* **2001**, *44*, 1313–1333.
- (76) Varma, M. V. S.; Feng, B.; Obach, R. S.; Troutman, M. D.; Chupka, J.; Miller, H. R.; El-Kattan, A. Physicochemical determinants of human renal clearance. *J. Med. Chem.* **2009**, *52*, 4844–4852.
- (77) Waring, M. J. Lipophilicity in drug discovery. *Expert Opin. Drug Discovery* **2010**, *5*, 235–248.
- (78) Zhang, Y.; Huo, M.; Solver, P. K. An add-in program for pharmacokinetic and pharmacodynamic data analysis in Microsoft Excel. *Comput. Methods Programs Biomed.* **2010**, *99*, 306–314.
- (79) Scharenberg, M.; Abgottspon, D.; Cicek, E.; Jiang, X.; Schwardt, O.; Rabbani, S.; Ernst, B. Cytometry-based assay for screening FimH antagonists. *Assay Drug Dev. Technol.* **2011**, *9*, 455–464.
- (80) Hooton, T. M. Fluoroquinolones and resistance in the treatment of uncomplicated urinary tract infection. *Int. J. Antimicrob. Agents* **2003**, *22*, 65–72.
- (81) Jakobsen, L.; Cattoir, V.; Jensen, K. S.; Hammerum, A. M.; Nordmann, P.; Frimodt-Møller, N. Impact of low-level fluoroquinolone resistance genes *qnrA1*, *qnrB19*, and *qnrS1* on ciprofloxacin treatment of isogenic *Escherichia coli* strains in a murine urinary tract infection model. *J. Antimicrob. Chemother.* **2012**, *67*, 2438–2444.
- (82) Mulvey, M. A. Adhesion and entry of uropathogenic *Escherichia coli*. *Cell. Microbiol.* **2002**, *4*, 257–271.
- (83) Ballatore, C.; Huryn, D. M.; Smith, A. B. Carboxylic acid (bio)isosteres in drug design. *ChemMedChem* **2013**, *8*, 385–395.
- (84) Justice, S. S.; Hung, C.; Theriot, J. A.; Fletcher, D. A.; Anderson, G. G.; Footer, M. J.; Hultgren, S. J. Differentiation and developmental pathways of uropathogenic *Escherichia coli* in urinary tract pathogenesis. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 1333–1338.
- (85) Mulvey, M. A.; Schilling, J. D.; Hultgren, S. J. Establishment of a persistent *Escherichia coli* reservoir during the acute phase of a bladder infection. *Infect. Immun.* **2001**, *69*, 4572–9.
- (86) Houtman, J. C.; Brown, P. C.; Bowden, B.; Yamaguchi, H.; Appella, E.; Samelson, L. E.; Schuck, P. Studying multisite binary and ternary protein interactions by global analysis of isothermal titration calorimetry data in SEDPHAT: application to adaptor protein complexes in cell signaling. *Protein Sci.* **2007**, *16*, 30–42.
- (87) Maestro, version 9.3; Schrödinger, LLC: New York, NY, 2012.
- (88) MacroModel, version 9.9; Schrödinger, LLC: New York, NY, 2012.
- (89) Hawkins, G. D.; Giesen, D. J.; Lynch, G. C.; Chambers, C. C.; Rossi, I.; Storer, J. W.; Li, J.; Thompson, J. D.; Winget, P.; Lynch, B. J.; Rinaldi, D.; Liotard, D. A.; Cramer, C. J.; Truhlar, D. G. AMSOL, version 7.1; University of Minnesota: Minneapolis, MN, 2003; based in part on the following: Liotard, D. A.; Healy, E. F.; Ruiz, J. M.; Dewar, M. J. S. AMPAC, version 2.1; Semichem, Inc.: Shawnee, KS.
- (90) Chambers, C. C.; Hawkins, G. D.; Cramer, C. J.; Truhlar, D. G. Model for aqueous solvation based on class IV atomic charges and first solvation shell effects. *J. Phys. Chem.* **1996**, *100*, 16385–16398.
- (91) Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P. AM1: a new general purpose quantum mechanical molecular model. *J. Am. Chem. Soc.* **1993**, *115*, 5348–5348 [Erratum to *J. Am. Chem. Soc.* **1985**, *107*, 3902–3909].
- (92) Mulvey, M. A.; Schilling, J. D.; Hultgren, S. J. Establishment of a persistent *Escherichia coli* reservoir during the acute phase of a bladder infection. *Infect. Immun.* **2001**, *69*, 4572–4579.
- (93) Kabsch, W. Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. *J. Appl. Crystallogr.* **1993**, *26*, 795–800.
- (94) Leslie, A. G. W. The integration of macromolecular diffraction data. *Acta Crystallogr. D* **2006**, *62*, 48–57.
- (95) Winn, M. D.; Ballard, C. C.; Cowtan, K. D.; Dodson, E. J.; Emsley, P.; Evans, P. R.; Keegan, R. M.; Krissinel, E. B.; Leslie, A. G. W.; McCoy, A.; McNicholas, S. J.; Murshudov, G. N.; Pannu, N. S.; Potterton, E. A.; Powell, H. R.; Read, R. J.; Vagin, A.; Wilson, K. S. Overview of the CCP4 suite and current developments. *Acta Crystallogr. D* **2011**, *67*, 235–242.
- (96) McCoy, A. J. Solving structures of protein complexes by molecular replacement with Phaser. *Acta Crystallogr. D* **2007**, *63*, 32–41.
- (97) Emsley, P.; Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D* **2004**, *60*, 2126–2132.
- (98) Adams, P. D.; Grosse-Kunstleve, R. W.; Hung, L.-W.; Ioerger, T. R.; McCoy, A. J.; Moriarty, N. W.; Read, R. J.; Sacchettini, J. C.; Sauter, N. K.; Terwilliger, T. C. PHENIX: building new software for automated crystallographic structure determination. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2002**, *58*, 1948–1954.
- (99) van Aalten, D. M. F.; Bywater, R.; Findlay, J. B. C.; Hendlich, M.; Hoof, R. W. W.; Vriend, G. PRODRG, a program for generating molecular topologies and unique molecular descriptors from coordinates of small molecules. *J. Comput.-Aided Mol. Des.* **1996**, *10*, 255–262.
- (100) Chen, V. B.; Arendall, W. B.; Headd, J. J.; Keedy, D. A.; Immormino, R. M.; Kapral, G. J.; Murray, L. W.; Richardson, J. S.; Richardson, D. C. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D* **2010**, *66*, 12–21.
- (101) Bezençon, J.; Wittwer, M. B.; Cutting, B.; Smiesko, M.; Wagner, B.; Kansy, M.; Ernst, B. pK<sub>a</sub> determination by <sup>1</sup>H NMR spectroscopy—an old methodology revisited. *J. Pharm. Biomed. Anal.* **2014**, *93*, 147–155.
- (102) (a) VCCLAB, Virtual Computational Chemistry Laboratory, 2005. <http://www.vcclab.org> (accessed August 14, 2012). (b) Tetko, I. V.; Gasteiger, J.; Todeschini, R.; Mauri, A.; Livingstone, D.; Ertl, P.; Palyulin, V. A.; Radchenko, E. V.; Zefirov, N. S.; Makarenko, A. S.; Tanchuk, V. Y.; Prokopenko, V. V. Virtual computational chemistry laboratory—design and description. *J. Comput.-Aided Mol. Des.* **2005**, *19*, 453–463.
- (103) Diehl, K.-H.; Hull, R. A. Good practice guide to the administration of substances and removal of blood, including routes and volumes. *J. Appl. Toxicol.* **2001**, *21*, 15–23.



## Manuscript 6

### KinITC – One method supports both thermodynamic and kinetic SARs

Pascal Zihlmann,<sup>†#</sup> Marleen Silbermann,<sup>†#</sup> Timothy Sharpe,<sup>†</sup> Xiaohua Jiang,<sup>†</sup>  
Tobias Mühlethaler,<sup>†</sup> Roman P. Jakob,<sup>§</sup> Said Rabbani,<sup>†</sup> Christoph P. Sager,<sup>†</sup>  
Priska Frei,<sup>†</sup> Timm Meier<sup>§</sup> and Beat Ernst<sup>†\*</sup>

<sup>#</sup>These authors contributed equally to the project

<sup>†</sup>University of Basel, Institute of Molecular Pharmacy,  
Klingelbergstr. 50, 4056 Basel, Switzerland

<sup>§</sup>University of Basel, Institute of Structural Biology,  
Klingelbergstr. 70, 4056 Basel, Switzerland

<sup>†</sup>University of Basel, Biophysics Facility,  
Klingelbergstr. 70, 4056 Basel, Switzerland

\*Corresponding author.

Tel.: 0041 (0)61 267 15 51; Fax: 0041 (0)61 267 15 52;  
E-mail: beat.ernst@unibas.ch

Contributions of Pascal Zihlmann:

- Manuscript preparation
- Kinetic profiling by ITC and data analysis



## Abstract

Affinity data, such as dissociation constants ( $K_D$ ) or inhibitory concentrations ( $IC_{50}$ ), are widely used in drug discovery. However, these parameters describe an equilibrium state, which is often not established *in vivo* due to pharmacokinetic effects and they are therefore not necessarily reliable for evaluating drug efficacy. More accurate indicators for pharmacological activity are the kinetics of binding processes, as they shed light on the rate of formation of drug-target complexes and their half-life. Nonetheless, although highly desirable for medicinal chemistry programs, studies on structure-kinetic relationships (SKR) are still rare. With the recently introduced analytical tool kinITC this situation may change, since not only thermodynamic but also kinetic information of the binding process can be deduced from isothermal titration calorimetry (ITC) experiments. Using kinITC, ITC data of 29 mannosides binding to the bacterial adhesin FimH were re-analyzed to make their binding kinetics accessible. To validate these kinetic data, surface plasmon resonance (SPR) experiments were conducted. The kinetic analysis by kinITC revealed that the nanomolar affinities of the FimH antagonists arise from both; (i) an optimized interaction between protein and ligand in the bound state (reduced off-rate constant,  $k_{off}$ ), and (ii) a stabilization of the transition state or a destabilization of the unbound state (increased on-rate constant,  $k_{on}$ ). Based on congeneric ligand modifications and structural input from co-crystal structures, a strong relationship between the formed hydrogen bond network and  $k_{off}$  could be concluded, whereas electrostatic interactions and conformational restrictions upon binding were found to have mainly an impact on  $k_{on}$ .

### Introduction

Although the theoretical foundation of binding kinetics in drug-target interactions was set by Paul Ehrlich more than a century ago (“*corpora non agunt nisi fixata*”), little attention has been paid to this concept until recently.<sup>(1,2)</sup> Only after Robert A. Copeland introduced the drug-target residence time concept in 2006, kinetic considerations gained increasing interest.<sup>(3)</sup> The key message of this model is that residence time ( $\tau = 1/k_{\text{off}}$ )<sup>(4)</sup> or half-life ( $t_{1/2} = \ln 2/k_{\text{off}}$ )<sup>(3,5)</sup> of a binary drug–target complex, and not the binding affinity (*e.g.* expressed by the dissociation constant  $K_D$ ), controls the *in vivo* pharmacological activity.

The launch of a new drug is associated with high financial costs and but also with a substantial risk of failure, mainly due to insufficient efficacy and unwanted side effects.<sup>(6)</sup> Several studies have shown that kinetic information is more reliable than affinity data to predict *in vivo* potency but also the resistance potential of a drug candidate.<sup>(7-9)</sup> One of many examples is the HIV reverse transcriptase inhibitor efavirenz that exhibits only a low affinity for its target (5  $\mu\text{M}$ ), but a long dissociation half-life ( $t_{1/2} = 2.8$  h).<sup>(10)</sup> Furthermore, kinetic studies of Maschera *et al.* on HIV-1 protease mutants revealed a correlation between drug resistance to the HIV protease inhibitor saquinavir and an increased dissociation rate of the drug-target complex.<sup>(11)</sup> Nevertheless, early-phase drug discovery still focuses mainly on the optimization of  $K_D$ s, although it has been recognized that kinetic rate constants for association ( $k_{\text{on}}$ ) and dissociation ( $k_{\text{off}}$ ) are a necessary precondition for a comprehensive description of the binding process.<sup>(12)</sup>

Long dissociation half-lives are a crucial feature of many small molecule drugs on the market, *e.g.* for the neuraminidase inhibitor oseltamivir (47 min)<sup>(13)</sup>, the selective COX-2 inhibitor rofecoxib (9 h)<sup>(14)</sup> and the HIV-1 protease inhibitor darunavir (> 240 h)<sup>(15)</sup>. Generally, a drug-target complex with a long dissociation half-life can compensate for unfavorable pharmacokinetics, *i.e.* a bimolecular complex can still exist while the unbound drug molecule is already cleared from the body.<sup>(16)</sup> This is of particular importance for substances with short plasma half-lives.<sup>(5)</sup> In contrast, in the case of drug toxicity fast off-rates are favored as it is the case for antipsychotics that occupy D2 receptors ( $t_{1/2} < 1$  min).<sup>(17,18)</sup> *In vivo*, fast on-rates play a central role in drug rebinding, which can be influenced for example by the local accumulation of the target and spatial

characteristics of the binding site.<sup>(19)</sup> In contrast to off-rates, which are independent of free ligand concentration, on-rates can be increased by administration of higher doses. However, this requires sufficient oral bioavailability and increases the risk of undesired side effects based on an elevated amount of drug in circulation.<sup>(20,21)</sup>

It is surprising that studies on the correlation of molecular structures and their binding kinetics, so-called structure-kinetic relationships (SKRs), are rare.<sup>(22)</sup> In a rough approximation, more rotatable bonds and higher molecular weights correlate with long complex half-lives.<sup>(23)</sup> Moreover, water-shielded hydrogen bonds also tend to improve the lifetime of drug-target complexes.<sup>(24)</sup> On the other hand, the on-rate is limited by diffusion and can be influenced by steric and electrostatic factors as well as conformational dynamics.<sup>(12)</sup>

Surface plasmon resonance (SPR) spectroscopy has evolved into the method of choice for measuring binding kinetics in drug discovery.<sup>(25)</sup> It monitors non-covalent interactions in real-time by detecting a mass-dependent change in the refractive index close to the sensor surface.<sup>(26)</sup> SPR is also applied to obtain thermodynamic information from affinity data as a function of temperature by van't Hoff analysis. However, since the heat capacity ( $\Delta C_p$ ) of macromolecular interactions with small molecules is mostly different from zero, enthalpy and entropy changes for binding are usually temperature dependent. This introduces curvatures in van't Hoff plots and can limit the accuracy of linear approximations compared to direct measurements.<sup>(27)</sup>

Therefore, the method of choice to directly determine the thermodynamics of a molecular binding event is isothermal titration calorimetry (ITC). In contrast to SPR, where one interaction partner has to be immobilized on the sensor chip surface, ITC measures the heat change of a binding interaction in solution. This heat change can be converted into a binding isotherm that allows the direct measurement of  $K_D$  and the change in enthalpy ( $\Delta H^\circ$ ), whereas the change in free energy ( $\Delta G^\circ$ ) ( $\Delta G^\circ = RT \ln K_D$ , with  $R$  being the universal gas constant and  $T$  the absolute temperature) and the change in entropy ( $\Delta S^\circ$ ) ( $\Delta G^\circ = \Delta H^\circ - T\Delta S^\circ$ ) are calculated.<sup>(28)</sup> Both techniques are the gold-standard in their field of application, SPR for the determination of kinetic and ITC for thermodynamic data.<sup>(29)</sup>

Recent publications from Dumas *et al.* have the potential to transform the role of ITC in drug discovery since it describes an approach to derive binding kinetics from ITC experiments (kinITC).<sup>(30,31)</sup> KinITC extracts kinetic information from the shape of each injection peak of an ITC experiment. A simplified version of this kinetic analysis, kinITC-ETC (Equilibration Time Curve), was recently implemented into the commercially available ITC analysis software AFFINImeter.<sup>(30)</sup> It determines the time necessary for the differential power curve to return to the baseline after the injection of an aliquot of ligand solution (equilibration time) as a function of the binding kinetics and the intrinsic response time constant of the calorimeter feedback circuit. Narrower peaks at the beginning and the end of the titration, as well as wider peaks around the inflection point, yield a bell-shaped ETC that can be analyzed in order to determine  $k_{\text{off}}$ . Obviously, kinITC-ETC greatly increases the value of ITC data, allowing the direct determination of all relevant biophysical constants describing a binding event within one experiment.

In the present study, the binding kinetics of a series of congeneric mannosides binding to the lectin domain of FimH (FimH<sub>LD</sub>) were determined. FimH is a virulence factor located at the tip of type 1 pili of uropathogenic *E. coli* (UPEC) strains.<sup>(32,33)</sup> It interacts with the highly mannosylated glycoprotein uroplakin Ia, which is part of the urothelial mucosa and thereby mediates bacterial adhesion to the bladder epithelium as the initial step of urinary tract infection (UTI).<sup>(34)</sup> An anti-adhesion therapy with FimH antagonists, which block the adhesion and thereby prevent the infection, could therefore be beneficial for patients suffering from recurrent UTI.<sup>(35,36)</sup> Reanalyzing data from numerous ITC<sup>(37-41)</sup> and SPR<sup>(42)</sup> studies we published in the last 5 years, enabled the validation of the kinITC-ETC approach and the correlation of kinetic parameters with structural properties of a large ligand dataset. The kinetic fingerprints of FimH antagonists offer the opportunity to further improve the binding characteristics essential for a clinical application.

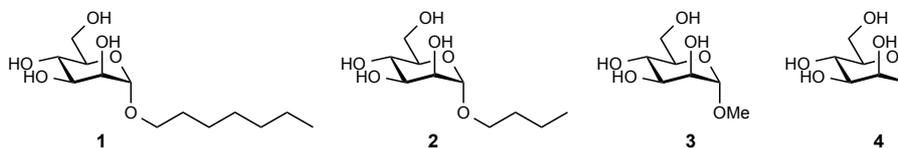
## Results and Discussion

The publication on kinITC by Burnouf *et al.* in 2012 gained wide interest, but the complexity of the analysis hampered a broader application.<sup>(31)</sup> However, with kinITC-ETC integrated into the commercially available ITC analysis software AFFINImeter<sup>(43)</sup> this hurdle could be overcome. It uses fully automated data processing to derive kinetic and thermodynamic information from ITC raw data avoiding the risk of user bias as can arise through manual evaluation of thermograms. We applied kinITC-ETC to deduce binding kinetics from ITC data of a large set of structurally diverse FimH antagonists (compounds **1-29**, see Table S1). This kinetic information proved to be extremely valuable as it provided a more detailed insight into the binding process of carbohydrate-based antagonists to the adhesin FimH<sub>LD</sub>.

**Validation of kinITC-ETC by SPR.** Although Dumas *et al.* proved the potential of kinITC with the evaluation of multi-step kinetic RNA folding,<sup>(31)</sup> there is, to the best of our knowledge, no independent comparison of data obtained by kinITC-ETC and those derived from alternative biophysical methods available to date. For this purpose, a subset of four mannose-based FimH antagonists (compounds **1-4**) was investigated by SPR. For the on-rate constants  $k_{\text{on}}$ , the comparison revealed an excellent correlation with absolute values differing by less than a factor of two (Table 1A), whereas up to 6-fold deviation was observed for  $k_{\text{off}}$  values (Table 1B). These differences might be a consequence of ligand rebinding on the SPR chip surface.<sup>(44)</sup> However, when we normalized the  $k_{\text{off}}$ -values obtained by SPR and kinITC-ETC to that of *n*-heptyl  $\alpha$ -D-mannopyranoside (**1**) (giving relative  $k_{\text{off}}$ ,  $r k_{\text{off}}$ ) differences of less than a factor of 2 arised (Table 1B). In their comparison of SPR and kinITC, Dumas *et al.* reported similar difference for thiamine pyrophosphate interacting with mRNA.<sup>(31)</sup>

## Section II. The Bacterial Adhesin FimH – Manuscript 6

**Table 1.** Comparison of kinetic data obtained by kinITC-ETC and SPR for the interaction of FimH<sub>LD</sub> with the mannose-derived FimH antagonists **1-4**. Absolute ( $k_{\text{on}}$  and  $k_{\text{off}}$ ) and relative ( $r k_{\text{on}}$  and  $r k_{\text{off}}$ ) kinetic parameters are reported for both methods (Table 1A for on-rates & Table 1B for off-rates). Relative values are normalized to *n*-heptyl  $\alpha$ -D-mannopyranoside (**1**), which is set to 1. The discrepancies between kinITC-ETC and SPR for  $k_{\text{on}}$  (Table 1A) and  $k_{\text{off}}$  (Table 1B) values are displayed in the rightmost columns.  $K_{\text{D}}$  values (Table S2), as well as confidence intervals for  $k_{\text{on}}$  and  $k_{\text{off}}$  (ITC, Table S4 and SPR, Table S5), are part of the supplementary information.



**Table 1A.** Association rate constants  $k_{\text{on}}$  of FimH<sub>LD</sub> with compound **1 – 4** determined by ITC and SPR

Compound	$k_{\text{on}}^{\text{kinITC-ETC}}$		$k_{\text{on}}^{\text{SPR}}$		$k_{\text{on}}^{\text{kinITC-ETC}} / k_{\text{on}}^{\text{SPR}}$	
	$k_{\text{on}}$ [M <sup>-1</sup> s <sup>-1</sup> ]	$r k_{\text{on}}$	$k_{\text{on}}$ [M <sup>-1</sup> s <sup>-1</sup> ]	$r k_{\text{on}}$	absolute	relative
<b>1</b>	3.32*10 <sup>4</sup>	1.00	2.45*10 <sup>4</sup>	1.00	1.35	1.00
<b>2</b>	9.71*10 <sup>3</sup>	0.29	5.57*10 <sup>3</sup>	0.23	1.74	1.29
<b>3</b>	2.31*10 <sup>3</sup>	0.07	2.45*10 <sup>3</sup>	0.10	0.94	0.70
<b>4</b>	2.00*10 <sup>3</sup>	0.06	2.67*10 <sup>3</sup>	0.11	0.75	0.55

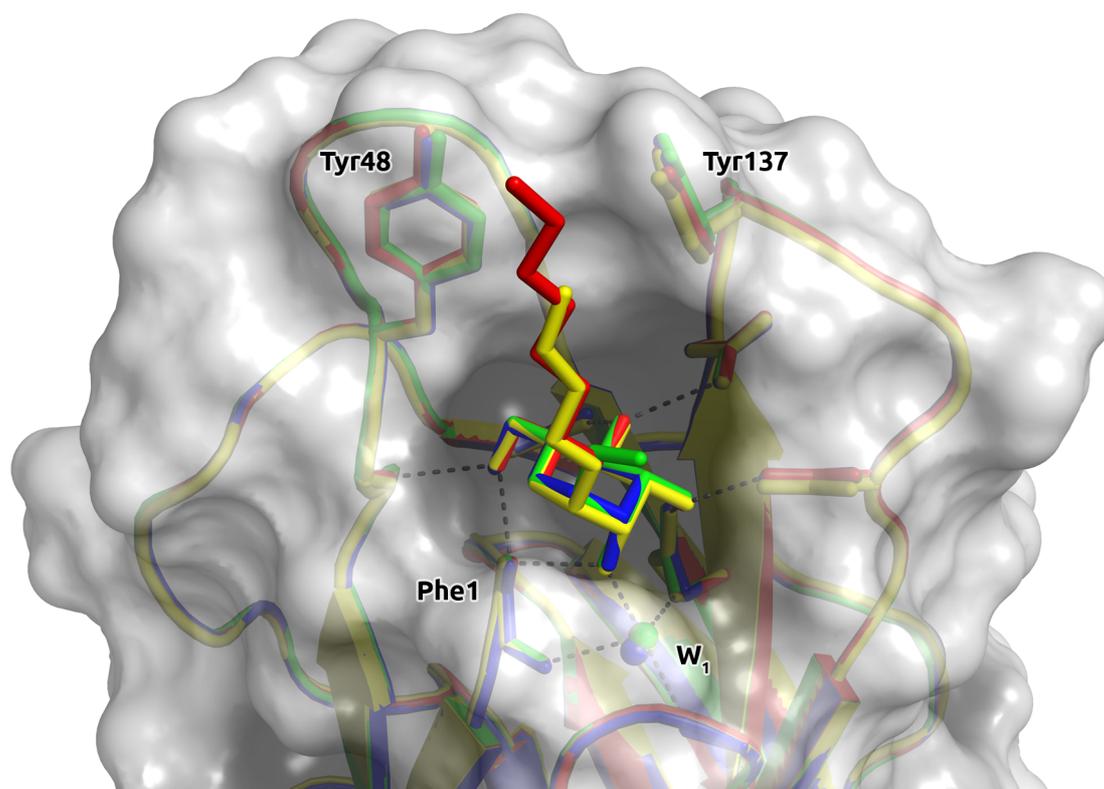
**Table 1B.** Dissociation rate constants  $k_{\text{off}}$  of FimH<sub>LD</sub> with compound **1 – 4** determined by ITC and SPR

Compound	$k_{\text{off}}^{\text{kinITC-ETC}}$			$k_{\text{off}}^{\text{SPR}}$			$k_{\text{off}}^{\text{kinITC-ETC}} / k_{\text{off}}^{\text{SPR}}$	
	$k_{\text{off}}$ [s <sup>-1</sup> ]	$r k_{\text{off}}$	$t_{1/2}$ [min]	$k_{\text{off}}$ [s <sup>-1</sup> ]	$r k_{\text{off}}$	$t_{1/2}$ [min]	absolute	relative
<b>1</b>	7.27*10 <sup>-4</sup>	1.00	15.9	1.54*10 <sup>-4</sup>	1.00	75.0	4.72	1.00
<b>2</b>	1.15*10 <sup>-3</sup>	1.58	10.1	1.89*10 <sup>-4</sup>	1.23	61.1	6.07	1.29
<b>3</b>	2.86*10 <sup>-3</sup>	3.93	4.1	1.11*10 <sup>-3</sup>	7.21	10.4	2.57	0.54
<b>4</b>	1.82*10 <sup>-3</sup>	2.50	6.3	6.51*10 <sup>-4</sup>	4.23	17.7	2.80	0.59

**Kinetics of FimH<sub>LD</sub> binding.** Only recently, when correlations between prolonged drug-target half-lives and clinical efficacy of drugs were reported, binding kinetics attracted the interest of the drug discovery community.<sup>(17,45,46)</sup> The initial assumption that association rate constants are only diffusion controlled and rather constant within a congeneric set of ligands for a specific target had to be revised.<sup>(22,23,47)</sup> Indeed, our test subset clearly indicates that the variations of the on-rate constants within the series of congeneric FimH antagonists **1-4** are more pronounced (for kinITC  $\approx$  1 : 16, Table 1A) than of the off-rate constants (for kinITC  $\approx$  1 : 4, Table 1B).

To exclude that the length of the aliphatic aglycone affects the extended hydrogen bond network formed by the mannose moieties of compounds **1-4**, co-crystal structures of

compound **3** (PDB-code: 5JCR) and **4** (PDB-code: 5MUC) were solved and compared with published structures of compound **1** (PDB-code: 4XO8)<sup>(48)</sup> and **2** (PDB-code: 1UWF)<sup>(49)</sup> (Table S3). For all four ligands, both protein structure and binding mode were found to be identical (Figure 1). Hence, the higher on-rate constants of antagonist **1** and **2** must be a result of their elongated aglycones. This finding is unexpected since the favorable interactions of the elongated aglycones of compound **1** and **2** with Tyr48 and Tyr137 (called tyrosine gate) were presumed to primarily lower the off-rate resulting in a prolonged complex half-life.



**Figure 1.** Binding mode of compounds **1-4** to FimH<sub>LD</sub>. The co-crystal structures of **1** (red, PDB-code: 4XO8), **2** (yellow, PDB-code: 1UWF), **3** (green, PDB-code: 5JCR) and **4** (blue, PDB-code: 5MUC) show coinciding binding modes. Their mannose moieties form identical hydrogen bond networks with FimH<sub>LD</sub> and a structural water molecule (W<sub>1</sub>), whereas only the aglycones of **1** and **2** interact with the two tyrosines 48 and 137.

A one-step model with ligand and protein in the unbound state (U), in the transition state (TS<sup>‡</sup>), and in the bound state (B) characterized by the kinetic rate constants  $k_{\text{on}}$  and  $k_{\text{off}}$  is the simplest way to describe a protein-ligand interaction (Figure 2-A). However, more common is a multistep binding mechanism with apparent rate constants composed of multiple elementary rate constants.<sup>(12)</sup> Rate constants depend upon the free energy difference  $\Delta G_{\text{TS-U}}$  (between TS<sup>‡</sup> and U) for  $k_{\text{on}}$  and the free energy difference  $\Delta G_{\text{TS-B}}$

(between  $TS^\ddagger$  and B) for  $k_{\text{off}}$ , respectively (Figure 2-A). The energy difference between B and U represents the equilibrium free energy of binding ( $\Delta G_{\text{B-U}}$ , or  $\Delta G^\circ$ ). Changes in the relative stability of these states due to structural differences of the ligand (or the protein as a consequence of mutations) provoke changes in one or both rate constants. A stabilization of U and B relative to  $TS^\ddagger$  (or destabilization of  $TS^\ddagger$  relative to U and B) decreases the on-rate and the off-rate, respectively. Ligand modifications leading to improved or worsened interaction with the target protein, affect the stability of the various states. Thus, for example, an additional interaction predominantly formed in B and to a lesser extent in  $TS^\ddagger$ , will increase  $\Delta G_{\text{TS-B}}$  (decrease  $k_{\text{off}}$ ) and to a lesser extent decrease  $\Delta G_{\text{TS-U}}$  (increase  $k_{\text{on}}$ ). Concomitantly, the change of  $\Delta G^\circ$  leads to tighter binding. Furthermore, structural modifications that affect long-range electrostatics of the ligand and thus its proper orientation can affect the rate constants by altered diffusion, significantly impeding an interpretation based only on changes in the relative thermodynamic stability of states.

The kinetic data of antagonists 1-4 obtained by both ITC and SPR demonstrate that the improvements of on-rates go hand in hand with the elongation of the aglycone, whereas off-rates are decreased, however to a lesser extent, leading to an overall improved  $K_{\text{D}}$ -value (Table S2). Hence, increased on-rates are related to favorable interactions realized in  $TS^\ddagger$  or a destabilization of U relative  $TS^\ddagger$ , whereas decreased off-rates imply an additional stabilization of B relative to  $TS^\ddagger$ . Although structural information regarding  $TS^\ddagger$  is not available, it seems plausible to speculate that a hydrophobic aglycone first establishes a contact with the surface-exposed tyrosine gate of FimH<sub>LD</sub>. Only then the protein-ligand complex relaxes from the  $TS^\ddagger$  to the bound state B by a successful completion of the hydrogen bond network within the binding site. A similar effect was described for the human carbonic anhydrase II (hCAII), where the on-rate increased in parallel with the chain length of interacting alkyl benzenesulfonamides.<sup>(50,51)</sup> In a pre-binding state, increasing hydrophobicity of the ligand was found to enhance the interaction with a hydrophobic patch of the enzymatic cavity.<sup>(52)</sup>

To investigate the kinetic behavior of FimH<sub>LD</sub> antagonists in more detail, its binding interaction with 29 antagonists (for structural details see supporting information, Table S1) covering a broad range of affinities (1 nM to 100  $\mu\text{M}$ ) were analyzed by kinITC-ETC.

Thermodynamic and kinetic fits (Table S1) and binding parameters (Table S3&S4) are listed in the supporting information.

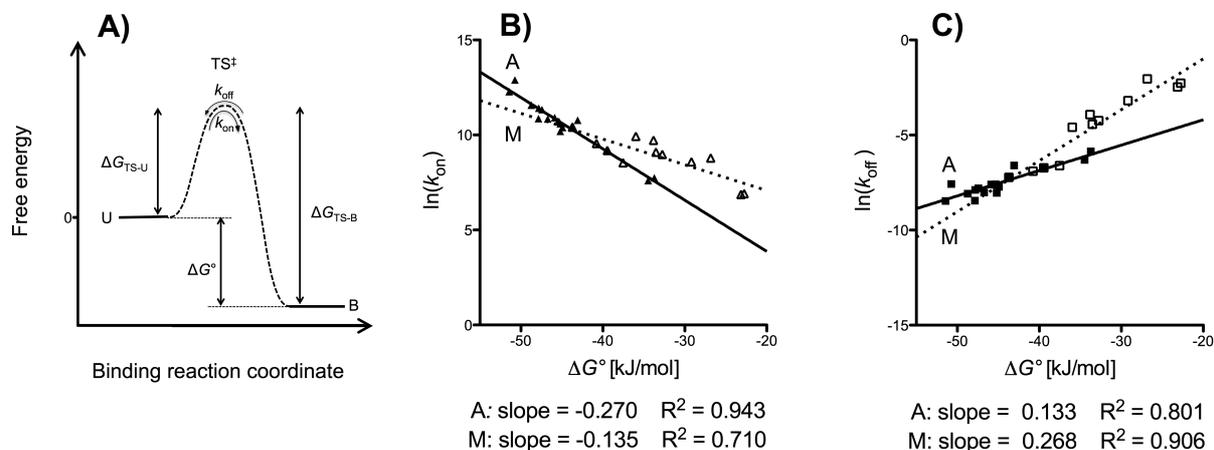
### Structure-kinetic relationship (SKR)

**Influence of the site of ligand modification.** With the set of 29 carbohydrate-based antagonists, we studied how the site of the structural modification affects the correlation between binding energies and kinetic binding rates. The ligands were divided into two subsets; one subset ( $n = 17$ ) containing all representatives with varied aglycones (compounds **2-4** with alkyl chains of different length; **10-21** with biphenyl aglycones, **22** with a squaric acid aglycone and **23** with an indolinyphenyl aglycone) and a second subset ( $n = 11$ ) representing compounds with modified mannose moieties (deoxy- and deoxy-halogeno derivatives **5-8** and **24-27**, septulose derivative **9** and two C-2 branched mannose derivatives **28** and **29** (see supporting information, Table S1). For the two subsets, significantly different correlations of  $\Delta G^\circ$  with  $k_{\text{on}}$  (Figure 2-B,  $p < 0.001$ ) and  $k_{\text{off}}$ -values (Figure 2-C,  $p < 0.001$ ) were obtained. Modifications of the mannose moiety, known to form an extended hydrogen bond network in the deep binding pocket of FimH<sub>LD</sub>, affect the  $k_{\text{off}}$ -values (slope = 0.268,  $R^2 = 0.906$ ) to a greater extent than those of the aglycone (slope = 0.133,  $R^2 = 0.801$ ). In contrast, modifications of the aglycone, forming beneficial contacts with the tyrosine gate in the bound state,<sup>(37)</sup> influence  $k_{\text{on}}$ -values (slope = -0.270,  $R^2 = 0.943$ ) to a larger degree than those of the mannose moiety (slope = -0.135,  $R^2 = 0.710$ ).

Thus, the mannose moiety establishes stronger interactions in the bound state B than in the transition state TS<sup>‡</sup>, resulting in the observed stronger effect on the off-rate constant  $k_{\text{off}}$  (slope = 0.268) than on the on-rate constant  $k_{\text{on}}$  (slope = -0.135). For structural modifications of the aglycone, however, the impact on  $k_{\text{on}}$ -values (slope = -0.270) is more pronounced than on  $k_{\text{off}}$ -values (slope = 0.133). This suggests that either the TS<sup>‡</sup> is stabilized relative to the bound state B (*e.g.* the aglycone establishes stronger interactions with the receptor in the TS<sup>‡</sup> than in the bound state), or that the unbound state U is destabilized relative to the TS<sup>‡</sup>.

The results observed for compounds **1-4** seamlessly fit into the trend observed for all studied FimH antagonists. However, these correlations are based on a rough ligand

classification (modifications of the glycan or aglycone), neglecting various factors, *e.g.* electrostatic interactions or flexibility of the ligand. Contributions from these factors can only be revealed by a closer look at individual structural details of the various ligands and at their interactions with the protein.

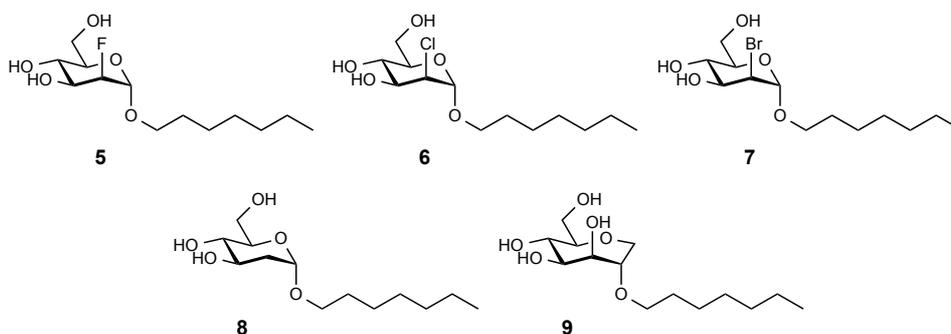


**Figure 2.** **A)** The energy diagram illustrates a simple one-step binding process with the transition state (TS<sup>‡</sup>) separating the unbound state (U) from the bound state (B).  $\Delta G_{TS-U}$  is the activation free energy of association,  $\Delta G_{TS-B}$  the activation free energy of dissociation,  $\Delta G^\circ$  the free energy of binding and  $k_{on}$  and  $k_{off}$  are the kinetic rate constants. **B)** The correlation of the kinetic parameters  $k_{on}$  and **C)**  $k_{off}$  with changes of  $\Delta G^\circ$  depends on the site of a ligand's modification. Ligands with modified aglycones (**2-4**; **10-23**; for structures and binding parameters see supporting information, full symbols, solid regression lines) are separated from ligands with modified mannose moieties (**5-9**; **24-29**; for structure and binding parameters see supporting information, hollow symbols, dotted regression lines).

**Hydrogen bonds, electrostatic interactions, and conformational changes.** To rationalize the trend that mannose modifications have a larger impact on off-rates than on on-rates (Figure 2-B & 2-C), interactions of ligands **5-9** with FimH<sub>LD</sub> were analyzed in more detail (Table 2). The 2-hydroxyl of *n*-heptyl  $\alpha$ -D-mannopyranoside (**1**) forms two hydrogen bonds with FimH<sub>LD</sub>, once as H-bond donor to a structural water molecule (W<sub>1</sub>) and once as H-bond acceptor from the positively charged N-terminal Phe (Figure 1). Substitution of the 2-hydroxy group with a fluorine (**1** → **5**) leads to a loss of hydrogen bonds, while the electrostatic interactions are largely preserved as oxygen and fluorine share a comparable polarity and a close isosteric relationship. Upon removal of the 2-hydroxy group (**1** → **8**) both the H-bonds and the electrostatic interactions are lost. In terms of kinetics, the loss of H-bonds (**1** → **5**) mainly affects the off-rate. In contrast, decreasing electronegativity reduces the electrostatic interaction with the N-terminus from fluorine to chlorine to bromine (**5** → **6** → **7**) and affects the  $k_{on}$ -value to a larger extent than the  $k_{off}$ -value. This complex behavior results from differential contributions: (i) electrostatic

guiding during the formation of the TS<sup>‡</sup> which enhances on- and off-rates and (ii) the formation of more specific interactions in the bound state which decreases off-rates.

**Table 2.** Kinetic binding parameters for the interaction of FimH<sub>LD</sub> with mannose derivatives of *n*-heptyl  $\alpha$ -D-mannopyranoside (**1**). Confidence intervals of the fitted parameters  $k_{\text{on}}$ ,  $k_{\text{off}}$ ,  $K_{\text{D}}$ , and the response time of the calorimeter feedback circuit are part of the supplementary information (Table S4). Relative changes ( $r k_{\text{on}}$ ,  $r k_{\text{off}}$ ) are compared to *n*-heptyl  $\alpha$ -D-mannopyranoside (**1**).



Compound	$k_{\text{on}}$ [M <sup>-1</sup> s <sup>-1</sup> ]	1/ $r k_{\text{on}}$	$k_{\text{off}}$ [s <sup>-1</sup> ]	$r k_{\text{off}}$	$t_{1/2}$ [min]
<b>1</b>	3.32*10 <sup>4</sup>	1.0	7.27*10 <sup>-4</sup>	1.0	15.9
<b>5</b> (2-F)	2.04*10 <sup>4</sup>	1.6	1.02*10 <sup>-2</sup>	14.0	1.1
<b>6</b> (2-Cl)	8.90*10 <sup>3</sup>	3.7	1.19*10 <sup>-2</sup>	16.4	1.0
<b>7</b> (2-Br)	7.90*10 <sup>3</sup>	4.2	1.45*10 <sup>-2</sup>	19.9	0.8
<b>8</b> (2-H)	5.35*10 <sup>3</sup>	6.2	4.13*10 <sup>-2</sup>	56.8	0.3
<b>9</b> (7-ring)	5.10*10 <sup>3</sup>	6.5	1.35*10 <sup>-3</sup>	1.9	8.6

However, mannose modifications may also affect the on-rate constant, as observed for the ring extended compound **9**.<sup>38</sup> Septulose **9** (PDB-code: 5CGB) and pyranoside **1** (PDB-code: 4XO8) establish an identical H-bond network with FimH<sub>LD</sub>. Upon binding to the lectin, the markedly increased flexibility of septulose **9** in solution leads to an entropy penalty due to a loss of conformational freedom upon binding, affecting thermodynamics as well as kinetics accordingly. The increase in flexibility of ligand **9** in solution stabilizes the U relative to TS<sup>‡</sup> and to a lesser extent to B where the flexibility of all ligands is largely restricted by binding interactions. This results in a considerably lowered on-rate and a relatively small increase in the off-rate (Table 2).

### Conclusion

Based on ITC measurements, the new analytical tool kinITC-ETC allowed the determination of kinetic data for 29 FimH antagonists. To test the reliability of kinITC-ETC, the kinetic rate constants for a subset of 4 ligands were compared with data measured by SPR. The systematically lower dissociation rate constant ( $k_{\text{off}}$ ) obtained by SPR may originate from a fundamental difference between the two approaches. ITC measures molecular interactions in free solution, whereas SPR requires the immobilization of one binding partner on a chip surface. The reduction of the apparent  $k_{\text{off}}$  in SPR measurements may be due to the analyte's rebinding to the immobilized partner. However, our attempts to mitigate this effect by using a low immobilization density had only a limited effect. Another rationale for the observed difference in  $k_{\text{off}}$  may arise from errors associated with the kinITC-ETC method, *i.e.* the observed kinetics of heat evolution in ITC experiments includes a contribution from the intrinsic instrumental response time ( $\tau_{\text{ITC}}$ ), which can vary between instruments and for the same instrument depending on cleanliness. Since we re-analyzed existing data for this study,  $\tau_{\text{ITC}}$  was included as a fitting parameter in the kinITC-ETC analysis and not measured for the instrument at the time of the experiment. Nevertheless, it should be noted that the systematic disagreement in off-rates between SPR and ITC measurements is still relatively small (less than an order of magnitude), and that the relative rate constants obtained with kinITC-ETC and SPR are in good agreement. Therefore, kinITC-ETC can be regarded as a reliable method to derive kinetic information from ITC experiments.

The insight into the binding kinetics of the 29 mannose-based antagonists strongly improved our understanding of their binding characteristics. Their enhanced  $K_{\text{D}}$  values not only stand for prolonged complex half-lives, but also for increased  $k_{\text{on}}$ -rates indicating a stabilization the transition state and destabilization of the unbound state. On a structural level, the kinetic contributions of the carbohydrate moiety could be separated from that of the aglycone. The structural variations of the sugar moiety, which forms an extended hydrogen bond network in the deep binding pocket of FimH<sub>LD</sub>, mainly influence the off-rate constant  $k_{\text{off}}$ , while modifications on the aglycone establishing hydrophobic interactions with the two tyrosines at the entrance to the mannose-binding pocket, predominantly affect  $k_{\text{on}}$ . The latter observation could be an indication that the aglycone initially establishes a contact with the tyrosine gate to facilitate the mannose

moiety to enter the carbohydrate recognition domain (CRD). This finding is in excellent agreement with previous observations from Gaspari *et al.* for the interaction of hCAII with hydrophobic ligands.<sup>(52)</sup> Initial hydrophobic interactions to stabilize protein-ligand interactions before the final binding mode is established might in general facilitate ligand recruitment from solution. Furthermore, electrostatic interactions may beneficially influence the on-rate as exemplified by the substitution of the 2-hydroxy group of the mannose moiety by fluorine, chlorine, bromine or hydrogen. Finally, due to its flexibility in solution, the unbound state of septulose derivative **9** is stabilized relative to its transition state, resulting in a reduced on-rate.

These kinetic considerations are essential for the success of a therapeutic treatment of UTI, because the half-life of the complex formed by a carbohydrate-based antagonist is a cardinal parameter. Only an extended half-life of the antagonist/FimH complex, prevents bacteria from interacting with the urothelial cells of the host for long enough to allow bacterial elimination by urination.

Our kinetic study with FimH<sub>LD</sub> is the first systematic analysis of a drug-target interaction using kinITC-ETC. Additional studies will improve the insight into the effects originating from ligand modifications upon binding kinetics and may lead to generally applicable rules. For numerous projects over past 20 years ITC data had been acquired and the hidden kinetic treasures can now be raised by kinITC-ETC. We therefore expect kinITC-ETC to become a popular instrument for “data mining.”

### **Acknowledgement**

This work was supported by Swiss National Science Foundation Grants 200020\_146202 and R'EQUIP 145023.

## References

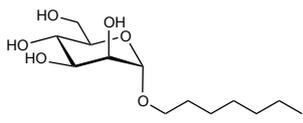
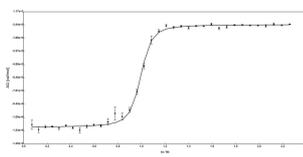
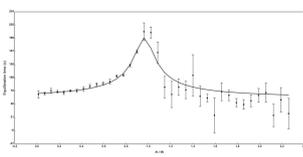
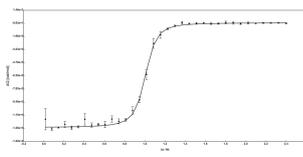
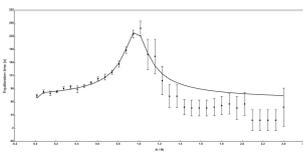
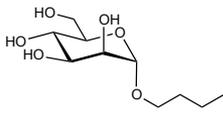
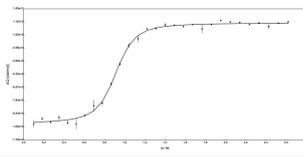
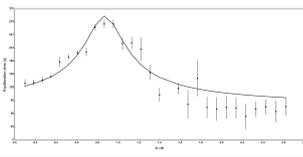
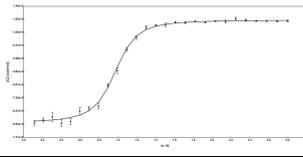
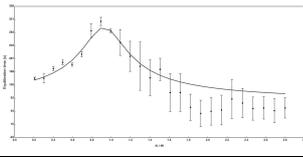
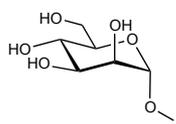
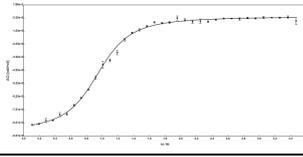
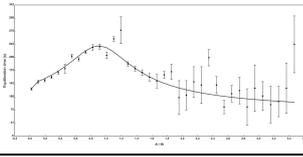
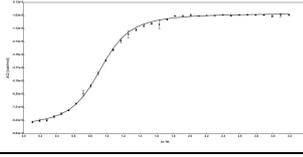
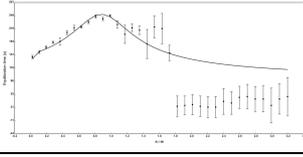
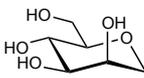
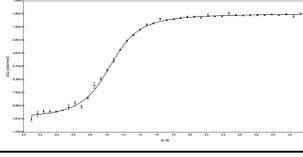
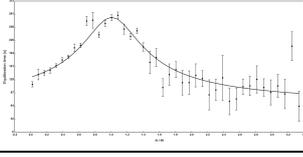
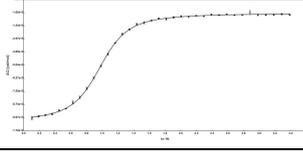
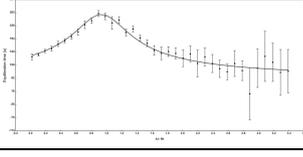
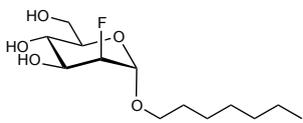
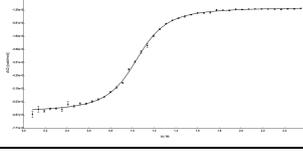
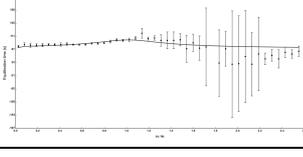
1. Ehrlich, P. (1913). Chemotherapeutics: Scientific Principles, Methods, and Results. *The Lancet*, 182(4694): 445-451.
2. Bosch, F. & Rosich, L. (2008). The contributions of Paul Ehrlich to pharmacology: a tribute on the occasion of the centenary of his Nobel Prize. *Pharmacology*, 82(3): 171-179.
3. Copeland, R.A., Pompliano, D.L., & Meek, T.D. (2006). Drug–target residence time and its implications for lead optimization. *Nat. Rev. Drug Discov.*, 5(9): 730-739.
4. Copeland, R.A. (2016). The drug-target residence time model: a 10-year retrospective. *Nat. Rev. Drug Discov.*, 15(2): 87-95.
5. Dahl, G. & Akerud, T. (2013). Pharmacokinetics and the drug–target residence time concept. *Drug Discov. Today*, 18(15): 697-707.
6. Hay, M., Thomas, D.W., Craighead, J.L., Economides, C., & Rosenthal, J. (2014). Clinical development success rates for investigational drugs. *Nat. Biotechnol.*, 32(1): 40-51.
7. Nishino, S., *et al.* (1997). Effects of thyrotropin-releasing hormone and its analogs on daytime sleepiness and cataplexy in canine narcolepsy. *J. Neurosci.*, 17(16): 6401-6408.
8. Vanhauwe, J.F., Ercken, M., van de Wiel, D., Jurzak, M., & Leysen, J.E. (2000). Effects of recent and reference antipsychotic agents at human dopamine D2 and D3 receptor signaling in Chinese hamster ovary cells. *Psychopharmacology*, 150(4): 383-390.
9. Garah, F.B.E., Stigliani, J.L., Coslédan, F., Meunier, B., & Robert, A. (2009). Docking studies of structurally diverse antimalarial drugs targeting PfATP6: no correlation between in silico binding affinity and in vitro antimalarial activity. *ChemMedChem*, 4(9): 1469-1479.
10. Braz, V.A., Holladay, L.A., & Barkley, M.D. (2009). Efavirenz binding to HIV-1 reverse transcriptase monomers and dimers. *Biochemistry*, 49(3): 601-610.
11. Maschera, B., *et al.* (1996). Human immunodeficiency virus mutations in the viral protease that confer resistance to saquinavir increase the dissociation rate constant of the protease-saquinavir complex. *J. Biol. Chem.*, 271(52): 33231-33235.
12. Pan, A.C., Borhani, D.W., Dror, R.O., & Shaw, D.E. (2013). Molecular determinants of drug–receptor binding kinetics. *Drug Discov. Today*, 18(13): 667-673.
13. Kati, W.M., *et al.* (2002). In vitro characterization of A-315675, a highly potent inhibitor of A and B strain influenza virus neuraminidases and influenza virus replication. *Antimicrob. Agents Chemother.*, 46(4): 1014-1021.
14. Chan, C., *et al.* (1999). Rofecoxib: A potent and orally active cyclooxygenase-2 inhibitor. Pharmacological and biochemical profiles. *J. Pharmacol. Exp. Ther.*, 290(2): 551-560.
15. Dierynck, I., *et al.* (2007). Binding kinetics of darunavir to human immunodeficiency virus type 1 protease explain the potent antiviral activity and high genetic barrier. *J. Virol.*, 81(24): 13845-13851.
16. Tummino, P.J. & Copeland, R.A. (2008). Residence time of receptor– ligand complexes and its effect on biological function. *Biochemistry*, 47(20): 5481-5492.
17. Lu, H. & Tonge, P.J. (2010). Drug–target residence time: critical information for lead optimization. *Curr. Opin. Chem. Biol.*, 14(4): 467-474.
18. Kapur, S. & Seeman, P. (2000). Antipsychotic agents differ in how fast they come off the dopamine D2 receptors. Implications for atypical antipsychotic action. *J. Psychiatry Neurosci.*, 25(2): 161.
19. Vauquelin, G. (2010). Rebinding: or why drugs may act longer in vivo than expected from their in vitro target residence time. *Expert Opin. Drug Discov.*, 5(10): 927-941.
20. Copeland, R.A. (2015). Drug-Target Residence Time. *Thermodynamics and Kinetics of Drug Binding*, eds Mannhold R, Kubinyi H, Folkers G, Keserü G, & Swinney DC, John Wiley & Sons, Weinheim. Vol. 65, pp. 157-167.
21. Muller, P.Y. & Milton, M.N. (2012). The determination and interpretation of the therapeutic index in drug development. *Nat. Rev. Drug Discov.*, 11(10): 751-761.
22. Klebe, G. (2015). The use of thermodynamic and kinetic data in drug discovery: decisive insight or increasing the puzzlement? *ChemMedChem*, 10(2): 229-231.
23. Miller, D.C., *et al.* (2012). Investigation of the effect of molecular properties on the binding kinetics of a ligand to its biological target. *MedChemComm*, 3(4): 449-452.
24. Schmidtke, P., Luque, F.J., Murray, J.B., & Barril, X. (2011). Shielded hydrogen bonds as structural determinants of binding kinetics: application in drug design. *J. Am. Chem. Soc.*, 133(46): 18903-18910.

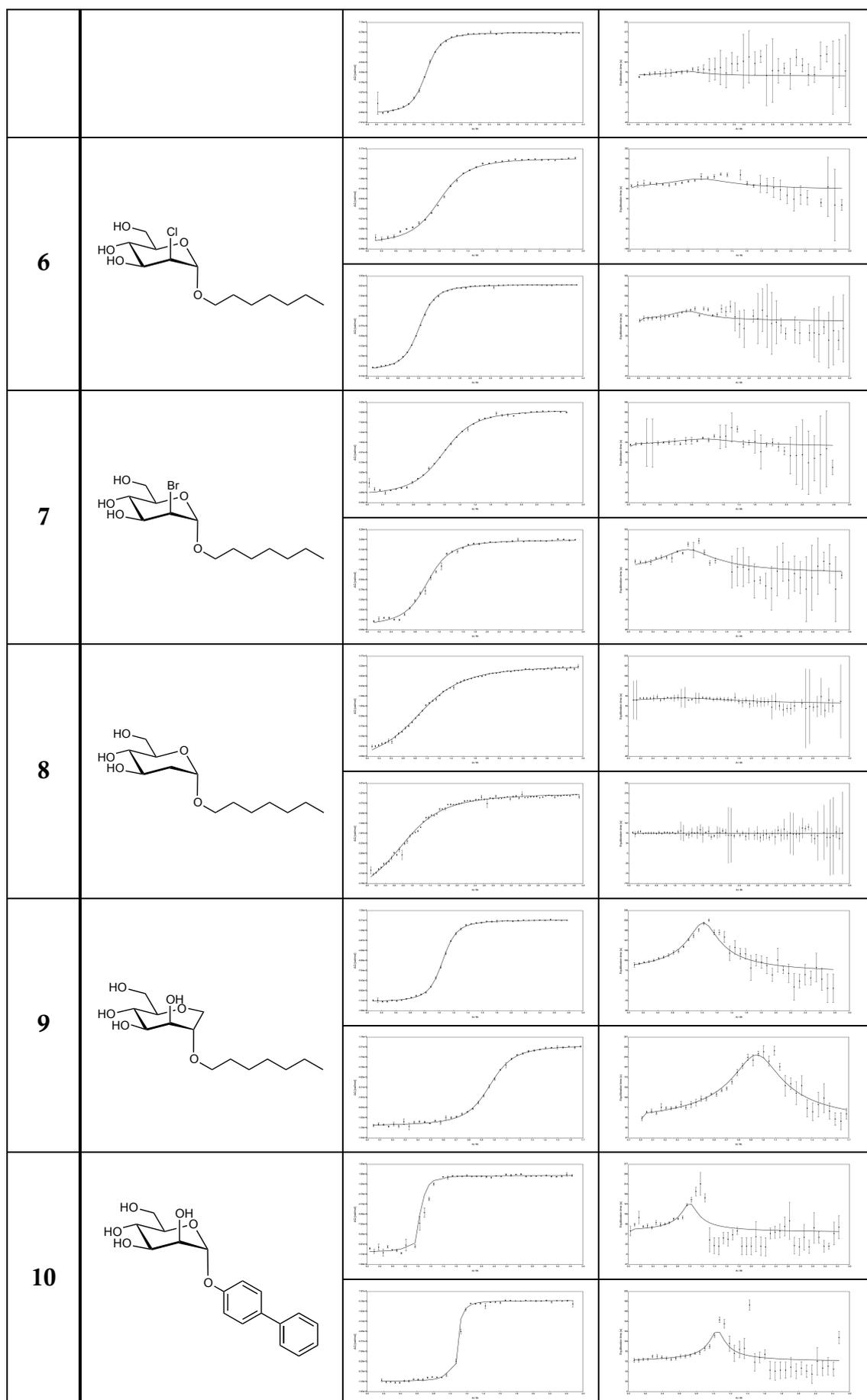
25. Jönsson, U., *et al.* (1991). Real-time biospecific interaction analysis using surface plasmon resonance and a sensor chip technology. *BioTechniques*, 11(5): 620-627.
26. Karlsson, R. (2004). SPR for molecular interaction analysis: a review of emerging application areas. *J. Mol. Recognit.*, 17(3): 151-161.
27. Holdgate, G.A. & Ward, W.H. (2005). Measurements of binding thermodynamics in drug discovery. *Drug Discov. Today*, 10(22): 1543-1550.
28. Turnbull, W.B. & Daranas, A.H. (2003). On the value of *c*: can low affinity systems be studied by isothermal titration calorimetry? *J. Am. Chem. Soc.*, 125(48): 14859-14866.
29. Holdgate, G., *et al.* (2013). Biophysical Methods in Drug Discovery from Small Molecule to Pharmaceutical. *Protein-Ligand Interactions, Methods in Molecular Biology*, eds Williams MA & Daviter T, Humana Press. Vol. 1008, pp. 327-355.
30. Dumas, P., *et al.* (2016). Extending ITC to Kinetics with kinITC. *Methods Enzymol.*, ed Andrew LF, Academic Press. Vol. Volume 567, pp. 157-180.
31. Burnouf, D., *et al.* (2012). kinITC: a new method for obtaining joint thermodynamic and kinetic data by isothermal titration calorimetry. *J. Am. Chem. Soc.*, 134(1): 559-565.
32. Sharon, N. (1987). Bacterial lectins, cell-cell recognition and infectious disease. *FEBS Lett.*, 217(2): 145-157.
33. Sokurenko, E.V., *et al.* (1998). Pathogenic adaptation of Escherichia coli by natural variation of the FimH adhesin. *Proc. Natl. Acad. Sci. U. S. A.*, 95(15): 8922-8926.
34. Wu, X.R., Sun, T.T., & Medina, J.J. (1996). In vitro binding of type 1-fimbriated Escherichia coli to uroplakins Ia and Ib: relation to urinary tract infections. *Proc. Natl. Acad. Sci. U. S. A.*, 93(18): 9630-9635.
35. Hartmann, M. & Lindhorst, T.K. (2011). The Bacterial Lectin FimH, a Target for Drug Discovery – Carbohydrate Inhibitors of Type 1 Fimbriae-Mediated Bacterial Adhesion. *Eur. J. Org. Chem.*, 2011(20-21): 3583-3609.
36. Abgottsporn, D. & Ernst, B. (2012). In vivo Evaluation of FimH Antagonists - A Novel Class of Antimicrobials for the Treatment of Urinary Tract Infection. *Chimia*, 66(4): 166-169.
37. Fiege, B., *et al.* (2015). The Tyrosine Gate of the Bacterial Lectin FimH: A Conformational Analysis by NMR Spectroscopy and X-ray Crystallography. *ChemBioChem*, 16(8): 1235-1246.
38. Sager, C.P., *et al.* (2018). The price of flexibility - a case study on septanoses as pyranose mimetics. *Chemical Science*, The Royal Society of Chemistry.
39. Kleeb, S., *et al.* (2015). FimH antagonists: bioisosteres to improve the in vitro and in vivo PK/PD profile. *J. Med. Chem.*, 58(5): 2221-2239.
40. Pang, L., *et al.* (2012). FimH antagonists: structure-activity and structure-property relationships for biphenyl alpha-D-mannopyranosides. *ChemMedChem*, 7(8): 1404-1422.
41. Schönemann, W., *et al.* (2016). 2-C-Branched mannosides as a novel family of FimH antagonists—Synthesis and biological evaluation. *Beilstein J. Org. Chem.*
42. Scharenberg, M., *et al.* (2014). Kinetic Properties of Carbohydrate–Lectin Interactions: FimH Antagonists. *ChemMedChem*, 9(1): 78-83.
43. AFFINImeter. Version 1.0415 - 1.1510; Software for Science Developments (S4Sd), Santiago de Compostela, Spain, 2015 - 2016.
44. Gopalakrishnan, M., *et al.* (2005). Ligand rebinding: self-consistent mean-field theory and numerical simulations applied to surface plasmon resonance studies. *Eur. Biophys. J.*, 34(7): 943-958.
45. Swinney, D.C. (2009). The role of binding kinetics in therapeutically useful drug action. *Curr. Opin. Drug Discov. Dev.*, 12(1): 31-39.
46. Copeland, R.A. (2010). The dynamics of drug-target interactions: drug-target residence time and its impact on efficacy and safety. *Expert Opin. Drug Discov.*, 5(4): 305-310.
47. Schneider, E.V., Böttcher, J., Huber, R., Maskos, K., & Neumann, L. (2013). Structure–kinetic relationship study of CDK8/CycC specific compounds. *Proc. Natl. Acad. Sci. U. S. A.*, 110(20): 8081-8086.
48. Sauer, M.M., *et al.* (2016). Catch-bond mechanism of the bacterial adhesin FimH. *Nat. Commun.*, 7: 10738.
49. Bouckaert, J., *et al.* (2005). Receptor binding studies disclose a novel class of high-affinity inhibitors of the Escherichia coli FimH adhesin. *Mol. Microbiol.*, 55(2): 441-455.
50. Taylor, P.W., King, R.W., & Burgen, A.S. (1970). Kinetics of complex formation between human carbonic anhydrases and aromatic sulfonamides. *Biochemistry*, 9(13): 2638-2645.

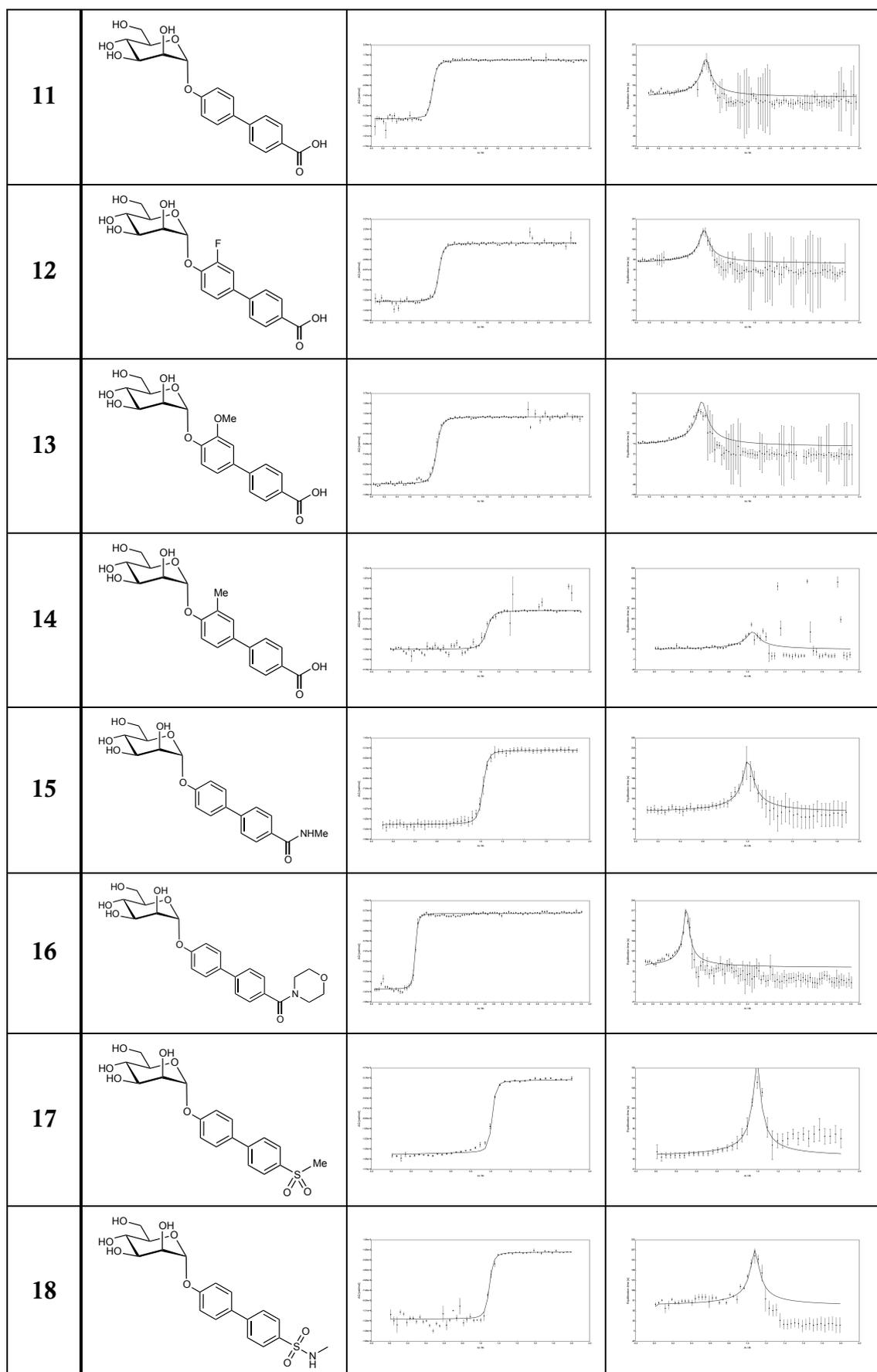
51. King, R. & Burgen, A. (1976). Kinetic aspects of structure-activity relations: the binding of sulphonamides by carbonic anhydrase. *Proceedings of the Royal Society of London B: Biological Sciences*, 193(1111): 107-125.
52. Gaspari, R., *et al.* (2016). Kinetic and Structural Insights into the Mechanism of Binding of Sulfonamides to Human Carbonic Anhydrase by Computational and Experimental Studies. *J. Med. Chem.*

## Supporting Information

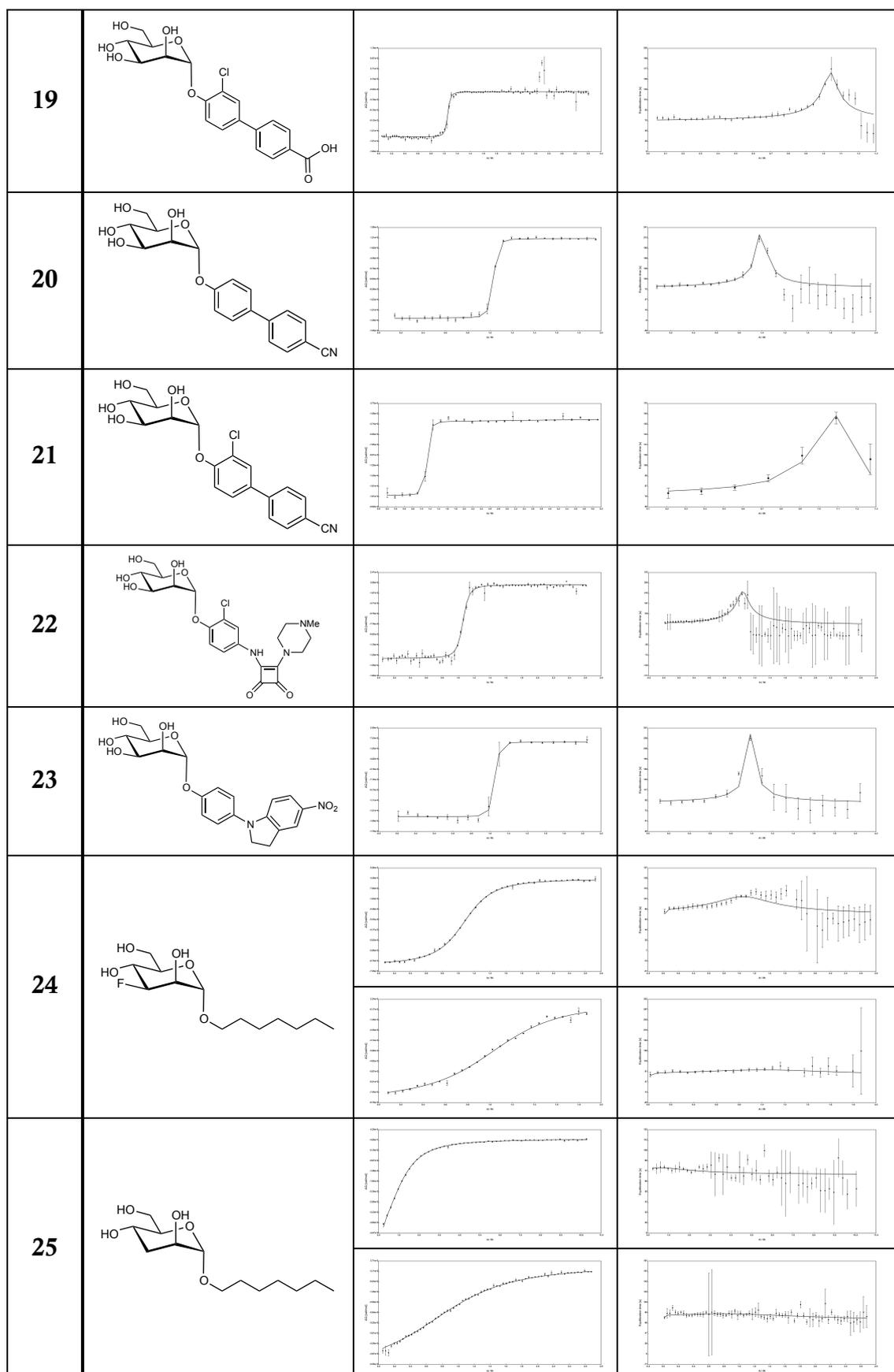
**Table S1:** Structure, thermodynamic and kinetic fits of FimH<sub>LD</sub> binding to mannoside ligands 1-29. All measurements were performed at 25°C in HEPES buffer adjusted to pH 7.4, containing 150 mM NaCl. Data were analyzed using the AFFINImeter software.

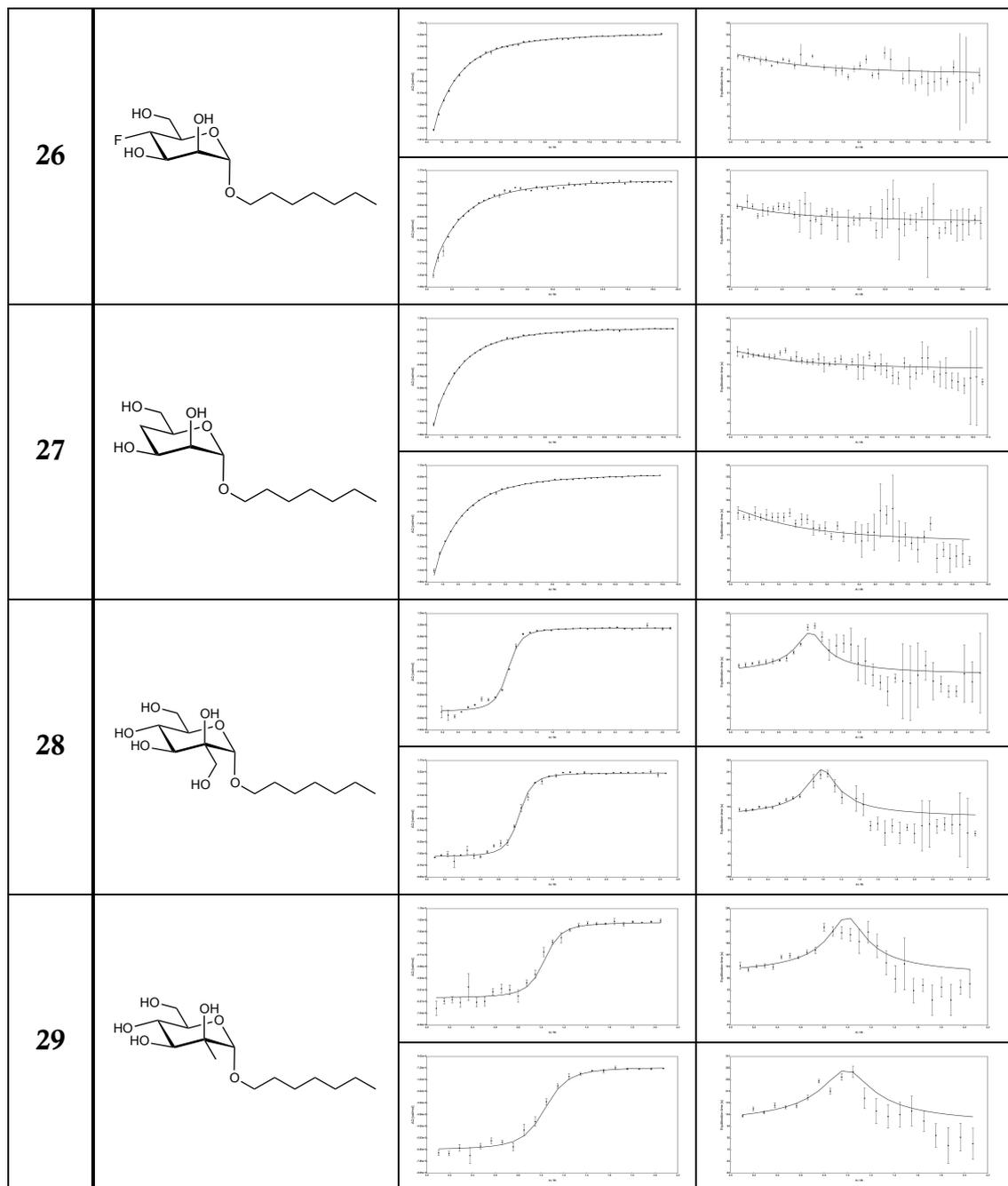
Cpd	Structure	Thermodynamic fit	Kinetic fit
1			
			
2			
			
3			
			
4			
			
5			





## Section II. The Bacterial Adhesin FimH – Manuscript 6





## Section II. The Bacterial Adhesin FimH – Manuscript 6

**Table S2:** Comparison of  $K_D$ -values obtained by kinITC-ETC and SPR for the interaction of FimH<sub>LD</sub> with mannose-derived FimH antagonists. Absolute ( $K_D$ ) and relative ( $rK_D$ ) parameters are shown for both methods. Relative values are normalized to *n*-heptyl  $\alpha$ -D-mannopyranoside (**1**), which is set to 1. The discrepancies between kinITC and SPR for  $K_D$  and  $rK_D$  is displayed in the right columns.

FimH <sub>LD</sub> ligand dissociation constants determined by ITC and SPR						
Compound	$K_D^{\text{kinITC-ETC}}$		$K_D^{\text{SPR}}$		$K_D^{\text{kinITC-ETC}}/K_D^{\text{SPR}}$	
	$K_D$ [M]	$rK_D$	$K_D$ [M]	$rK_D$	absolute	relative
<b>1</b>	$2.19 \times 10^{-8}$	1.00	$6.28 \times 10^{-9}$	1.00	3.48	1.00
<b>2</b>	$1.18 \times 10^{-7}$	5.40	$3.39 \times 10^{-8}$	5.39	3.49	1.00
<b>3</b>	$1.24 \times 10^{-6}$	56.47	$4.54 \times 10^{-7}$	72.21	2.72	0.78
<b>4</b>	$9.13 \times 10^{-7}$	41.71	$2.44 \times 10^{-7}$	38.78	3.75	1.08

**Table S3:** Thermodynamic parameters of FimH<sub>LD</sub> binding to mannose ligands **1-29**. All ITC measurements were performed at 25°C in HEPES buffer adjusted to pH 7.4, containing 150 mM NaCl. Data were analyzed using the AFFINImeter software.

Cpd	$K_D$ [M]	c.i. $K_D$ [M]	$\Delta G^\circ$ [kJ/mol]	$\Delta H^\circ$ [kJ/mol]	c.i. $\Delta H^\circ$ [kJ/mol]	$-T\Delta S^\circ$ [kJ/mol]	$N$	$c$
1	$1.93 \times 10^{-8}$	$2.45 \times 10^{-9}$	-44.0	-48.8	0.2	4.8	0.97	518
1	$2.45 \times 10^{-8}$	$2.81 \times 10^{-9}$	-43.4	-52.1	0.2	8.7	0.98	351
<b>1</b>	<b><math>2.19 \times 10^{-8}</math></b>	<b><math>2.63 \times 10^{-9}</math></b>	<b>-43.7</b>	<b>-50.5</b>	<b>0.2</b>	<b>6.7</b>		
2	$1.08 \times 10^{-7}$	$7.38 \times 10^{-9}$	-39.8	-42.7	0.3	3.0	0.88	104
2	$1.28 \times 10^{-7}$	$1.35 \times 10^{-8}$	-39.3	-41.1	0.6	1.8	0.94	74
<b>2</b>	<b><math>1.18 \times 10^{-7}</math></b>	<b><math>1.04 \times 10^{-8}</math></b>	<b>-39.5</b>	<b>-41.9</b>	<b>0.4</b>	<b>2.4</b>		
3	$1.39 \times 10^{-6}$	$4.87 \times 10^{-8}$	-33.4	-36.3	0.2	2.8	0.95	21
3	$1.08 \times 10^{-6}$	$4.55 \times 10^{-8}$	-34.0	-36.7	0.2	2.6	0.92	27
<b>3</b>	<b><math>1.24 \times 10^{-6}</math></b>	<b><math>4.71 \times 10^{-8}</math></b>	<b>-33.7</b>	<b>-36.5</b>	<b>0.2</b>	<b>2.8</b>		
4	$8.06 \times 10^{-7}$	$3.69 \times 10^{-8}$	-34.8	-41.1	0.2	6.3	1.04	34
4	$1.02 \times 10^{-6}$	$3.73 \times 10^{-8}$	-34.2	-42.3	0.2	8.1	0.97	29
<b>4</b>	<b><math>9.13 \times 10^{-7}</math></b>	<b><math>3.71 \times 10^{-8}</math></b>	<b>-34.5</b>	<b>-41.7</b>	<b>0.2</b>	<b>7.2</b>		
5 <sup>b</sup>	$5.65 \times 10^{-7}$	$2.04 \times 10^{-8}$	-35.7	-27.5	0.1	-8.2	1.02	41
<b>5</b>	<b><math>4.99 \times 10^{-7}</math></b>	<b><math>1.30 \times 10^{-8}</math></b>	<b>-36.0</b>	<b>-28.0</b>	<b>0.1</b>	<b>-8.0</b>	1.04	49
6	$1.68 \times 10^{-6}$	$3.36 \times 10^{-8}$	-33.0	-24.9	0.1	-8.1	1.04	28
6	$9.95 \times 10^{-7}$	$1.80 \times 10^{-8}$	-34.3	-24.1	0.1	-10.1	1.02	40
<b>6</b>	<b><math>1.34 \times 10^{-6}</math></b>	<b><math>2.58 \times 10^{-8}</math></b>	<b>-33.5</b>	<b>-24.5</b>	<b>0.1</b>	<b>-9.0</b>		
7	$2.14 \times 10^{-6}$	$3.81 \times 10^{-8}$	-32.4	-21.1	0.1	-11.2	1.00	23
7	$1.52 \times 10^{-6}$	$5.48 \times 10^{-8}$	-33.2	-20.5	0.1	-12.7	1.03	21
<b>7</b>	<b><math>1.83 \times 10^{-6}</math></b>	<b><math>4.65 \times 10^{-8}</math></b>	<b>-32.7</b>	<b>-20.8</b>	<b>0.1</b>	<b>-11.9</b>		
8 <sup>b</sup>	$1.18 \times 10^{-5}$	$1.16 \times 10^{-7}$	-28.1	-20.3	0.0	-7.9	1.01	5.3
<b>8</b>	<b><math>7.71 \times 10^{-6}</math></b>	<b><math>8.07 \times 10^{-8}</math></b>	<b>-29.2</b>	<b>-18.4</b>	<b>0.0</b>	<b>-10.7</b>	1.00	5.4
9	$2.59 \times 10^{-7}$	$8.84 \times 10^{-9}$	-37.6	-47.5	0.1	9.9	1.04	110
9	$2.71 \times 10^{-7}$	$9.85 \times 10^{-8}$	-37.5	-48.6	0.1	11.2	0.97	110
<b>9</b>	<b><math>2.65 \times 10^{-7}</math></b>	<b><math>5.37 \times 10^{-8}</math></b>	<b>-37.5</b>	<b>-48.0</b>	<b>0.1</b>	<b>10.5</b>		
10	$3.62 \times 10^{-8}$	$3.09 \times 10^{-9}$	-42.5	-45.9	0.2	3.4	1.00	290
10	$2.06 \times 10^{-8}$	$1.33 \times 10^{-9}$	-43.9	-43.9	0.1	0.0	1.05	681
<b>10</b>	<b><math>2.84 \times 10^{-8}</math></b>	<b><math>2.21 \times 10^{-9}</math></b>	<b>-43.1</b>	<b>-44.9</b>	<b>0.2</b>	<b>1.8</b>		
11	$1.21 \times 10^{-8}$	$9.65 \times 10^{-9}$	-45.2	-46.0	0.2	0.8	1.05	802
12	$9.21 \times 10^{-9}$	$7.20 \times 10^{-10}$	-45.9	-49.5	0.1	3.6	1.04	1053
13	$1.35 \times 10^{-8}$	$9.50 \times 10^{-10}$	-44.9	-51.2	0.1	6.3	1.00	717
14	$1.08 \times 10^{-8}$	$1.24 \times 10^{-9}$	-45.5	-56.2	0.3	10.8	1.06	1046
15 <sup>a</sup>	$6.49 \times 10^{-9}$	$1.13 \times 10^{-9}$	-46.7	-52.0	0.5	5.3	1.00	1757
16 <sup>a</sup>	$4.19 \times 10^{-9}$	$8.47 \times 10^{-10}$	-47.8	-56.2	0.2	8.3	0.98	2016
17 <sup>a</sup>	$1.21 \times 10^{-8}$	$3.49 \times 10^{-9}$	-45.2	-49.6	0.1	4.4	0.99	1056
18 <sup>a</sup>	$2.91 \times 10^{-9}$	$1.01 \times 10^{-9}$	-48.7	-50.1	0.3	1.3	1.08	3056

19 <sup>a</sup>	4.90*10 <sup>-9</sup>	8.57*10 <sup>-10</sup>	-47.4	-53.9	0.2	6.4	1.04	1980
20 <sup>a</sup>	4.13*10 <sup>-9</sup>	8.35*10 <sup>-10</sup>	-47.9	-56.1	0.3	8.2	0.98	2544
21 <sup>a</sup>	1.29*10 <sup>-9</sup>	2.61*10 <sup>-10</sup>	-50.7	-60.8	0.5	10.1	1.05	7737
22	1.15*10 <sup>-8</sup>	8.48*10 <sup>-10</sup>	-45.3	-60.7	0.2	15.4	1.04	758
23 <sup>a</sup>	9.74*10 <sup>-10</sup>	3.70*10 <sup>-10</sup>	-51.4	-59.5	0.5	8.0	1.00	5134
24	9.38*10 <sup>-7</sup>	1.54*10 <sup>-8</sup>	-34.4	-29.3	0.1	-5.1	1.09	33
24	1.45*10 <sup>-6</sup>	4.40*10 <sup>-8</sup>	-33.3	-33.4	0.2	0.1	1.08	14
24	1.19*10 <sup>-6</sup>	2.97*10 <sup>-8</sup>	-33.8	-31.4	0.1	-2.4		
25	2.09*10 <sup>-5</sup>	1.52*10 <sup>-7</sup>	-26.7	-25.2	0.1	-1.5	1.00	2.1
25	1.89*10 <sup>-5</sup>	1.81*10 <sup>-7</sup>	-27.0	-24.4	0.1	-2.6	1.00	3.8
25	1.99*10 <sup>-5</sup>	1.67*10 <sup>-7</sup>	-26.8	-24.8	0.1	-2.1		
26 <sup>c</sup>	9.19*10 <sup>-5</sup>	1.21*10 <sup>-6</sup>	-23.0	-19.3	0.1	-3.7	1.00	0.5
26 <sup>c</sup>	1.15*10 <sup>-4</sup>	1.46*10 <sup>-6</sup>	-22.5	-21.7	0.1	-0.8	1.00	0.5
26	1.03*10 <sup>-4</sup>	1.34*10 <sup>-6</sup>	-22.7	-20.5	0.1	-2.2		
27 <sup>c</sup>	8.57*10 <sup>-5</sup>	5.74*10 <sup>-7</sup>	-23.2	-19.6	0.0	-3.7	1.00	0.5
27 <sup>c</sup>	9.11*10 <sup>-5</sup>	6.60*10 <sup>-7</sup>	-23.1	-20.1	0.0	-3.0	1.00	0.5
27	8.84*10 <sup>-5</sup>	6.17*10 <sup>-7</sup>	-23.1	-19.8	0.0	-3.3		
28	1.69*10 <sup>-7</sup>	4.48*10 <sup>-9</sup>	-38.7	-35.0	0.1	-3.7	1.00	164
28	7.64*10 <sup>-8</sup>	4.27*10 <sup>-9</sup>	-40.6	-33.1	0.1	-7.5	1.00	107
28	1.22*10 <sup>-7</sup>	4.38*10 <sup>-9</sup>	-39.5	-34.0	0.1	-5.4		
29	5.05*10 <sup>-8</sup>	7.83*10 <sup>-9</sup>	-41.6	-26.1	4.1	-15.5	0.98	77
29	9.33*10 <sup>-8</sup>	9.66*10 <sup>-8</sup>	-40.1	-28.8	0.3	-11.3	0.99	94
29	7.19*10 <sup>-8</sup>	5.22*10 <sup>-8</sup>	-40.8	-27.5	2.2	-13.3		

<sup>a</sup>  $K_D$  fixed from a competitive experiment evaluated with SEDPHAT due to  $c$ -values above 1000.

<sup>b</sup> Excluded from data analysis due to large errors of the kinetic fit.

<sup>c</sup> Stoichiometry fixed to 1 due to  $c$ -values below 1.

**Table S4:** Kinetic parameters of FimH<sub>LD</sub> binding to mannoside ligands 1-29. All ITC measurements were performed at 25°C in HEPES buffer adjusted to pH 7.4, containing 150 mM NaCl. Data were analyzed using AFFINImeter software.  $k_{on}$  was calculated from the fitted parameters  $K_D$  and  $k_{off}$ .

Cpd	$K_D$ [M]	$\ln(K_D)$	$k_{off}$ [s <sup>-1</sup> ]	$\ln(k_{off})$	c.i. $k_{off}$ [s <sup>-1</sup> ]	$k_{on}$ [M <sup>-1</sup> s <sup>-1</sup> ]	$\ln(k_{on})$	$t_{1/2}$ [min]	Response time [s]
1	1.93*10 <sup>-8</sup>	-17.76	7.47*10 <sup>-4</sup>	-7.20	6.10*10 <sup>-5</sup>	3.87*10 <sup>4</sup>	10.56	15.46	12.9
1	2.45*10 <sup>-8</sup>	-17.53	7.06*10 <sup>-4</sup>	-7.26	5.73*10 <sup>-5</sup>	2.88*10 <sup>4</sup>	10.27	16.37	12.0
1	2.19*10 <sup>-8</sup>	-17.64	7.27*10 <sup>-4</sup>	-7.23	5.91*10 <sup>-5</sup>	3.32*10 <sup>4</sup>	10.41	15.90	
2	1.08*10 <sup>-7</sup>	-16.04	1.06*10 <sup>-3</sup>	-6.85	5.98*10 <sup>-5</sup>	9.84*10 <sup>3</sup>	9.19	10.85	12.1
2	1.28*10 <sup>-7</sup>	-15.87	1.23*10 <sup>-3</sup>	-6.70	9.32*10 <sup>-5</sup>	9.61*10 <sup>3</sup>	9.17	9.39	13.2
2	1.18*10 <sup>-7</sup>	-15.95	1.15*10 <sup>-3</sup>	-6.77	7.65*10 <sup>-5</sup>	9.71*10 <sup>3</sup>	9.18	10.07	
3	1.39*10 <sup>-6</sup>	-13.49	2.92*10 <sup>-3</sup>	-5.84	1.21*10 <sup>-4</sup>	2.10*10 <sup>3</sup>	7.65	3.96	11.2
3	1.08*10 <sup>-6</sup>	-13.74	2.79*10 <sup>-3</sup>	-5.88	1.04*10 <sup>-4</sup>	2.58*10 <sup>3</sup>	7.85	4.14	16.3
3	1.24*10 <sup>-6</sup>	-13.60	2.86*10 <sup>-3</sup>	-5.86	1.12*10 <sup>-4</sup>	2.31*10 <sup>3</sup>	7.75	4.05	
4	8.06*10 <sup>-7</sup>	-14.03	1.73*10 <sup>-3</sup>	-6.36	7.43*10 <sup>-5</sup>	2.14*10 <sup>3</sup>	7.67	6.69	13.0
4	1.02*10 <sup>-6</sup>	-13.80	1.92*10 <sup>-3</sup>	-6.26	9.53*10 <sup>-5</sup>	1.88*10 <sup>3</sup>	7.54	6.03	11.9
4	9.13*10 <sup>-7</sup>	-13.91	1.82*10 <sup>-3</sup>	-6.31	5.14*10 <sup>-4</sup>	2.00*10 <sup>3</sup>	7.60	6.34	
5 <sup>b</sup>	5.65*10 <sup>-7</sup>	-14.39	5.33*10 <sup>-2</sup>	-2.93	3.13*10 <sup>-2</sup>	9.43*10 <sup>4</sup>	11.45	0.22	15.0
5	4.99*10 <sup>-7</sup>	-14.51	1.02*10 <sup>-2</sup>	-4.59	1.22*10 <sup>-3</sup>	2.04*10 <sup>4</sup>	9.92	1.14	13.6
6	1.68*10 <sup>-6</sup>	-13.30	1.15*10 <sup>-2</sup>	-4.47	1.50*10 <sup>-3</sup>	6.84*10 <sup>3</sup>	8.83	1.01	11.8
6	9.95*10 <sup>-7</sup>	-13.82	1.23*10 <sup>-2</sup>	-4.40	1.42*10 <sup>-3</sup>	1.24*10 <sup>4</sup>	9.42	0.94	12.6
6	1.34*10 <sup>-6</sup>	-13.53	1.19*10 <sup>-2</sup>	-4.43	1.46*10 <sup>-3</sup>	8.90*10 <sup>3</sup>	9.09	0.97	
7	2.14*10 <sup>-6</sup>	-13.05	2.19*10 <sup>-2</sup>	-3.82	3.49*10 <sup>-3</sup>	1.02*10 <sup>4</sup>	9.23	0.53	13.7
7	1.52*10 <sup>-6</sup>	-13.40	7.00*10 <sup>-3</sup>	-4.96	5.01*10 <sup>-4</sup>	4.62*10 <sup>3</sup>	8.44	1.65	8.9
7	1.83*10 <sup>-6</sup>	-13.21	1.45*10 <sup>-2</sup>	-4.24	1.99*10 <sup>-3</sup>	7.90*10 <sup>3</sup>	8.97	0.80	
8 <sup>b</sup>	1.18*10 <sup>-5</sup>	-11.34	4.98*10 <sup>-1</sup>	-0.70	8.21*10 <sup>-1</sup>	4.21*10 <sup>4</sup>	10.65	0.02	14.6
8	7.71*10 <sup>-6</sup>	-11.77	4.13*10 <sup>-2</sup>	-3.19	1.17*10 <sup>-2</sup>	5.35*10 <sup>3</sup>	8.59	0.28	13.1
9	2.59*10 <sup>-7</sup>	-15.17	1.29*10 <sup>-3</sup>	-6.65	3.88*10 <sup>-5</sup>	4.98*10 <sup>3</sup>	8.51	8.95	10.3
9	2.71*10 <sup>-7</sup>	-15.12	1.41*10 <sup>-3</sup>	-6.56	4.72*10 <sup>-5</sup>	5.20*10 <sup>3</sup>	8.56	8.20	12.2
9	2.65*10 <sup>-7</sup>	-15.14	1.35*10 <sup>-3</sup>	-6.61	4.30*10 <sup>-5</sup>	5.10*10 <sup>3</sup>	8.54	8.56	

## Section II. The Bacterial Adhesin FimH – Manuscript 6

10	3.62*10 <sup>-8</sup>	-17.13	1.42*10 <sup>-3</sup>	-6.56	9.98*10 <sup>-5</sup>	3.91*10 <sup>4</sup>	10.57	8.16	11.1
10	2.06*10 <sup>-8</sup>	-17.70	1.31*10 <sup>-3</sup>	-6.64	8.56*10 <sup>-5</sup>	6.36*10 <sup>4</sup>	11.06	8.83	12.9
<b>10</b>	<b>2.84*10<sup>-8</sup></b>	-17.38	<b>1.36*10<sup>-3</sup></b>	-6.60	<b>9.27*10<sup>-5</sup></b>	<b>4.80*10<sup>4</sup></b>	10.78	<b>8.48</b>	
11	1.21*10 <sup>-8</sup>	-18.23	4.83*10 <sup>-4</sup>	-7.63	2.85*10 <sup>-5</sup>	4.00*10 <sup>4</sup>	10.60	<b>23.90</b>	12.2
12	9.21*10 <sup>-9</sup>	-18.50	5.03*10 <sup>-4</sup>	-7.60	2.94*10 <sup>-5</sup>	5.45*10 <sup>4</sup>	10.91	<b>22.99</b>	13.1
13	1.35*10 <sup>-8</sup>	-18.12	4.40*10 <sup>-4</sup>	-7.73	2.19*10 <sup>-5</sup>	3.25*10 <sup>4</sup>	10.39	<b>26.26</b>	12.1
14	1.08*10 <sup>-8</sup>	-18.34	4.96*10 <sup>-4</sup>	-7.61	4.01*10 <sup>-5</sup>	4.59*10 <sup>4</sup>	10.73	<b>23.29</b>	13.3
15 <sup>a</sup>	6.49*10 <sup>-9</sup>	-18.85	3.31*10 <sup>-4</sup>	-8.01	3.51*10 <sup>-5</sup>	5.10*10 <sup>4</sup>	10.84	<b>34.89</b>	12.4
16 <sup>a</sup>	4.19*10 <sup>-9</sup>	-19.29	3.77*10 <sup>-4</sup>	-7.88	1.07*10 <sup>-5</sup>	9.02*10 <sup>4</sup>	11.41	<b>30.61</b>	12.4
17 <sup>a</sup>	1.21*10 <sup>-8</sup>	-18.23	3.23*10 <sup>-4</sup>	-8.04	1.36*10 <sup>-5</sup>	2.66*10 <sup>4</sup>	10.19	<b>35.81</b>	11.1
18 <sup>a</sup>	2.91*10 <sup>-9</sup>	-19.66	3.08*10 <sup>-4</sup>	-8.08	1.42*10 <sup>-5</sup>	1.06*10 <sup>5</sup>	11.57	<b>37.47</b>	12.7
19 <sup>a</sup>	4.90*10 <sup>-9</sup>	-19.13	4.10*10 <sup>-4</sup>	-7.80	1.56*10 <sup>-5</sup>	8.36*10 <sup>4</sup>	11.33	<b>28.20</b>	13.8
20 <sup>a</sup>	4.13*10 <sup>-9</sup>	-19.31	2.14*10 <sup>-4</sup>	-8.45	9.02*10 <sup>-6</sup>	5.19*10 <sup>4</sup>	10.86	<b>53.96</b>	12.8
21 <sup>a</sup>	1.29*10 <sup>-9</sup>	-20.47	5.14*10 <sup>-4</sup>	-7.57	4.30*10 <sup>-5</sup>	3.98*10 <sup>5</sup>	12.89	<b>22.48</b>	12.3
22	1.15*10 <sup>-8</sup>	-18.28	5.04*10 <sup>-4</sup>	-7.59	2.97*10 <sup>-5</sup>	4.39*10 <sup>4</sup>	10.69	<b>22.94</b>	12.8
23 <sup>a</sup>	9.74*10 <sup>-10</sup>	-20.75	2.10*10 <sup>-4</sup>	-8.47	6.27*10 <sup>-6</sup>	2.16*10 <sup>5</sup>	12.28	<b>55.00</b>	13.2
24	9.38*10 <sup>-7</sup>	-13.88	1.07*10 <sup>-2</sup>	-4.54	8.04*10 <sup>-4</sup>	1.14*10 <sup>4</sup>	9.34	1.08	12.8
24	1.45*10 <sup>-6</sup>	-13.44	2.90*10 <sup>-2</sup>	-3.54	8.16*10 <sup>-3</sup>	2.00*10 <sup>4</sup>	9.91	0.40	11.9
<b>24</b>	<b>1.19*10<sup>-6</sup></b>	-13.64	<b>1.98*10<sup>-2</sup></b>	-3.92	<b>4.48*10<sup>-3</sup></b>	<b>1.66*10<sup>4</sup></b>	9.72	<b>0.58</b>	
25	2.09*10 <sup>-5</sup>	-10.78	1.85*10 <sup>-1</sup>	-1.69	5.66*10 <sup>-2</sup>	8.85*10 <sup>3</sup>	9.09	0.06	13.6
25	1.89*10 <sup>-5</sup>	-10.88	7.36*10 <sup>-2</sup>	-2.61	1.86*10 <sup>-2</sup>	3.90*10 <sup>3</sup>	8.27	0.16	13.4
<b>25</b>	<b>1.99*10<sup>-5</sup></b>	-10.83	<b>1.29*10<sup>-1</sup></b>	-2.05	<b>3.76*10<sup>-2</sup></b>	<b>6.50*10<sup>3</sup></b>	8.78	<b>0.09</b>	
26 <sup>c</sup>	9.19*10 <sup>-5</sup>	-9.29	1.03*10 <sup>-1</sup>	-2.27	1.17*10 <sup>-2</sup>	1.12*10 <sup>3</sup>	7.02	0.11	13.5
26 <sup>c</sup>	1.15*10 <sup>-4</sup>	-9.07	1.04*10 <sup>-1</sup>	-2.26	1.48*10 <sup>-2</sup>	9.04*10 <sup>2</sup>	6.81	0.11	12.7
<b>26</b>	<b>1.03*10<sup>-4</sup></b>	-9.18	<b>1.03*10<sup>-1</sup></b>	-2.27	<b>1.32*10<sup>-2</sup></b>	<b>1.00*10<sup>3</sup></b>	6.91	<b>0.11</b>	
27 <sup>c</sup>	8.57*10 <sup>-5</sup>	-9.36	8.65*10 <sup>-2</sup>	-2.45	8.33*10 <sup>-3</sup>	1.01*10 <sup>3</sup>	6.92	0.13	10.9
27 <sup>c</sup>	9.11*10 <sup>-5</sup>	-9.30	8.25*10 <sup>-2</sup>	-2.50	7.34*10 <sup>-3</sup>	9.05*10 <sup>2</sup>	6.81	0.14	10.9
<b>27</b>	<b>8.84*10<sup>-5</sup></b>	-9.33	<b>8.45*10<sup>-2</sup></b>	-2.47	<b>7.84*10<sup>-3</sup></b>	<b>9.56*10<sup>2</sup></b>	6.86	<b>0.14</b>	
28	1.69*10 <sup>-7</sup>	-15.60	1.52*10 <sup>-3</sup>	-6.49	6.98*10 <sup>-5</sup>	9.01*10 <sup>3</sup>	9.11	7.61	10.8
28	7.64*10 <sup>-8</sup>	-16.39	8.77*10 <sup>-4</sup>	-7.04	5.35*10 <sup>-5</sup>	1.15*10 <sup>4</sup>	9.35	13.18	12.5
<b>28</b>	<b>1.22*10<sup>-7</sup></b>	-15.92	<b>1.20*10<sup>-3</sup></b>	-6.73	<b>6.16*10<sup>-5</sup></b>	<b>9.78*10<sup>3</sup></b>	9.19	<b>9.65</b>	
29	5.05*10 <sup>-8</sup>	-16.80	8.93*10 <sup>-4</sup>	-7.02	1.03*10 <sup>-4</sup>	1.77*10 <sup>4</sup>	9.78	12.94	16.0
29	9.33*10 <sup>-8</sup>	-16.19	1.13*10 <sup>-3</sup>	-6.79	9.01*10 <sup>-5</sup>	1.21*10 <sup>4</sup>	9.40	10.24	15.0
<b>29</b>	<b>7.19*10<sup>-8</sup></b>	-16.45	<b>1.01*10<sup>-3</sup></b>	-6.90	<b>9.67*10<sup>-5</sup></b>	<b>1.41*10<sup>4</sup></b>	9.55	<b>11.43</b>	

<sup>a</sup>  $K_D$  fixed from a competitive experiment evaluated with SEDPHAT due to  $c$ -values above 1000.

<sup>b</sup> Excluded from data analysis due to large errors of the kinetic fit.

<sup>c</sup> Stoichiometry fixed to 1 due to  $c$ -values below 1.

**Table S5.** Kinetic parameters of FimH<sub>LD</sub> binding to mannoside ligands 1-4. All SPR measurements were performed at 25°C in HBS-EP buffer adjusted to pH 7.4, containing 150 mM NaCl, 3 mM EDTA and 0,005% P20. Two independent measurements were performed and the best fit (Biacore T200 Evaluation software version 1) was averaged (bold). Standard deviations are given in parentheses.  $K_D$  was calculated from the fitted parameters  $k_{on}$  and  $k_{off}$ .

Cpd	$K_D$ [M]	$k_{off}$ [s <sup>-1</sup> ]	$k_{on}$ [M <sup>-1</sup> s <sup>-1</sup> ]	$t_{1/2}$ [min]	Response Units
1	6.10*10 <sup>-9</sup>	1.68*10 <sup>-4</sup>	2.76*10 <sup>4</sup>	68.6	5.7
1	6.53*10 <sup>-9</sup>	1.39*10 <sup>-4</sup>	2.14*10 <sup>4</sup>	82.9	5.9
<b>1</b>	<b>6.28*10<sup>-9</sup></b> (3.05*10 <sup>-10</sup> )	<b>1.54*10<sup>-4</sup></b> (2.05*10 <sup>-5</sup> )	<b>2.45*10<sup>4</sup></b> (4.44*10 <sup>3</sup> )	<b>75.1</b>	
2	3.04*10 <sup>-8</sup>	1.71*10 <sup>-4</sup>	5.62*10 <sup>3</sup>	67.7	4.8
2	3.75*10 <sup>-8</sup>	2.07*10 <sup>-4</sup>	5.52*10 <sup>3</sup>	55.9	4.9
<b>2</b>	<b>3.39*10<sup>-8</sup></b> (5.05*10 <sup>-9</sup> )	<b>1.89*10<sup>-4</sup></b> (2.56*10 <sup>-5</sup> )	<b>5.57*10<sup>3</sup></b> (7.42*10 <sup>1</sup> )	<b>61.2</b>	
3	4.57*10 <sup>-7</sup>	1.21*10 <sup>-3</sup>	2.65*10 <sup>3</sup>	9.50	5.2
3	4.50*10 <sup>-7</sup>	1.01*10 <sup>-3</sup>	2.25*10 <sup>3</sup>	11.4	4.1
<b>3</b>	<b>4.54*10<sup>-7</sup></b> (5.23*10 <sup>-9</sup> )	<b>1.11*10<sup>-3</sup></b> (1.44*10 <sup>-4</sup> )	<b>2.45*10<sup>3</sup></b> (2.88*10 <sup>2</sup> )	<b>10.4</b>	
4	1.77*10 <sup>-7</sup>	4.91*10 <sup>-4</sup>	2.77*10 <sup>3</sup>	23.6	3.8
4	3.16*10 <sup>-7</sup>	8.11*10 <sup>-4</sup>	2.57*10 <sup>3</sup>	14.2	3.8
<b>4</b>	<b>2.44*10<sup>-7</sup></b> (9.82*10 <sup>-8</sup> )	<b>6.51*10<sup>-4</sup></b> (2.27*10 <sup>-4</sup> )	<b>2.67*10<sup>3</sup></b> (1.44*10 <sup>2</sup> )	<b>17.8</b>	

**Table S6:** Data collection and refinement statistics of FimH<sub>LD</sub> ligand complexes.

	FimH <sub>LD</sub> /3	FimH <sub>LD</sub> /4
PDB ID	5JCR	5MUC
Wavelength (Å)	1.00004	1.87287
Resolution range (Å)	55.92 - 1.70 (1.76 - 1.70)*	49.5 - 2.6 (2.70 - 2.60)*
Space group	P 21 21 21	P 21 21 21
Unit cell	63.0 69.0 95.6	62.4 68.6 95.7
a, β, γ (°)	90 90 90	90 90 90
Unique reflections	46026 (4525)	12481 (1322)
Multiplicity	6.5 (6.4)	4.8 (4.8)
Completeness (%)	99.8 (97.8)	94.9 (96.2)
Mean I/sigma(I)	12.3 (2.1)	6.0 (2.0)
Wilson B-factor	14.7	27.8
R-merge	0.101 (0.505)	0.196 (0.79)
CC1/2	0.999 (0.959)	0.990 (0.707)
R-work	0.2114 (0.2542)	0.235 (0.304)
R-free	0.2330 (0.2729)	0.260 (0.356)
Number of atoms	5475	5024
macromolecules	4732	2392
ligands	54	44
water	689	266
Protein residues	316	316
RMS(bonds)	0.004	0.003
RMS(angles)	1.10	0.71
Ramachandran favored (%)	97.1	96.5
Ramachandran allowed (%)	2.9	3.5
Clashscore	0.7	0.6
Average B-factor	19.0	36.0

\*Values in parentheses are for the highest resolution shell.

### Experimental section

**Protein cloning, expression, purification, and biotinylation.** For all ITC experiments FimH<sub>LD</sub> of *E. coli* K-12 strain was expressed with a C-terminal thrombin cleavage site and a 6His-tag (FimH<sub>LD</sub>-Th-6His, 173 residues) following a previously published protocol.<sup>(1)</sup> For SPR experiments a FimH<sub>LD</sub>-AVITag-6His gene construct was cloned by overlap extension PCR using the FimH<sub>LD</sub>-Th-6His<sup>(1)</sup> gene construct and an AVITag-6His gene sequence as templates and was subsequently ligated into the plasmid *pNT*.<sup>(2)</sup> The *histidine-tagged* recombinant protein was expressed in the protease-deficient *E. coli* strain HM125 and purified with *Ni-NTA affinity chromatography*.<sup>(1)</sup> AVITag is a specific 15-amino acid peptide sequence (GLNDIFEAQKIEWHE) that can be biotinylated by the *E. coli biotin ligase BirA*. To keep the *BirA* enzyme active, the protein was dialyzed with Slide-A-Lyzer MINI Dialysis Units (3,500K MWCO, Thermo Scientific) in 50 mM bicine, pH 8.3. Biotinylation was done according to the *manufacturer's protocol* (Avidity). In order to remove excessive biotin the protein sample was dialyzed in HBS-EP, 7.4, overnight.

**Surface plasmon resonance (SPR).** Surface plasmon resonance-based experiments were performed using a Biacore T200 instrument (GE Healthcare). Biotinylated FimH<sub>LD</sub> (100 nM) was immobilized on the surface of a streptavidin chip (sensor chip SA) using the “aim for immobilized level wizard” (1300 RU, 5  $\mu$ L/min). A reference surface with Amino PEG biotin (50  $\mu$ M, Polypure) was prepared (time and flow rate: 60 s and 10  $\mu$ L/min, respectively) to correct for unspecific binding events of the glycomimetics with streptavidin on the sample surface. Kinetic experiments were run at 25°C using HBS-EP (0.01 M HEPES pH 7.4, 0.15 M NaCl, 3 mM EDTA, 0.005% surfactant P20, from GE Healthcare) as running buffer at a flow rate of 30  $\mu$ L/min. Since a convenient regeneration condition, which keeps the protein active, was not found, single cycle kinetics (SCKs) without regeneration steps was used. Instead of surface regeneration with chemical agents after each injection, a complete dissociation of the respective compound from the protein was allowed, before a new run was started. Blank injections (HBS-EP running buffer) were conducted under equal conditions in order to apply double referencing. Binding data was evaluated using *Biacore* T200 Evaluation software version 1 (GE Healthcare).

**Isothermal titration calorimetry.** Standard ITC experiments were performed at 25°C using a VP-ITC (Malvern Instruments, Worcestershire, UK) with an injection volume between 3  $\mu$ l and 15  $\mu$ l, a reference power of 10  $\mu$ cal/sec, a stirring speed of 307 rpm, in *high feedback* mode and with a filter period of 2 sec. Preceding the measurements, FimH<sub>LD</sub>-Th-His6 was dialyzed against a 10 mM HEPES buffer adjusted to pH 7.4, containing 150 mM NaCl. Ligand and protein were dissolved in the same buffer. Protein concentration was determined by NanoDrop ND-1000 Spectrophotometer (Thermo Scientific, MA, USA) using an extinction coefficient of 24'180 M<sup>-1</sup> cm<sup>-1</sup>. The active protein concentration was determined by an ITC experiment with FimH<sub>LD</sub> binding to compound 1.<sup>(3)</sup> The thermodynamic parameters  $K_A$  (association constant) and  $\Delta H^\circ$  (change in enthalpy) and the kinetic parameter  $k_{\text{off}}$  (dissociation rate constant) are measured by ITC. All parameters were evaluated using the fully automated analysis software package from AFFINImeter (Software for Science Developments, Santiago de Compostela, Spain).<sup>(4,5)</sup> The parameters  $\Delta G^\circ$  (free energy of binding) and  $\Delta S^\circ$  (change in

entropy) were calculated from equation S1, and  $k_{\text{on}}$  (association rate constant) was calculated according to equation S2:

$$\Delta G^\circ = \Delta H^\circ - T\Delta S^\circ = -RT \ln K_A \quad (\text{eq. S1})$$

$$K_A = \frac{k_{\text{on}}}{k_{\text{off}}} = \frac{1}{K_D} \quad (\text{eq. S2})$$

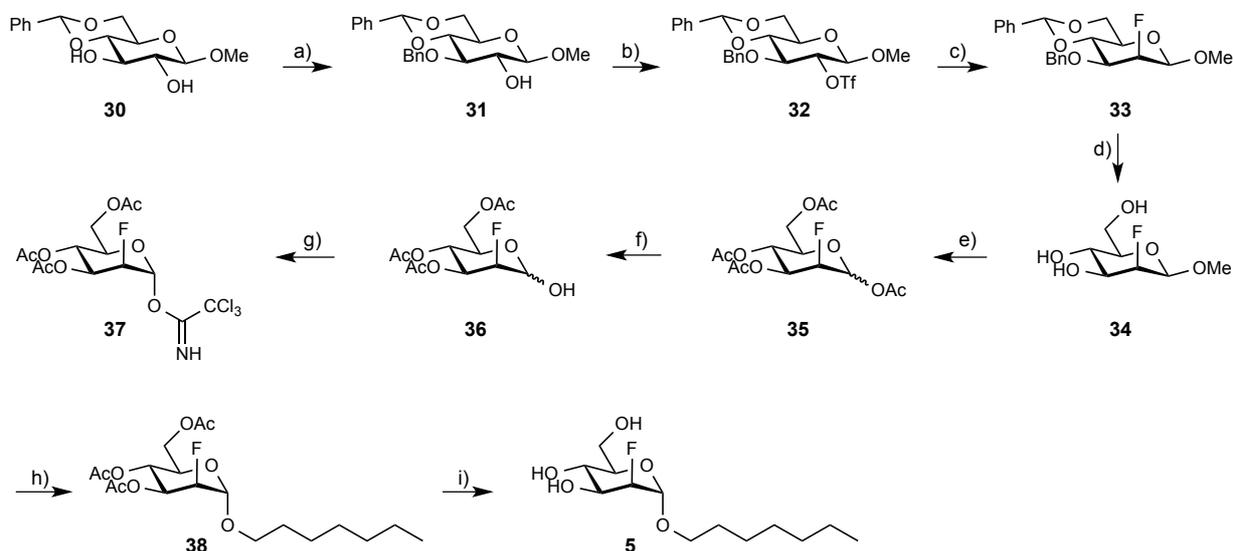
with T being the absolute temperature and R the universal gas constant (8.314 Jmol<sup>-1</sup>K<sup>-1</sup>). The thermodynamic and kinetic parameters of compounds **1-10** and **24-29** were calculated from the average of two independent experiments. The  $c$ -values of compounds **15-21** and **23** were significantly above 1'000 for the direct titrations, therefore additional competitive ITC experiments were performed to achieve valid  $K_D$ -values.<sup>(6)</sup> These compounds were titrated into protein, which was preincubated with excess of the weak binding compound **8**, resulting in sigmoidal titration curves. The experimental conditions and the analysis method with Origin 7.0 (OriginLab, Northampton, MA, USA) and SEDPHAT version 10.4 (National Institute of Health)<sup>(7)</sup> for the competitive experiments of compounds **19** and **23** are described in Fiege *et al.* 2015<sup>(8)</sup> and for compounds **15-18**, **20**, and **21** in Kleeb *et al.* 2015.<sup>(9)</sup> These independently derived  $K_D$ -values from competitive experiments of compounds **15-21** and **23** were manually fixed for the kinetic analysis with AFFINImeter. For the low-affinity compounds **26** and **27**, the  $c$ -value was below 1. To extract reliable thermodynamic and kinetic data from low  $c$ -value experiments the stoichiometry was manually fixed to 1.<sup>(10)</sup> Correlations of the kinetic parameters  $k_{\text{on}}$  (B) and  $k_{\text{off}}$  (C) with the change of free energy of binding ( $\Delta G^\circ$ ), as well as the F-test to check if the slopes of the two subgroups (modified mannose moiety vs. modified aglycone moiety) are identical or different, were calculated with Prism version 5.0c from GraphPad, Inc. (La Jolla, CA, USA).<sup>(11)</sup>

**Crystallization and structure determination of compound 3 and 4 in complex with FimH<sub>LD</sub>.** FimH<sub>LD</sub>/3 and FimH<sub>LD</sub>/4 were crystallized by sitting-drop vapor diffusion at 4°C. FimH<sub>LD</sub> (residues 1–158) was used at a final concentration of 10 mg/mL<sup>-1</sup> (ca. 0.5 mM) with a fivefold molar ligand excess in HEPES buffer (pH 7.4, 20 mM). After several months, plate like crystals appeared in 1.6 M (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 0.1 M Hepes pH 7.5, 1% PEG 3350 (w/v) and were flash-cooled to 100 K after a quick soak in 2.5 M Li<sub>2</sub>SO<sub>4</sub>.<sup>(12)</sup> Data were collected at the PX beamline (X06SA) of the Swiss Light Source (Paul Scherrer Institute, Switzerland), indexed, integrated, and scaled with XDS<sup>(13)</sup>. The structures were solved by molecular replacement with PHASER<sup>(14)</sup> with 4X50.pdb<sup>(8)</sup> as a search model. The models were built in COOT<sup>(15,16)</sup> and refined with the PHENIX software<sup>(17,18)</sup>. Geometric ligand restraints were generated with PRODRG<sup>(19)</sup>. The atomic coordinates and structure factors are deposited in the Protein Data Bank with PDB code 5JCR (**3**) and 5MUC (**4**).

**Molecular Modeling.** Protein-Ligand complexes of the crystal structures (PDB accession codes: 1UWF, 4XO8, 5JCR, 5MUC) were processed with the Protein Preparation Wizard.<sup>(20)</sup> Figures were produced using PyMOL.<sup>(21)</sup>

## Synthesis

**General methods.** Commercially available reagents were purchased from Aldrich, Merck or Alfa Aesar. Methanol was dried by distillation from sodium methoxide. Dichloromethane ( $\text{CH}_2\text{Cl}_2$ ) was dried by filtration through  $\text{Al}_2\text{O}_3$  (Fluka, basic; 0.05-0.15 mm). Toluene was dried by distillation from sodium/benzophenone. Optical rotations were measured at 20 °C on a Perkin Elmer 341 polarimeter with a path length of 1 dm. Concentrations are given in g/100 mL. NMR spectra were obtained on a Bruker Avance 500 UltraShield spectrometer at 500.13 MHz ( $^1\text{H}$ ) or 125.76 MHz ( $^{13}\text{C}$ ). Chemical shifts are given in ppm and were calibrated on residual solvent peaks or to tetramethyl silane as internal standard. Multiplicities are specified as s (singlet), d (doublet), dd (doublet of a doublet), t (triplet), q (quartet), or m (multiplet). Assignment of the  $^1\text{H}$  and  $^{13}\text{C}$  NMR spectra was achieved using 2D methods (COSY, HSQC). ESI mass spectra were recorded on a Waters micromass ZQ instrument. High resolution mass spectra were obtained on an ESI Bruker Daltonics micrOTOF spectrometer equipped with a TOF hexapole detector. Reactions were monitored by TLC using glass plates coated with silica gel 60 F<sub>254</sub> and visualized by using UV light and/or by charring with a molybdate solution (a 0.02 M solution of ammonium cerium sulfate dihydrate and ammonium molybdate tetrahydrate in aqueous 10%  $\text{H}_2\text{SO}_4$ ) with heating to 140 °C for 5 min. Column chromatography was performed on a CombiFlash Companion (ISCO, Inc.) using RediSep normal phase disposable flash columns (silica gel). Reversed phase chromatography was performed on LiChroprep®RP-18 (Merck, 40- 63  $\mu\text{m}$ ).

Synthesis of 5 (*n*-heptyl 2-fluoro-2-deoxy- $\alpha$ -D-mannopyranoside)

**Scheme 1.** a) Dibutyltin oxide, MeOH, reflux, 3 h, then BnBr, CsF, DMF, 40 °C, overnight, (51%); b)  $\text{Tf}_2\text{O}$ , 2,6-di-tertbutyl-4-methylpyridine, DCM, -15 °C-rt, 2.5 h, (50-70%); c) 1M TBAF in THF, 50 °C, 2 h, (78%); d) i. HOAc/Water (V:V 4:1), 80 °C, 1 h, quantitative; ii. Pd/C, MeOH/ $\text{H}_2\text{O}$ , catalytic amount HOAc, 4.8 mbar, hydrogenation, overnight, (91%); e) 4%  $\text{H}_2\text{SO}_4/\text{Ac}_2\text{O}$ , rt, 3 h, (94%); f)  $\text{NH}_2\text{NH}_2 \cdot \text{HOAc}$ , DMF, rt, 3 h, (77%); g)  $\text{Cl}_3\text{CCN}$ , DCM, NaH, 2 h, (60%); h) *n*-Heptanol, TMSOTf, DCM, rt; i)  $\text{CH}_3\text{ONa}/\text{MeOH}$ , rt, 2 h, (two steps 77%).

**Methyl 3-O-benzyl-4,6-O-benzylidene- $\beta$ -D-glucopyranoside (31)<sup>(22)</sup>**

The suspension of **30** (1.998 g, 7.07 mmol), dibutyltin oxide (2.026 g, 8.14 mmol) in dry methanol (30 mL) was refluxing at 80°C for 3 h. The solution was then concentrated to dryness and CsF (1.181 g, 7.78 mmol) and BnBr (0.923 mL, 7.78 mmol), DMF (17 mL) were added. The reaction mixture was stirred at 40°C overnight and then diluted with EE, filtered, the filtrate was washed with sat. NaHCO<sub>3</sub>, brine and dried over Na<sub>2</sub>SO<sub>4</sub>. The solvent was removed under vacuo and the residue was purified by flash chromatography on Silica gel (PE:EA 3:1-3:2) to afford **31** (1.338 g, 51%) as a white fluffy solid. <sup>1</sup>H-NMR (500 MHz, CDCl<sub>3</sub>):  $\delta$  7.49 (m, 2H), 7.41-7.28 (m, 8H), 5.59 (s, 1H, PhCH), 4.98 (d,  $J$  = 11.5 Hz, 1H, OCH<sub>2</sub>Ph), 4.79 (d,  $J$  = 11.5 Hz, 1H, OCH<sub>2</sub>Ph), 4.37 (dd,  $J$  = 10.5, 5.0 Hz, 1H, H-6a), 4.33 (d,  $J$  = 7.5 Hz, 1H, H-1), 3.81 (t,  $J$  = 10.5 Hz, 1H, H-6b), 3.72 (t,  $J$  = 9.5 Hz, 1H, H-4), 3.68 (t,  $J$  = 9.5 Hz, 1H, H-3), 3.59 (s, 3H, OMe), 3.56 (m, 1H, H-2), 3.46 (td,  $J$  = 9.5, 4.5 Hz, 1H, H-5), 2.47 (dd,  $J$  = 7.5, 2.0 Hz, 1H, 2-OH); <sup>13</sup>C-NMR (125 MHz, CDCl<sub>3</sub>):  $\delta$  138.22, 137.18, 129.00, 128.45, 128.26, 128.05, 127.85, 125.98, 104.17 (C-1), 101.24 (PhCH), 81.42 (C-4), 80.18 (C-3), 74.62 (OCH<sub>2</sub>Ph), 74.18 (C-2), 68.69 (C-6), 66.38 (C-5), 57.44 (OCH<sub>3</sub>).

**Methyl 3-O-Benzyl-4,6-O-benzylidene-2-O-(trifluoromethanesulfonyl)- $\beta$ -D-glucopyranoside (32)**

To a solution of **31** (550 mg, 1.477 mmol), 2,6-Di-tert-butyl-4-methylpyridine (606 mg, 2.95 mmol) in dry DCM (4.0 mL) was added trifluoromethanesulfonic anhydride (0.294 mL, 1.77 mmol) slowly at -15°C. The reaction mixture was stirred at -15°C for 0.5 h, then removed the cooling bath and stirred for another 1.5 h at rt. The reaction mixture was diluted with DCM, washed with 5% NaHCO<sub>3</sub>, brine and dried over Na<sub>2</sub>SO<sub>4</sub>. The solvent was removed under vacuo and the residue was purified by flash chromatography on silica gel (PE:EA 15:1-10:1) to afford **32** (527 mg, 71%) as a white solid. <sup>1</sup>H-NMR (500 MHz, CDCl<sub>3</sub>):  $\delta$  7.48 (m, 2H), 7.42-7.29 (m, 8H), 5.58 (s, 1H, PhCH), 4.90 (d,  $J$  = 11.0 Hz, 1H, OCH<sub>2</sub>Ph), 4.79 (d,  $J$  = 10.5 Hz, 1H, OCH<sub>2</sub>Ph), 4.61 (t,  $J$  = 8.5 Hz, 1H, H-2), 4.52 (d,  $J$  = 7.5 Hz, 1H, H-1), 4.40 (dd,  $J$  = 10.5, 5.0 Hz, 1H, H-6a), 3.88 (t,  $J$  = 9.0 Hz, 1H, H-3), 3.81 (t,  $J$  = 10.5 Hz, 1H, H-6b), 3.76 (t,  $J$  = 9.5 Hz, 1H, H-4), 3.58 (s, 3H, OMe), 3.48 (m, 1H, H-5); <sup>13</sup>C-NMR (125 MHz, CDCl<sub>3</sub>):  $\delta$  136.93, 136.71, 129.21, 128.45, 128.35, 128.01, 125.92, 101.38 (PhCH), 101.30 (C-1), 84.59 (C-2), 81.58 (C-4), 77.34 (C-3), 74.95 (OCH<sub>2</sub>Ph), 68.36 (C-6), 66.18 (C-5), 57.70 (OCH<sub>3</sub>); ESI-MS Calcd for [M+H]<sup>+</sup>, 505.11, found 505.13, [M+Na]<sup>+</sup>, 527.10, found 527.05.

**Methyl 3-O-Benzyl-4,6-O-benzylidene-2-fluoro-2-deoxy- $\beta$ -D-mannopyranoside (33)<sup>(23)</sup>**

A solution of **32** (529 mg, 1.047 mmol) in 1M TBAF/THF (5.4 mL) was stirred at 50°C for 2 h. Then concentrated and the residue was purified by flash chromatography on silica gel (PE:EA 5:1-3:1) to afford **33** (307 mg, 78%) as a fluffy white solid. <sup>1</sup>H-NMR (500 MHz, CDCl<sub>3</sub>):  $\delta$  7.51 (m, 2H), 7.42-7.28 (m, 8H), 5.63 (s, 1H, PhCH), 4.87 (d,  $J$  = 12.0 Hz, 1H, OCH<sub>2</sub>Ph), 4.78 (d,  $J$  = 12.0 Hz, 1H, OCH<sub>2</sub>Ph), 4.80 (dd,  $J$  = 50.0, 2.5 Hz, 1H, H-2), 4.43 (d,  $J$  = 19.0 Hz, 1H, H-1), 4.36 (dd,  $J$  = 10.5, 5.0 Hz, 1H, H-6a), 4.11 (td,  $J$  = 12.0, 2.0 Hz, 1H, H-4), 3.92 (t,  $J$  = 10.5 Hz, 1H, H-6b), 3.67 (ddd,  $J$  = 26.5, 9.5, 2.5 Hz, 1H, H-3), 3.57 (s, 3H, OMe), 3.39 (td,  $J$  = 10.0, 5.0 Hz, 1H, H-5); <sup>13</sup>C-NMR (125 MHz, CDCl<sub>3</sub>):  $\delta$  137.61, 137.20, 129.03, 128.48, 128.26, 127.92, 127.90, 126.01, 101.61 (PhCH), 100.42 (d,  $J$  = 15.75 Hz, C-1), 88.47 (d,  $J$  = 188.5 Hz, C-2), 78.32 (d,  $J$  = 2.0 Hz, C-4), 75.56 (d,  $J$  = 16.75 Hz, C-3), 72.69 (OCH<sub>2</sub>Ph), 68.47 (C-6), 67.07 (C-5), 57.62 (OCH<sub>3</sub>); <sup>19</sup>F-NMR (470 MHz, CDCl<sub>3</sub>): -218.677; ESI-MS Calcd for [M+Na]<sup>+</sup>, 397.14, found 397.05.

**1,3,4,6-Tetra-O-acetyl-2-fluoro-2-deoxy- $\alpha$ -D-mannopyranoside (35)<sup>(24)</sup>**

A solution of **33** (306 mg, 0.817 mmol) in 80% aqueous HOAc (7.5 mL) was heated at 80°C for 1 h and then concentrated to dryness. The residue was hydrogenolyzed in methanol/water (V:V 9:1, 10 mL) containing catalytic amount of HOAc in the presence of 5-10% Pd/C (150 mg) under 4.8 bar hydrogen at rt overnight. Then filtered the reaction suspension, the filtrate was concentrated to dryness to provide **34** (147 mg, two steps yield 91%). To a suspension of **34** (147 mg, 0.748 mmol) in acetic anhydride (1.5 mL) was added H<sub>2</sub>SO<sub>4</sub> (95-97%) (60  $\mu$ L) at rt. The reaction mixture was stirred at rt for 3 h. Then diluted with DCM, neutralized carefully with sat. aqueous NaHCO<sub>3</sub> at 0°C. The aqueous layer was extracted with DCM for two times, the combined organic layers were washed with brine and dried over Na<sub>2</sub>SO<sub>4</sub>. Solvent was removed under vacuo and the residue was purified by flash chromatography on silica gel (PE:EA 3:1-7:3) to afford **35** (248 mg, 94%) as a white foam. <sup>1</sup>H-NMR (500 MHz, CDCl<sub>3</sub>):  $\delta$  6.28 (dd,  $J$  = 7.0, 2.0 Hz, 1H, H-1), 5.41 (t,  $J$  = 10.0 Hz, 1H, H-4), 5.26 (ddd,  $J$  = 28.0, 10.0, 2.5 Hz, 1H, H-3), 4.75 (dt,  $J$  = 48.5, 2.5 Hz, 1H, H-2), 4.28 (dd,  $J$  = 12.5, 4.5 Hz, 1H, H-6a), 4.11 (dd,  $J$  = 12.5, 2.5 Hz, 1H, H-6b), 4.04 (m, 1H, H-5), 2.17, 2.11, 2.09, 2.05 (4xs, 4 COCH<sub>3</sub>); <sup>13</sup>C-NMR (125 MHz, CDCl<sub>3</sub>):  $\delta$  170.72, 170.23, 169.28, 168.04 (4xCOCH<sub>3</sub>), 90.08 (d,  $J$  = 30.88 Hz, C-1), 85.88 (d,  $J$  = 180.88 Hz, C-2), 70.69 (C-5), 69.44 (d,  $J$  = 16.88 Hz, C-3), 65.09 (C-4), 61.69 (C-6), 20.83, 20.70, 20.69, 20.58 (4xC, COCH<sub>3</sub>).

**3,4,6-Tri-O-acetyl-2-fluoro-2-deoxy- $\alpha$ -D-mannopyranosyl trichloroacetimidate (37)<sup>(24)</sup>**

To a solution of **35** (260 mg, 0.74 mmol) in dry DMF (4.0 mL) was added hydrazine acetate (82 mg, 0.89 mmol) and the reaction mixture was stirred at rt for 3 h. Diluted with DCM and acetic acid and washed with water, aqueous NaHCO<sub>3</sub>, brine and dried over Na<sub>2</sub>SO<sub>4</sub>. Solvent was removed under vacuo and the residue was purified by flash chromatography on silica gel (PE:EA 3:2-1:1) to afford **36** (176 mg, 77%). To a solution of **36** (176 mg, 0.57 mmol) and trichloroacetonitrile (0.57 mL, 5.7 mmol) in dry DCM (5.0 mL) was added NaH (60%) (4 mg) at rt. The reaction mixture was stirred at rt for 2 h, then Silica gel was added and the solvent was removed *in vacuo*. The residue was purified by flash chromatography on silica gel (PE:EA 6:1-5:1) to afford the **37** (155 mg, 60%) as a colorless syrup. <sup>1</sup>H-NMR (500 MHz, CDCl<sub>3</sub>):  $\delta$  8.23 (s, 1H, NHCCCl<sub>3</sub>), 6.47 (dd,  $J$  = 6.0, 2.0 Hz, 1H, H-1), 5.47 (t,  $J$  = 10.0 Hz, 1H, H-4), 5.32 (ddd,  $J$  = 27.5, 10.0, 2.5 Hz, 1H, H-3), 4.97 (dt,  $J$  = 48.5, 2.5 Hz, 1H, H-2), 4.28 (dd,  $J$  = 12.5, 4.5 Hz, 1H, H-6a), 4.20-4.16 (m, 2H, H-6b, H-5), 2.12, 2.08, 2.07 (3xs, 3 COCH<sub>3</sub>); <sup>13</sup>C-NMR (125 MHz, CDCl<sub>3</sub>):  $\delta$  170.65, 170.08, 169.36 (3xCOCH<sub>3</sub>), 159.75 (NHCCCl<sub>3</sub>), 93.92 (d,  $J$  = 31.38 Hz, C-1), 85.34 (d,  $J$  = 181 Hz, C-2), 71.21 (C-5), 69.56 (d,  $J$  = 16.75 Hz, C-3), 64.92 (C-4), 61.63 (C-6), 20.70, 20.68, 20.61 (3xC, COCH<sub>3</sub>); ESI-MS Calcd for [M+Na]<sup>+</sup>, 473.99, found 473.96.

***n*-Heptyl 3,4,6-tri-O-acetyl-2-fluoro-2-deoxy- $\alpha$ -D-mannopyranoside (38)**

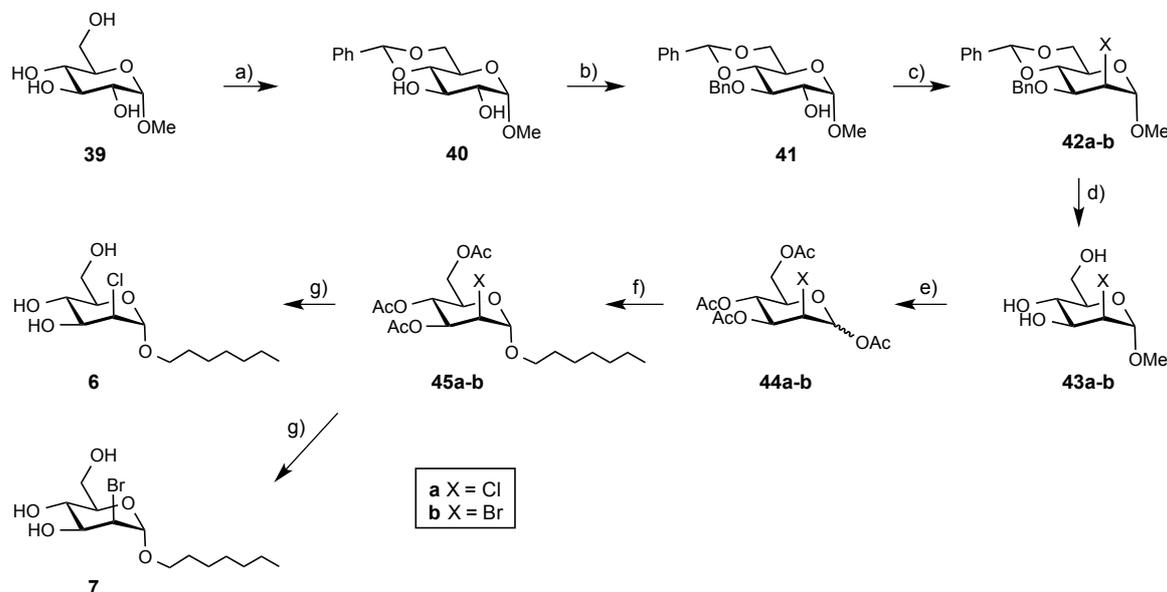
A mixture of **37** (150 mg, 0.33 mmol), 4Å MS and *n*-heptanol (94  $\mu$ L, 0.66 mmol) in dry DCM (4 mL) was stirred at rt for 0.5 h before the addition of TMSOTf (9.0  $\mu$ L, 0.0495 mmol) at rt under Argon. The reaction mixture was stirred under these conditions for 2 h, then neutralized with Et<sub>3</sub>N, and concentrated to dryness. The residue was purified by flash chromatography on silica gel (PE:EA 5:1) to afford **38** (quantitative yield due to containing some *n*-heptanol) as a colorless oil. <sup>1</sup>H-NMR (500 MHz, CDCl<sub>3</sub>):  $\delta$  5.31 (t,  $J$  = 10.0 Hz, 1H, H-4), 5.25 (ddd,  $J$  = 28.0, 10.0, 2.5 Hz, 1H, H-3), 4.98 (dd,  $J$  = 7.0, 1.5 Hz, 1H, H-1), 4.73 (dt,  $J$  = 50.0, 2.5 Hz, 1H, H-2), 4.27 (dd,  $J$  = 12.0, 5.0 Hz, 1H, H-6a), 4.12

(dd,  $J = 12.0, 2.0$  Hz, 1H, H-6b), 3.97 (ddd,  $J = 9.5, 4.5, 2.0$  Hz, 1H, H-5), 3.71 (dt,  $J = 9.5, 6.5$  Hz, 1H, H-OCH<sub>2</sub>C<sub>6</sub>H<sub>13</sub>), 3.47 (dt,  $J = 9.5, 6.5$  Hz, 1H, H-OCH<sub>2</sub>C<sub>6</sub>H<sub>13</sub>), 2.10, 2.09, 2.04 (3xs, 3 COCH<sub>3</sub>), 1.58 (m, 2H), 1.31 (m, 8H), 0.87 (m, 3H, CH<sub>3</sub>); <sup>13</sup>C-NMR (125 MHz, CDCl<sub>3</sub>):  $\delta$  170.75, 170.17, 169.53 (3xCOCH<sub>3</sub>), 97.12 (d,  $J = 28.63$  Hz, C-1), 87.05 (d,  $J = 178.25$  Hz, C-2), 69.99 (d,  $J = 16.5$  Hz, C-3), 68.67 (OCH<sub>2</sub>C<sub>6</sub>H<sub>13</sub>), 68.41 (C-5), 65.91 (C-4), 62.18 (C-6), 32.79, 31.71, 28.98, 25.99, 22.58, 14.06 (OCH<sub>2</sub>C<sub>6</sub>H<sub>13</sub>), 20.78, 20.74, 20.65 (3xC, COCH<sub>3</sub>); HR-MS Calcd. for C<sub>19</sub>H<sub>31</sub>FO<sub>8</sub> [M+Na]<sup>+</sup>, 429.1901, found 429.1905.

### *n*-Heptyl 2-deoxy-2-fluoro- $\alpha$ -D-mannopyranoside (**5**)

To a solution of **38** (134 mg, 0.33 mmol) in dry methanol (4.0 mL) was added 0.5 M CH<sub>3</sub>ONa/MeOH solution (99  $\mu$ L) at rt. The reaction mixture was stirred for 2 h and then neutralized with Amberlyst 15, filtered and the filtrate was concentrated under vacuo. The residue was purified by flash chromatography on silica gel (DCM:MeOH 20:1-16:1) to afford **5** (71.3 mg, two steps yield 77%) as colorless oil.  $[\alpha]_D^{20} +75.4$  (c 0.4, DCM); <sup>1</sup>H-NMR (500 MHz, CDCl<sub>3</sub>):  $\delta$  4.93 (dd,  $J = 7.5, 1.5$  Hz, 1H, H-1), 4.66 (d,  $J = 50.0$  Hz, 1H, H-2), 3.92-3.82 (m, 4H, H-6a, H-6b, H-4, H-3), 3.68 (dt,  $J = 9.5, 6.5$  Hz, 1H, H-OCH<sub>2</sub>C<sub>6</sub>H<sub>13</sub>), 3.59 (m, 1H, H-5), 3.41 (dt,  $J = 9.5, 6.5$  Hz, 1H, H-OCH<sub>2</sub>C<sub>6</sub>H<sub>13</sub>), 4.21, 3.03, 2.10 (3,4,6-OH), 1.56 (m, 2H), 1.27 (m, 8H), 0.88 (t,  $J = 7.0$  Hz, 3H, CH<sub>3</sub>); <sup>13</sup>C-NMR (125 MHz, CDCl<sub>3</sub>):  $\delta$  97.41 (d,  $J = 29.13$  Hz, C-1), 89.71 (d,  $J = 172.63$  Hz, C-2), 71.91 (C-5), 70.82 (d,  $J = 17.5$  Hz, C-3), 68.22 (OCH<sub>2</sub>C<sub>6</sub>H<sub>13</sub>), 67.66 (C-4), 61.73 (C-6), 31.73, 29.30, 29.02, 25.98, 22.59, 14.06 (OCH<sub>2</sub>C<sub>6</sub>H<sub>13</sub>); HR-MS Calcd for C<sub>13</sub>H<sub>25</sub>FO<sub>5</sub> [M+Na]<sup>+</sup>, 303.1584, found 303.1585.

### Synthesis of **6** (*n*-heptyl 2-chloro-2-deoxy- $\alpha$ -D-mannopyranoside) and **7** (*n*-heptyl 2-bromo-2-deoxy- $\alpha$ -D-mannopyranoside)



**Scheme 2.** a) PhCH(OMe)<sub>2</sub>, camphorsulfonic acid, CH<sub>3</sub>CN, 50°C, overnight (81%); b) i. nBu<sub>2</sub>SnO, toluene, reflux, 3h; ii. BnBr, CsF, DMF, rt, overnight (50%); c) i. Tf<sub>2</sub>O, pyridine, DCM, -20°C, 2h; ii. LiCl, NMP, rt, 2 days for **42a**; TBAB, DMF, 60°C, 2 days for **42b**; d) i. 80% AcOH/H<sub>2</sub>O, 60°C; ii. 10% Pd/C, H<sub>2</sub> (g), MeOH/EtOAc/DCM (3:1:1), rt, 1h, **43a-b** (90-94%); e) conc. H<sub>2</sub>SO<sub>4</sub>, Ac<sub>2</sub>O, rt, 3h, **44a-b** (98% to

quant.); f) i.  $\text{NH}_4\text{OAc}$ , DMF, rt, overnight; ii. NaH, trichloroacetonitrile, DCM, rt, 2h; iii. *n*-heptanol, TMSOTf, DCM, 0°C to rt, 6h, **45a-b** (70-80%); g) NaOMe, MeOH, rt, 4h, **6** and **7** (67-70%).

### General procedure A for the preparation of mannosides **45a-b**:

To a solution of **44a-b** (1.0 equiv) in DMF (2 mL),  $\text{NH}_4\text{OAc}$  (2.0 equiv) was added. The mixture was stirred at RT overnight. The reaction mixture was extracted with EtOAc (10 mL) and washed with water (10 mL). The organic layer was dried over  $\text{Na}_2\text{SO}_4$  and concentrated in vacuo. The residue was dissolved with DCM (2 mL), and added with 60% NaH in mineral oil (0.15 equiv) and trichloroacetonitrile (10 equiv) at RT. The reaction mixture was stirred at RT for 1.5h. The mixture was concentrated and dried in vacuo to give the crude trichloroacetimidate donor. To a suspension of the crude trichloroacetimidate, *n*-heptanol (2 equiv), and molecular sieves (4 Å, 600 mg) in dry DCM (5 mL), TMSOTf (0.15 equiv) was added dropwise under argon. The mixture was stirred at RT for 5h, then filtered over Celite and concentrated. The residue was purified by MPLC on silica gel (PE/EtOAc) to yield **45a-b**.

### General procedure B for deacetylation:

To a solution of **45a-b**, (1.0 equiv) in dry MeOH (5 mL) was added freshly prepared 1M NaOMe/MeOH (0.1 equiv) under argon. The mixture was stirred at RT until the reaction was complete (monitored by TLC), then neutralized with Amberlyst-15 ( $\text{H}^+$ ) ion-exchange resin, filtered and concentrated in vacuo. The residue was purified by MPLC on silica gel (DCM/MeOH 10:1-8:1) to afford **6** and **7**.

### Methyl 4,6-*O*-benzylidene- $\alpha$ -D-glucopyranoside (**40**).

Prepared according to the similar procedure as described in literature.  $^1\text{H}$  NMR (500 MHz,  $\text{CDCl}_3$ ):  $\delta$  7.49 (d,  $J = 5.6$  Hz, 2H, Ph), 7.37 (d,  $J = 4.7$  Hz, 3H, Ph), 5.53 (s, 1H,  $\text{PhCHO}_2$ ), 4.79 (d,  $J = 2.9$  Hz, 1H, H-1), 4.29 (dd,  $J = 9.8, 4.3$  Hz, 1H, H-6a), 3.93 (t,  $J = 9.2$  Hz, 1H, H-3), 3.86 - 3.77 (m, 2H, H-5 and H-6b), 3.74 (t,  $J = 10.2$  Hz, 1H, H-4), 3.63 (t,  $J = 7.5$  Hz, 1H, H-2), 3.53 - 3.43 (m, 4H, H-4,  $-\text{OCH}_3$ ), 2.77 (s, 1H, OH), 2.31 (d,  $J = 9.4$  Hz, 1H, OH). The proton NMR was consistent with literature data.<sup>(25)</sup>

### Methyl 3-*O*-benzyl-4,6-*O*-benzylidene- $\alpha$ -D-glucopyranoside (**41**).

Prepared according to the similar procedure as described in literature.  $^1\text{H}$  NMR (500 MHz,  $\text{CDCl}_3$ ):  $\delta$  7.49 (d,  $J = 5.8$  Hz, 2H, Ph), 7.42 - 7.28 (m, 8H, Ph), 5.51 (s, 1H,  $\text{PhCHO}_2$ ), 4.78 (d,  $J = 12.2$  Hz, 1H,  $\text{PhCHH}$ ), 4.70 (d,  $J = 12.2$  Hz, 1H,  $\text{PhCHH}$ ), 4.61 (d,  $J = 3.3$  Hz, 1H, H-1), 4.26 (dd,  $J = 10.1, 4.8$  Hz, 1H, H-6a), 4.15 (t,  $J = 9.3$  Hz, 1H, H-3), 3.81 (td,  $J = 9.9, 4.8$  Hz, 1H, H-5), 3.70 (t,  $J = 10.3$  Hz, 1H, H-6b), 3.54 - 3.44 (m, 2H, H-2, H-4), 3.37 (s, 3H,  $\text{OCH}_3$ ). The proton NMR was consistent with literature data.<sup>(26)</sup>

### Methyl 2-deoxy-2-chloro- $\alpha$ -D-mannopyranoside (**43a**):

According to the similar procedure for the preparation of **33**, **42a** was prepared by treatment of the triflate intermediate (478 mg, 0.95 mmol) with LiCl (201 mg, 4.75 mmol) in *N*-methyl-2-pyrrolidone (2 mL) at RT overnight. The reaction mixture was then diluted with EtOAc (20 mL), and washed with water (20 mL) and brine (20 mL). The organic layer was dried over  $\text{Na}_2\text{SO}_4$ , concentrated and purified by MPLC (PE/EtOAc 8:1) to give crude **42a** (393 mg). The crude **42a** was treated with 80% AcOH in water (5 mL) at 60°C for 1h, and then concentrated in vacuo. The residue was dissolved in a mixed solvent of MeOH/EtOAc/DCM (3:1:1), added with 10% Pd/C (60 mg), and

stirred under hydrogen atmosphere at RT for 1h. The reaction mixture was filtered and concentrated. The residue was purified by MPLC (DCM/MeOH 9:1) to yield **43a** (60 mg, 32% over two steps) as colorless oil.  $[\alpha]_D = +80.22$  ( $c = 1.5$ , MeOH);  $^1\text{H NMR}$  (500 MHz, MeOD):  $\delta$  4.85 (s, 1H, H-1), 4.22 (d,  $J = 2.0$  Hz, 1H, H-2), 3.99 (dd,  $J = 9.1$ , 3.3 Hz, 1H, H-3), 3.87 (d,  $J = 11.8$  Hz, 1H, H-6a), 3.71 (dd,  $J = 11.8$ , 6.1 Hz, 1H, H-6b), 3.66 (t,  $J = 9.5$  Hz, 1H, H-4), 3.60 - 3.53 (m, 1H, H-5), 3.42 (s, 3H);  $^{13}\text{C NMR}$  (126 MHz, MeOD):  $\delta$  102.43(C-1), 75.22 (C-5), 71.00 (C-3), 68.30 (C-4), 63.00 (C-2), 62.84 (C-6), 55.58 (OCH<sub>3</sub>); ESI-MS:  $m/z$ : calcd for C<sub>7</sub>H<sub>13</sub>ClNaO<sub>5</sub> [M+Na]<sup>+</sup>: 235.03, found: 234.76.

#### **Methyl 2-deoxy-2-bromo- $\alpha$ -D-mannopyranoside (43b):**

According to the similar procedure for the preparation of **33**, **42b** was prepared by treatment of the triflate intermediate (540 mg, 1.07 mmol) with nBu<sub>4</sub>NBr (700 mg, 2.14 mmol) in DMF (5 mL) at 60°C for 48h. The reaction mixture was then diluted with EtOAc (50 mL), and washed with water (50 mL) and brine (50 mL). The organic layer was dried over Na<sub>2</sub>SO<sub>4</sub>, concentrated and purified by MPLC (PE/EtOAc 5:1) to give the crude **42b** (280 mg). The crude product **42b** was treated with 80% AcOH in water (5 mL) at 60°C for 1h, and then concentrated in vacuo. The residue was dissolved in a mixed solvent of MeOH/EtOAc/DCM (3:1:1), added with 10% Pd/C (150 mg), and stirred under hydrogen atmosphere at RT for 1h. The reaction mixture was filtered and concentrated. The residue was purified by MPLC (DCM/MeOH 10:1) to yield **43b** (150 mg, 55% over two steps) as colorless oil.  $[\alpha]_D = +53.47$  ( $c = 0.8$ , EtOAc);  $^1\text{H NMR}$  (500 MHz, MeOD):  $\delta$  4.96 (s, 1H), 4.31 (dd,  $J = 3.8$ , 1.3 Hz, 1H, H-2), 3.87 (dd,  $J = 11.8$ , 2.3 Hz, 1H, H-6a), 3.82 (dd,  $J = 9.0$ , 3.9 Hz, 1H, H-3), 3.71 (dd,  $J = 11.8$ , 6.1 Hz, 1H, H-6b), 3.66 (t,  $J = 9.4$  Hz, 1H, H-4), 3.62 - 3.57 (m, 1H, H-5), 3.42 (s, 3H, CH<sub>3</sub>);  $^{13}\text{C NMR}$  (126 MHz, MeOD):  $\delta$  102.65 (C-1), 75.35 (C-5), 70.44 (C-3), 69.12 (C-4), 62.88 (C-6), 56.37 (C-6), 55.60 (C-2); ESI-MS:  $m/z$ : calcd for C<sub>7</sub>H<sub>13</sub>BrNaO<sub>5</sub> [M+Na]<sup>+</sup>: 278.98, found 278.79.

#### **2-Deoxy-2-chloro-1,3,4,6-tetra-*O*-acetyl- $\alpha$ -D-mannopyranoside (44a):**

**43a** (20 mg, 0.09 mmol) was mixed with acetic anhydride (0.25 mL) and conc. H<sub>2</sub>SO<sub>4</sub> (7.5  $\mu$ l, 0.14 mmol). The reaction mixture was stirred at RT overnight, and then diluted with DCM (10 mL). After being washed with sat. NaHCO<sub>3</sub> aq. solution (10 mL) and brine (10 mL), the organic layer was dried over Na<sub>2</sub>SO<sub>4</sub> and concentrated in vacuo. The residue was purified by MPLC (PE/EtOAc 2:1) to yield **44a** (34 mg, 99%) as colorless oil.  $[\alpha]_D = +27.51$  ( $c = 1.7$ , EtOAc);  $^1\text{H NMR}$  (500 MHz, CDCl<sub>3</sub>):  $\delta$  6.24 (s, 1H, H-1), 5.48 (t,  $J = 9.9$  Hz, 1H, H-4), 5.36 (dd,  $J = 9.8$ , 3.1 Hz, 1H, H-3), 4.40 (s, 1H, H-2), 4.25 (dd,  $J = 12.4$ , 4.3 Hz, 1H, H-6a), 4.14 (d,  $J = 12.7$  Hz, 1H, H-6b), 4.09 (d,  $J = 10.0$  Hz, 1H, H-5), 2.18, 2.11, 2.10, 2.07 (4 s, 12H, 4 COOCH<sub>3</sub>);  $^{13}\text{C NMR}$  (126 MHz, CDCl<sub>3</sub>):  $\delta$  170.68, 170.12, 169.27, 168.10 (4 CO), 92.91 (C-1), 71.19 (C-5), 69.24 (C-3), 64.91 (C-4), 61.88 (C-6), 56.36 (C-2), 20.87, 20.71, 20.69, 20.61 (4C, COCH<sub>3</sub>). ESI-MS:  $m/z$ : calcd for C<sub>14</sub>H<sub>19</sub>ClNaO<sub>9</sub> [M+Na]<sup>+</sup>: 389.06, found: 388.97.

#### **2-Deoxy-2-bromo-1,3,4,6-tetra-*O*-acetyl- $\alpha$ -D-mannopyranoside (44b):**

Prepared according to the similar procedure of **44a** from **43b** (100mg, 0.39 mmol). Yield: 136 mg (85%).  $[\alpha]_D = +20.13$  ( $c = 0.9$ , EtOAc);  $^1\text{H NMR}$  (500 MHz, CDCl<sub>3</sub>):  $\delta$  6.32 (d,  $J = 1.7$  Hz, 1H, H-1), 5.49 (t,  $J = 9.9$  Hz, 1H, H-4), 5.20 (dd,  $J = 9.7$ , 4.0 Hz, 1H, H-3), 4.44 (dd,  $J = 4.0$ , 1.7 Hz, 1H, H-2), 4.24 (dd,  $J = 12.4$ , 4.6 Hz, 1H, H-6a), 4.16 (dd,  $J = 12.4$ , 2.4 Hz, 1H, H-6b), 4.10 (ddd,  $J = 10.0$ , 4.5, 2.6 Hz, 1H, H-5), 2.17, 2.11, 2.11, 2.07 (4 s, 12H, 4 COOCH<sub>3</sub>);  $^{13}\text{C NMR}$  (126 MHz, CDCl<sub>3</sub>):  $\delta$  170.70, 170.07, 169.28, 168.12

(4 CO), 93.19 (C-1), 71.33 (C-5), 68.82 (C-3), 65.63 (C-4), 61.89 (C-6), 47.82 (C-2), 20.88, 20.77, 20.71, 20.62 (4C, COCH<sub>3</sub>). ESI-MS: *m/z*: calcd for C<sub>14</sub>H<sub>19</sub>BrNaO<sub>9</sub> [M+Na]<sup>+</sup>: 433.01, found: 432.91.

**Heptyl 2-deoxy-2-chloro-3,4,6-tetra-O-acetyl- $\alpha$ -D-mannopyranoside (45a):**

Prepared according to general procedure A from **44a** (60 mg, 0.28mmol). Yield: 42 mg (35%) as colorless oil.  $[\alpha]_D = +35.0$  ( $c = 1.3$ , EtOAc); <sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>):  $\delta$  5.43 - 5.34 (m, 2H, H-3, H-4), 4.96 (s, 1H, H-1), 4.38 (s, 1H, H-2), 4.25 (dd,  $J = 12.2, 4.9$  Hz, 1H, H-6a), 4.14 (d,  $J = 12.1$  Hz, 1H, H-6b), 3.99 (s, 1H, H-5), 3.68 (dd,  $J = 16.3, 6.9$  Hz, 1H, -OCH- of heptyl), 3.48 (dd,  $J = 16.1, 6.7$  Hz, 1H, -OCH- of heptyl), 2.10, 2.09, 2.05 (3 s, 9H, 3COOCH<sub>3</sub>), 1.67 - 1.55 (m, 2H, CH<sub>2</sub>), 1.30 (d,  $J = 12.6$  Hz, 8H, 4 CH<sub>2</sub>), 0.89 (d,  $J = 6.9$  Hz, 3H, CH<sub>3</sub>); <sup>13</sup>C NMR (126 MHz, CDCl<sub>3</sub>):  $\delta$  170.72, 170.07, 169.50 (3 CO), 99.72 (C-1), 69.77 (C-3), 68.88 (CH<sub>2</sub>), 68.83 (C-5), 65.60 (C-4), 62.34 (C-6), 57.70 (C-2), 31.74, 29.28, 29.00, 26.03, 22.60 (5C, CH<sub>2</sub>), 20.79, 20.73, 20.67 (3C, COCH<sub>3</sub>), 14.07 (CH<sub>3</sub>); ESI-MS: *m/z*: calcd for C<sub>19</sub>H<sub>31</sub>ClNaO<sub>8</sub> [M+Na]<sup>+</sup>: 445.16, found: 445.12.

**Heptyl 2-deoxy-2-bromo-3,4,6-tetra-O-acetyl- $\alpha$ -D-mannopyranoside (45b):**

Prepared according to general procedure A from **44b** (207 mg, 0.52 mmol). Yield: 60 mg (26%) as colorless oil.  $[\alpha]_D = + 38.2$  ( $c = 0.4$ , EtOAc); <sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>):  $\delta$  5.40 (t,  $J = 9.9$  Hz, 1H, H-4), 5.23 (dd,  $J = 9.7, 4.0$  Hz, 1H, H-3), 5.06 (s, 1H, H-1), 4.44 (dd,  $J = 4.0, 1.3$  Hz, 1H, H-2), 4.24 (dd,  $J = 12.2, 5.0$  Hz, 1H, H-6a), 4.14 (dd,  $J = 12.2, 2.4$  Hz, 1H, H-6b), 4.00 (ddd,  $J = 10.0, 4.9, 2.4$  Hz, 1H, H-5), 3.68 (dt,  $J = 9.6, 6.8$  Hz, 1H, -OCH-), 3.47 (dt,  $J = 9.6, 6.6$  Hz, 1H, -OCH-), 2.11, 2.09, 2.05 (3 s, 9H, 3 COOCH<sub>3</sub>), 1.60 (dd,  $J = 14.1, 6.9$  Hz, 2H, CH<sub>2</sub>), 1.37 - 1.27 (m, 8H, 4 CH<sub>2</sub>), 0.89 (t,  $J = 6.9$  Hz, 3H, CH<sub>3</sub>); <sup>13</sup>C NMR (126 MHz, CDCl<sub>3</sub>):  $\delta$  170.72, 170.01, 169.50 (3 CO), 99.95 ( $J = 9.43$  Hz, C-1), 69.34 (C-3), 68.98 (C-5), 68.87 (OCH<sub>2</sub>), 66.28 (C-4), 62.34 (C-6), 49.71 (C-2), 31.74, 29.31, 29.01, 26.04, 22.60 (5C, 5 CH<sub>2</sub>), 20.83, 20.74, 20.67 (3C, COCH<sub>3</sub>), 14.07 (CH<sub>3</sub>); ESI-MS: *m/z*: calcd for C<sub>19</sub>H<sub>31</sub>BrNaO<sub>8</sub> [M+Na]<sup>+</sup>: 489.11, found: 489.04.

**Heptyl 2-deoxy-2-chloro- $\alpha$ -D-mannopyranoside (6):**

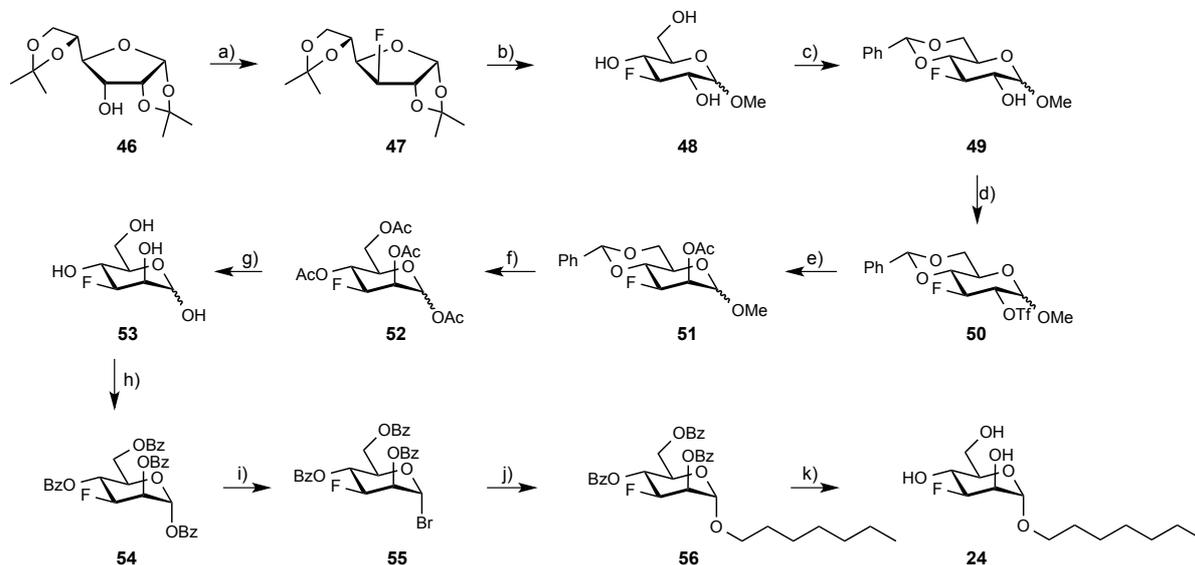
Prepared according to general procedure B from **45a** (30 mg, 0.07 mmol). Yield: 14 mg (67%) as a white solid.  $[\alpha]_D = + 60.9$  ( $c = 0.6$ , MeOH); <sup>1</sup>H NMR (500 MHz, MeOD):  $\delta$  4.94 (s, 1H, H-1), 4.22 (s, 1H, H-2), 4.02 (dd,  $J = 8.9, 2.9$  Hz, 1H, H-3), 3.85 (d,  $J = 11.7$  Hz, 1H, H-6a), 3.77 (dd,  $J = 15.8, 6.9$  Hz, 1H, -OCH- of heptyl), 3.74 - 3.63 (m, 2H, H-4, H-6b), 3.63 - 3.56 (m, 1H, H-5), 3.47 (dd,  $J = 15.6, 6.5$  Hz, 1H, -OCH- of heptyl), 1.62 (d,  $J = 6.7$  Hz, 2H, CH<sub>2</sub>), 1.45 - 1.28 (m, 8H, 4 CH<sub>2</sub>), 0.93 (t,  $J = 6.3$  Hz, 3H, CH<sub>3</sub>); <sup>13</sup>C NMR (126 MHz, MeOD):  $\delta$  101.31 (C-1), 75.33 (C-5), 71.09 (C-3), 69.10 (CH<sub>2</sub>), 68.35 (C-4), 63.24 (C-6), 62.86 (C-2), 32.99, 30.62, 30.25, 27.30, 23.70 (5 CH<sub>2</sub>), 14.44 (CH<sub>3</sub>); HR-MS: *m/z*: calcd for C<sub>13</sub>H<sub>25</sub>ClNaO<sub>5</sub> [M+Na]<sup>+</sup>: 319.1283, found: 319.1285.

**Heptyl 2-deoxy-2-bromo- $\alpha$ -D-mannopyranoside (7):**

Prepared according to general procedure B from **45b** (60 mg, 0.13 mmol). Yield: 26 mg (60 %) as a white solid.  $[\alpha]_D = +67.5$  ( $c = 0.8$ , MeOH); <sup>1</sup>H NMR (500 MHz, MeOD):  $\delta$  5.05 (s, 1H, H-1), 4.31 (dd,  $J = 3.8, 1.3$  Hz, 1H, H-2), 3.88 - 3.82 (m, 2H, H-3, H-6a), 3.76 (dt,  $J = 9.5, 6.7$  Hz, 1H, -OCH-), 3.73 - 3.60 (m, 3H, H-4, H-5, H-6b), 3.47 (dt,  $J = 9.5, 6.4$  Hz, 1H, -OCH-), 1.67 - 1.58 (m, 2H, CH<sub>2</sub>), 1.44 - 1.30 (m, 8H, 4 CH<sub>2</sub>), 0.93 (t,  $J = 7.0$  Hz, 3H, CH<sub>3</sub>); <sup>13</sup>C NMR (126 MHz, MeOD):  $\delta$  101.54 (C-1), 75.46 (C-5), 70.53 (C-

3), 69.18 (C-4), 69.15 (OCH<sub>2</sub>), 62.90 (C-6), 56.66 (C-2), 33.00, 30.66, 30.25, 27.30, 23.70 (5 CH<sub>2</sub>), 14.45 (CH<sub>3</sub>); HR-MS: *m/z*: calcd for C<sub>13</sub>H<sub>25</sub>BrNaO<sub>5</sub> [M+Na]<sup>+</sup>: 363.0778, found: 363.0780.

### Synthesis of 24 (*n*-heptyl 3-deoxy-3-fluoro- $\alpha$ -D-mannopyranoside)



**Scheme 3.** a) DAST, DCM, pyr, (46-53%); b) acetyl chloride, MeOH-benzene, reflux, quantitative; c) benzaldehyde dimethylacetal, *p*-TsOH, CH<sub>3</sub>CN, reflux, 1 h, (69%); d) Tf<sub>2</sub>O, pyr, DCM, -15°C-rt, 1.5 h, (96%); e) CsOAc, 18-crown-6, toluene, reflux, 2 h, (92%); f) H<sub>2</sub>SO<sub>4</sub> / Ac<sub>2</sub>O, (92%); g) CH<sub>3</sub>ONa/CH<sub>3</sub>OH, rt, quantitative; h) BzCl, pyr, DMAP, (94%); i) HBr/HOAc, DCM, (58-81%); j) HgBr<sub>2</sub>, Hg(CN)<sub>2</sub>, *n*-heptanol, DCM, (27%); k) CH<sub>3</sub>ONa/CH<sub>3</sub>OH, rt, (86%).

### 3-Deoxy-3-fluoro-1,2:5,6-di-O-isopropylidene- $\alpha$ -D-glucopyranose (47)

To a solution of 1,2:5,6-di-O-isopropylidene- $\alpha$ -D-allofuranose **46** (1 g, 3.84 mmol) and pyridine (1 mL, 12.29 mmol) in dry DCM (7 mL) was added DAST (95%) (0.79 mL, 6.14 mmol) dropwise at 0°C under argon. The reaction mixture was stirred at 0°C for 0.5 h and stirred at rt overnight. Then diluted with DCM, quenched with NaHCO<sub>3</sub> carefully at 0°C, the aqueous layer was extracted with DCM for 2 times, the combined organic layers were washed with water, brine and dried over Na<sub>2</sub>SO<sub>4</sub>. The solvent was removed under vacuo and the residue was purified by flash chromatography on silica gel (PE:EA 10:1-9:1) to afford **47** (532 mg, 53%) as a colorless oil. <sup>1</sup>H-NMR (500 MHz, CDCl<sub>3</sub>):  $\delta$  5.94 (d, *J* = 4.0 Hz, 1H, H-1), 5.0 (dd, *J* = 49.5, 2.0 Hz, 1H, H-3), 4.69 (dd, *J* = 11.0, 4.0 Hz, 1H, H-2), 4.28 (m, 1H, H-5), 4.14-4.06 (m, 2H, H-6a, H-4), 4.03 (dd, *J* = 9.0, 5.0 Hz, 1H, H-6b), 1.49, 1.44, 1.36, 1.32 (4xs, 4CH<sub>3</sub>); <sup>13</sup>C-NMR (125 MHz, CDCl<sub>3</sub>):  $\delta$  112.35, 109.49 (2xacetal), 105.16 (C-1), 93.79 (d, *J* = 182.75 Hz, C-3), 82.51 (d, *J* = 32.63 Hz, C-2), 80.62 (d, *J* = 18.88 Hz, C-4), 71.86 (d, *J* = 7.13 Hz, C-5), 67.15 (d, *J* = 0.75 Hz, C-6), 26.83, 26.67, 26.15, 25.13 (4xC, CH<sub>3</sub>); <sup>19</sup>F-NMR (470 MHz, CDCl<sub>3</sub>):  $\delta$  -207.60.

### Methyl 3-deoxy-3-fluoro-D-glucoside (48)

Acetyl chloride (6.5 mL) was added to a solution of 3-deoxy-3-fluoro-1,2:5,6-di-O-isopropylidene- $\alpha$ -D-glucopyranose (1.28 g) in a mixture of MeOH (45 mL) and benzene (18 mL) at 0°C. The solution was heated overnight under reflux with the exclusion of atmospheric moisture. Solvent was removed under *vacuo* and the residue was purified by

flash chromatography on Silica gel (DCM:MeOH 10:1-8:1) to afford **48** as unseparated  $\alpha/\beta$  anomers in quantitative yield. ESI-MS Calcd for  $[M+Na]^+$ , 219.06, found 218.63.

#### Methyl 3-deoxy-3-fluoro-4,6-O-benzylidene-D-glucopyranoside (**49**)

A mixture of **48** (1.04 g, 5.3 mmol) and benzaldehyde dimethyl acetal (3.17 mL, 21.2 mmol) in dry acetonitrile (10 mL) was heated to reflux and *p*-toluenesulfonic acid (60 mg) was added. The reaction mixture was refluxed for 1 h and then neutralized with triethylamine. The reaction mixture was then concentrated to dryness and the residue was purified by flash chromatography on Silica gel (PE:EE 3:1-3:2) to afford **49** (1.04 g, 69%) ( $\alpha:\beta = 2.2:1$ , by NMR) as a white solid.  $\alpha$  isomer:  $^1\text{H-NMR}$  (500 MHz,  $\text{CDCl}_3$ ):  $\delta$  7.51 (m, 2H), 7.37 (m, 3H), 5.55 (s, 1H, PhCH), 4.83 (t,  $J = 3.5$  Hz, 1H, H-1), 4.69 (dt,  $J = 54.5, 9.0$  Hz, 1H, H-3), 4.32 (m, 1H, H-6a), 3.87-3.69 (m, 4H, H-2, H-4, H-6b, H-5), 3.45 (s, 3H, OMe), 2.51 (m, 1H, 2-OH);  $^{13}\text{C-NMR}$  (125 MHz,  $\text{CDCl}_3$ ):  $\delta$  136.70, 129.18, 128.25, 126.17, 101.54 (PhCH), 99.99 (d,  $J = 9.88$  Hz, C-1), 91.66 (d,  $J = 185.5$  Hz, C-3), 79.12 (d,  $J = 16.88$  Hz, C-4), 71.42 (d,  $J = 17.88$  Hz, C-2), 68.69 (d,  $J = 0.63$  Hz, C-6), 55.61 (OCH<sub>3</sub>);  $\beta$  isomer:  $^1\text{H-NMR}$  (500 MHz,  $\text{CDCl}_3$ ):  $\delta$  7.51 (m, 2H), 7.37 (m, 3H), 5.56 (s, 1H, PhCH), 4.62 (dt,  $J = 54.5, 9.0$  Hz, 1H, H-3), 4.39 (ddd,  $J = 10.5, 5.0, 2.0$  Hz, 1H, H-6a), 4.32 (m, 1H, H-1), 3.87-3.69 (m, 3H, H-2, H-4, H-6b), 3.59 (s, 3H, OMe), 3.43 (m, 1H, H-5), 2.88 (m, 1H, 2-OH);  $^{13}\text{C-NMR}$  (125 MHz,  $\text{CDCl}_3$ ):  $\delta$  136.60, 129.21, 128.26, 126.14, 103.75 (d,  $J = 10.75$  Hz, C-1), 101.58 (PhCH), 92.10 (d,  $J = 186.5$  Hz, C-3), 78.80 (d,  $J = 16.88$  Hz, C-4), 73.34 (d,  $J = 17.88$  Hz, C-2), 68.42 (d,  $J = 1.38$  Hz, C-6), 57.63 (OCH<sub>3</sub>); ESI-MS Calcd for  $\text{C}_{14}\text{H}_{17}\text{FO}_5$   $[M+Na]^+$ , 307.0958, found 307.0957.

#### Methyl 3-deoxy-3-fluoro-4,6-O-benzylidene-2-O-(trifluoromethanesulfonyl)-D-glucopyranoside (**50**)

To a solution of **49** (970 mg, 3.41 mmol) and pyridine (0.55 mL, 6.82 mmol) in dry DCM (10 mL) was added trifluoromethanesulfonic anhydride (0.68 mL, 4.09 mmol) slowly at  $-15^\circ\text{C}$ . The reaction mixture was stirred at  $-15^\circ\text{C}$  for 0.5 h, then removed the cooling bath and stirred for another 1.5 h at rt. The reaction mixture was diluted with DCM, washed with 5%  $\text{NaHCO}_3$ , brine and dried over  $\text{Na}_2\text{SO}_4$ . Solvent was removed under vacuo and the residue was purified by flash chromatography on silica gel (PE:EA 15:1-10:1) to afford **50** ( $\alpha, \beta$  isomers, 1.37 g, 96%) as a pale yellow solid.  $^1\text{H-NMR}$  (500 MHz,  $\text{CDCl}_3$ ):  $\delta$  7.49 (m, 2H), 7.39 (m, 3H), 5.57 (s, 1H, PhCH), 5.02 (t,  $J = 3.5$  Hz, 1H, H-1), 5.00 (dt,  $J = 54.0, 9.0$  Hz, 1H, H-3), 4.82 (ddd,  $J = 12.5, 9.0, 3.5$  Hz, 1H, H-2), 4.35 (ddd,  $J = 10.5, 4.5, 2.0$  Hz, 1H, H-6a), 3.90 (m, 1H, H-5), 3.81 (m, 2H, H-6b, H-4), 3.49 (s, 3H, OMe);  $^{13}\text{C-NMR}$  (125 MHz,  $\text{CDCl}_3$ ):  $\delta$  136.27, 129.42, 128.37, 126.14, 101.83 (PhCH), 97.90 (d,  $J = 7.88$  Hz, C-1), 87.05 (d,  $J = 192.13$  Hz, C-3), 82.07 (d,  $J = 17.38$  Hz, C-2), 79.10 (d,  $J = 16.88$  Hz, C-4), 68.46 (C-6), 61.53 (d,  $J = 7.0$  Hz, C-5), 56.05 (OCH<sub>3</sub>); ESI-MS Calcd for  $[M+Na]^+$ , 439.0, found 438.85, Calcd for  $[M+H]^+$ , 417.06, found 416.98.

#### Methyl 2-O-acetyl-3-deoxy-3-fluoro-4,6-O-benzylidene-D-mannopyranoside (**51**)

A mixture of **50** (1.327 g, 3.187 mmol), cesium acetate (918 mg, 4.78 mmol) and 18-crown-6 (1.263 g, 4.78 mmol) in dry toluene (20 mL) was refluxing for 2.5 h. After cooling down to rt, toluene was removed under vacuo and the residue was redissolved in DCM and washed with water, brine and dried over  $\text{Na}_2\text{SO}_4$ . Solvent was removed in vacuo and the residue was purified by flash chromatography on silica gel (PE:EA 5:1-4:1) to afford **51** ( $\alpha, \beta$  isomers, 0.955 g, 92%) as white solid.  $\alpha$  isomer:  $^1\text{H-NMR}$  (500 MHz,

CDCl<sub>3</sub>):  $\delta$  7.51 (m, 2H), 7.37 (m, 3H), 5.64 (s, 1H, PhCH), 5.19 (ddd,  $J = 5.5, 4.0, 1.5$  Hz, 1H, H-2), 4.98 (ddd,  $J = 49.5, 10.0, 4.0$  Hz, 1H, H-3), 4.74 (dd,  $J = 4.0, 1.5$  Hz, 1H, H-1), 4.32 (m, 1H, H-6a), 4.18 (q,  $J = 10.0$  Hz, 1H, H-4), 3.89 (t,  $J = 10.0$  Hz, 1H, H-6b), 3.84 (m, 1H, H-5), 3.40 (s, 3H, OMe), 2.18 (s, 3H, COCH<sub>3</sub>); <sup>13</sup>C-NMR (125 MHz, CDCl<sub>3</sub>):  $\delta$  169.89 (COCH<sub>3</sub>), 136.85, 129.22, 128.32, 126.17, 101.91 (PhCH), 99.74 (d,  $J = 6.38$  Hz, C-1), 86.18 (d,  $J = 191.75$  Hz, C-3), 77.02 (d,  $J = 22.5$  Hz, C-4), 70.05 (d,  $J = 15.75$  Hz, C-2), 68.58 (d,  $J = 1.3$  Hz, C-6), 62.88 (d,  $J = 7.5$  Hz, C-5), 55.32 (OCH<sub>3</sub>), 20.88 (COCH<sub>3</sub>); <sup>19</sup>F-NMR (470 MHz, CDCl<sub>3</sub>):  $\delta$  -205.93; HR-MS Calcd for C<sub>16</sub>H<sub>19</sub>FO<sub>6</sub> [M+Na]<sup>+</sup>, 349.1063, found 349.1060;  $\beta$  isomer: <sup>1</sup>H-NMR (500 MHz, CDCl<sub>3</sub>):  $\delta$  7.50 (m, 2H), 7.37 (m, 3H), 5.71 (m, 1H, H-2), 5.62 (s, 1H, PhCH), 4.76 (ddd,  $J = 48.5, 10.0, 4.0$  Hz, 1H, H-3), 4.56 (m, 1H, H-1), 4.40 (ddd,  $J = 10.5, 5.0, 2.0$  Hz, 1H, H-6a), 4.14 (q,  $J = 9.5$  Hz, 1H, H-4), 3.94 (t,  $J = 10.5$  Hz, 1H, H-6b), 3.55 (s, 3H, OMe), 3.41 (m, 1H, H-5), 2.20 (s, 3H, COCH<sub>3</sub>); <sup>13</sup>C-NMR (125 MHz, CDCl<sub>3</sub>):  $\delta$  169.91 (COCH<sub>3</sub>), 136.70, 129.26, 128.33, 126.11, 101.79 (PhCH), 100.09 (d,  $J = 8.38$  Hz, C-1), 87.64 (d,  $J = 194$  Hz, C-3), 76.86 (C-4), 68.90 (d,  $J = 16.63$  Hz, C-2), 68.29 (d,  $J = 1.75$  Hz, C-6), 65.79 (d,  $J = 8.0$  Hz, C-5), 57.68 (OCH<sub>3</sub>), 20.79 (COCH<sub>3</sub>); <sup>19</sup>F-NMR (470 MHz, CDCl<sub>3</sub>):  $\delta$  -201.82. ESI-MS Calcd for [M+Na]<sup>+</sup>, 349.11, found 348.92.

### 1,2,4,6-Tetra-O-acetyl-3-deoxy-3-fluoro-D-mannopyranoside (52)

To a mixture of **51** (780 mg, 2.387 mmol) in acetic anhydride (9.9 mL) was added H<sub>2</sub>SO<sub>4</sub> (95-97%) (397  $\mu$ L) at rt. The reaction mixture was stirred at rt for 3 h. Then diluted with DCM, neutralized carefully with Sat. aqueous NaHCO<sub>3</sub> at 0°C. The aqueous layer was extracted with DCM for two times, the combined organic layers were washed with brine and dried over Na<sub>2</sub>SO<sub>4</sub>. Solvent was removed under *vacuo* and the residue was purified by flash chromatography on silica gel (PE:EA 3:1-7:3) to afford **52** (770 mg, 92%) as a white foam as unseparated  $\alpha/\beta$  anomers.

### 3-Deoxy-3-fluoro-D-mannose (53)

To a solution of **52** (296 mg, 0.845 mmol) in dry methanol (5 mL) was added 0.5 M CH<sub>3</sub>ONa/MeOH (0.34 mL). The reaction mixture was stirred at rt for 3 h. Then neutralized with Amberlyst 15, filtered, concentrated to dryness to afford 3-deoxy-3-fluoro-D-mannose as a mixture of unseparated  $\alpha/\beta$  anomers **53** in quantitative yield which was used for next step without further purification.

### 1,2,4,6-Tetra-O-benzoyl-3-deoxy-3-fluoro- $\alpha$ -D-mannopyranoside (54)

To a suspension **53** (153 mg, 0.845 mmol) in pyridine (4.0 mL) was added DMAP (5 mg) and benzoyl chloride (0.59 mL) dropwise at 0°C. The reaction mixture was stirred at rt overnight and then diluted with DCM, washed with ice-water, Sat. NaHCO<sub>3</sub> and brine, dried over Na<sub>2</sub>SO<sub>4</sub>. Solvent was removed under *vacuo* and the residue was purified by flash chromatography on silica gel (PE:EA 5:1-4:1) to afford **54** (500 mg, 98%) as a white foam. <sup>1</sup>H-NMR (500 MHz, CDCl<sub>3</sub>):  $\delta$  8.14-8.05 (m, 8H), 7.69-7.35 (m, 12H), 6.61 (dd,  $J = 4.5, 2.0$  Hz, 1H, H-1), 6.20 (q,  $J = 10.0$  Hz, 1H, H-4), 5.87 (ddd,  $J = 5.5, 3.5, 2.0$  Hz, 1H, H-2), 5.40 (ddd,  $J = 48.0, 9.5, 3.5$  Hz, 1H, H-3), 4.69 (dd,  $J = 10.0, 2.0$  Hz, 1H, H-6a), 4.47-4.41 (m, 2H, H-6b, H-5); <sup>13</sup>C-NMR (125 MHz, CDCl<sub>3</sub>):  $\delta$  166.00, 165.15, 165.10, 163.56 (4xCOPh), 148.86, 136.81, 134.11, 133.72, 133.67, 133.26, 133.07, 130.03, 130.01, 129.97, 129.90, 129.72, 128.78, 128.61, 128.52, 128.18, 128.16 (Ar-C), 91.41 (d,  $J = 6.75$  Hz, C-1), 87.49 (d,  $J = 193.38$  Hz, C-3), 70.60 (d,  $J = 6.5$  Hz, C-5),

69.09 (d,  $J = 16.25$  Hz, C-2), 67.03 (d,  $J = 19.38$  Hz, C-4), 62.11 (C-6); HR-MS Calcd for  $C_{34}H_{27}FO_9 [M+Na]^+$ , 621.1537, found 621.1537.

### 2,4,6-Tri-O-benzoyl-3-deoxy-3-fluoro- $\alpha$ -D-mannopyranosyl bromide (55)

To a solution of **54** (766 mg, 1.28 mmol) in dry DCM (2.0 mL) was added 33% HBr/HOAc (2.5 mL) at 0°C and then reaction mixture was stirred at rt overnight. The reaction mixture was diluted with DCM, washed with ice-water, Sat.  $NaHCO_3$  and brine, dried over  $Na_2SO_4$ . Solvent was removed under *vacuo* and the residue was purified by flash chromatography on silica gel (PE:EA 8:1-6:1) to afford **55** (580 mg, 81%) as a white foam.  $^1H$ -NMR (500 MHz,  $CDCl_3$ ):  $\delta$  8.07 (m, 6H), 7.59 (m, 3H), 7.48 (t,  $J = 7.5$  Hz, 2H), 7.40 (t,  $J = 7.5$  Hz, 4H), 6.55 (dd,  $J = 5.0, 1.5$  Hz, 1H, H-1), 6.26 (m, 1H, H-4), 5.85 (ddd,  $J = 5.0, 3.5, 1.5$  Hz, 1H, H-2), 5.62 (ddd,  $J = 48.0, 9.5, 3.5$  Hz, 1H, H-3), 4.73 (dt,  $J = 12.0, 1.5$  Hz, 1H, H-6a), 4.51-4.44 (m, 2H, H-5, H-6b);  $^{13}C$ -NMR (125 MHz,  $CDCl_3$ ):  $\delta$  165.90, 165.08, 165.07 (3xCOPh), 133.86, 133.80, 133.21, 129.98, 129.97, 129.77, 129.50, 128.72, 128.65, 128.60, 128.53, 128.46 (Ar-C), 86.55 (d,  $J = 192.88$  Hz, C-3), 82.97 (d,  $J = 6.75$  Hz, C-1), 72.73 (d,  $J = 6.87$  Hz, C-5), 72.58 (d,  $J = 17.13$  Hz, C-2), 66.68 (d,  $J = 19.75$  Hz, C-4), 61.53 (C-6);  $^{19}F$ -NMR (470 MHz,  $CDCl_3$ ):  $\delta$  -204.34.

### *n*-Heptyl 2,4,6-tri-O-benzoyl-3-deoxy-3-fluoro- $\alpha$ -D-mannopyranoside (56)

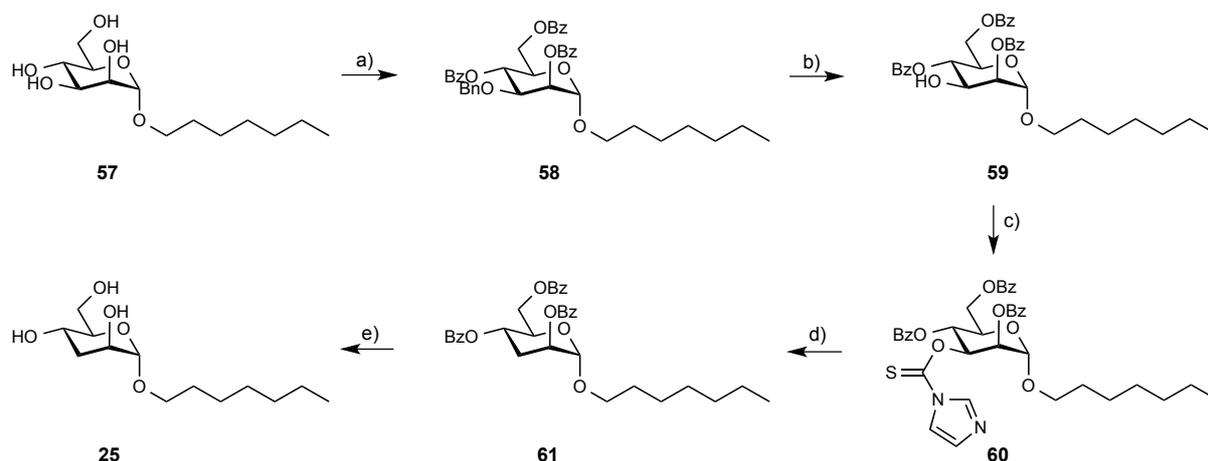
To a suspension of **55** (425 mg, 0.76 mmol), *n*-Heptanol (0.212 mL, 1.52 mmol) and 4Å MS in dry DCM (7 mL) was added  $Hg(CN)_2$  (232 mg, 0.918 mmol),  $HgBr_2$  (331 mg, 0.918 mmol) and was stirred at rt for 5 h under argon. The reaction mixture was filtered through a pad of celite, the celite was washed with DCM thoroughly and the filtrate was washed with  $NaHCO_3$ , brine, dried over  $Na_2SO_4$ . Solvent was removed under *vacuo* and the residue was purified by flash chromatography on silica gel (PE:EA 15:1) to **56** (123 mg, 27%) as colorless syrup.  $[\alpha]_D^{20} +9.65$  (c 1.54,  $CHCl_3$ );  $^1H$ -NMR (500 MHz,  $CDCl_3$ ):  $\delta$  8.08 (m, 6H), 7.58 (m, 3H), 7.47-7.38 (m, 6H), 6.02 (m, 1H, H-4), 5.64 (ddd,  $J = 6.0, 4.5, 2.0$  Hz, 1H, H-2), 5.23 (ddd,  $J = 48.0, 9.5, 4.0$  Hz, 1H, H-3), 5.08 (dd,  $J = 4.5, 2.0$  Hz, 1H, H-1), 4.69 (ddd,  $J = 12.0, 2.0, 1.5$  Hz, 1H, H-6a), 4.44 (dd,  $J = 12.0, 4.5$  Hz, 1H, H-6b), 4.25 (ddd,  $J = 10.0, 4.5, 2.5$  Hz, 1H, H-5), 3.76 (dt,  $J = 10.0, 7.0$  Hz, 1H, H-OCH<sub>2</sub>C<sub>6</sub>H<sub>13</sub>), 3.52 (dt,  $J = 10.0, 7.0$  Hz, 1H, H-OCH<sub>2</sub>C<sub>6</sub>H<sub>13</sub>), 1.67 (m, 2H), 1.34 (m, 8H), 0.91 (t,  $J = 7.0$  Hz, 3H, CH<sub>3</sub>);  $^{13}C$ -NMR (125 MHz,  $CDCl_3$ ):  $\delta$  166.10, 165.51, 165.32 (3xCOPh), 133.53, 133.48, 133.06, 129.89, 129.74, 129.68, 129.16, 129.10, 128.51, 128.48, 128.38 (Ar-C), 97.63 (d,  $J = 7.13$  Hz, C-1), 87.70 (d,  $J = 191.25$  Hz, C-3), 70.33 (d,  $J = 15.5$  Hz, C-2), 68.76 (OCH<sub>2</sub>C<sub>6</sub>H<sub>13</sub>), 68.31 (d,  $J = 6.75$  Hz, C-5), 67.76 (d,  $J = 19.13$  Hz, C-4), 62.74 (d,  $J = 0.88$  Hz, C-6), 31.74, 29.29, 28.98, 25.99, 22.57, 14.08 (OCH<sub>2</sub>C<sub>6</sub>H<sub>13</sub>);  $^{19}F$ -NMR (470 MHz,  $CDCl_3$ ):  $\delta$  -203.69; HR-MS Calcd for  $C_{34}H_{37}FO_8 [M+Na]^+$ , 615.2370, found 615.2360.

### *n*-Heptyl 3-deoxy-3-fluoro- $\alpha$ -D-mannopyranoside (24)

To a solution of **56** (68.8 mg, 0.116 mmol) in dry methanol (3.0 mL) was added 0.5 M  $CH_3ONa/MeOH$  (30  $\mu$ L) at rt. The reaction mixture was stirred for 4 h, then neutralized with Amberlyst 15, filtered, concentrated to dryness. The residue was purified by flash chromatography on silica gel (PE:EA 1:1 to 1:2) to afford **24** (28 mg, 86%) as a colorless syrup.  $[\alpha]_D^{20} = +66.51$  (c 0.78, DCM);  $^1H$ -NMR (500 MHz,  $CDCl_3$ ):  $\delta$  4.86 (d,  $J = 3.0$  Hz, 1H, H-1), 4.71 (ddd,  $J = 49.5, 9.5, 3.0$  Hz, 1H, H-3), 4.60 (d,  $J = 5.0$  Hz, 1H, 2-OH), 4.41 (d,  $J = 4.0$  Hz, 1H, 4-OH), 4.24 (m, 1H, H-4), 4.13 (m, 1H, H-2), 4.06 (m, 1H, 6-OH), 3.97 (m, 1H, H-6a), 3.82 (m, 1H, H-6b), 3.56 (m, 1H, H-5), 3.63 (dt,  $J = 9.5, 6.5$

Hz, 1H, H-OCH<sub>2</sub>C<sub>6</sub>H<sub>13</sub>), 3.38 (dt,  $J = 9.5, 6.5$  Hz, 1H, H-OCH<sub>2</sub>C<sub>6</sub>H<sub>13</sub>), 1.54 (m, 2H), 1.28 (m, 8H), 0.88 (t,  $J = 7.0$  Hz, 3H, CH<sub>3</sub>); <sup>13</sup>C-NMR (125 MHz, CDCl<sub>3</sub>): δ 99.88 (d,  $J = 8.13$  Hz, C-1), 93.18 (d,  $J = 180.25$  Hz, C-3), 71.17 (d,  $J = 7.13$  Hz, C-5), 69.50 (d,  $J = 16.13$  Hz, C-2), 68.11 (OCH<sub>2</sub>C<sub>6</sub>H<sub>13</sub>), 64.61 (d,  $J = 18.5$  Hz, C-4), 60.85 (C-6), 31.72, 29.29, 29.02, 26.00, 22.59, 14.06 (OCH<sub>2</sub>C<sub>6</sub>H<sub>13</sub>); <sup>19</sup>F-NMR (470 MHz, CDCl<sub>3</sub>): δ -203.25; HR-MS Calcd for C<sub>13</sub>H<sub>25</sub>FO<sub>5</sub> [M+Na]<sup>+</sup>, 303.1584, found 303.1582.

### Synthesis of 25 (*n*-heptyl 3-deoxy- $\alpha$ -D-mannopyranoside)



**Scheme 4.** a) i. Dibutyltin oxide, toluene, TBAB, BnBr, reflux, (44%); ii. Benzoyl chloride, pyr, DMAP, (68%); b) Pd(OH)<sub>2</sub>/H<sub>2</sub>, dioxane, rt, 3 bar, (80%); c) TCDI, DCE, reflux, (94%); d) Bu<sub>3</sub>SnH/toluene, reflux, overnight, (99%); e) CH<sub>3</sub>ONa/CH<sub>3</sub>OH, rt, 6 h, (83%).

### *n*-Heptyl 2,4,6-O-tribenzoyl-3-O-benzyl- $\alpha$ -D-mannopyranoside (58)

The mixtures of **57** (380 mg, 1.365 mmol) and dibutyltin oxide (374 mg, 1.5 mmol) were dissolved in dry MeOH (10 mL). The reaction mixture was refluxed for 4 h, concentrated to dryness under reduced pressure. To a solution of the product generated above in dry toluene (10 mL) was added tetrabutylammonium bromide (TBAB) (484 mg, 1.5 mmol) and benzyl bromide (194  $\mu$ L, 1.638 mmol). The mixture was stirred at 80°C overnight, concentrated to dryness, and purified by chromatography on silica gel (PE-EA 3:1-7:3) to give heptyl 3-O-benzyl- $\alpha$ -D-mannopyranoside (220 mg, 44%) as colorless oil. The above product was dissolved in pyridine (5.0 mL) and benzoyl chloride (0.42 mL, 3.58 mmol), DMAP (3.6 mg) was added to the solution at rt. The reaction mixture was stirred at rt for 2 h, then diluted with ethyl acetate, washed with 5% NaHCO<sub>3</sub>, brine and the organic layer was dried over Na<sub>2</sub>SO<sub>4</sub>. The solvent was removed in vacuo and the residue was purified by flash chromatography on silica gel (PE-EA 9:1-4:1) to provide **58** (277 mg, 68%) as colorless syrup. <sup>1</sup>H NMR (CDCl<sub>3</sub>, 500 MHz): δ 8.08 (m, 4H), 7.99 (m, 2H), 7.60-7.54 (m, 3H), 7.45-7.16 (m, 6H), 7.17-7.06 (m, 5H), 5.86 (t,  $J = 10.0$  Hz, 1H, H-4), 5.65 (m, 1H, H-2), 5.01 (s, 1H, H-1), 4.69 (d,  $J = 12.5$  Hz, 1H, OCH<sub>2</sub>Ph), 4.66 (dd,  $J = 12.0, 2.5$  Hz, 1H, H-6a), 4.51 (d,  $J = 12.5$  Hz, 1H, OCH<sub>2</sub>Ph), 4.39 (dd,  $J = 12.0, 5.0$  Hz, 1H, H-6b), 4.21 (ddd,  $J = 10.0, 4.5, 2.0$  Hz, 1H, H-5), 4.16 (dd,  $J = 10.0, 3.5$  Hz, 1H, H-3), 3.73 (m, 1H, OCH<sub>2</sub>C<sub>6</sub>H<sub>13</sub>), 3.49 (dt,  $J = 9.5, 6.5$  Hz, 1H, OCH<sub>2</sub>C<sub>6</sub>H<sub>13</sub>), 1.62 (m, 2H, OCH<sub>2</sub>CH<sub>2</sub>C<sub>5</sub>H<sub>11</sub>), 1.30 (m, 8H, OC<sub>2</sub>H<sub>4</sub>C<sub>4</sub>H<sub>8</sub>CH<sub>3</sub>), 0.90 (t,  $J = 7.0$  Hz, 3H, OC<sub>2</sub>H<sub>4</sub>C<sub>4</sub>H<sub>8</sub>CH<sub>3</sub>); <sup>13</sup>C NMR (CDCl<sub>3</sub>, 125 MHz) δ 166.25, 165.83, 165.45 (3x C=O), 137.53, 133.28, 133.26, 132.99, 129.97, 129.95, 129.76, 128.47, 128.39, 128.21, 127.86, 127.57 (Ar-C), 97.86 (C-1), 74.38 (C-3), 71.04 (OCH<sub>2</sub>Ph), 68.87 (C-5), 68.80 (C-2), 68.49

(OCH<sub>2</sub>C<sub>6</sub>H<sub>13</sub>), 68.30 (C-4), 63.27 (C-6), 31.81, 29.35, 29.04, 26.04, 22.62, 14.15 (OCH<sub>2</sub>C<sub>6</sub>H<sub>13</sub>); ESI-MS Calcd for C<sub>41</sub>H<sub>44</sub>O<sub>9</sub> [M+Na]<sup>+</sup>, 703.29, found 703.33.

#### ***n*-Heptyl 2,4,6-O-tribenzoyl- $\alpha$ -D-mannopyranoside (59)**

Hydrogenolysis of **58** (266 mg, 0.39 mmol) in dioxane (6.0 mL) in the presence of 10% Pd(OH)<sub>2</sub> (20 mg) and catalytic amount of HOAc under hydrogen (3 bar) at rt overnight. Then filtered the reaction suspension through celite, the filtrate was concentrated under vacuo. The residue was purified by flash chromatography on silica gel (PE-EA 9:1-5:1) to provide **59** (184 mg, 80%) as colorless syrup. <sup>1</sup>H NMR (CDCl<sub>3</sub>, 500 MHz):  $\delta$  8.09-8.05 (m, 6H), 7.61-7.55 (m, 3H), 7.46 (t, *J* = 8.0 Hz, 2H), 7.40 (m, 4H), 5.69 (t, *J* = 10.0 Hz, 1H, H-4), 5.42 (dd, *J* = 3.0, 1.5 Hz, 1H, H-2), 5.04 (d, *J* = 1.5 Hz, 1H, H-1), 4.68 (dd, *J* = 12.0, 2.0 Hz, 1H, H-6a), 4.46 (dd, *J* = 12.0, 4.5 Hz, 1H, H-6b), 4.41 (m, 1H, H-3), 4.30 (ddd, *J* = 10.0, 4.5, 2.5 Hz, 1H, H-5), 3.77 (dt, *J* = 9.5, 6.5 Hz, 1H, OCH<sub>2</sub>C<sub>6</sub>H<sub>13</sub>), 3.52 (dt, *J* = 9.5, 6.5 Hz, 1H, OCH<sub>2</sub>C<sub>6</sub>H<sub>13</sub>), 2.47 (d, *J* = 8.0 Hz, 1H, OH), 1.64 (m, 2H, OCH<sub>2</sub>CH<sub>2</sub>C<sub>5</sub>H<sub>11</sub>), 1.31 (m, 8H, OC<sub>2</sub>H<sub>4</sub>C<sub>4</sub>H<sub>8</sub>CH<sub>3</sub>), 0.90 (t, *J* = 7.0 Hz, 3H, OC<sub>2</sub>H<sub>4</sub>C<sub>4</sub>H<sub>8</sub>CH<sub>3</sub>); <sup>13</sup>C NMR (CDCl<sub>3</sub>, 125 MHz):  $\delta$  166.80, 166.14, 165.92 (3x C=O), 133.59, 133.49, 133.06, 129.91, 129.83, 129.80, 129.68, 129.31, 129.16, 128.54, 128.51, 128.39 (Ar-C), 97.36 (C-1), 73.01 (C-2), 70.43 (C-4), 69.11 (C-3), 68.58 (OCH<sub>2</sub>C<sub>6</sub>H<sub>13</sub>), 68.42 (C-5), 63.10 (C-6), 31.77, 29.37, 29.01, 26.03, 22.59, 14.10 (OCH<sub>2</sub>C<sub>6</sub>H<sub>13</sub>); ESI-MS Calcd for C<sub>34</sub>H<sub>38</sub>O<sub>9</sub> [M+Na]<sup>+</sup>, 613.24, found 613.17.

#### ***n*-Heptyl 2,4,6-O-tribenzoyl-3-O-thiocarbonylimidazolyl- $\alpha$ -D-mannopyranoside (60)**

A mixture of **59** (168 mg, 0.286 mmol) and N,N'-thiocarbonyldiimidazole (153 mg, 0.86 mmol) in 1,2-dichloroethane (5.0 mL) was refluxed overnight. The solution was concentrated in vacuo and the residue was diluted with DCM, washed with 1N HCl and brine. The organic layer was dried over Na<sub>2</sub>SO<sub>4</sub> and the solvent was removed in vacuo, the residue was purified by chromatography on silica gel (PE-EA 4:1-3:1) to afford **60** (188 mg, 94%) as pale yellow viscous solid. <sup>1</sup>H NMR (CDCl<sub>3</sub>, 500 MHz):  $\delta$  8.17 (s, 1H), 8.08 (d, *J* = 7.0 Hz, 2H), 8.02 (d, *J* = 7.0 Hz, 2H), 7.97 (d, *J* = 7.0 Hz, 2H), 7.62-7.54 (m, 3H), 7.44-7.37 (m, 7H), 6.87 (s, 1H), 6.34 (dd, *J* = 10.0, 3.0 Hz, 1H, H-3), 6.19 (t, *J* = 10.0 Hz, 1H, H-4), 5.84 (dd, *J* = 3.0, 2.0 Hz, 1H, H-2), 5.11 (s, 1H, H-1), 4.74 (m, 1H, H-6a), 4.50-4.46 (m, 2H, H-6b, H-5), 3.84 (dt, *J* = 9.5, 6.5 Hz, 1H, OCH<sub>2</sub>C<sub>6</sub>H<sub>13</sub>), 3.60 (dt, *J* = 9.5, 6.5 Hz, 1H, OCH<sub>2</sub>C<sub>6</sub>H<sub>13</sub>), 1.72 (m, 2H, OCH<sub>2</sub>CH<sub>2</sub>C<sub>5</sub>H<sub>11</sub>), 1.41-1.33 (m, 8H, OC<sub>2</sub>H<sub>4</sub>C<sub>4</sub>H<sub>8</sub>CH<sub>3</sub>), 0.91 (t, *J* = 7.0 Hz, 3H, OC<sub>2</sub>H<sub>4</sub>C<sub>4</sub>H<sub>8</sub>CH<sub>3</sub>); <sup>13</sup>C NMR (CDCl<sub>3</sub>, 125 MHz):  $\delta$  182.54 (C=S), 166.07, 165.32, 165.28 (3x C=O), 137.05, 133.77, 133.17, 130.88, 129.89, 129.71, 129.68, 128.78, 128.72, 128.58, 128.48, 117.82 (Ar-C), 97.54 (C-1), 78.24 (C-3), 69.38 (C-2), 68.99 (OCH<sub>2</sub>C<sub>6</sub>H<sub>13</sub>), 68.74 (C-5), 66.35 (C-4), 62.60 (C-6), 31.78, 29.29, 29.01, 25.99, 22.60, 14.11 (OCH<sub>2</sub>C<sub>6</sub>H<sub>13</sub>); ESI-MS Calcd for C<sub>38</sub>H<sub>40</sub>N<sub>2</sub>O<sub>9</sub>S [M+H]<sup>+</sup>, 701.25, found 701.28, Calcd for [M+Na]<sup>+</sup>, 723.24, found 723.26.

#### ***n*-Heptyl 2,4,6-O-tribenzoyl-3-deoxy- $\alpha$ -D-mannopyranoside (61)**

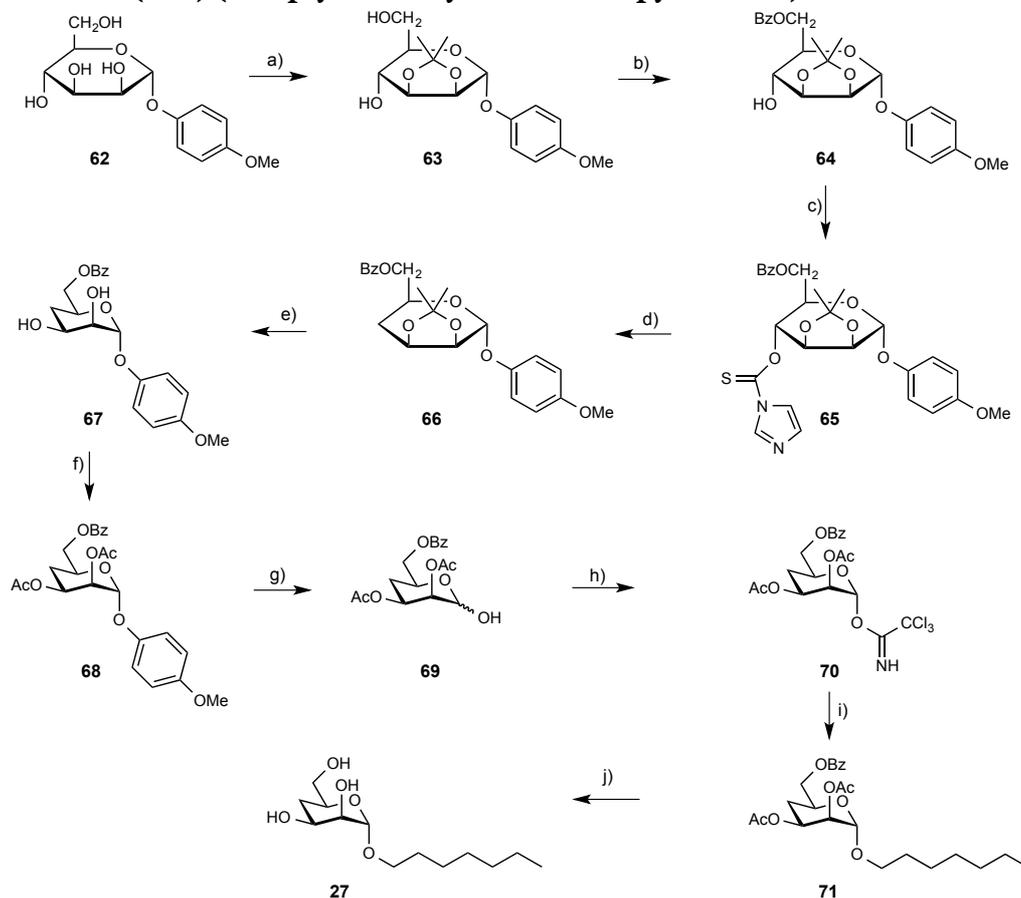
A solution of **60** (178 mg, 0.254 mmol) in dry toluene (2 mL) was added dropwise over 10 min to a stirred solution of refluxing toluene (3 mL) and tributylstanne (0.103 mL, 0.381 mmol) under argon. After the reaction mixture was refluxed overnight, the solvent was removed in vacuo and the residue was purified by chromatography on silica gel (PE-EA 10:1) to afford **61** (145 mg, 99%) as colorless oil. <sup>1</sup>H NMR (CDCl<sub>3</sub>, 500 MHz):  $\delta$  8.10- 8.02 (m, 6H), 7.60-7.54 (m, 3H), 7.46-7.37 (m, 6H), 5.57 (td, *J* = 10.5, 4.5 Hz, 1H, H-4), 5.24 (d, *J* = 1.0 Hz, 1H, H-2), 4.92 (s, 1H, H-1), 4.65 (dd, *J* = 12.0, 2.5 Hz, 1H, H-6a), 4.46 (dd, *J* = 12.0, 5.5 Hz, 1H, H-6b), 4.34 (ddd, *J* = 10.0, 5.5, 2.5 Hz, 1H, H-5),

3.80 (dt,  $J = 9.5, 7.0$  Hz, 1H,  $\text{OCH}_2\text{C}_6\text{H}_{13}$ ), 3.54 (dt,  $J = 9.5, 6.5$  Hz, 1H,  $\text{OCH}_2\text{C}_6\text{H}_{13}$ ), 2.54 (dt,  $J = 13.0, 3.5$  Hz, 1H, H-3a), 2.30 (m, 1H, H-3e), 1.66 (m, 2H,  $\text{OCH}_2\text{CH}_2\text{C}_5\text{H}_{11}$ ), 1.30 (m, 8H,  $\text{OC}_2\text{H}_4\text{C}_4\text{H}_8\text{CH}_3$ ), 0.92 (t,  $J = 7.0$  Hz, 3H,  $\text{OC}_2\text{H}_4\text{C}_4\text{H}_8\text{CH}_3$ );  $^{13}\text{C}$  NMR ( $\text{CDCl}_3$ , 125 MHz):  $\delta$  166.26, 165.56, 165.39 ( $3\times\text{C}=\text{O}$ ), 133.33, 133.28, 132.99, 129.91, 129.81, 129.72, 129.68, 129.65, 129.59, 128.46, 128.34 (Ar-C), 96.27 (C-1), 70.05 (C-2), 68.83 (C-5), 68.16 ( $\text{OCH}_2\text{C}_6\text{H}_{13}$ ), 65.46 (C-4), 63.68 (C-6), 29.04 (C-3), 31.78, 29.48, 29.43, 26.08, 22.59, 14.10 ( $\text{OCH}_2\text{C}_6\text{H}_{13}$ ); ESI-MS Calcd for  $\text{C}_{34}\text{H}_{38}\text{O}_8$   $[\text{M}+\text{Na}]^+$ , 597.25, found 597.21.

### *n*-Heptyl 3-deoxy- $\alpha$ -D-mannopyranoside (25)

To a solution of **61** (138 mg, 0.24 mmol) in methanol (4 mL) was added 0.5 M  $\text{CH}_3\text{ONa}/\text{MeOH}$  (96  $\mu\text{L}$ ) at rt. The reaction mixture was stirred at rt for 6 h, then neutralized with amberlyst 15, filtered and the solvent was concentrated to dryness. The residue was purified by flash chromatography on silica gel with (DCM-MeOH 11:1) to give **25** (52 mg, 83%) as colorless oil.  $[\alpha]_{\text{D}}^{20} +90.24$  (c 0.34, DCM);  $^1\text{H}$  NMR ( $\text{CD}_3\text{OD}$ , 500 MHz):  $\delta$  4.57 (s, 1H, H-1), 3.81 (dd,  $J = 11.5, 2.0$  Hz, 1H, H-6a), 3.78-3.74 (m, 3H, H-2, H-4,  $\text{OCH}_2\text{C}_6\text{H}_{13}$ ), 3.67 (dd,  $J = 11.5, 6.0$  Hz, 1H, H-6b), 3.52 (ddd,  $J = 9.0, 6.0, 2.0$  Hz, 1H, H-5), 3.42 (dt,  $J = 9.5, 6.5$  Hz, 1H,  $\text{OCH}_2\text{C}_6\text{H}_{13}$ ), 1.98 (dt,  $J = 13.0, 3.5$  Hz, 1H, H-3a), 1.82 (m, 1H, H-3e), 1.58 (m, 2H,  $\text{OCH}_2\text{CH}_2\text{C}_5\text{H}_{11}$ ), 1.40-1.32 (m, 8H,  $\text{OC}_2\text{H}_4\text{C}_4\text{H}_8\text{CH}_3$ ), 0.91 (t,  $J = 7.0$  Hz, 3H,  $\text{OC}_2\text{H}_4\text{C}_4\text{H}_8\text{CH}_3$ );  $^{13}\text{C}$  NMR ( $\text{CD}_3\text{OD}$ , 125 MHz):  $\delta$  100.17 (C-1), 75.50 (C-5), 69.23 (C-2), 68.34 ( $\text{OCH}_2\text{C}_6\text{H}_{13}$ ), 63.13 (C-4), 62.89 (C-6), 35.93 (C-3), 33.02, 30.69, 30.29, 27.38, 23.70, 14.43 ( $\text{OCH}_2\text{C}_6\text{H}_{13}$ ); HR-MS Calcd for  $\text{C}_{13}\text{H}_{26}\text{NaO}_5$   $[\text{M}+\text{Na}]^+$  285.1678, found 285.1682.

### Synthesis of **27** (4-H) (*n*-heptyl 4-deoxy- $\alpha$ -D-mannopyranoside)



**Scheme 5.** a) DMP, acetone, *p*-toluenesulfonic acid, water, rt, 3 h, (69%); b) DCM, benzoyl chloride, <15°C, 10 min, (87%); c) TCDI, DCE, reflux, (96%); d) Bu<sub>3</sub>SnH/toluene, reflux, overnight, (95%); e) HOAc/H<sub>2</sub>O (4:1), 70°C, 0.5 h, quantitative; f) (Ac)<sub>2</sub>O/pyr, DMAP, 4 h, (96%); g) CAN, CH<sub>3</sub>CN/H<sub>2</sub>O (4:1), 2 h, (91%); h) Cl<sub>3</sub>CCN, DCM, NaH, (71%); i) TMSOTf, toluene, (52%); j) CH<sub>3</sub>ONa/CH<sub>3</sub>OH, rt, (61%).

#### 4-Methoxyphenyl $\alpha$ -D-mannopyranoside (**62**)

To a mixture of 1,2,3,4,6-penta-O-acetyl- $\alpha$ -D-mannopyranoside (3.3 g, 8.45 mmol) and 4-methoxyphenol (1.57 g, 12.68 mmol) in 1,2-dichloroethane (30 mL) was added TMSOTf (0.18 mL, 1.01 mmol) at 0°C. The reaction mixture was stirred at 0°C for 1 h and then stirred at rt overnight. Diluted with EtOAc, washed with NaHCO<sub>3</sub>, water and brine. The organic layer was dried over Na<sub>2</sub>SO<sub>4</sub>, filtered and concentrated. The residue was purified by flash chromatography on silica gel (PE-EA 4:1-2:1) to give the desired compound (3.14 g, 82%) as a white solid. <sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>):  $\delta$  7.02 (d, *J* = 9.0 Hz, 2H), 6.82 (d, *J* = 9.0 Hz, 2H), 5.55 (dd, *J* = 10.0, 3.5 Hz, 1H, H-3), 5.43 (m, 1H, H-2), 5.41 (s, 1H, H-1), 5.36 (t, *J* = 10.0 Hz, 1H, H-4), 4.28 (dd, *J* = 12.5 Hz, 5.0 Hz, 1H, H-6a), 4.14 (ddd, *J* = 10.0, 5.5, 1.0 Hz, 1H, H-5), 4.11 (dd, *J* = 12.5, 2.0 Hz, 1H, H-6b), 3.77 (s, 3H, OCH<sub>3</sub>), 2.10, 2.06, 2.05, 2.03 (4xs, 12H, 4xCOCH<sub>3</sub>). To a solution of 4-methoxyphenyl 2,3,4,6-tetra-O-acetyl- $\alpha$ -D-mannopyranoside (2.7 g, 5.94 mmol) in MeOH (25 mL) was added 0.5 M CH<sub>3</sub>ONa/MeOH (1.2 mL) at rt. The reaction mixture was stirred at rt overnight, then neutralized Amberlyst 15. The reaction mixture was filtered and the residue was washed thoroughly with MeOH. The filtrate was concentrated in vacuo to give compound **62** (1.7 g, quantitative) as a white solid, which was used for next step without further purification. <sup>1</sup>H NMR (500 MHz, CD<sub>3</sub>OD):  $\delta$  7.04 (d, *J* = 9.0 Hz, 2H), 6.83 (d, *J* = 9.0 Hz, 2H), 5.33 (d, *J* = 1.5 Hz, 1H, H-1), 3.99 (dd, *J* = 3.5, 1.5 Hz, 1H, H-2), 3.88 (dd, *J* = 9.5, 3.5 Hz, 1H, H-3), 3.78 (dd, *J* = 12.0, 2.5 Hz, 1H, H-6a), 3.74 (s, 3H, OCH<sub>3</sub>), 3.74-3.70 (m, 2H, H-4, H-6b), 3.65 (m, 1H, H-5); <sup>13</sup>C NMR (CD<sub>3</sub>OD, 125 MHz):  $\delta$  156.60, 152.04, 119.20, 115.60, 101.17 (C-1), 75.23 (C-5), 72.45 (C-3), 72.13 (C-2), 68.44 (C-4), 62.72 (C-6), 56.04 (OCH<sub>3</sub>).

#### Methoxyphenyl 2,3-isopropylidene- $\alpha$ -D-mannopyranoside (**63**)

Compound **62** (1 g), acetone (10 mL), DMP (10 mL), and *p*-toluenesulfonic acid (0.2 g) were stirred together until the solid dissolved (about 15 minutes). Water (20 mL) was added and stirring was continued for 3 h. The reaction mixture was diluted with DCM, the aqueous layer was extracted with DCM for 3 times, the combined organic layers were washed with water, brine and dried over Na<sub>2</sub>SO<sub>4</sub>. Concentrated and the residue was purified with (PE-EA 2:1-1:1) to afford **63** (790 mg, 69%) as a white solid. <sup>1</sup>H NMR (500 MHz, CD<sub>3</sub>OD):  $\delta$  7.05 (d, *J* = 9.0 Hz, 2H), 6.85 (d, *J* = 9.0 Hz, 2H), 5.60 (s, 1H, H-1), 4.36 (dd, *J* = 5.5, 0.5 Hz, 1H, H-2), 4.18 (m, 1H, H-3), 3.75 (s, 3H, OCH<sub>3</sub>), 3.72 (m, 1H, H-6a), 3.69 (m, 1H, H-5), 3.66 (m, 2H, H-4, H-6b), 1.52, 1.38 (2xs, 6H, CH<sub>3</sub>CCH<sub>3</sub>); <sup>13</sup>C NMR (125 MHz, CD<sub>3</sub>OD):  $\delta$  156.70, 151.76, 119.46, 115.56, 110.65, 98.21 (C-1), 80.16 (C-3), 77.19 (C-2), 72.53 (C-5), 69.79 (C-4), 62.16 (C-6), 56.01 (OCH<sub>3</sub>), 28.37, 26.64 (CH<sub>3</sub>CCH<sub>3</sub>).

#### Methoxyphenyl 6-O-benzoyl-2,3-isopropylidene- $\alpha$ -D-mannopyranoside (**64**)

A solution of **63** (990 mg, 3.04 mmol) in DCM (7 mL) and pyridine (0.7 mL) was stirred at < 15°C while benzoyl chloride (0.37 mL, 3.188 mmol) was added. After 10 minutes, TLC shows no **64** left. Then methanol was added to quench the BzCl, the reaction mixture was diluted with DCM and then washed with 1N HCl, brine. The organic layer

was dried over  $\text{Na}_2\text{SO}_4$ , concentrated and the residue was purified by chromatography on silica gel (PE-EA 4:1-3:2) to afford **64** (1.14g, 87%) as a white solid.  $[\alpha]_{\text{D}}^{20} = +38.95$  ( $c = 0.44$ , DCM);  $^1\text{H}$  NMR ( $\text{CDCl}_3$ , 500 MHz):  $\delta$  7.88 (dd,  $J = 8.5, 1.5$  Hz, 1H), 7.56 (t,  $J = 7.5$  Hz, 1H), 7.38 (t,  $J = 8.0$  Hz, 2H), 7.02 (d,  $J = 9.0$  Hz, 2H), 6.76 (d,  $J = 9.0$  Hz, 2H), 5.71 (s, 1H, H-1), 4.60 (dd,  $J = 12.0, 6.0$  Hz, 1H, H-6a), 4.50 (dd,  $J = 12.0, 2.0$  Hz, 1H, H-6b), 4.39 (d,  $J = 6.0$  Hz, 1H, H-2), 4.33 (t,  $J = 6.0$  Hz, 1H, H-3), 4.03 (ddd,  $J = 10.0, 6.0, 2.0$  Hz, 1H, H-5), 3.73 (s, 3H,  $\text{OCH}_3$ ), 3.69 (ddd,  $J = 11.0, 7.0, 4.0$  Hz, 1H, H-4), 3.04 (d,  $J = 4.5$  Hz, 1H, 4-OH), 1.52, 1.40 (2xs, 6H,  $\text{CH}_3\text{CCH}_3$ );  $^{13}\text{C}$  NMR ( $\text{CDCl}_3$ , 125 MHz):  $\delta$  169.93, 154.96, 149.67, 133.13, 129.78, 129.52, 128.26, 117.65, 114.56, 109.94, 95.98 (C-1), 78.27 (C-3), 75.59 (C-2), 69.42 (C-4), 68.94 (C-5), 63.88 (C-6), 55.50 ( $\text{OCH}_3$ ), 28.01, 26.21 ( $\text{CH}_3\text{CCH}_3$ ); ESI-MS Calcd for  $\text{C}_{23}\text{H}_{26}\text{O}_8$   $[\text{M}+\text{Na}]^+$ , 453.15, found 453.14.

#### **Methoxyphenyl 6-O-benzoyl-2, 3-isopropylidene-4-O-thiocarbonylimidazolyl- $\alpha$ -D-mannopyranoside (65)**

A mixture of **64** (570 mg, 1.324 mmol) and  $\text{N,N}'$ -thiocarbonyldiimidazole (708 mg, 3.97 mmol) in 1,2-dichloroethane (6.0 mL) was refluxed for 5 h. The solution was concentrated in vacuo and the residue was diluted with DCM, washed with 1N HCl and brine. The organic layer was dried over  $\text{Na}_2\text{SO}_4$  and the solvent was removed in vacuo, the residue was purified by chromatography on silica gel (PE-EA 3:1-1:1) to afford **65** (693 mg, 97%) as an off-white solid.  $[\alpha]_{\text{D}}^{20} = +75.4$  ( $c = 0.615$ , MeOH);  $^1\text{H}$  NMR ( $\text{CDCl}_3$ , 500 MHz):  $\delta$  8.35 (s, 1H), 7.79 (d,  $J = 7.5$  Hz, 2H), 7.63 (s, 1H), 7.52 (t,  $J = 7.5$  Hz, 1H), 7.35 (t,  $J = 10.5$  Hz, 2H), 7.02 (m, 3H), 6.73 (d,  $J = 9.5$  Hz, 2H), 6.04 (t,  $J = 7.5$  Hz, 1H, H-4), 5.75 (s, 1H, H-1), 4.64 (t,  $J = 6.5$  Hz, 1H, H-3), 4.48 (d,  $J = 5.5$  Hz, 1H, H-2), 4.46-4.38 (m, 3H, H-6a, H-5, H-6b), 3.71 (s, 3H,  $\text{OCH}_3$ ), 1.66, 1.42 (2xs, 6H,  $\text{CH}_3\text{CCH}_3$ );  $^{13}\text{C}$  NMR ( $\text{CDCl}_3$ , 125 MHz):  $\delta$  183.46 (C=S), 165.98 (C=O), 155.24, 149.53, 136.98, 133.09, 131.16, 129.63, 129.42, 128.23, 118.12, 117.60, 114.67, 110.90, 95.82 (C-1), 78.40 (C-4), 75.75 (C-2), 75.17 (C-3), 66.91 (C-5), 63.20 (C-6), 55.49 ( $\text{OCH}_3$ ), 27.53, 26.29 ( $\text{CH}_3\text{CCH}_3$ ); HR-MS Calcd for  $\text{C}_{27}\text{H}_{28}\text{N}_2\text{O}_8\text{S}$   $[\text{M}+\text{Na}]^+$ , 563.1461, found 563.1462.

#### **Methoxyphenyl 6-O-benzoyl-4-deoxy-2,3-isopropylidene- $\alpha$ -D-mannopyranoside (66)**

A solution of **65** (1.378 g, 2.55 mmol) in dry toluene (20 mL) was added dropwise over 30 min to a stirred solution of refluxing toluene (60 mL) and tributylstanne (1.05 mL, 3.83 mmol) under argon. After the reaction mixture was refluxed for 6 h, the solvent was removed in vacuo and the residue was purified by chromatography on silica gel (PE-EA 9:1-4:1) to afford **66** (1.046 g, 99%) as a pale yellow solid.  $[\alpha]_{\text{D}}^{20} = +95.0$  ( $c = 0.502$ ,  $\text{CHCl}_3$ );  $^1\text{H}$  NMR ( $\text{CDCl}_3$ , 500 MHz):  $\delta$  7.85 (m, 2H), 7.52 (m, 1H), 7.38 (t,  $J = 8.0$  Hz, 2H), 7.01 (d,  $J = 9.0$  Hz, 2H), 6.72 (d,  $J = 9.0$  Hz, 2H), 5.71 (s, 1H, H-1), 4.56 (m, 1H, H-3), 4.33 (m, 2H, H-6a, 6b), 4.21 (m, 2H, H-2, H-5), 3.71 (s, 3H,  $\text{OCH}_3$ ), 2.04 (ddd,  $J = 13.0, 6.5, 2.5$  Hz, 1H, H-4a), 1.69 (ddd,  $J = 13.0, 10.5, 9.0$  Hz, 1H, H-4b), 1.57, 1.40 (2xs, 6H,  $\text{CH}_3\text{CCH}_3$ );  $^{13}\text{C}$  NMR ( $\text{CDCl}_3$ , 125 MHz):  $\delta$  166.20 (C=O), 154.87, 150.03, 132.94, 129.83, 129.69, 128.21, 117.77, 114.51, 109.35, 96.61 (C-1), 73.01 (C-2), 70.43 (C-3), 66.74 (C-6), 65.27 (C-5), 55.49 ( $\text{OCH}_3$ ), 29.98 (C-4), 28.09, 26.25 ( $\text{CH}_3\text{CCH}_3$ ); HR-MS Calcd for  $\text{C}_{23}\text{H}_{26}\text{O}_7$   $[\text{M}+\text{Na}]^+$ , 437.1576, found 437.1576.

#### **Methoxyphenyl 6-O-benzoyl-4-deoxy- $\alpha$ -D-mannopyranoside (67)**

A solution of **66** (322 mg) in 80% aqueous HOAc (5.0 mL) was heated at  $75^\circ\text{C}$  for 0.5 h and then concentrated to afford **67** as a white solid in quantitative yield. It was used for next step without further purification.  $^1\text{H}$  NMR ( $\text{CD}_3\text{OD}$ , 500 MHz):  $\delta$  7.87 (dd,  $J = 8.0$ ,

1.0 Hz, 2H), 7.58 (m, 1H), 7.43 (t,  $J = 8.0$  Hz, 2H), 7.00 (d,  $J = 9.0$  Hz, 2H), 6.71 (d,  $J = 9.5$  Hz, 2H), 5.43 (d,  $J = 1.0$  Hz, 1H, H-1), 4.35-4.22 (m, 3H, H-6a, H-6b, H-5), 4.17 (m, 1H, H-3), 3.89 (m, 1H, H-2), 3.69 (s, 3H, OCH<sub>3</sub>), 1.84 (q,  $J = 12.0$  Hz, 1H, H-4a), 1.74 (m, 1H, H-4b).

#### Methoxyphenyl 6-O-benzoyl-4-deoxy-2,3-O-acetyl- $\alpha$ -D-mannopyranoside (68)

To a solution of **67** (290 mg, 0.77 mmol) in dry pyridine (2.0 mL) was added acetic anhydride (1.5 mL) and DMAP (15 mg) at rt. The reaction mixture was stirred at rt for 4 h then concentrated to dryness. The residue was purified by flash chromatography on silica gel (PE-EA 6:1-3:1) to afford **68** (340 mg, 96%) as a white solid.  $[\alpha]_{\text{D}}^{20} = +66.7$  ( $c = 0.8$ , CHCl<sub>3</sub>); <sup>1</sup>H NMR (CDCl<sub>3</sub>, 500 MHz):  $\delta$  7.95 (dd,  $J = 8.5$ , 1.0 Hz, 2H), 7.56 (t,  $J = 7.5$  Hz, 1H), 7.41 (t,  $J = 8.0$  Hz, 2H), 7.01 (d,  $J = 9.0$  Hz, 2H), 6.73 (d,  $J = 9.0$  Hz, 2H), 5.53 (ddd,  $J = 12.0$ , 5.5, 3.5 Hz, 1H, H-5), 5.46 (d,  $J = 1.5$  Hz, 1H, H-1), 5.30 (m, 1H, H-2), 4.38 (m, 3H, H-3, H-6a, H-6b), 3.73 (s, 3H, OCH<sub>3</sub>), 2.16, 2.06 (2xs, 6H, 2xCOCH<sub>3</sub>), 1.97 (m, 2H, H-4); <sup>13</sup>C NMR (CDCl<sub>3</sub>, 125 MHz):  $\delta$  170.07, 169.96, 166.13 (3xC=O), 155.16, 149.86, 133.09, 129.76, 129.70, 128.29, 117.99, 114.52, 97.22 (C-1), 67.71 (C-2), 66.91 (C-3), 66.30 (C-5), 66.16 (C-6), 55.54 (OCH<sub>3</sub>), 27.84 (C-4), 20.96, 20.88 (CH<sub>3</sub>CO); HR-MS Calcd. for C<sub>24</sub>H<sub>26</sub>O<sub>9</sub> [M+Na]<sup>+</sup>, 481.1475, found 481.1447.

#### 6-O-Benzoyl-2,3-di-O-acetyl-4-deoxy- $\alpha$ -D-mannopyranosyl trichloroacetimidate (70)

To a solution of **68** (151 mg, 0.328 mmol) in CH<sub>3</sub>CN/H<sub>2</sub>O (v:v = 4:1, 20.0 mL) was added CAN (539 mg, 0.983 mmol) in portions at 0°C. The mixture was stirred at 0°C for 0.5 h, continuously stirred at rt for additional 1 h. Then diluted with EA, the organic layers were washed with water, brine and dried over Na<sub>2</sub>SO<sub>4</sub> and concentrated in vacuum, the residue was purified by chromatography on silica gel (PE-EA 6:1-3:1) to afford a mixture of **69** (106 mg, 91%). To a solution of this mixture in dry DCM (6.0 mL) was added NaH (60%) (1.2 mg) at rt. The mixture was stirred at rt for 4 h, then added silica gel and concentrated, the residue was purified by chromatography on silica gel (PE-EA 9:1-4:1) to afford **70** (106 mg, 71%) as a white solid.  $[\alpha]_{\text{D}}^{20} = +54.02$  ( $c = 0.81$ , CHCl<sub>3</sub>); <sup>1</sup>H NMR (CDCl<sub>3</sub>, 500 MHz):  $\delta$  8.71 (s, 1H, NH), 8.03 (dd,  $J = 8.5$ , 1.0 Hz, 2H), 7.57 (t,  $J = 7.5$  Hz, 1H), 7.44 (t,  $J = 7.5$  Hz, 2H), 6.32 (d,  $J = 1.5$  Hz, 1H, H-1), 5.41 (m, 1H, H-5), 5.32 (m, 1H, H-2), 4.46 (m, 1H, H-3), 4.43 (m, 2H, H-6a, H-6b), 2.17, 2.06 (2xs, 6H, 2xCOCH<sub>3</sub>), 2.08 (m, 2H, H-4); <sup>13</sup>C NMR (CDCl<sub>3</sub>, 125 MHz):  $\delta$  169.84, 169.77, 166.13 (3xC=O), 159.97 (NHCCl<sub>3</sub>), 133.19, 129.75, 129.68, 128.37, 95.46 (C-1), 69.17 (C-3), 66.12, 66.09, 65.85 (C-2, C-5, C-6), 27.73 (C-4), 20.87, 20.76 (2xCOCH<sub>3</sub>); HR-MS Calcd. for C<sub>19</sub>H<sub>20</sub>Cl<sub>3</sub>NO<sub>6</sub> [M+Na]<sup>+</sup>, 518.0152, found 518.0156.

#### *n*-Heptyl 6-O-benzoyl-4-deoxy-2,3-O-acetyl- $\alpha$ -D-mannopyranoside (71)

To a solution of **70** (100 mg, 0.2 mmol) and *n*-heptanol (85  $\mu$ L) in dry toluene (4 mL) was added a solution of TMSOTf/toluene (20  $\mu$ L/480  $\mu$ L) (50  $\mu$ L) at rt. The reaction mixture was stirred at rt for 3 h, then diluted with EtOAc, washed with 5% NaHCO<sub>3</sub>, brine and dried over Na<sub>2</sub>SO<sub>4</sub> and the solvent was removed in vacuo, the residue was purified by chromatography on silica gel (PE-EA 9:1-6:1) to afford **71** (47 mg, 52%) as colorless oil.  $[\alpha]_{\text{D}}^{20} = +41.15$  ( $c = 0.88$ , CHCl<sub>3</sub>); <sup>1</sup>H NMR (CDCl<sub>3</sub>, 500 MHz):  $\delta$  8.06 (d,  $J = 7.5$ , 1.0 Hz, 2H), 7.57 (t,  $J = 7.0$  Hz, 1H), 7.44 (t,  $J = 7.5$  Hz, 2H), 5.32 (m, 1H, H-3), 5.09 (s, 1H, H-2), 4.84 (d,  $J = 1.0$  Hz, 1H, H-1), 4.40 (m, 2H, H-6), 4.23 (m, 1H, H-5), 3.67 (dt,  $J = 9.5$ , 6.5 Hz, 1H, OCH<sub>2</sub>C<sub>6</sub>H<sub>13</sub>), 3.42 (dt,  $J = 9.5$ , 6.5 Hz, 1H, OCH<sub>2</sub>C<sub>6</sub>H<sub>13</sub>), 2.14, 2.02 (2xs, 6H, 2xCOCH<sub>3</sub>), 1.92 (m, 2H, H-4), 1.57 (m, 2H), 1.28 (m, 8H), 0.88 (t,  $J$

= 7.0 Hz, 3H);  $^{13}\text{C}$  NMR ( $\text{CDCl}_3$ , 125 MHz):  $\delta$  170.21, 169.99, 166.24 (3xCO), 133.14, 129.85, 129.64, 128.37, 98.08 (C-1), 68.13 ( $\text{OCH}_2\text{C}_6\text{H}_{13}$ ), 68.04 (C-2), 66.60 (C-3), 66.43 (C-6), 66.16 (C-5), 28.09 (C-4), 32.79, 31.80, 29.08, 25.68, 22.59, 14.07 ( $\text{OCH}_2\text{C}_6\text{H}_{13}$ ), 21.00, 20.96 (2x $\text{CH}_3\text{CO}$ ); HR-MS Calcd. for  $\text{C}_{24}\text{H}_{34}\text{O}_8$   $[\text{M}+\text{Na}]^+$ , 473.2151, found 473.2149.

#### ***n*-Heptyl 4-deoxy $\alpha$ -D-mannopyranoside (27)**

To a solution of **71** (47 mg, 0.1 mmol) in methanol (4 mL) was added 0.5 M  $\text{CH}_3\text{ONa}/\text{MeOH}$  (40  $\mu\text{L}$ ) at rt. The reaction mixture was stirred at rt overnight, then neutralized with HOAc. The solvent was concentrated to dryness and the residue was purified by flash chromatography on silica gel (PE-EA 1:1-0:1) to **27** (16 mg, 61%) as a white solid.  $[\alpha]_{\text{D}}^{20} = +86.14$  ( $c = 0.19$ ,  $\text{CHCl}_3$ );  $^1\text{H}$  NMR ( $\text{CDCl}_3$ , 500 MHz):  $\delta$  4.89 (s, 1H, H-1), 4.05 (m, 1H, H-3), 3.84 (m, 1H, H-5), 3.77 (m, 1H, H-2), 3.71-3.64 (m, 2H, H-6a, H- $\text{OCH}_2\text{C}_6\text{H}_{13}$ ), 3.60 (m, 1H, H-6b), 3.39 (dt,  $J = 9.5, 6.5$  Hz, 1H,  $\text{OCH}_2\text{C}_6\text{H}_{13}$ ), 2.59 (d,  $J = 5.5$  Hz, 1H, 2-OH), 2.30 (m, 2H, 3-OH, 6-OH), 1.68 (m, 2H, H-4), 1.55 (m, 2H), 1.28 (m, 8H), 0.88 (t,  $J = 7.0$  Hz, 3H);  $^{13}\text{C}$  NMR ( $\text{CDCl}_3$ , 125 MHz):  $\delta$  100.12 (C-1), 69.20 (C-2), 68.36 (C-5), 67.80 ( $\text{OCH}_2\text{C}_6\text{H}_{13}$ ), 65.59 (C-3), 65.32 (C-6), 30.08 (C-4), 31.76, 29.42, 29.06, 26.09, 22.59, 14.07 ( $\text{OCH}_2\text{C}_6\text{H}_{13}$ ); HR-MS Calcd for  $\text{C}_{13}\text{H}_{26}\text{O}_5$   $[\text{M}+\text{Na}]^+$  285.1678, found 285.1676.

#### **Synthesis of compounds previously published**

Compound **1** and **2** by Bouckaert *et al.* 2005<sup>(27)</sup>; compound **10** by Han *et al.* 2010<sup>(28)</sup>; compound **11** and **20** by Klein *et al.* 2010<sup>(2)</sup>; compound **12** - **14** by Pang *et al.* 2012<sup>(29)</sup>; compound **15** by Scharenberg *et al.* 2012<sup>(30)</sup>; compound **16** - **19** and **21** - **22** by Kleeb *et al.* 2015<sup>(9)</sup>; compound **23** by Jiang *et al.* 2012<sup>(31)</sup>; compound **8** and **26** by Fiege *et al.* 2015<sup>(8)</sup>; compound **9** by Sager *et al.* 2018<sup>(32)</sup> and compound **28** and **29** by Schönemann *et al.* 2018<sup>(33)</sup>. Compound **4** was purchased from TRC Canada (Toronto, Canada) and compound **3** was obtained from Sigma Aldrich (Buchs, Switzerland).

#### **References**

1. Rabbani, S., Jiang, X., Schwardt, O., & Ernst, B. (2010). Expression of the carbohydrate recognition domain of FimH and development of a competitive binding assay. *Anal. Biochem.*, 407(2): 188-195.
2. Klein, T., *et al.* (2010). FimH antagonists for the oral treatment of urinary tract infections: from design and synthesis to in vitro and in vivo evaluation. *J. Med. Chem.*, 53(24): 8627-8641.
3. Edelhoch, H. (1967). Spectroscopic determination of tryptophan and tyrosine in proteins. *Biochemistry*, 6(7): 1948-1954.
4. AFFINImeter. Version 1.0415 - 1.1510; Software for Science Developments (S4Sd), Santiago de Compostela, Spain, 2015 - 2016.
5. Burnouf, D., *et al.* (2012). kinITC: a new method for obtaining joint thermodynamic and kinetic data by isothermal titration calorimetry. *J. Am. Chem. Soc.*, 134(1): 559-565.
6. Sigurskjold, B.W. (2000). Exact analysis of competition ligand binding by displacement isothermal titration calorimetry. *Anal. Biochem.*, 277(2): 260-266.

7. Houtman, J.C., *et al.* (2007). Studying multisite binary and ternary protein interactions by global analysis of isothermal titration calorimetry data in SEDPHAT: application to adaptor protein complexes in cell signaling. *Protein Sci.*, 16(1): 30-42.
8. Fiege, B., *et al.* (2015). The Tyrosine Gate of the Bacterial Lectin FimH: A Conformational Analysis by NMR Spectroscopy and X-ray Crystallography. *ChemBioChem*, 16(8): 1235-1246.
9. Kleeb, S., *et al.* (2015). FimH antagonists: bioisosteres to improve the in vitro and in vivo PK/PD profile. *J. Med. Chem.*, 58(5): 2221-2239.
10. Turnbull, W.B. & Daranas, A.H. (2003). On the value of c: can low affinity systems be studied by isothermal titration calorimetry? *J. Am. Chem. Soc.*, 125(48): 14859-14866.
11. Prism 5 for Mac OS X. Version 5.0c; GraphPad Softwar, Inc., La Jolla, CA, USA, 1994 - 2010.
12. Rubinson, K.A., Ladner, J.E., Tordova, M., & Gilliland, G.L. (2000). Cryosalts: suppression of ice formation in macromolecular crystallography. *Acta Crystallogr. Sect. D.-Biol. Crystallogr.*, 56(8): 996-1001.
13. Kabsch, W. (2010). Integration, scaling, space-group assignment and post-refinement. *Acta Crystallogr. D Biol. Crystallogr.*, 66(Pt 2): 133-144.
14. McCoy, A.J., *et al.* (2007). Phaser crystallographic software. *J. Appl. Crystallogr.*, 40(Pt 4): 658-674.
15. Emsley, P. & Cowtan, K. (2004). Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.*, 60(Pt 12 Pt 1): 2126-2132.
16. Emsley, P., Lohkamp, B., Scott, W.G., & Cowtan, K. (2010). Features and development of Coot. *Acta Crystallogr. Sect. D.-Biol. Crystallogr.*, 66(4): 486-501.
17. Adams, P.D., *et al.* (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.*, 66(Pt 2): 213-221.
18. Adams, P.D., *et al.* (2002). PHENIX: building new software for automated crystallographic structure determination. *Acta Crystallogr. Sect. D.-Biol. Crystallogr.*, 58(11): 1948-1954.
19. Schuttelkopf, A.W. & Van Aalten, D.M. (2004). PRODRG: a tool for high-throughput crystallography of protein–ligand complexes. *Acta Crystallogr. Sect. D.-Biol. Crystallogr.*, 60(8): 1355-1363.
20. Sastry, G.M., Adzhigirey, M., Day, T., Annabhimoju, R., & Sherman, W. (2013). Protein and ligand preparation: parameters, protocols, and influence on virtual screening enrichments. *J. Comput. Aided Mol. Des.*, 27(3): 221-234.
21. Bosch, F. & Rosich, L. (2008). The contributions of Paul Ehrlich to pharmacology: a tribute on the occasion of the centenary of his Nobel Prize. *Pharmacology*, 82(3): 171-179.
22. van der Ven, J.G., Wijkmans, J.C., Kamerling, J.P., & Vliegthart, J.F. (1994). Synthesis of three tetrasaccharides containing 3-O-methyl- $\alpha$ -D-mannose, as model compounds for xylose-containing carbohydrate chains from N-glycoproteins. *Carbohydr. Res.*, 253: 121-139.
23. Haradahira, T., Maeda, M., Omae, H., Yano, Y., & Kojima, M. (1984). Synthesis of 2-deoxy-2-fluoro-D-mannose using fluoride ion. *Chem. Pharm. Bull.*, 32(12): 4758-4766.
24. Spijker, N.M., Slief, J.-W., & Van Boeckel, C.A. (1993). Synthesis of Modified Di-and Trisaccharide Fragments of N-Glycoproteins. *J. Carbohydr. Chem.*, 12(8): 1017-1041.
25. Elchert, B., *et al.* (2004). Application of the synthetic aminosugars for glycodiversification: synthesis and antimicrobial studies of pyranmycin. *J. Org. Chem.*, 69(5): 1513-1523.
26. Matwiejuk, M. & Thiem, J. (2011). Defining oxyanion reactivities in base-promoted glycosylations. *Chem. Commun.*, 47(29): 8379-8381.
27. Bouckaert, J., *et al.* (2005). Receptor binding studies disclose a novel class of high-affinity inhibitors of the Escherichia coli FimH adhesin. *Mol. Microbiol.*, 55(2): 441-455.
28. Han, Z., *et al.* (2010). Structure-Based Drug Design and Optimization of Mannoside Bacterial FimH Antagonists. *J. Med. Chem.*, 53(12): 4779-4792.
29. Pang, L., *et al.* (2012). FimH antagonists: structure-activity and structure-property relationships for biphenyl  $\alpha$ -D-mannopyranosides. *ChemMedChem*, 7(8): 1404-1422.
30. Scharenberg, M., Schwardt, O., Rabbani, S., & Ernst, B. (2012). Target Selectivity of FimH Antagonists. *J. Med. Chem.*, 55(22): 9810-9816.
31. Jiang, X., *et al.* (2012). Antiadhesion Therapy for Urinary Tract Infections—A Balanced PK/PD Profile Proved To Be Key for Success. *J. Med. Chem.*, 55(10): 4700-4713.
32. Sager, C.P., *et al.* (2018). The price of flexibility - a case study on septanoses as pyranose mimetics. *Chemical Science*, The Royal Society of Chemistry.
33. Schönemann, W., *et al.* (2016). 2-C-Branched mannosides as a novel family of FimH antagonists—Synthesis and biological evaluation. *Beilstein J. Org. Chem.*

# **Section III.**

## **E-selectin**



## 1. Introduction

### 1.1. A brief overview of the selectin family members

The selectin family consists of E-, P-, and L-selectin, three structurally closely related carbohydrate binding cell-adhesion molecules (CAMs) belonging to the C-type lectins. All selectins are transmembrane proteins, expressed on the surface of vascular endothelial cells (E-selectin, P-selectin), leukocytes (L-selectin) and platelets (P-selectin).<sup>(1)</sup> They play an important role in the initiation of an inflammation by mediating tethering and rolling of leukocytes on vascular endothelial cells.<sup>(2)</sup> Table 1 gives an overview of the most important characteristics of E-, P-, and L-selectin.

**Table 1. Characteristics of the three selectin family members L-, E-, and P-selectin.** Historical nomenclature, structural traits, features of expression, and binding targets.

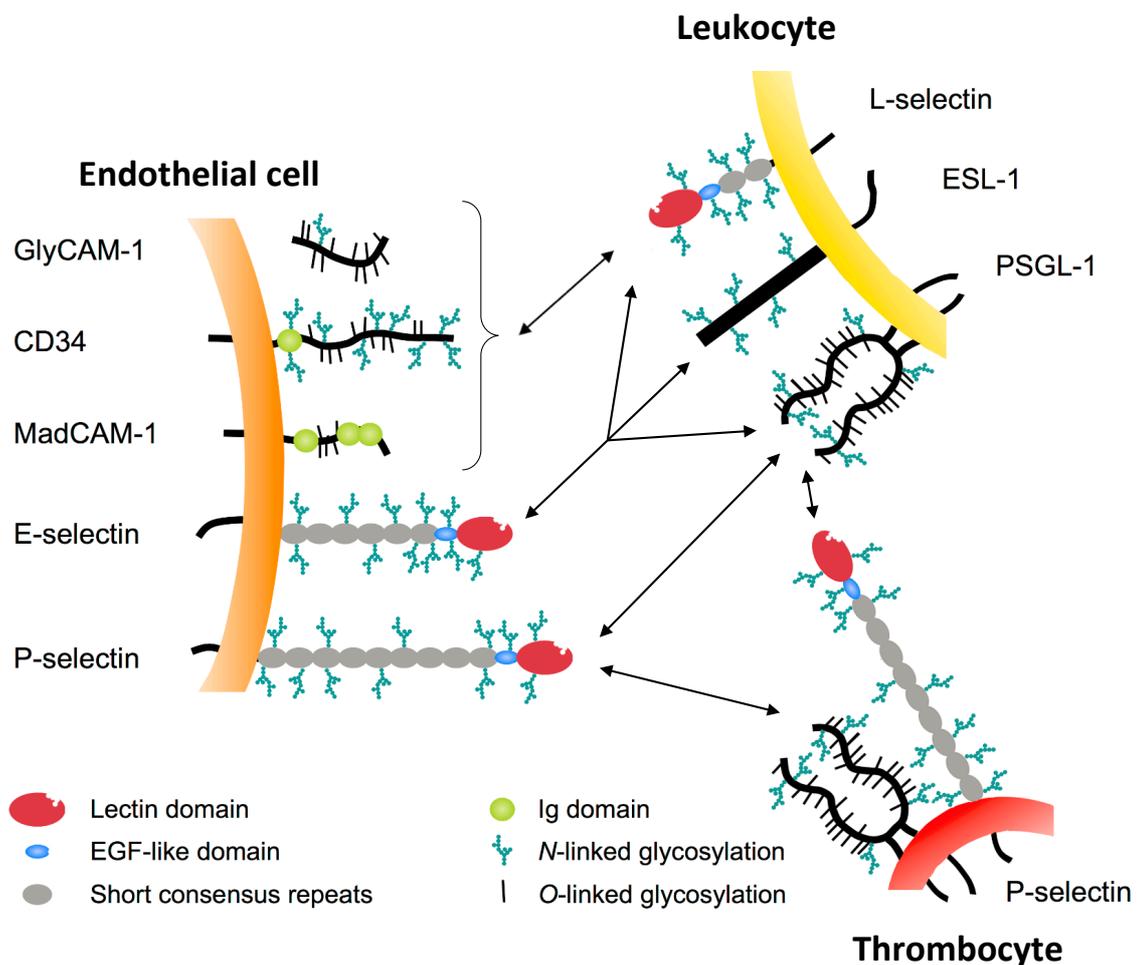
	<b>L-Selectin</b>	<b>E-Selectin</b>	<b>P-Selectin</b>
Historical synonyms	CD62L, LAM-1, LECAM-1	CD62E, ELAM-1, LECAM-2	CD62P, GMP-140, LECAM-3
Molecular weight (protein)	42 kDa	64 kDa	86 kDa
Molecular weight (glycosylated)	90 kDa <sup>(3)</sup>	115 kDa <sup>(4)</sup>	140 kDa <sup>(5)</sup>
Number glycosylation sites	7	11	12
Number of consensus repeats	2	6	9
Expressing cells	Leukocytes	Endothelial cells	Endothelial cells, Thrombocytes
Activating molecules	Constitutively expressed	LPS, TNF- $\alpha$ , IL-1	Histamine, Thrombin
Activation time	-	4-6 hrs	10 mins
Target cells	Endothelial cells	Leukocytes	Leukocytes
Target structures	Various Sialomucins, PSGL-1	ESL-1, PSGL-1	PSGL-1

### 1.2. Structural features

Selectins are heavily glycosylated and share an overall sequence identity of approx. 50% with a higher similarity in the domains relevant for ligand binding. Furthermore, selectins share a similar overall architecture (Figure 1): An N-terminal lectin domain (~120 amino acids) points into the vascular lumen where it is responsible for binding interactions. It harbors a structural Ca<sup>2+</sup> ion, which stabilizes the protein conformation and coordinates with the carbohydrate ligand.<sup>(6)</sup> Although not in direct contact with the ligand, the EGF-like domain (~40 amino acids) C-terminally connected to the lectin domain is essential for ligand binding.<sup>(7,8)</sup> A major structural difference between the

### Section III. E-selectin – Introduction

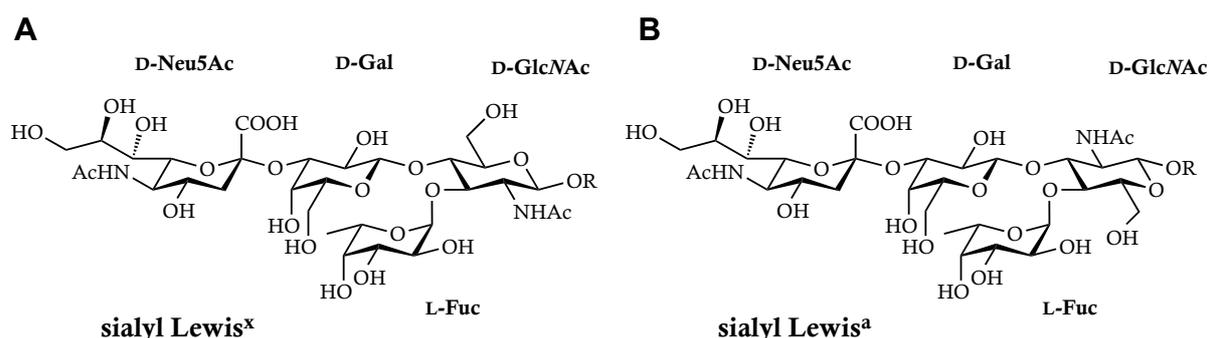
members of the selectin family is the number of short consensus repeat (SCR) domains (~60 amino acids) connecting the EGF-like domain to the transmembrane domain.<sup>(9)</sup> SCR domains are thought to act as spacers between the binding interface and the endothelial surface<sup>(10)</sup>, but they were furthermore found to influence binding affinity and specificity.<sup>(11,12)</sup> However, some publications did not confirm these findings, which are therefore still debatable.<sup>(13)</sup> Selectins are anchored to the cell by a short  $\alpha$ -helical transmembrane region (23 amino acids). A C-terminal cytoplasmic tail (17-35 amino acids) is involved in signal transduction<sup>(14,15)</sup> and interacts with the cytoskeleton.<sup>(16)</sup>



**Figure 1. The structure of selectins and their natural ligands.** The overall architecture of the selectins is conserved. A lectin domain is followed by an epidermal growth factor (EGF)-like domain, a variable number of short consensus repeats, a transmembrane anchor, and a C-terminal cytosolic tail. L-selectin ligands are sulfated sialomucins and PSGL-1 that binds to all selectins. ESL-1 binds to E-selectin. (Picture modified from Vestweber and Blanks (1999)).<sup>(9)</sup>

### 1.3. Selectin ligands

The natural ligands of selectins are expressed on the surface of vascular endothelial cells, leukocytes, and thrombocytes (Figure 1). They are heavily glycosylated proteins which bear sialylated and fucosylated terminal glycan epitopes.<sup>(17)</sup> The common minimal binding motif for all selectins is the tetrasaccharide sialyl Lewis<sup>x</sup> (sLe<sup>x</sup>) and its regioisomer sialyl Lewis<sup>a</sup> (sLe<sup>a</sup>) (Figure 2).<sup>(18,19)</sup>



**Figure 2. Minimal carbohydrate epitopes binding to E-, P-, and L-selectin.** Sialyl Lewis<sup>x</sup> (Neu5Ac(α2-3)Gal(β1-4)[Fuc(α1-3)]GlcNAc) and its regioisomer sialyl Lewis<sup>a</sup> (Neu5Ac(α2-3)Gal(β1-3)[Fuc(α1-4)]GlcNAc). R is the linkage site to the glycan structure of a glycoprotein.

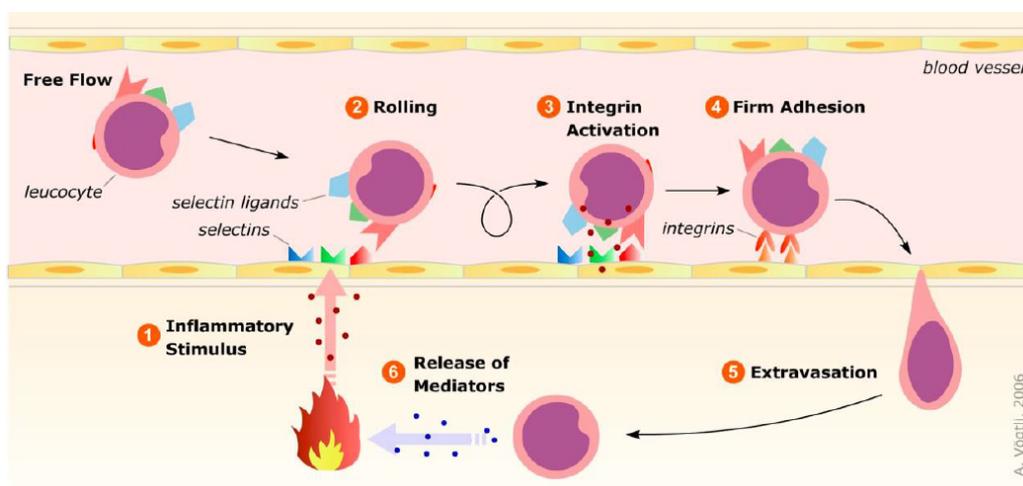
The interaction between the selectins and this minimal binding epitope is very specific but exhibits a rather weak affinity: sLe<sup>x</sup> binds with an affinity of  $0.8 \pm 0.1$  mM to E-selectin,  $7.8 \pm 1.0$  mM to P-selectin and  $3.9 \pm 0.6$  mM to L-selectin.<sup>(20,21)</sup> As an exception for a much higher affinity, *P-selectin glycoprotein ligand-1 (PSGL-1)* presented on leukocytes and thrombocytes and P-selectin show an affinity of in the nanomolar range ( $320$  nM<sup>(22)</sup>;  $826$  nM<sup>(23)</sup>;  $80$  nM<sup>(24)</sup>). Because PSGL-1 harbors multiple glycans presenting sLe<sup>x</sup> the avidity of binding might be increased due to multivalency.<sup>(25)</sup> However, PSGL-1 additionally contains sulfated tyrosines that form interactions to a secondary binding site on P-selectin.<sup>(23,26)</sup> Like P-selectin, L-selectin requires sulfation on its counterreceptors, while for E-selectin the minimal binding epitope sLe<sup>x</sup> is sufficient for binding.<sup>(27,28)</sup> PSGL-1 is, therefore, a ligand of all selectins,<sup>(20)</sup> but the affinity to E-selectin is weaker as it lacks the secondary binding site for sulfated tyrosines.<sup>(29)</sup> The 150 kDa glycoprotein ligand *E-selectin ligand-1 (ESL-1)* is not sulfated and consequently binds only to E-selectin with an affinity of  $60$  μM.<sup>(20,30)</sup> E-selectin is furthermore reported to bind glycosphingolipids on human neutrophils<sup>(31)</sup> and P- and L-selectin are reported to bind to the anticoagulant heparin, which is a heterogeneous mixture of sulfated

glycosaminoglycan.<sup>(32)</sup> For L-selectin currently nine confirmed ligands are known: glycosylation-dependent cell adhesion molecule-1 (GlyCAM-1), mucosal vascular addressin cell adhesion molecule-1 (MAdCAM-1), CD34, podocalyxin, nepmucin, Sgp200, endomucin, endoglycan, and PSGL-1.<sup>(33)</sup> Most of these ligands are sialo-mucins expressed by high endothelial venule (HEV) cells that mediate the homing of lymphocytes to lymph nodes.

### 1.4. The Role of Selectins in Inflammation

#### 1.4.1. Inflammatory cascade

The inflammatory cascade is a *response of the body to cell injuries or infections*. An infection triggers the inflammatory response by pathogen-associated molecular patterns (PAMPs) (e.g. bacterial LPS, viral dsRNA).<sup>(34)</sup> Damage-associated molecular patterns (DAMPs) are molecules located inside the cell in healthy tissue that are released upon injury (e.g. DNA, RNA, ATP).<sup>(35)</sup> PAMPs and DAMPs are recognized by pattern recognition receptors (PRR) on sentinel cells such as macrophages or dendritic cells.<sup>(36)</sup> Activated sentinel cells release *proinflammatory cytokines*<sup>(37)</sup> that cause the five classical signs of an acute inflammation, namely redness, swelling, heat, pain, and loss of function by vasodilatation, increased permeability of the vascular endothelium, and hyperalgesia. Furthermore, these cytokines *initiate the inflammatory cascade* as a part of the non-specific (innate) immune response (Figure 3).<sup>(1)</sup>



**Figure 3. The role of selectins in the inflammatory cascade.** Selectins are exposed on the surface upon an inflammatory stimulus. By initiating the tethering and rolling of leukocytes along vascular endothelial cells they allow leukocytes to reduce their velocity in the bloodstream. Finally, integrins provide firm adhesion of decelerated leukocytes and enable extravasation to the site of inflammation. (Picture from A. Vögtli)

*Leukocytes*, which circulate in the bloodstream, are cells of the immune system with executing functions (e.g. phagocytosis, release of reactive oxygen species) to defend the body against infections. However, to do so they have to extravasate from the bloodstream to the site of inflammation. As a response to the proinflammatory cytokines cell adhesion molecules (CAMs) such as proteins of the selectin and the integrin family are expressed that mediate the process of extravasation.<sup>(38,39)</sup> Initially, *P-selectin*, stored in Weibel-Palade bodies within the vascular endothelial cells, is rapidly transported to the cell surface as a response to the proinflammatory mediators thrombin and histamine.<sup>(40-42)</sup> This process is completed within a few minutes. The expression of *E-selectin* is regulated on the transcription level by the NFκB-pathway and has to be expressed *de novo*. It is induced by lipopolysaccharides (LPS), tumor necrosis factor-α (TNF-α), and interleukin-1β (IL-1β).<sup>(43,44)</sup> The highest expression level of E-selectin on vascular endothelial cells is reached 4-6 hours after the stimulus.<sup>(45,46)</sup> In contrast, *L-selectin* is constitutively expressed on leukocytes.<sup>(47)</sup> It interacts with a broad variety of sialomucins on endothelial cells and furthermore with PSGL-1. This allows leukocytes to interact with other leukocytes (*secondary tethering*)<sup>(33,48)</sup> and to form clusters with thrombocytes.<sup>(49-51)</sup> Interactions between the selectins and their carbohydrate ligands are characterized by fast association and dissociation kinetics to allow leukocytes rolling along the vascular endothelium.<sup>(30)</sup> While L- and P-selectin redundantly mediate rolling at high velocities, E-selectin decelerates to slow motions.<sup>(52)</sup> Together, the selectins decelerate the mobility of the leukocytes from 1-10 mm/s down to 5 μm/s.<sup>(52)</sup> However, it is the *integrin* mediated interaction that finally provides firm adhesion and enables diapedesis.<sup>(53,54)</sup> The interplay between selectins and integrins is crucial since knockout mice of either system suffer from an impaired leukocyte extravasation.<sup>(46)</sup> Leukocytes either transmigrate via a paracellular (cell junctions) or transcellular pathway through the endothelial cell layer to the site of inflammation where they execute their protective function.<sup>(55,56)</sup>

#### 1.4.2. Role of selectins in disease

The proper reaction of the body to an injury or infection is an acute inflammation that leads to a complete cure. An acute inflammation has to be distinguished from a chronic inflammation, which is established in case defense mechanisms are insufficient to eradicate the causative stimulus or an acute inflammation is not downregulated after resolution of the initial trauma. In both forms of inflammation, acute and chronic,

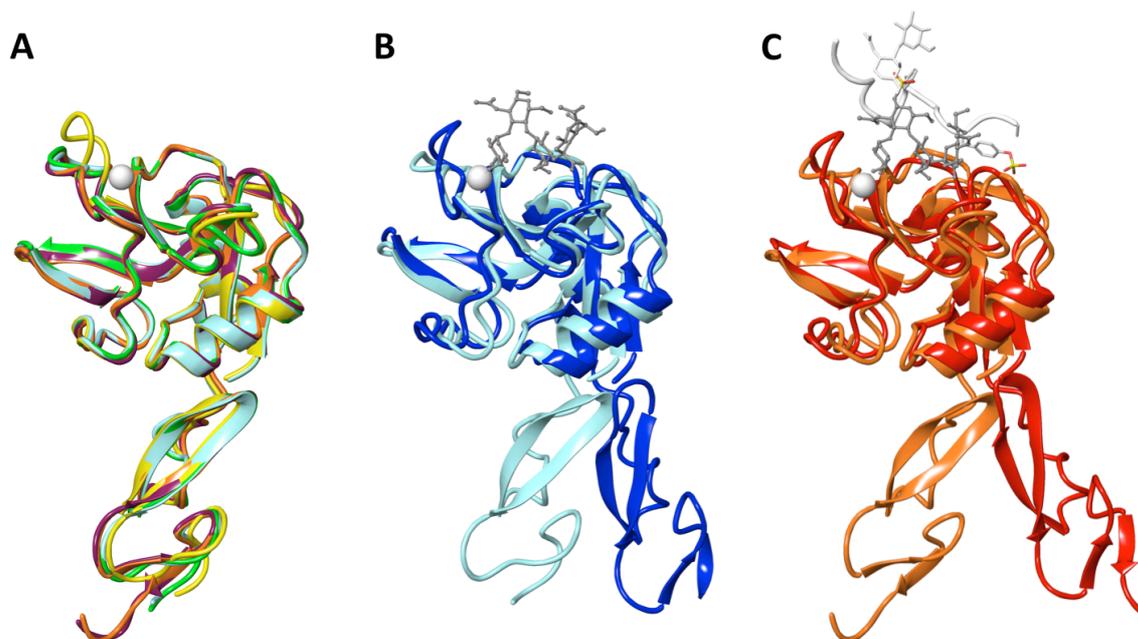
excessive leukocyte extravasation may lead to the destruction of healthy tissue.<sup>(57)</sup> Although selectins are not the cause, their dysregulation during immune reactions can be correlated with a broad variety of diseases: Selectins play a major role in *cardiovascular diseases* like atherosclerosis,<sup>(58)</sup> myocardial infarction,<sup>(59)</sup> hypertension,<sup>(60)</sup> and reperfusion injury.<sup>(61)</sup> Furthermore, they are relevant for *immune diseases* like rheumatoid arthritis<sup>(62)</sup> and asthma bronchiale<sup>(63)</sup> and they are associated with *graft rejection*,<sup>(64,65)</sup> *sickle cell disease*,<sup>(66)</sup> and *carcinoma* cell metastasis.<sup>(67,68)</sup> Therefore, the carbohydrate epitopes sLe<sup>x</sup> and sLe<sup>a</sup> binding selectins are tumor-associated antigens and can be used as predictive biomarkers for the diagnosis of breast and colon cancer.<sup>(57,69)</sup> Selectin-deficient mice have no developmental defects, but they show impaired inflammatory responses.<sup>(70,71)</sup> Similar symptoms could be reproduced by blocking the selectins with antibodies in mice and rats.<sup>(72,73)</sup> Hence, their key role as initiators of leukocyte extravasation make the selectins target of utmost importance for drug discovery.<sup>(74)</sup>

#### 1.5. Structural binding properties

Since Graves *et al.* published the first structure of apo E-selectin in 1994<sup>(75)</sup>, many more selectin structures were solved and provided information of inestimable value for numerous drug discovery programs (Table 2). The selectins share a high sequence identity (61-65% for the lectin and EGF-like domain) and consequently a similar 3-D architecture (Figure 4A). However, upon ligand binding the co-crystallized structures of P- and E-selectin (4CSY, 4C16, and 1G1S) unveil a conformational change of great importance for the biological function of selectins (Figures 4B and 4C), as well as for the rational design of selectin antagonists.

**Table 2. Overview of the selectin structures deposited in the Protein Data Bank (PDB).** Resolution for X-ray structures is given in Ångström (Å).

PDB-Code	Construct	Ligand	Resolution	Remarks	Ref
1ESL	E-selectin (LecEGF)	none	2.0	apo	<sup>(75)</sup>
1G1T	E-selectin (LecEGF)	sLe <sup>x</sup>	1.5	soaked	<sup>(23)</sup>
4CSY	E-selectin (LecEGF SCR2)	sLe <sup>x</sup>	2.4	co-crystal	<sup>(76)</sup>
4C16	E-selectin (LecEGF SCR2)	DS04-115	1.9	co-crystal	<sup>(76)</sup>
1G1Q	P-selectin (LecEGF)	none	2.4	apo	<sup>(23)</sup>
1G1R	P-selectin (LecEGF)	sLe <sup>x</sup>	3.4	soaked	<sup>(23)</sup>
1G1S	P-selectin (LecEGF)	SGP-3 (glycosylated PSGL-1 fragment)	1.9	co-crystal	<sup>(23)</sup>
3CFW	L-selectin (LecEGF)	none	2.2	apo	



**Figure 4. Superimposed selectin structures.** (A) Apo structures of E-selectin (green, PDB-Code: 1ESL),<sup>(75)</sup> P-selectin (purple, PDB-Code: 1G1Q),<sup>(23)</sup> and L-selectin (yellow, PDB-Code: 3CFW) share a high structural similarity to the selectins soaked with sLe<sup>x</sup> [E-selectin (light blue, PDB-Code: 1G1T)<sup>(77)</sup> and P-selectin (orange, PDB-Code: 1G1R)]<sup>(77)</sup>. (B) The co-crystallized structures of E-selectin with sLe<sup>x</sup> (dark blue, PDB-Code: 4CSY)<sup>(76)</sup> and P-selectin with SGP-3 (C) (red, PDB-Code: 1G1S)<sup>(23)</sup> exhibit major structural changes in the binding site and the relative orientation between the lectin domain and the EGF-like domain compared to their soaked analogue. Ligands are shown in grey and the calcium ion in white. Two consensus repeats of 4CSY were removed.

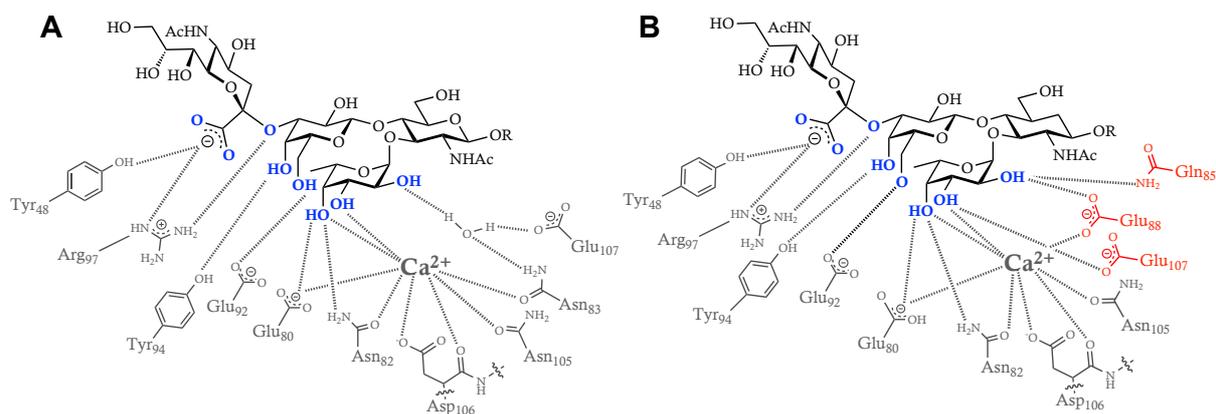
### 1.5.1. E-selectin

Before the first crystal structure of E-selectin in complex with sLe<sup>x</sup> was published, pharmacophores had to be elucidated by systematic *structure-activity relationship* (SAR) studies. Hence, all hydroxyl groups of L-fucose (L-Fuc) were identified as crucial for binding.<sup>(78,79)</sup> On the D-galactose (D-Gal) only the hydroxyl groups in position C-4 and C-6 (4-OH and 6-OH), but not the 2-OH was found to interact with the protein.<sup>(80)</sup> From the sialic acid moiety (Neu5Ac) solely the carboxylate group appeared essential for binding.<sup>(18,79)</sup> Finally, N-acetylglucosamin (D-GlcNAc) does not interact with the protein, but rather serve as a spacer to keep the pharmacophores of D-Gal and L-Fuc in the correct spatial orientation.<sup>(81,82)</sup>

The *crystal structure of E-selectin soaked with sLe<sup>x</sup>* published by Somers *et al.* in 2000<sup>(23)</sup> was very similar to the apo structure previously published by Graves *et al.* six years earlier.<sup>(75)</sup> The bound crystal structure could largely confirm previous findings of the pharmacophores of sLe<sup>x</sup> (Figure 5A):<sup>(83)</sup>

### Section III. E-selectin – Introduction

- $\text{Ca}^{2+}$  coordinates with 3-OH and 4-OH of L-Fuc, and 4-OH furthermore forms hydrogen bonds with Glu80 and Asn82.
- A salt bridge is formed between Arg97 and the carboxylic acid of Neu5Ac, which additionally forms a hydrogen bond with Tyr48.
- 4-OH and 6-OH of D-Gal each are involved in one hydrogen bond each.
- D-GlcNAc does not interact with the protein, but pre-organizes D-Gal and L-Fuc in their bioactive conformation.
- The oxygen in the glycosidic bond between Neu5Ac and D-Gal accepts a hydrogen bond from Arg97.
- The bioactive conformation of  $s\text{Le}^x$  observed in the crystal structure is nearly identical to its conformation in solution.<sup>(84-86)</sup>



**Figure 5. Schematic representation of the interaction between  $s\text{Le}^x$  and E-selectin.** Differences between ligand soaking (A) (PDB-Code 1G1T) and co-crystallization (B) (PDB-code 4CSY) are shown in red. Pharmacophores are highlighted in blue. (Picture modified from R. Preston).

Surprisingly, the hydroxyl group in 2-position of L-Fuc that was found to be crucial for binding by SAR studies, is not involved in a direct interaction with the protein. This discrepancy could only recently be solved when Preston *et al.* (2015)<sup>(76)</sup> published a *co-crystal structure of E-selectin in complex with  $s\text{Le}^x$* , disclosing substantial conformational changes compared to the soaked crystal structure from Somers *et al.* (Figures 4B and 5B). The changes were not only restricted to the binding site, but furthermore the relative position between the lectin domain and the EGF-like domain altered significantly:

- Loop 81-89 in the binding site moves up to 10 Å towards the binding site. Glu88 displaces Asn83 and coordinates with  $\text{Ca}^{2+}$  and forms a

hydrogen bond with 2-OH of L-Fuc. Furthermore, Gln85 and Glu107 shift into hydrogen bonding distance of 2-OH and 3-OH of L-Fuc, respectively.

- The angle between the lectin and the EGF-like domain around the pivot residue Trp1 is 154° and therefore more stretched ('extended conformation') than in the unbound structure, where the angle between the domains is 120° ('bent conformation').
- Two loops (residues 52-72) bridge the loop movements between the binding site and the pivot region extending the overall alterations to a total distance of 35 Å.

It is known that this induced fit is part of a catch-bond mechanism, which is described in section 1.5.3. The conformational change could not be observed in the soaked X-ray structure as the protein in a grown crystal is locked due to close protein contacts.

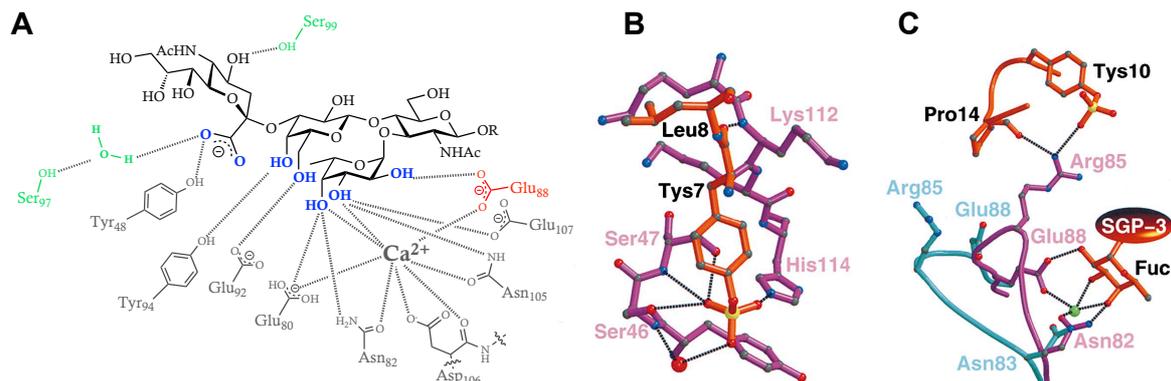
### 1.5.2. P-selectin

The binding sites of E- and P-selectin resemble each other but exhibit some important differences (Figure 6A). Compared to E-selectin, P-selectin has a 10-fold reduced affinity towards sLe<sup>x</sup>.<sup>(21)</sup> However, three sulfated tyrosines on the natural ligand PSGL-1 significantly improve the affinity. Somers et al. (2000)<sup>(23)</sup> therefore expressed an O-glycosylated peptide fragment (SGP-3) of PSGL-1 containing the sulfotyrosine sequence (Tys5, Tys7, Tys10). SGP-3 showed a similar affinity to P-selectin as PSGL-1. Furthermore, both, SGP-3 containing no sulfated tyrosines and SGP-3 containing no glycan revealed a massive drop in affinity. SGP-3 could be co-crystallized with P-selectin (Figures 6B and 6C) and induced a conformational change in a similar manner as observed with E-selectin (Figure 4C).<sup>(83)</sup>

- Arg97 forming a salt bridge to the carboxylic acid of Neu5Ac in E-selectin is substituted by Ser97 in P-selectin and there forms a water-mediated hydrogen bond.
- Lys99 that points away from the binding site in E-selectin is substituted by Ser99, which builds a hydrogen bond to 4-OH of Neu5Ac.

## Section III. E-selectin – Introduction

- Glu88 swings in and simultaneously coordinates with  $\text{Ca}^{2+}$  and forms a hydrogen bond to 2-OH of L-Fuc.
- 13 of totally 19 amino acids of the peptide fragment SGP-3 interact with the protein.
- All three oxygen atoms of the Tys7 sulfate group form direct or water-mediated hydrogen bonds to P-selectin (Figure 6B).



**Figure 6. Schematic representation of the interaction between SGP-3 and P-selectin.** (A) The terminal  $\text{sLe}^x$  motif of O-glycosylated SGP-3 is interacting with the binding site of P-selectin (PDB-Code: 1G1S). A difference to the binding site interacting with soaked  $\text{sLe}^x$  (PDB-Code: 1G1R) is highlighted in red, and differences to E-selectin co-crystallized with  $\text{sLe}^x$  (PDB-Code: 4CSY) are highlighted in green. The interactions of the sulfated tyrosines Tys7 (B) and Tys10 (C) of the protein backbone of SGP-3 contribute beneficially to the binding affinity to P-selectin. (Picture 6B and 6C modified after Somers et al. (2000)<sup>(23)</sup>; reprinted with permission from Elsevier)

The different amino acids in position 97 (Arg97 vs. Ser97) are mainly responsible for the reduced affinity of  $\text{sLe}^x$  to P-selectin, compared to E-selectin. Additional contacts of Pro98 and Ser99 formed to P-selectin cannot compensate for the salt bridge formed in case of E-selectin. Instead, a large cluster of positive electrostatic potential adjacent to the  $\text{sLe}^x$  binding site supports the binding of anionic molecules to P-selectin, such as the sulfotyrosines of PSGL-1. This cluster might also promote the binding of heparin, which was reported to bind P-selectin but not E-selectin.<sup>(32)</sup>

### 1.5.3. Catch-bond mechanism

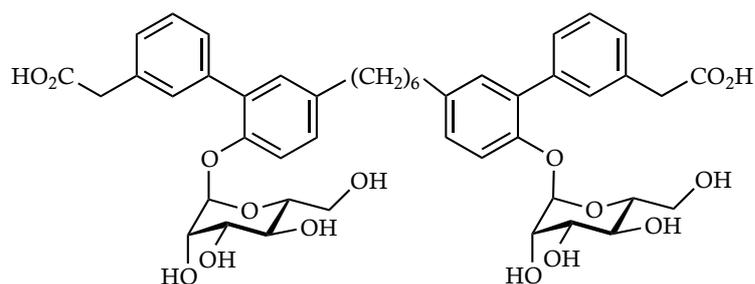
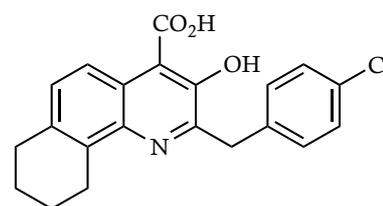
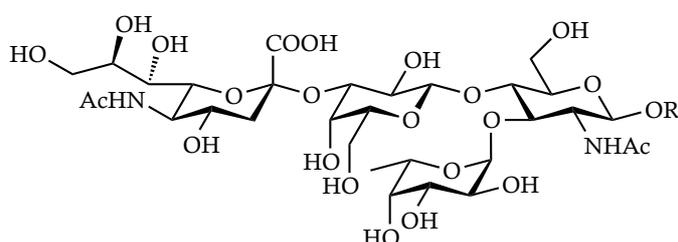
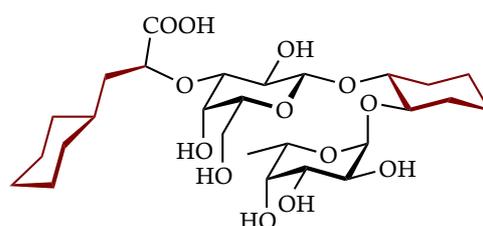
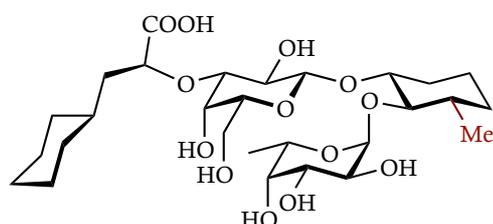
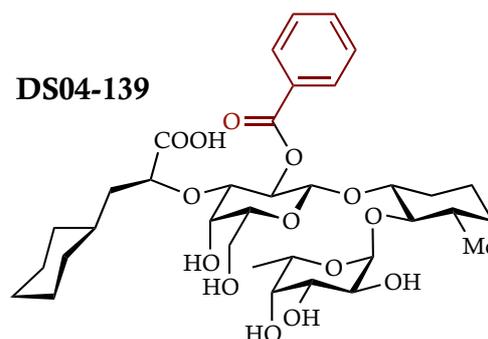
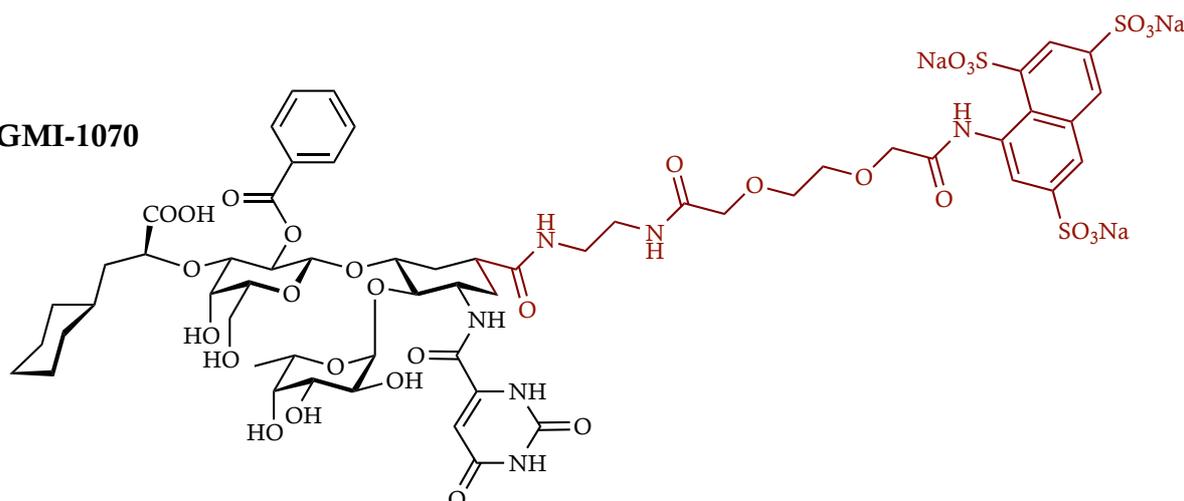
Like FimH, selectins show an increased bond lifetime under tensile force conditions.<sup>(87)</sup> This *catch-bond* behavior could be demonstrated first for P-selectin<sup>(88)</sup>, later for L-selectin<sup>(89)</sup> and E-selectin<sup>(90)</sup> with atomic force microscopy and by performing binding experiments in flow chambers. The *sliding-rebinding model* is an attempt to link the catch-

bond behavior to the existence of a bent and extended selectin conformation observed in crystal structures (Figure 4): Mechanical force is thought to induce the extended conformation aligning the bound selectins along the direction of flow. This allows leukocytes to slide on the endothelial surface by continuously forming transient new contacts with unbound selectins in bent conformation before completely unbinding.<sup>(91)</sup> Other models assume an *allosteric regulation* by the EGF-like domain.<sup>(87,92)</sup> Indeed, the removal of the EGF-like domain prevents P-selectin from binding PSGL-1.<sup>(93)</sup> Moreover, stabilizing the extended conformation of P-selectin increases the affinity to PSGL-1 at low force, but impairs the ability to mediate leukocyte rolling.<sup>(94)</sup> On the basis of allosteric regulation and observations in crystal structures, a *two-state-two-pathway model* was suggested.<sup>(76,95)</sup> It assumes the bent and extended conformation in equilibrium and the ligand able to bind to both of them, although with different rates. Tensile force shifts the equilibrium towards the extended conformation and increases the average bond lifetime.

### 1.6. Selectin Antagonists

A pan-selectin antagonist is the aim of many drug development programs for more than 20 years. However, this is a very ambitious aim, since tight binding of ligands seems against the nature of selectins. Nevertheless, some considerable achievements were made and three candidates are currently in clinical evaluation (Figure 7). The preclinical and clinical selectin antagonists were recently reviewed by Ernst and Magnani.<sup>(74)</sup>

Most approaches started with sLe<sup>x</sup> as lead structure, but other leads (e.g glycyrrhizin<sup>(96)</sup>, efomycin M<sup>(97)</sup>, quinic salicylic acid<sup>(98)</sup>) have been explored as well. Quinic salicylic acid was discovered by a high throughput ELISA screen at Wyeth and was suggested as a non-carbohydrate sLe<sup>x</sup> mimic due to hydroxyl groups similarly oriented as on L-Fuc.<sup>(77,98)</sup> The affinity towards P-selectin was determined to be in the millimolar range.<sup>(77)</sup> Subsequent optimization led to the clinical candidate **PSI-697** with an improved affinity of 125  $\mu$ M and drug-like properties allowing oral application.<sup>(98)</sup> It completed clinical phase II in 2008 but the successful reduction of platelet-monocyte aggregates *ex vivo*<sup>(99)</sup> could not be confirmed *in vivo*.<sup>(100)</sup> An approach to mimic sialyl di-Lewis<sup>x</sup> by biphenyl substituted D-mannose resulted in the divalent molecule **Bimosiamose (TBC1269)** developed by Revotar Biopharmaceuticals AG (Hennigsdorf, DE).<sup>(101)</sup> It was successfully

**TBC1269**

**PSI-697**

**sLe<sup>x</sup>**

**GCP69669A**

**DS04-115**

**DS04-139**

**GMI-1070**


**Figure 7. Chemical structures of selectin antagonists in clinical trials and glycomimetics of sLe<sup>x</sup>.** TBC1269 (Bimosiamose), PSI-697 and GMI-1070 (Rivipansel) are currently in clinical evaluation. Rivipansel is a glycomimetic of sLe<sup>x</sup> and was rationally designed. GCP69669A, DS04-115, and DS04-139 are important intermediate steps towards a low micromolar E-selectin antagonist (beneficial modifications are shown in brown). Linking sulfonated naphthalene to a glycomimetic of sLe<sup>x</sup> leads to Rivipansel.

tested for the treatment of patients with chronic obstructive pulmonary disease (COPD) by inhalation in phase II clinical trials.<sup>(102)</sup> So far the most successful approach was the systematic reduction and substitution of hydrophilic parts of sLe<sup>x</sup> not directly involved in binding interactions. The GlcNAc spacer of sLe<sup>x</sup> between L-Fuc and D-Gal was substituted by cyclohexane-1,2-diol and the highly polar sialic acid was replaced by S-cyclohexylactic acid resulting in **GCP69669** with a 10 fold improved affinity towards E-selectin (IC<sub>50</sub> = 80 μM).<sup>(103)</sup> A further increase in affinity was achieved by stabilizing the core conformation of sLe<sup>x</sup>. Steric repulsion between L-Fuc and the 3'-methyl substituent of the adjacent cyclohexane-1,2-diol (**DS04-115**), as well as an improved pre-organization of the acid orientation by 2'-benzoylation on D-Gal (**DS04-139**) improved the affinity towards E-selectin by a factor of 15.<sup>(104,105)</sup> These findings finally led to the rationally designed pan-selectin antagonist **GMI-1070** (Rivipansel, GlycoMimetics Inc.), where the sLe<sup>x</sup> mimetic was linked to a threefold sulfated naphthalene fragment that is required to increase the affinity towards P- and L-selectin.<sup>(106)</sup> It was successfully tested for the reduction in time to resolution of a vaso-occlusive crisis in patients with sickle cell disease in phase II clinical trials<sup>(107)</sup> and recently entered phase III.

### 2. References

1. Carlos, T.M. & Harlan, J.M. (1994). Leukocyte-endothelial adhesion molecules. *Blood*, 84(7): 2068-2101.
2. Ley, K. & Kansas, G.S. (2004). Selectins in T-cell recruitment to non-lymphoid tissues and sites of inflammation. *Nat. Rev. Immunol.*, 4(5): 325-336.
3. Lasky, L.A., *et al.* (1989). Cloning of a lymphocyte homing receptor reveals a lectin domain. *Cell*, 56(6): 1045-1055.
4. Bevilacqua, M.P., Stengelin, S., Gimbrone, M.A., & Seed, B. (1989). Endothelial leukocyte adhesion molecule 1: an inducible receptor for neutrophils related to complement regulatory proteins and lectins. *Science*, 243(4895): 1160-1165.
5. Johnston, G.I., Cook, R.G., & McEver, R.P. (1989). Cloning of GMP-140, a granule membrane protein of platelets and endothelium: sequence similarity to proteins involved in cell adhesion and inflammation. *Cell*, 56(6): 1033-1044.
6. Erbe, D.V., *et al.* (1992). Identification of an E-selectin region critical for carbohydrate recognition and cell adhesion. *J. Cell Biol.*, 119(1): 215-227.
7. Kansas, G.S., *et al.* (1994). A role for the epidermal growth factor-like domain of P-selectin in ligand recognition and cell adhesion. *J. Cell Biol.*, 124(4): 609-618.
8. Pigott, R., Needham, L.A., Edwards, R.M., Walker, C., & Power, C. (1991). Structural and functional studies of the endothelial activation antigen endothelial leucocyte adhesion molecule-1 using a panel of monoclonal antibodies. *J. Immunol.*, 147(1): 130-135.
9. Vestweber, D. & Blanks, J.E. (1999). Mechanisms that regulate the function of the selectins and their ligands. *Physiol. Rev.*, 79(1): 181-213.
10. Patel, K.D., Nollert, M.U., & McEver, R.P. (1995). P-selectin must extend a sufficient length from the plasma membrane to mediate rolling of neutrophils. *J. Cell Biol.*, 131(6): 1893-1902.
11. Li, S.H., *et al.* (1994). Consensus repeat domains of E-selectin enhance ligand binding. *J. Biol. Chem.*, 269(6): 4431-4437.
12. Tu, L., *et al.* (1996). L-selectin binds to P-selectin glycoprotein ligand-1 on leukocytes: interactions between the lectin, epidermal growth factor, and consensus repeat domains of the selectins determine ligand binding specificity. *J. Immunol.*, 157(9): 3995-4004.
13. Kolbinger, F., *et al.* (1996). The carbohydrate-recognition domain of E-selectin is sufficient for ligand binding under both static and flow conditions. *Biochemistry*, 35(20): 6385-6392.
14. Lorenzon, P., *et al.* (1998). Endothelial cell E-and P-selectin and vascular cell adhesion molecule-1 function as signaling receptors. *J. Cell Biol.*, 142(5): 1381-1391.
15. Hu, Y., Szente, B., Kiely, J.-M., & Gimbrone, M.A. (2001). Molecular events in transmembrane signaling via E-selectin - SHP2 association, adaptor protein complex formation and ERK1/2 activation. *J. Biol. Chem.*, 276(51): 48549-48553.
16. Yoshida, M., *et al.* (1996). Leukocyte adhesion to vascular endothelium induces E-selectin linkage to the actin cytoskeleton. *J. Cell Biol.*, 133(2): 445-455.
17. Varki, A. (1997). Selectin ligands: will the real ones please stand up? *J. Clin. Invest.*, 99(2): 158.
18. Tyrrell, D., *et al.* (1991). Structural requirements for the carbohydrate ligand of E-selectin. *Proc. Natl. Acad. Sci. U. S. A.*, 88(22): 10372-10376.
19. Phillips, M.L., *et al.* (1990). ELAM-1 mediates cell adhesion by recognition of a carbohydrate ligand, sialyl-Lex. *Science*, 250(4984): 1130-1132.
20. Kneuer, C., Ehrhardt, C., Radomski, M.W., & Bakowsky, U. (2006). Selectins—potential pharmacological targets? *Drug Discov. Today*, 11(21): 1034-1040.
21. Poppe, L., Brown, G.S., Philo, J.S., Nikrad, P.V., & Shah, B.H. (1997). Conformation of sLe<sup>x</sup> tetrasaccharide, free in solution and bound to E-, P-, and L-selectin. *J. Am. Chem. Soc.*, 119(7): 1727-1736.
22. Mehta, P., Cummings, R.D., & McEver, R.P. (1998). Affinity and kinetic analysis of P-selectin binding to P-selectin glycoprotein ligand-1. *J. Biol. Chem.*, 273(49): 32506-32513.
23. Somers, W.S., Tang, J., Shaw, G.D., & Camphausen, R.T. (2000). Insights into the molecular basis of leukocyte tethering and rolling revealed by structures of P-and E-selectin bound to SLe<sup>x</sup> and PSGL-1. *Cell*, 103(3): 467-479.
24. Croce, K., Freedman, S.J., Furie, B.C., & Furie, B. (1998). Interaction between soluble P-selectin and soluble P-selectin glycoprotein ligand 1: equilibrium binding analysis. *Biochemistry*, 37(47): 16472-16480.

25. Lawrence, M.B. & Springer, T.A. (1991). Leukocytes roll on a selectin at physiologic flow rates: distinction from and prerequisite for adhesion through integrins. *Cell*, 65(5): 859-873.
26. Wilkins, P.P., Moore, K.L., McEver, R.P., & Cummings, R.D. (1995). Tyrosine sulfation of P-selectin glycoprotein ligand-1 is required for high affinity binding to P-selectin. *J. Biol. Chem.*, 270(39): 22677-22680.
27. Lami, Y., Lasky, L.A., & Rosen, S.D. (1993). Sulphation requirement for GlyCAM-1, an endothelial ligand for L-selectin. *Nature*, 361: 555-557.
28. Hemmerich, S., Butcher, E.C., & Rosen, S.D. (1994). Sulfation-dependent recognition of high endothelial venules (HEV)-ligands by L-selectin and MECA 79, and adhesion-blocking monoclonal antibody. *J. Exp. Med.*, 180(6): 2219-2226.
29. Li, F., *et al.* (1996). Post-translational modifications of recombinant P-selectin glycoprotein ligand-1 required for binding to P-and E-selectin. *J. Biol. Chem.*, 271(6): 3255-3264.
30. Wild, M.K., Huang, M.-C., Schulze-Horsel, U., van der Merwe, P.A., & Vestweber, D. (2001). Affinity, kinetics, and thermodynamics of E-selectin binding to E-selectin ligand-1. *J. Biol. Chem.*, 276(34): 31602-31612.
31. Nimrichter, L., *et al.* (2008). E-selectin receptors on human leukocytes. *Blood*, 112(9): 3744-3752.
32. Nelson, R.M., *et al.* (1993). Heparin oligosaccharides bind L-and P-selectin and inhibit acute inflammation. *Blood*, 82: 3253-3258.
33. Steeber, D., Subramanian, H., Grailer, J., Conway, R., & Storey, T. (2007). L-selectin-mediated leukocyte adhesion and migration. *Adhesion Molecules: Function and Inhibition*, Progress in Inflammation Research, ed Ley K, Birkhäuser Basel. pp. 27-70.
34. Janeway, C.A. (1989). Approaching the asymptote? Evolution and revolution in immunology. *Cold Spring Harb. Symp. Quant. Biol.*, Cold Spring Harbor Laboratory Press, pp. 1-13.
35. Seong, S.-Y. & Matzinger, P. (2004). Hydrophobicity: an ancient damage-associated molecular pattern that initiates innate immune responses. *Nat. Rev. Immunol.*, 4(6): 469-478.
36. Creagh, E.M. & O'Neill, L.A. (2006). TLRs, NLRs and RLRs: a trinity of pathogen sensors that co-operate in innate immunity. *Trends Immunol.*, 27(8): 352-357.
37. Iwasaki, A. & Medzhitov, R. (2004). Toll-like receptor control of the adaptive immune responses. *Nat. Immunol.*, 5(10): 987-995.
38. Stocker, C.J., *et al.* (2000). TNF- $\alpha$ , IL-4, and IFN- $\gamma$  regulate differential expression of P-and E-selectin expression by porcine aortic endothelial cells. *J. Immunol.*, 164(6): 3309-3315.
39. Hynes, R.O. (1992). Integrins: versatility, modulation, and signaling in cell adhesion. *Cell*, 69(1): 11-25.
40. Rondajij, M.G., Bierings, R., Kragt, A., van Mourik, J.A., & Voorberg, J. (2006). Dynamics and plasticity of Weibel-Palade bodies in endothelial cells. *Arterioscler. Thromb. Vasc. Biol.*, 26(5): 1002-1007.
41. Hattori, R., Hamilton, K., Fugate, R., McEver, R., & Sims, P. (1989). Stimulated secretion of endothelial von Willebrand factor is accompanied by rapid redistribution to the cell surface of the intracellular granule membrane protein GMP-140. *J. Biol. Chem.*, 264(14): 7768-7771.
42. Geng, J.-G., *et al.* (1990). Rapid neutrophil adhesion to activated endothelium mediated by GMP-140. *Nature*, 343: 757-760.
43. Wyble, C.W., *et al.* (1997). TNF- $\alpha$  and IL-1 upregulate membrane-bound and soluble E-selectin through a common pathway. *J. Surg. Res.*, 73(2): 107-112.
44. Bevilacqua, M.P., Pober, J.S., Mendrick, D.L., Cotran, R.S., & Gimbrone, M.A. (1987). Identification of an inducible endothelial-leukocyte adhesion molecule. *Proc. Natl. Acad. Sci. U. S. A.*, 84(24): 9238-9242.
45. Bevilacqua, M., Ph. D, Michael P, Nelson, P.D., Richard M, Mannori, M., Ph. D, Gianna, & Cecconi, M., Oliviero (1994). Endothelial-leukocyte adhesion molecules in human disease. *Annu. Rev. Med.*, 45(1): 361-378.
46. Bullard, D.C., *et al.* (1996). Infectious susceptibility and severe deficiency of leukocyte rolling and recruitment in E-selectin and P-selectin double mutant mice. *J. Exp. Med.*, 183(5): 2329-2336.
47. Bevilacqua, M.P. & Nelson, R.M. (1993). Selectins. *J. Clin. Invest.*, 91(2): 379-387.
48. Mitchell, D.J., Li, P., Reinhardt, P.H., & Kubes, P. (2000). Importance of L-selectin-dependent leukocyte-leukocyte interactions in human whole blood. *Blood*, 95(9): 2954-2959.
49. Larsen, E., *et al.* (1989). PADGEM protein: a receptor that mediates the interaction of activated platelets with neutrophils and monocytes. *Cell*, 59(2): 305-312.
50. Theilmeyer, G., *et al.* (1999). Circulating activated platelets assist THP-1 monocytoid/endothelial cell interaction under shear stress. *Blood*, 94(8): 2725-2734.

51. Burger, P.C. & Wagner, D.D. (2003). *Platelet P-selectin facilitates atherosclerotic lesion development*. pp. 2661-2666.
52. Kunkel, E.J. & Ley, K. (1996). Distinct Phenotype of E-Selectin-Deficient Mice E-Selectin Is Required for Slow Leukocyte Rolling In Vivo. *Circ. Res.*, 79(6): 1196-1204.
53. Ley, K., Laudanna, C., Cybulsky, M.I., & Nourshargh, S. (2007). Getting to the site of inflammation: the leukocyte adhesion cascade updated. *Nat. Rev. Immunol.*, 7(9): 678-689.
54. Springer, T.A. (1995). Traffic signals on endothelium for lymphocyte recirculation and leukocyte emigration. *Annu. Rev. Physiol.*, 57(1): 827-872.
55. Carman, C.V. & Springer, T.A. (2004). A transmigratory cup in leukocyte diapedesis both through individual vascular endothelial cells and between them. *J. Cell Biol.*, 167(2): 377-388.
56. Dejana, E. (2006). The transcellular railway: insights into leukocyte diapedesis. *Nat. Cell Biol.*, 8(2): 105-107.
57. Ulbrich, H., Eriksson, E.E., & Lindbom, L. (2003). Leukocyte and endothelial cell adhesion molecules as targets for therapeutic interventions in inflammatory disease. *Trends Pharmacol. Sci.*, 24(12): 640-647.
58. Davies, M.J., *et al.* (1993). The expression of the adhesion molecules ICAM-1, VCAM-1, PECAM, and E-selectin in human atherosclerosis. *J. Pathol.*, 171(3): 223-229.
59. Pellegatta, F., *et al.* (1997). Soluble E-selectin and intercellular adhesion molecule-1 plasma levels increase during acute myocardial infarction. *J. Cardiovasc. Pharmacol.*, 30(4): 455-460.
60. Blann, A.D., Tse, W., Maxwell, S.J., & Waite, M.A. (1994). Increased levels of the soluble adhesion molecule E-selectin in essential hypertension. *J. Hypertens.*, 12(8): 925-928.
61. Takada, M., Nadeau, K.C., Shaw, G.D., Marquette, K.A., & Tilney, N.L. (1997). The cytokine-adhesion molecule cascade in ischemia/reperfusion injury of the rat kidney. Inhibition by a soluble P-selectin ligand. *J. Clin. Invest.*, 99(11): 2682.
62. Koch, A.E., *et al.* (1991). Immunolocalization of endothelial and leukocyte adhesion molecules in human rheumatoid and osteoarthritic synovial tissues. *Lab. Invest.*, 64(3): 313-320.
63. Kobayashi, T., *et al.* (1994). Elevation of serum soluble intercellular adhesion Molecule-1 (sICAM-1) and sE-Selectin levels in bronchial asthma. *Clin. Exp. Immunol.*, 96(1): 110-115.
64. Rosen, S.D. (1999). Endothelial ligands for L-selectin: from lymphocyte recirculation to allograft rejection. *Am. J. Pathol.*, 155(4): 1013-1020.
65. Taylor, P.A., *et al.* (2004). L-Selectin<sup>hi</sup> but not the L-selectin<sup>lo</sup> CD4<sup>+</sup> 25<sup>+</sup> T-regulatory cells are potent inhibitors of GVHD and BM graft rejection. *Blood*, 104(12): 3804-3812.
66. Natarajan, M., Udden, M.M., & McIntire, L. (1996). Adhesion of sickle red blood cells and damage to interleukin-1 beta stimulated endothelial cells under flow in vitro. *Blood*, 87(11): 4845-4852.
67. St Hill, C.A. (2010). Interactions between endothelial selectins and cancer cells regulate metastasis. *Front. Biosci.*, 16: 3233-3251.
68. Kannagi, R., Izawa, M., Koike, T., Miyazaki, K., & Kimura, N. (2004). Carbohydrate-mediated cell adhesion in cancer metastasis and angiogenesis. *Cancer Sci.*, 95(5): 377-384.
69. Ross, R. (1993). The pathogenesis of atherosclerosis: a perspective for the 1990s. *Nature*, 362: 801-809.
70. Briaud, S.A., *et al.* (2001). Leukocyte trafficking and myocardial reperfusion injury in ICAM-1/P-selectin-knockout mice. *Am. J. Physiol. Heart Circ. Physiol.*, 280(1): H60-H67.
71. McEver, R.P., Moore, K.L., & Cummings, R.D. (1995). Leukocyte trafficking mediated by selectin-carbohydrate interactions. *J. Biol. Chem.*, 270(19): 11025-11028.
72. Mulligan, M.S., *et al.* (1991). Role of endothelial-leukocyte adhesion molecule 1 (ELAM-1) in neutrophil-mediated lung injury in rats. *J. Clin. Invest.*, 88(4): 1396.
73. Ley, K., *et al.* (1995). Sequential contribution of L-and P-selectin to leukocyte rolling in vivo. *J. Exp. Med.*, 181(2): 669-675.
74. Ernst, B. & Magnani, J.L. (2009). From carbohydrate leads to glycomimetic drugs. *Nat. Rev. Drug Discov.*, 8(8): 661-677.
75. Graves, B.J., *et al.* (1994). Insight into E-selectin/ligand interaction from the crystal structure and mutagenesis of the lec/EGF domains. *Nature*, 367(6463): 532-538.
76. Preston, R.C., *et al.* (2015). E-selectin ligand complexes adopt an extended high-affinity conformation. *J. Mol. Cell. Biol.*: mjev046.
77. Kaila, N., *et al.* (2005). Quinic Acid Derivatives as Sialyl Lewisx-Mimicking Selectin Inhibitors: Design, Synthesis, and Crystal Structure in Complex with E-Selectin. *J. Med. Chem.*, 48(13): 4346-4357.

78. Ramphal, J.Y., *et al.* (1994). Structure-activity relationships of sialyl Lewis x-containing oligosaccharides. 1. Effect of modifications of the fucose moiety. *J. Med. Chem.*, 37(21): 3459-3463.
79. Brandley, B.K., *et al.* (1993). Structure—function studies on selectin carbohydrate ligands. Modifications to fucose, sialic acid and sulphate as a sialic acid replacement. *Glycobiology*, 3(6): 633-641.
80. Stahl, W., Sprengard, U., Kretzschmar, G., & Kunz, H. (1994). Synthesis of Deoxy Sialyl Lewis<sup>x</sup> Analogues, Potential Selectin Antagonists. *Angew. Chem. Int. Ed.*, 33(20): 2096-2098.
81. Wada, Y., *et al.* (1996). Studies on Selectin Blockers. 2. Novel Selectin Blocker as Potential Therapeutics for Inflammatory Disorders. *J. Med. Chem.*, 39(10): 2055-2059.
82. DeFrees, S.A., Gaeta, F.C., Lin, Y.C., Ichikawa, Y., & Wong, C.H. (1993). Ligand recognition by E-selectin: Analysis of conformation and activity of synthetic monomeric and bivalent sialyl Lewis X analogs. *J. Am. Chem. Soc.*, 115(16): 7549-7550.
83. Preston, R.C. (2014). *Lectins as drug targets : functional, structural, and pharmacological insights into E-selectin and FimH*. PhD University of Basel, Basel.
84. Lin, Y.C., *et al.* (1992). Conformational studies of sialyl Lewis X in aqueous solution. *J. Am. Chem. Soc.*, 114(13): 5452-5454.
85. Ichikawa, Y., *et al.* (1992). Chemical-enzymic synthesis and conformational analysis of sialyl Lewis X and derivatives. *J. Am. Chem. Soc.*, 114(24): 9283-9298.
86. Rutherford, T., Spackman, D., Simpson, P., & Homans, S. (1994). 5 Nanosecond molecular dynamics and NMR study of conformational transitions in the sialyl-Lewis X antigen. *Glycobiology*, 4(1): 59-68.
87. Thomas, W.E., Vogel, V., & Sokurenko, E. (2008). Biophysics of catch bonds. *Biophysics*, 37.
88. Marshall, B.T., *et al.* (2003). Direct observation of catch bonds involving cell-adhesion molecules. *Nature*, 423(6936): 190-193.
89. Sarangapani, K.K., *et al.* (2004). Low force decelerates L-selectin dissociation from P-selectin glycoprotein ligand-1 and endoglycan. *J. Biol. Chem.*, 279(3): 2291-2298.
90. Snook, J. & Guilford, W. (2010). The Effects of Load on E-Selectin Bond Rupture and Bond Formation. *Cell. Mol. Bioeng.*, 3(2): 128-138.
91. Lou, J. & Zhu, C. (2007). A structure-based sliding-rebinding mechanism for catch bonds. *Biophys. J.*, 92(5): 1471-1485.
92. Springer, T.A. (2009). Structural basis for selectin mechanochemistry. *Proc. Natl. Acad. Sci. U. S. A.*, 106(1): 91-96.
93. Kansas, G.S. (1996). Selectins and their ligands: current concepts and controversies. *Blood*, 88(9): 3259-3287.
94. Waldron, T.T. & Springer, T.A. (2009). Transmission of allostery through the lectin domain in selectin-mediated cell adhesion. *Proc. Natl. Acad. Sci. U. S. A.*, 106(1): 85-90.
95. Evans, E., Leung, A., Heinrich, V., & Zhu, C. (2004). Mechanical switching and coupling between two dissociation pathways in a P-selectin adhesion bond. *Proc. Natl. Acad. Sci. U. S. A.*, 101(31): 11281-11286.
96. Rao, B., *et al.* (1994). Sialyl Lewis X mimics derived from a pharmacophore search are selectin inhibitors with anti-inflammatory activity. *J. Biol. Chem.*, 269(31): 19663-19666.
97. Schon, M.P., *et al.* (2002). Efomycine M, a new specific inhibitor of selectin, impairs leukocyte adhesion and alleviates cutaneous inflammation. *Nat. Med.*, 8(4): 366-372.
98. Kaila, N., *et al.* (2007). Synthesis and Biological Evaluation of Quinoline Salicylic Acids As P-Selectin Antagonists. *J. Med. Chem.*, 50(1): 21-39.
99. Chelliah, R., *et al.* (2009). P-selectin antagonism reduces thrombus formation in humans. *J. Thromb. Haemost.*, 7(11): 1915-1919.
100. Japp, A.G., *et al.* (2013). Effect of PSI-697, a Novel P-Selectin Inhibitor, on Platelet–Monocyte Aggregate Formation in Humans. *J. Am. Heart Assoc.*, 2(1): e006007.
101. Kogan, T.P., *et al.* (1998). Novel Synthetic Inhibitors of Selectin-Mediated Cell Adhesion: Synthesis of 1,6-Bis[3-(3-carboxymethylphenyl)-4-(2- $\alpha$ -d-mannopyranosyloxy)phenyl]hexane (TBC1269). *J. Med. Chem.*, 41(7): 1099-1111.
102. Kirsten, A., *et al.* (2011). Efficacy of the pan-selectin antagonist Bimosiamose on ozone-induced airway inflammation in healthy subjects – A double blind, randomized, placebo-controlled, cross-over clinical trial. *Pulm. Pharmacol. Ther.*, 24(5): 555-558.
103. Norman, K.E., Anderson, G.P., Kolb, H.C., Ley, K., & Ernst, B. (1998). Sialyl Lewis<sup>x</sup> (sLe<sup>x</sup>) and an sLe<sup>x</sup> Mimetic, CGP69669A, Disrupt E-Selectin–Dependent Leukocyte Rolling In Vivo. *Blood*, 91(2): 475-483.

### Section III. E-selectin – Introduction

---

104. Schwizer, D., *et al.* (2012). Pre-organization of the Core Structure of E-Selectin Antagonists. *Chem. Eur. J.*, 18(5): 1342-1351.
105. Thoma, G., Magnani, J.L., Patton, J.T., Ernst, B., & Jahnke, W. (2001). Preorganization of the Bioactive Conformation of Sialyl LewisX Analogues Correlates with Their Affinity to E-Selectin. *Angew. Chem. Int. Ed.*, 40(10): 1941-1945.
106. Chang, J., *et al.* (2010). GMI-1070, a novel pan-selectin antagonist, reverses acute vascular occlusions in sickle cell mice. *Blood*, 116(10): 1779-1786.
107. Telen, M.J., *et al.* (2015). Randomized phase 2 study of GMI-1070 in SCD: reduction in time to resolution of vaso-occlusive events and decreased opioid use. *Blood*, 125(17): 2656-2664.

## Manuscript 7

# Microscale Thermophoresis: A Powerful Opportunity for Carbohydrate-Based Drug Discovery

Pascal Zihlmann,<sup>†</sup> Christoph P. Sager,<sup>†</sup> Tobias Mühlethaler,<sup>†</sup> Katrin Lemme,<sup>†</sup>  
Michael Misev,<sup>†</sup> Roland C. Preston,<sup>†</sup> Timothy Sharpe<sup>§</sup> and Beat Ernst<sup>†\*</sup>

<sup>†</sup>University of Basel, Institute of Molecular Pharmacy,  
Klingelbergstr. 50, 4056 Basel, Switzerland

<sup>§</sup> University of Basel, Biophysics Facility,  
Klingelbergstr. 70, 4056 Basel, Switzerland

\*Corresponding author.

Tel.: 0041 (0)61 267 15 51; Fax: 0041 (0)61 267 15 52;  
E-mail: beat.ernst@unibas.ch

Contributions of Pascal Zihlmann:

- Manuscript preparation
- MST and ITC experiments and data analysis



**Abstract**

Carbohydrate-lectin interactions are generally of rather low affinity. As a consequence, high ligand concentrations are required to saturate a lectin and to accurately determine the dissociation constants ( $K_D$ ). Probably the most prominent lectin targets in drug discovery are the selectins. These cell-adhesion macromolecules play an important role in early stages of inflammation by mediating tethering and rolling of leukocytes on vascular endothelial cells and are therefore correlated with a broad spectrum of diseases. Their minimally required binding epitope is the tetrasaccharide sialyl Lewis<sup>x</sup> (sLe<sup>x</sup>) with an affinity in the millimolar range. Nevertheless, it was selected as a lead structure for the development of therapeutic antagonists because extensive random screens did not reveal acceptable hits.

To reduce the amounts of required protein and ligand for the determination of  $K_D$  values an assay using microscale thermophoresis (MST) was developed. The method is based on the phenomenon that bound molecules move with different velocity along a temperature gradient than unbound molecules. Here, we compare  $K_D$  values determined by MST and isothermal titration calorimetry (ITC), which is the gold standard method for the evaluation of molecular interactions. Both methods were found to be in good agreement. Furthermore, the effect of DMSO – a frequently used co-solvent for solubility reasons - on the binding properties of E-selectin was investigated. Surprisingly, DMSO was found to improve the affinity of E-selectin ligands.

### Introduction

In recent years, lectins, such as selectins, siglecs or DC-SIGN have received increasing attention as drug targets.<sup>(1)</sup> They play important roles in biological processes like cell adhesion,<sup>(2)</sup> signal transduction,<sup>(3)</sup> cell recognition,<sup>(4,5)</sup> or inflammation<sup>(2)</sup>. Among the lectins, selectins are the most extensively studied family. As a response to an injury, pro-inflammatory cytokines initiate the presentation of selectins on the surface of endothelial cells (E- and P-selectin) and leukocytes (L-selectin).<sup>(6)</sup> Selectins have been found to recognize the carbohydrate motifs sialyl Lewis<sup>x</sup> (sLe<sup>x</sup>) and sialyl Lewis<sup>a</sup> (sLe<sup>a</sup>) as part of the glycan of their physiological glycoprotein ligands.<sup>(7,8)</sup> This rather weak interaction mediates the initial tethering and rolling of leukocytes on endothelial surfaces, which is followed by firm adhesion and transmigration to the injured tissue.<sup>(9)</sup> Dysregulation of selectins by immune reactions could be correlated with a variety of cardiovascular diseases (e.g. atherosclerosis,<sup>(10)</sup> myocardial infarction,<sup>(11)</sup> reperfusion injury<sup>(12)</sup>), immune diseases (e.g. rheumatoid arthritis,<sup>(13)</sup> psoriasis,<sup>(14)</sup> asthma bronchiale<sup>(15)</sup>), metastasis<sup>(16)</sup> and sickle cell disease.<sup>(17)</sup> Hence, its role as an initial mediator of the inflammatory cascade makes E-selectin a valuable target for drug discovery.<sup>(18,19)</sup>

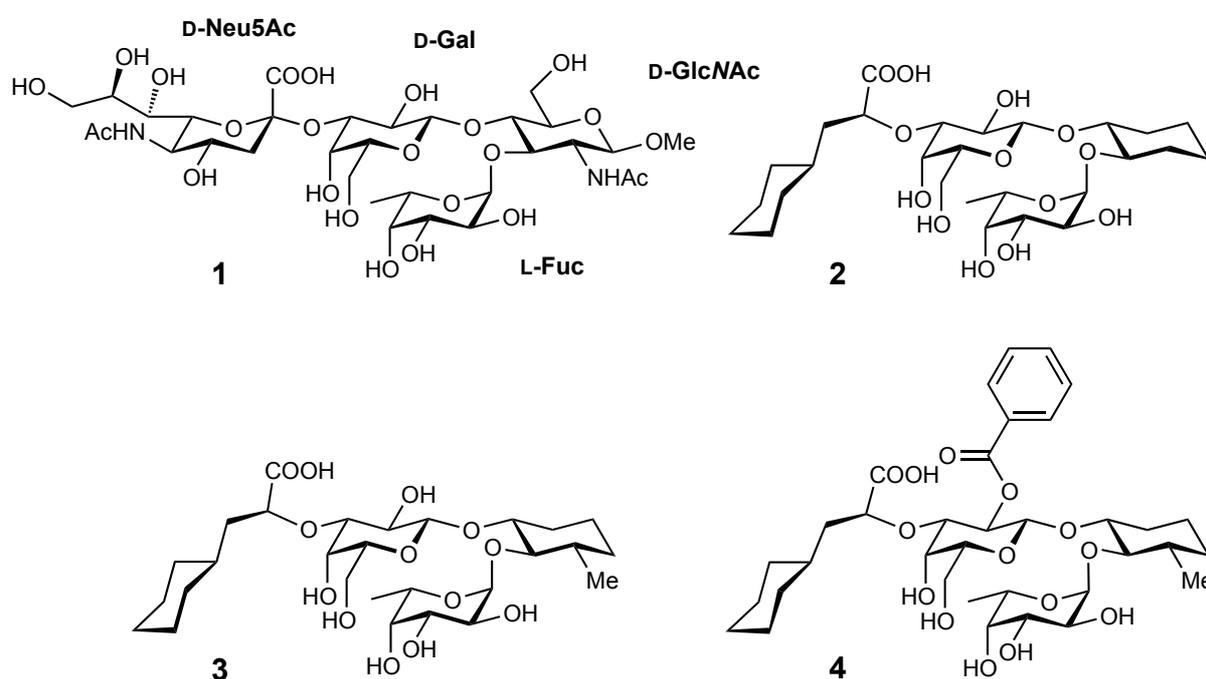
Lectin-carbohydrate interactions are characterized by complex hydrogen bond networks formed by the hydroxyl groups on the carbohydrate residues.<sup>(20)</sup> The polar hydroxyl groups establish strong interactions with water molecules. Hence, the interaction of a carbohydrate with its lectin target is accompanied by a high enthalpic penalty for water displacement.<sup>(21)</sup> As a consequence, carbohydrate-lectin interactions tend to be highly specific but at the same time in average weaker than biomolecular interactions are in general.<sup>(22)</sup> Because the determination of affinities of weak interactions requires high sample concentrations it is important to run the binding experiment at the lowest possible volume. Microscale thermophoresis (MST) is a new method with low sample consumption and short measurement time, allowing the determination of the dissociation constants ( $K_D$ ).<sup>(23)</sup> Therefore, the movement of molecules along a temperature gradient is monitored. An infrared laser induces this temperature gradient within capillaries containing the interacting molecules. Their movement is monitored by a change in fluorescence requiring labeling or intrinsic fluorescence of one of the interaction partners. The velocity of the molecule moving in the temperature gradient

depends on its size, charge, and solvation.<sup>(24)</sup> Upon a binding interaction of a ligand and a protein these parameters are changed.

In this study, we evaluated MST as a new method for lead optimization of E-selectin antagonists by a comparison with results achieved by isothermal titration calorimetry (ITC). Thereby, we additionally found the widely used co-solvent dimethylsulfoxide (DMSO) to have a significant influence on the  $K_D$ . Several groups reported the influences of DMSO on proteins, but to our knowledge, only two publications are examining the influence on affinity.<sup>(25,26)</sup> Both reported a decrease in affinity considering DMSO to act as a competing ligand. Here, we report the first protein-ligand interaction with an improved affinity in the presence of DMSO.

## Results

**Measurements of reference compounds with ITC.** The four reference compounds (Figure 1) represent a typical carbohydrate lead optimization process in drug discovery starting from the natural binding epitope sLe<sup>x</sup> (**1**) with a binding affinity in the millimolar range. Rational substitution of the polar non-binding parts of the molecule leads to compound **2** with a more than 10 fold improved affinity (Table 1).<sup>(27)</sup> A further increase in affinity is achieved by stabilizing the core conformation of sLe<sup>x</sup>. Steric repulsion between L-Fuc and the 3'-methyl substituent of the adjacent cyclohexane-1,2-diol (**3**),<sup>(28)</sup> as well as a hydrophobic collapse between the cyclohexane on the *S*-cyclohexylactic acid site and the benzoate in the 2'-position of the D-Gal moiety (**4**)<sup>(29,30)</sup> pre-organizes the acid in its bioactive conformation and improves the affinity towards E-selectin by a factor of 15 (Table 1). Compounds 1-3 were previously studied by ITC and these data are already published.<sup>(31)</sup> Since ITC is the “gold standard” for the evaluation of molecular interaction, we used it to validate the results obtained by MST.<sup>(32)</sup>



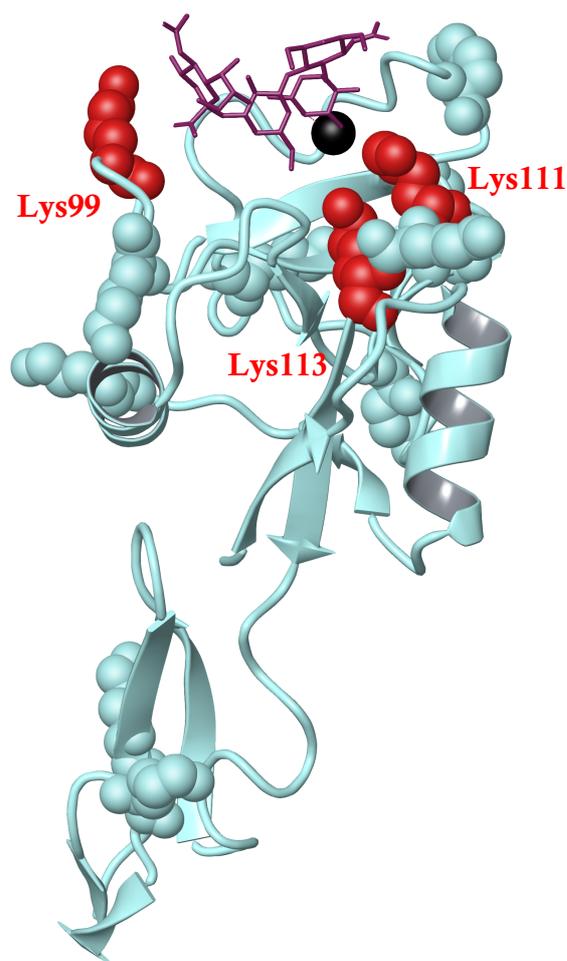
**Figure 1. Structures of the reference compounds tested with ITC and MST for their affinity to E-selectin.** Sialyl Lewis<sup>x</sup> (**1**) is the natural binding epitope, while ligands **2-4** represent rationally designed antagonists with improved binding affinity to E-selectin.

**Table 1. Comparison of dissociation constants achieved by ITC and MST in the presence or absence of DMSO.** Standard errors and 95% confidence intervals are reported in Table S1 (MST) and Table S2 (ITC).

Compound	ITC (0% DMSO)		MST (0% DMSO)			MST (10% DMSO)		
	$K_D$ [ $\mu\text{M}$ ]	$rK_D$	$K_{D \text{ app}}$ [ $\mu\text{M}$ ]	$rK_{D \text{ app}}$	Fluorescence change [%]	$K_{D \text{ app}}$ [ $\mu\text{M}$ ]	$rK_{D \text{ app}}$	Fluorescence change [%]
1	778	1.000	695	1.000	5.14	392	1.000	9.71
2	60.7	0.078	57.1	0.082	5.58	35.7	0.091	11.1
3	17.8	0.022	12.9	0.019	5.59	8.69	0.022	11.2
4	3.99	0.005	3.93	0.006	5.19	2.65	0.007	8.34

**Construct.** Attempts to perform MST experiments with a full-length construct of E-selectin containing the lectin domain, the EGF-like domain, six short consensus repeats, and the Fc part of an IgG (E-selectin<sub>LEC6</sub>-IgG<sub>Fc</sub>, previously used for the ITC measurements) failed, presumably due to the high molecular weight. This construct is heavily glycosylated and forms dimers resulting in a molecular weight of approx. 300 kDa. Hence, it could not be sufficiently mobilized in the temperature gradient of the MST device. We therefore expressed a smaller construct consisting of the lectin domain, the EGF-like domain, and two short consensus repeats (E-selectin<sub>LEC2</sub>). This construct is monomeric and has a molecular weight of approx. 60 kDa.<sup>(33)</sup> Performing ITC measurements revealed that antagonist **4** binds to both constructs with the same affinity (Table 2), but the smaller E-selectin<sub>LEC2</sub> construct has an improved thermophoresis compared to E-selectin<sub>LEC6</sub>-IgG<sub>Fc</sub>.

**Labeling procedure.** Following the standard labeling procedure recommended for the amine reactive protein labeling kit resulted in a strongly fluorescent protein. However, affinities of compounds **2** and **4** measured by MST differed by a constant factor of 3 compared to the ITC measurements (Table 2). Although not directly interacting with the ligand, three lysines (Lys99, Lys111, Lys113) are located in close proximity to the binding site (PDB: 4C16) (Figure 2). Labeling of one of these lysines could influence the lectin-carbohydrate interaction. To prevent labeling of lysines close to the binding site, E-selectin was pre-incubated with antagonist **4** prior to the labeling procedure. After the labeling reaction, free dye and ligand were removed by dialysis. As a result, labeled E-selectin was obtained delivering comparable affinity data by MST and ITC (Table 2).



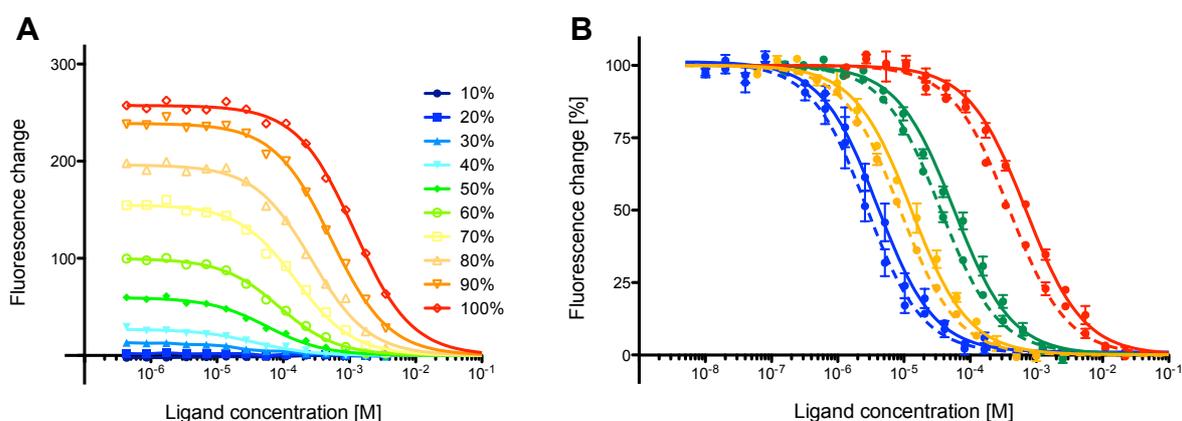
**Figure 2. Crystal structure of E-selectin bound to compound 1.** Lys99, Lys111, and Lys113 (highlighted in red) are pointing towards the binding site and potentially influence the binding affinity when labeled. Compound 1 is shown in purple, the  $\text{Ca}^{2+}$  ion is depicted by a black sphere, and the short consensus repeats were removed for clarity (PDB-Code: 4CSY).<sup>(33)</sup>

**Table 2. Comparison of different protein constructs used for ITC and MST.** Both, the E-selectin<sub>LEC6</sub>-IgG<sub>Fc</sub> and the E-selectin<sub>LEC2</sub>, show comparable affinity to reference compound 4. The protein labeling method is crucial for the validity of the MST assay. Standard errors and 95% confidence intervals are reported in Table S2 (ITC) and Table S3 (MST).

Measuring technique	E-selectin construct	Labeling procedure	$K_D$ [ $\mu\text{M}$ ] Compound 2	$K_D$ [ $\mu\text{M}$ ] Compound 4
ITC	LEC6-IgG <sub>Fc</sub>	Unlabeled	60.7	3.99
ITC	LEC2	Unlabeled	n.d.	3.58
MST	LEC2	Unblocked binding site	166	12.2
MST	LEC2	Blocked binding site	57.1	3.93

**Laser power.** Laser power is an adjustable parameter in the setting of MST measurements and defines the temperature gradient established in the capillaries containing the interacting molecules. Evaluating the optimal laser power for the

interaction of antagonist **2** with E-selectin revealed a strong dependence of the apparent  $K_D$  ( $K_{D,app}$ ) on the temperature gradient (Figure 3A). Most likely, this dependence is caused by a combination of convection and thermal diffusion leading to local accumulations of ligand. The effect is comparable to the ‘Clusius tube’ effect, which was described by Baaske *et al.* (2007).<sup>(34)</sup> Suitable measuring conditions are encountered using lower laser powers as the accumulation grows exponentially with an increasing temperature difference (Table S4). The conditions at a laser power of 50% are a good trade-off between a reduced accumulation effect (lower laser power) and an improved signal to noise ratio (higher laser power).

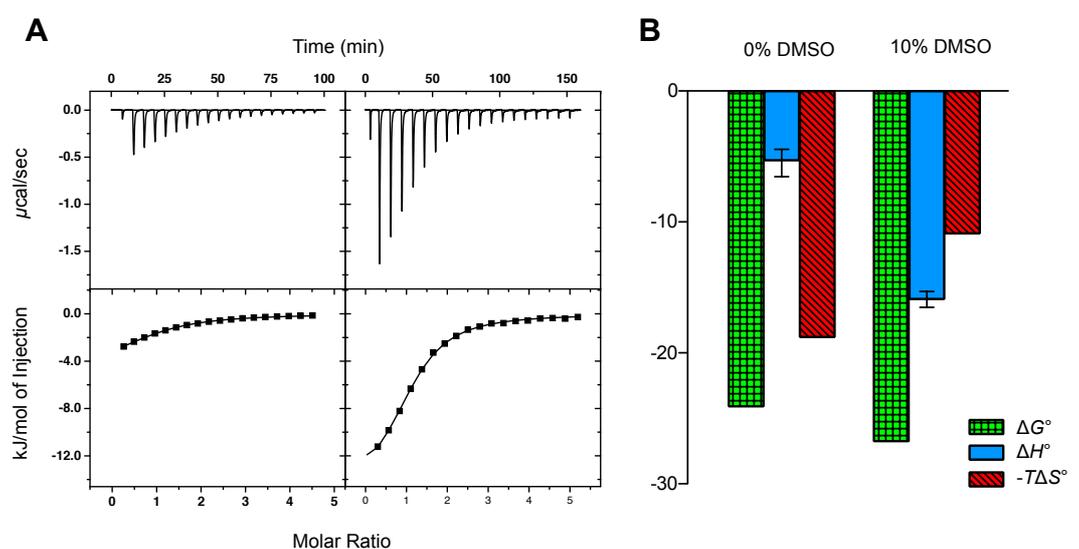


**Figure 3. Apparent binding constants ( $K_{D,app}$ ) of E-selectin<sub>LEC2</sub> interacting with compound **2** in dependence of the laser power and the presence of DMSO.** (A) A temperature dependent accumulation effect is observed at laser power  $\geq 40\%$ . (B) The presence of 10% DMSO improves the binding affinity for all four antagonists (**1** in red, **2** in green, **3** in yellow, and **4** in blue). Best-fit values are reported in Table S4 (Laser power) and Table S1 (DMSO).

**Validation of MST results.** With the E-selectin<sub>LEC2</sub> construct applying a laser power of 50%, highly reliable affinity data characterized by a small standard deviation were obtained by MST (Table S1). For validation, the  $K_{D,app}$  values determined by MST were compared to those determined by ITC (Table 1). In the absence of DMSO, the absolute values of both methods are in very good agreement. In the presence of DMSO, the  $K_D$  values varied significantly, however, the relative affinities ( $rK_D$ ) correspond well.

**DMSO improves the signal to noise ratio and the binding affinity.** The influence of DMSO was of interest since it is frequently used as a co-solvent for substances with low solubility. Thereby two notable effects were observed. First, the fluorescence change (signal to noise ratio) was beneficially increased in the presence of DMSO (Table 1).

Second, the affinity of E-selectin<sub>LEC2</sub> to all tested ligands was significantly increased (by a factor of about 1.5 to 1.8) in the presence of 10% DMSO (Table 1, Figure 3B). Furthermore, up to 20% DMSO the affinities for compound **2** are constantly increasing (Table S5). To exclude an MST specific effect due to unspecific changes in the solvation shell of the protein, an ITC measurement was performed to confirm and furthermore revealed a dramatic change in the thermodynamic fingerprint (Table 3, Figure 4): The former mainly entropy driven interaction becomes enthalpy driven in the presence of DMSO.



**Figure 4. Thermodynamic fingerprint for antagonist 2 binding to E-selectin<sub>LEC6</sub>-IgG<sub>Fc</sub> in presence or absence of DMSO.** (A) Enthalpograms of antagonist **2** in the absence (left) and presence of 10% (right) DMSO. (B) The thermodynamic fingerprint reveals a dramatic change in enthalpy and entropy.

**Table 3. Thermodynamic parameters of compound 2 binding to E-Selectin in presence or absence of DMSO.** A substantial enthalpy-entropy compensation leads only a minor effect on the free energy of binding. 95% confidence intervals are reported in parentheses.

	$K_D$ [ $\mu$ M]	$\Delta G^\circ$ [kJ/mol]	$\Delta H^\circ$ [kJ/mol]	$-T\Delta S^\circ$ [kJ/mol K]
<b>0% DMSO</b>	60.7 (43.4 – 84.8)	-24.1	-5.3 (-4.5 - -6.5)	-18.8
<b>10% DMSO</b>	20.7 (18.8 – 22.8)	-26.7	-15.9 (-15.3 - -16.5)	-10.9

## Discussion

Although sialyl Lewis<sup>x</sup> (**1**) binds to all selectins<sup>(35)</sup> with affinities in the millimolar range, it was selected as lead structure for the search of glycomimetic antagonists.<sup>(36)</sup> Knowledge of its pharmacophores allowed glycomimetics with largely improved affinities in the low micromolar range.<sup>(29,30,37)</sup> Still, for ITC measurements in this affinity range, high ligand and protein concentrations are required. This can be problematic and expensive for synthetically cumbersome ligands and/or protein production. Since for obtaining an accurate  $K_D$  ligand concentrations cannot be reduced, the total volume used in a binding assay has to be diminished.

Therefore, MST became the preferred assay format in our research group for several reasons. First, only low amounts of protein and ligand are required. Compared to ITC, ligand consumption could be reduced by a factor of ten and protein consumption was even lowered by more than 10'000 times. Second, high reliability of this assay format allows the comparison of very small changes, which is often decisive in lead optimization. Third, short measuring times and the straightforward measuring procedure enable a more project orientated approach enabling immediate feedback on activities of newly synthesized compounds.

However, drawbacks such as ligand accumulation at higher laser powers and the requirement for protein labeling must not be ignored and have to be solved in advance. Furthermore, MST seems limited for macromolecules with high molecular weights. For E-selectin<sub>LEC2</sub>, approaches like label-free MST, specific N-terminal labeling, or labeling of a free cysteine were explored but not suitable. We could demonstrate that blocking the binding site with a competitive ligand is a simple and efficient way to prevent labeling next to the binding site.

Since DMSO is often used for solubility reasons, its influence on affinities was studied. We discovered that DMSO increases the affinity of E-selectin-ligand interactions for all tested cases by a constant factor. Nevertheless, relative affinities did not differ in the presence or absence of DMSO and therefore  $K_D$  values of related ligands can still be used for ranking compounds in cases the use of DMSO cannot be avoided. However, the fact that DMSO increases the affinity of E-selectin ligands is remarkable and requires further

investigation, especially since to our knowledge no publication reports an improved  $K_D$  in the presence of DMSO. Tjenberg *et al.* (2006)<sup>(25)</sup> and Cubrilovic & Zenobi (2013)<sup>(26)</sup> investigated the influence of DMSO on protein-ligand interactions, but both reported a decrease in affinity. They speculated that DMSO may act as a competitive binder and furthermore considered a protein destabilization and a thereof resulting destruction of the binding pocket as possible explanations. Yet, none of these explanations is consistent with our observations. To rule out an MST method specific effect, an insight into the thermodynamics of the binding reaction in the presence of DMSO was gained by ITC measurements. A spectacular change of the thermodynamic fingerprint was observed. In the presence of DMSO, the enthalpy term became much more beneficial and the gain in entropy decreased significantly. We speculate, that DMSO alters the hydrogen bond network of water molecules in the binding site of unbound E-selectin or around the ligands. Hence, binding becomes enthalpically more favorable as a consequence of the decreased desolvation penalty. On the other hand, displacing fewer water molecules to the bulk is associated with a reduced gain in entropy. This example demonstrates the interplay of ITC and MST where MST was used as a fast standard method to determine the  $K_D$  of an interaction and the more laborious ITC was used as a reference method to answer questions, which require in-depth information.

In summary, we have shown the application of MST to the lectin E-selectin, established a simple way to overcome the obstacle of labeled lysines in the vicinity of the binding site and report a functional E-selectin construct that is suitable for MST. By comparing the  $K_D$  of four reference compounds with the affinities determined by ITC the MST assay was validated. Furthermore, we demonstrated that a combination of MST and ITC might help to overcome drawbacks of both individual methods, which entails valuable synergies and guarantees opportunity for carbohydrate-based drug discovery.

## References

1. Ernst, B. & Magnani, J.L. (2009). From carbohydrate leads to glycomimetic drugs. *Nat. Rev. Drug Discov.*, 8(8): 661-677.
2. Ley, K., Laudanna, C., Cybulsky, M.I., & Nourshargh, S. (2007). Getting to the site of inflammation: the leukocyte adhesion cascade updated. *Nat. Rev. Immunol.*, 7(9): 678-689.
3. Aplin, A.E., Howe, A., Alahari, S.K., & Juliano, R.L. (1998). Signal transduction and signal modulation by cell adhesion receptors: the role of integrins, cadherins, immunoglobulin-cell adhesion molecules, and selectins. *Pharmacol. Rev.*, 50(2): 197-263.
4. Sharon, N. & Lis, H. (1995). Lectins-proteins with a sweet tooth: functions in cell recognition. *Essays Biochem.*, 30: 59-75.
5. Weis, W.I. & Drickamer, K. (1996). Structural basis of lectin-carbohydrate recognition. *Annu. Rev. Biochem.*, 65(1): 441-473.
6. Bevilacqua, M.P., Pober, J.S., Mendrick, D.L., Cotran, R.S., & Gimbrone, M.A. (1987). Identification of an inducible endothelial-leukocyte adhesion molecule. *Proc. Natl. Acad. Sci. U. S. A.*, 84(24): 9238-9242.
7. Phillips, M.L., *et al.* (1990). ELAM-1 mediates cell adhesion by recognition of a carbohydrate ligand, sialyl-Lex. *Science*, 250(4984): 1130-1132.
8. Walz, G., Aruffo, A., Kolanus, W., Bevilacqua, M., & Seed, B. (1990). Recognition by ELAM-1 of the sialyl-Lex determinant on myeloid and tumor cells. *Science*, 250(4984): 1132-1135.
9. Bevilacqua, M.P. & Nelson, R.M. (1993). Selectins. *J. Clin. Invest.*, 91(2): 379-387.
10. Ross, R. (1999). Atherosclerosis-an inflammatory disease. *N. Engl. J. Med.*, 340(2): 115-126.
11. Pellegatta, F., *et al.* (1997). Soluble E-selectin and intercellular adhesion molecule-1 plasma levels increase during acute myocardial infarction. *J. Cardiovasc. Pharmacol.*, 30(4): 455-460.
12. Weyrich, A.S., Ma, X.Y., Lefer, D.J., Albertine, K.H., & Lefer, A.M. (1993). In vivo neutralization of P-selectin protects feline heart and endothelium in myocardial ischemia and reperfusion injury. *J. Clin. Invest.*, 91(6): 2620-2629.
13. Cronstein, B.N. & Weissmann, G. (1993). The adhesion molecules of inflammation. *Arthritis Rheum.*, 36(2): 147-157.
14. Schon, M.P., Drewniok, C., & Boehncke, W.H. (2004). Targeting selectin functions in the therapy of psoriasis. *Curr. Drug Targets Inflamm. Allergy*, 3(2): 163-168.
15. Gundel, R.H., *et al.* (1991). Endothelial leukocyte adhesion molecule-1 mediates antigen-induced acute airway inflammation and late-phase airway obstruction in monkeys. *J. Clin. Invest.*, 88(4): 1407-1411.
16. Kannagi, R., Izawa, M., Koike, T., Miyazaki, K., & Kimura, N. (2004). Carbohydrate-mediated cell adhesion in cancer metastasis and angiogenesis. *Cancer Sci.*, 95(5): 377-384.
17. Matsui, N.M., *et al.* (2001). P-selectin mediates the adhesion of sickle erythrocytes to the endothelium. *Blood*, 98(6): 1955-1962.
18. Ulbrich, H., Eriksson, E.E., & Lindbom, L. (2003). Leukocyte and endothelial cell adhesion molecules as targets for therapeutic interventions in inflammatory disease. *Trends Pharmacol. Sci.*, 24(12): 640-647.
19. Barthel, S.R., Gavino, J.D., Descheny, L., & Dimitroff, C.J. (2007). Targeting selectins and selectin ligands in inflammation and cancer. *Expert Opin. Ther. Targets*, 11(11): 1473-1491.
20. Quioco, F.A. (1989). Protein-carbohydrate interactions: basic molecular features. *Pure Appl. Chem.*, 61(7): 1293-1306.
21. Cabani, S., Gianni, P., Mollica, V., & Lepori, L. (1981). Group contributions to the thermodynamic properties of non-ionic organic solutes in dilute aqueous solution. *J. Solution Chem.*, 10(8): 563-595.
22. Lundquist, J.J. & Toone, E.J. (2002). The Cluster Glycoside Effect. *Chem. Rev.*, 102(2): 555-578.
23. Jerabek-Willemsen, M., Wienken, C.J., Braun, D., Baaske, P., & Duhr, S. (2011). Molecular interaction studies using microscale thermophoresis. *Assay Drug Dev. Technol.*, 9(4): 342-353.
24. Duhr, S. & Braun, D. (2006). Why molecules move along a temperature gradient. *Proc. Natl. Acad. Sci. U. S. A.*, 103(52): 19678-19682.
25. Tjernberg, A., Markova, N., Griffiths, W.J., & Hallen, D. (2006). DMSO-related effects in protein characterization. *J. Biomol. Screen.*, 11(2): 131-137.
26. Cubrilovic, D. & Zenobi, R. (2013). Influence of dimethylsulfoxide on protein-ligand binding affinities. *Anal. Chem.*, 85(5): 2724-2730.

27. Norman, K.E., Anderson, G.P., Kolb, H.C., Ley, K., & Ernst, B. (1998). Sialyl Lewis<sup>x</sup> (sLe<sup>x</sup>) and an sLe<sup>x</sup> Mimetic, CGP69669A, Disrupt E-Selectin-Dependent Leukocyte Rolling In Vivo. *Blood*, 91(2): 475-483.
28. Schwizer, D., *et al.* (2012). Pre-organization of the Core Structure of E-Selectin Antagonists. *Chem. Eur. J.*, 18(5): 1342-1351.
29. Schwizer, D., *et al.* (2012). Pre-organization of the core structure of E-selectin antagonists. *Chem. Eur. J.*, 18(5): 1342-1351.
30. Thoma, G., Magnani, J.L., Patton, J.T., Ernst, B., & Jahnke, W. (2001). Preorganization of the Bioactive Conformation of Sialyl LewisX Analogues Correlates with Their Affinity to E-Selectin. *Angew. Chem. Int. Ed.*, 40(10): 1941-1945.
31. Binder, F.P., Lemme, K., Preston, R.C., & Ernst, B. (2012). Sialyl Lewis<sup>x</sup>: A “Pre-Organized Water Oligomer”? *Angew. Chem. Int. Ed.*, 51(29): 7327-7331.
32. Cooper, M.A. (2003). Label-free screening of bio-molecular interactions. *Anal. Bioanal. Chem.*, 377(5): 834-842.
33. Preston, R.C., *et al.* (2015). E-selectin ligand complexes adopt an extended high-affinity conformation. *J. Mol. Cell. Biol.*: mjev046.
34. Baaske, P., *et al.* (2007). Extreme accumulation of nucleotides in simulated hydrothermal pore systems. *Proc. Natl. Acad. Sci. U. S. A.*, 104(22): 9346-9351.
35. Foxall, C., *et al.* (1992). The three members of the selectin receptor family recognize a common carbohydrate epitope, the sialyl Lewis(x) oligosaccharide. *J. Cell Biol.*, 117(4): 895-902.
36. Simanek, E.E., McGarvey, G.J., Jablonowski, J.A., & Wong, C.H. (1998). Selectin-Carbohydrate Interactions: From Natural Ligands to Designed Mimics. *Chem. Rev.*, 98(2): 833-862.
37. Kolb, H.C. & Ernst, B. (1997). Development of tools for the design of selectin antagonists. *Chem. Eur. J.*, 3(10): 1571-1578.

## Supporting Information

**Table S1. Best-fit values and standard errors of compounds 1-4 binding to E-selectin<sub>LEC2</sub> in the absence and presence of 10% DMSO measured by MST.** For each compound a global analysis was performed over a total of three independent measurements using GraphPad Prism software version 5.0.

	0% DMSO				10% DMSO			
	Compound 1	Compound 2	Compound 3	Compound 4	Compound 1	Compound 2	Compound 3	Compound 4
<b>Best-fit values</b>								
Ymax	0.005299	-0.005346	-0.05184	-1.324	-0.004238	-0.03934	-0.04724	-0.9876
Yoff	99.94	99.83	99.92	99.10	99.97	99.57	100.1	99.47
pKd	3.158	4.243	4.890	5.406	3.406	4.447	5.061	5.577
kd [ $\mu$ M]	694.9	57.11	12.88	3.928	392.2	35.74	8.694	2.649
<b>Std. Error</b>								
pKd	0.02469	0.03236	0.02848	0.04650	0.02232	0.01665	0.01175	0.04888
<b>95% Conf. Int.</b>								
pKd	3.108 to 3.208	4.178 to 4.309	4.832 to 4.948	5.312 to 5.500	3.361 to 3.452	4.413 to 4.481	5.037 to 5.085	5.478 to 5.676
kd [ $\mu$ M]	619.4 to 779.6	49.13 to 66.39	11.28 to 14.71	3.165 to 4.876	353.5 to 435.2	33.07 to 38.62	8.231 to 9.182	2.110 to 3.324
<b>Goodness of Fit</b>								
R square	0.9941	0.9898	0.9923	0.9796	0.9951	0.9973	0.9988	0.9778
Sum of Squares	328.7	656.5	471.0	1369	300.3	166.8	82.01	1558
<b>Constraints</b>								
Conc. Protein	2.500e-007	2.500e-007	2.000e-007	2.000e-007	2.500e-007	2.500e-007	2.000e-007	2.000e-007
Points analyzed	42	44	41	45	42	42	43	45

**Table S2. Thermodynamic parameters of compounds 1-4 binding to E-selectin.** Standard measurements were carried out at 25°C using the E-selectin<sub>LFC6-IgG<sub>FC</sub></sub> construct in the absence of DMSO. Compound 2 was additionally measured in the presence of 10% DMSO and the binding properties of E-selectin<sub>LEG2</sub> were evaluated with compound 4. For the evaluation of experiments with low c-values (compounds 1 and 2) the stoichiometry (*N*) had to be fixed to 1.

Ligand	Protein	Ligand [μM]	Protein [μM]	DMSO [%]	K <sub>D</sub> [μM]	ΔG° [kJ/mol]	ΔH° [kJ/mol]	-TΔS° [kJ/mol K]	N	c-value
1	E-selectin <sub>LFC6-IgG<sub>FC</sub></sub>	12000	180	0	779 (740 – 822)	-17.7	4.8 (5.0 – 4.6)	-22.6	1 (fix)	0.2
1	E-selectin <sub>LFC6-IgG<sub>FC</sub></sub>	7600	100	0	873 (750 – 1024)	-17.5	5.6 (6.5 – 4.8)	-23.0	1 (fix)	0.1
<b>1</b>					<b>778 (722 – 842)</b>	<b>-17.7</b>	<b>4.9 (5.2 – 4.6)</b>	<b>-22.6</b>		
2	E-selectin <sub>LFC6-IgG<sub>FC</sub></sub>	2000	90.9	0	74.0 (64.0 – 85.4)	-23.6	-5.7 (-5.2 – -6.2)	-17.9	1 (fix)	1.2
2	E-selectin <sub>LFC6-IgG<sub>FC</sub></sub>	1500	78.0	0	62.7 (38.8 – 102.2)	-24.0	-5.6 (-4.4 – -7.7)	-18.4	1 (fix)	1.2
<b>2</b>					<b>60.7 (43.4 – 84.8)</b>	<b>-24.1</b>	<b>-5.3 (-4.5 – -6.5)</b>	<b>-18.8</b>		
<b>2</b>	E-selectin <sub>LFC6-IgG<sub>FC</sub></sub>	1800	71.0	10	<b>20.7 (18.8 – 22.8)</b>	<b>-26.7</b>	<b>-15.9 (-15.3 – -16.5)</b>	<b>-10.9</b>	1.06	3.4
3	E-selectin <sub>LEG6-IgG<sub>FC</sub></sub>	1350	82.5	0	16.1 (13.5 – 19.3)	-27.4	-5.7 (-5.4 – -6.1)	-21.6	0.97	5.1
3	E-selectin <sub>LEG6-IgG<sub>FC</sub></sub>	1350	82.5	0	22.2 (18.6 – 26.7)	-26.6	-6.4 (-6.0 – -6.9)	-20.2	0.97	3.7
<b>3</b>					<b>17.8 (15.1 – 21.1)</b>	<b>-27.1</b>	<b>-5.9 (-5.6 – -6.3)</b>	<b>-21.2</b>		
4	E-selectin <sub>LFC6-IgG<sub>FC</sub></sub>	1000	39.0	0	4.03 (3.42 – 4.73)	-30.8	-12.6 (-12.1 – -13.1)	-18.2	0.97	9.7
4	E-selectin <sub>LFC6-IgG<sub>FC</sub></sub>	475	30.0	0	3.97 (3.56 – 4.42)	-30.8	-12.4 (-12.1 – -12.8)	-18.4	1.08	7.6
<b>4</b>					<b>3.99 (3.56 – 4.48)</b>	<b>-30.8</b>	<b>-12.5 (-12.2 – -12.8)</b>	<b>-18.3</b>		
4	E-selectin <sub>LEG2</sub>	900	72.0	0	3.70 (3.47 – 3.95)	-31.0	-12.2 (-12.0 – -12.3)	-18.8	1.15	19.4
4	E-selectin <sub>LEG2</sub>	600	34.2	0	4.07 (3.54 – 4.67)	-30.8	-13.2 (-12.8 – -13.7)	-17.5	1.05	8.4
<b>4</b>					<b>3.58 (3.12 – 4.11)</b>	<b>-31.1</b>	<b>-12.3 (-12.0 – -12.7)</b>	<b>-18.7</b>		

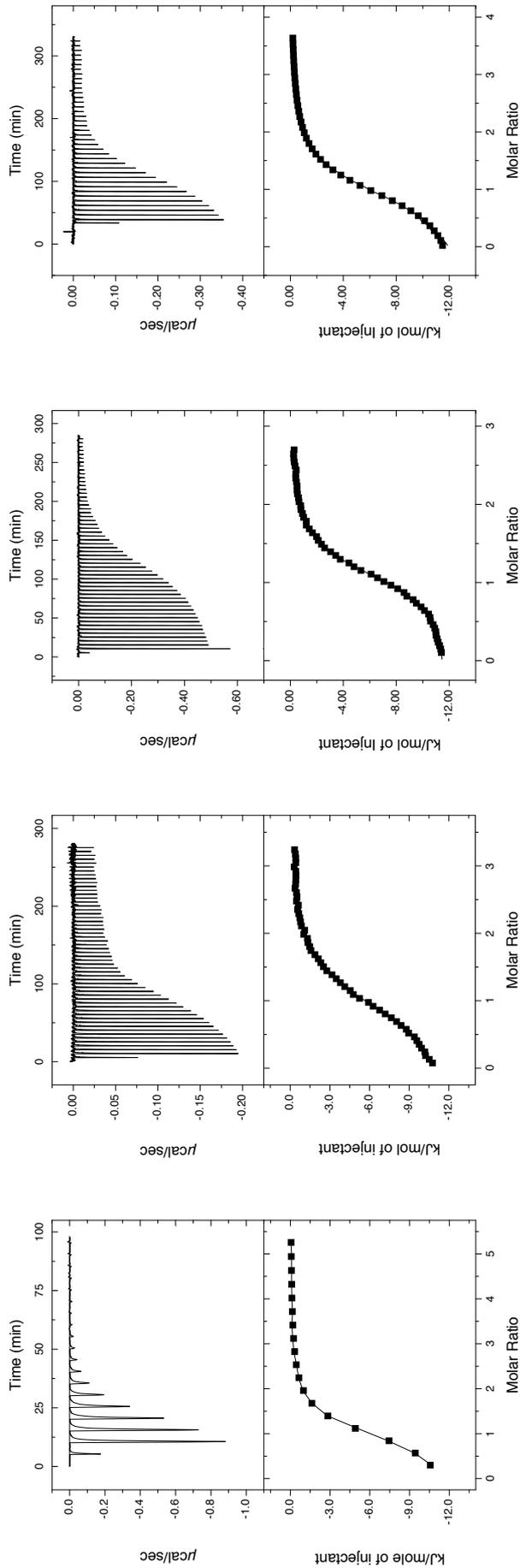


Figure S1. Calorimetric titrations of compound 4 binding to E-selectin<sub>LECC</sub>IgG<sub>Fc</sub> (the two enthalpograms on the left) and E-selectin<sub>LECC</sub> (the two enthalpograms on the right).

Table S3. Best-fit values and standard errors of compounds 2 and 4 binding to E-selectin<sub>LEC2</sub> labeled with blocked and unblocked binding site measured by MST. For each compound a global analysis was performed over a total of three independent measurements using GraphPad Prism software version 5.0.

	Compound 2		Compound 4	
	unblocked	blocked	unblocked	blocked
<b>Best-fit values</b>				
Ymax	0.1152	-1.324	0.01936	-0.005346
Yoff	99.71	99.10	99.96	99.83
pKd	4.913	5.406	3.781	4.243
kd [ $\mu$ M]	12.23	3.928	165.6	57.11
<b>Std. Error</b>				
pKd	0.05406	0.04650	0.04416	0.03236
<b>95% Conf. Int.</b>				
pKd	4.804 to 5.022	5.312 to 5.500	3.692 to 3.870	4.178 to 4.309
kd [ $\mu$ M]	9.508 to 15.72	3.165 to 4.876	134.8 to 203.4	49.13 to 66.39
<b>Goodness of Fit</b>				
R square	0.9724	0.9796	0.9818	0.9898
Sum of Squares	1783	1369	1065	656.5
<b>Constraints</b>				
Conc. Protein	2.500e-007	2.000e-007	2.500e-007	2.500e-007
Points analyzed	44	45	42	44

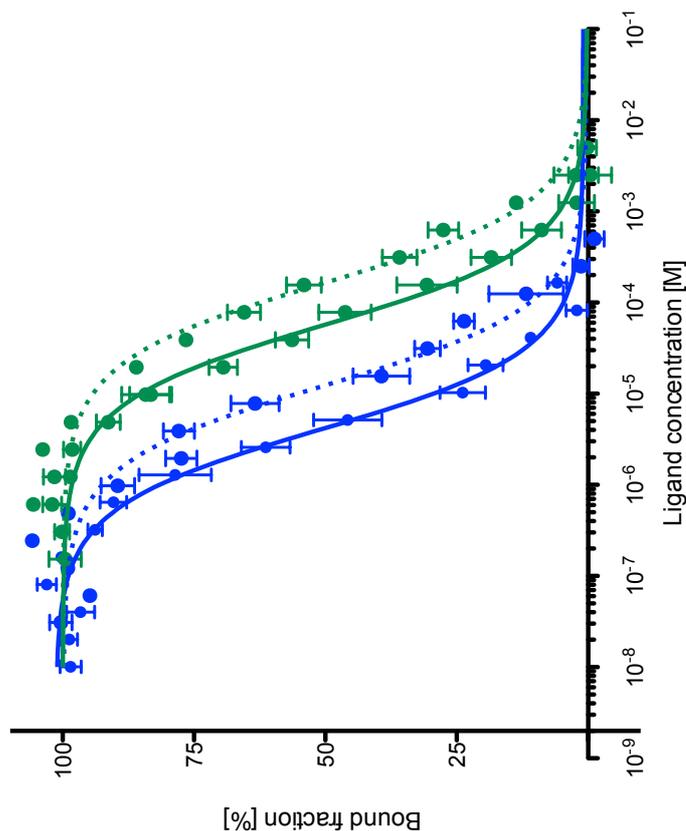


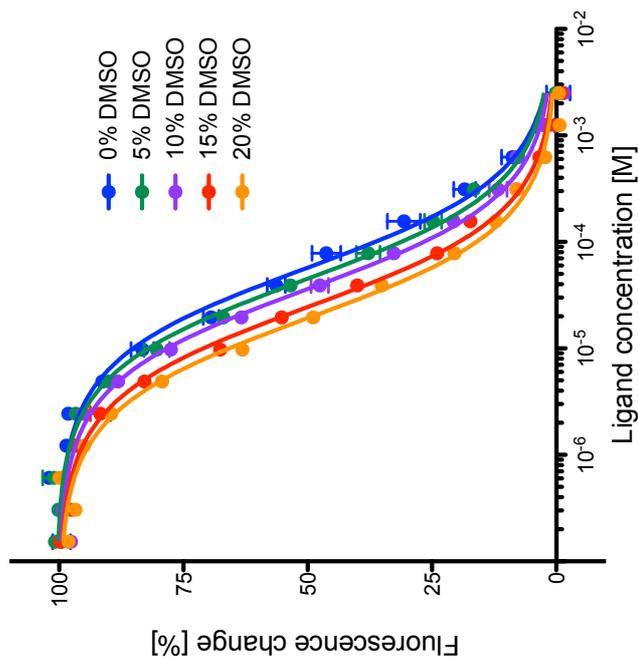
Figure S2. Binding curves of compounds 2 (green) and 4 (blue) binding to E-selectin<sub>LEC2</sub> labeled with blocked (solid line) and unblocked binding site (dashed line) measured by MST. The error bars represent the standard deviation over a total of three independent measurements.

Table S4. Best-fit values and standard errors of compounds 2 binding to E-selectin<sub>LEC2</sub> with varying laser power (LP) measured by MST.

Compound 2										
	20% LP	30% LP	40% LP	50% LP	60% LP	70% LP	80% LP	90% LP	100% LP	
<b>Best-fit values</b>										
Ymax	0.004448	-0.003269	-0.001728	0.0002237	0.005382	0.006234	0.001053	0.001360	0.003634	
Yoff	2.142	12.99	27.02	59.58	99.79	155.1	196.2	239.0	257.4	
pKd	4.301	4.450	4.432	4.244	4.057	3.809	3.570	3.273	2.927	
kd [ $\mu$ M]	49.95	35.48	36.96	57.06	87.67	155.1	269.2	533.0	1184	
<b>Std. Error</b>										
pKd	0.4019	0.08400	0.07784	0.04991	0.03182	0.04313	0.05110	0.03528	0.04479	
<b>95% Conf. Int.</b>										
pKd	3.417 to 5.186	4.265 to 4.635	4.261 to 4.604	4.134 to 4.354	3.987 to 4.127	3.714 to 3.904	3.457 to 3.682	3.196 to 3.351	2.828 to 3.025	
kd [ $\mu$ M]	6.517 to 382.9	23.18 to 54.32	24.91 to 54.83	44.31 to 73.48	74.61 to 103.0	124.7 to 193.0	207.8 to 348.8	445.8 to 637.4	943.7 to 1486	
<b>Goodness of Fit</b>										
R square	0.6973	0.9814	0.9840	0.9933	0.9973	0.9951	0.9933	0.9970	0.9961	
Sum of Squares	4.111	6.659	48.99	53.91	217.9	419.2	215.5	211.1	48.99	
<b>Constraints</b>										
Conc. Protein	2.500e-007									
Points analyzed	14	14	14	14	14	14	14	14	14	

**Table S5. Best-fit values and standard errors of compound 2 binding to E-selectin<sub>LEC2</sub> with varying amounts of DMSO measured by MST.** For each compound a global analysis was performed over a total of three independent measurements using GraphPad Prism software version 5.0.

	Compound 2				
	0% DMSO	5% DMSO	10% DMSO	15% DMSO	20% DMSO
<b>Best-fit values</b>					
Ymax	-0.005346	-0.5178	-0.03934	0.005598	-0.001698
Yoff	99.83	99.29	99.57	100.0	99.99
pKd	4.243	4.356	4.447	4.613	4.718
kd [ $\mu$ M]	57.11	44.03	35.74	24.38	19.13
<b>Std. Error</b>					
pKd	0.03236	0.02298	0.01665	0.01662	0.01583
<b>95% Conf. Int.</b>					
pKd	4.178 to 4.309	4.310 to 4.403	4.413 to 4.481	4.579 to 4.646	4.686 to 4.750
kd [ $\mu$ M]	49.13 to 66.39	39.56 to 48.99	33.07 to 38.62	22.57 to 26.34	17.77 to 20.59
<b>Goodness of Fit</b>					
R square	0.9898	0.9949	0.9973	0.9973	0.9975
Sum of Squares	656.5	324.0	166.8	191.4	173.9
<b>Constraints</b>					
Conc. Protein	2.500e-007	2.500e-007	2.500e-007	2.500e-007	2.500e-007
Points analyzed	44	43	42	45	45



**Figure S3. Binding curves of compound 2 binding to E-selectin<sub>LEC2</sub> in the presence of varying amounts of DMSO measured by MST.** The error bars represent the standard deviation over a total of three independent measurements

## Experimental Section

**Reagents, devices and compounds.** VP-ITC apparatus was purchased from MicroCal Inc. (Uppsala, Sweden). Monolith™ NT.115, Protein Labeling Kit BLUE-NHS (Amine Reactive), and Standard Treated Capillaries were purchased from Nanotemper Technologies GmbH (Munich, Germany). HEPES, NaOH, NaCl, CaCl<sub>2</sub>·2H<sub>2</sub>O, DMSO and Tween 20 were purchased from Sigma-Aldrich Chemie GmbH (Steinheim, Germany). Slide-A-Lyzer cassettes (10 kDa MWCO) and bovine serum albumin (BSA) standard ampules were obtained from Thermo Fisher Scientific (Rockford, IL, USA) Compound **1** was obtained from Carbosynth (San Diego, CA, USA) and compounds **2-4** were synthesized as according to Schwizer *et al.*<sup>(1)</sup>

**E-selectin production.** Cloning, transfection, expression and purification were previously described for the E-selectin<sub>LEC6</sub>-IgG<sub>Fc</sub> construct by Jahnke *et al.*(1997)<sup>(2)</sup> and for the E-selectin<sub>LEC2</sub> construct by Preston *et al.* (2015).<sup>(3)</sup>

**Labeling.** E-selectin<sub>LEC2</sub> was labeled using the amine reactive protein labeling kit BLUE-NHS. Buffer exchange and labeling were performed according to the manufacturer's protocol. To protect the lysines in the binding site from being labeled, the protein was saturated with 600 μM of compound **4**. The labeled protein was dialyzed over night against assay buffer (10mM HEPES pH 7.4, 150 mM NaCl, 1 mM CaCl<sub>2</sub>) using Slide-A-Lyzer dialysis cassettes (10 kDa MWCO). Protein concentration was determined by HPLC-UV against a BSA standard.<sup>(4,5)</sup> Figure 2 was illustrated with Maestro software version 9.7 (Schrödinger, LLC, NY, USA).

**Microscale Thermophoresis.** MST experiments were carried out at 25° C with 100% LED power, 10-100% laser power (standard condition 50% laser power), a laser on time of 30 sec, and a laser off time of 5 sec using standard treated capillaries. Ligands were dissolved in assay buffer supplemented with 0.05% v/v Tween 20 (and 0-20% v/v DMSO) at a concentration of 88 mM (**1**), 10 mM (**2**), 4 mM (**3**), and 1 mM (**4**). Titration series were generated diluting the ligand stocks 15 times 1:1 with the corresponding buffer. The dilution series of ligand was 1:1 mixed with a solution of 0.2 to 0.25 μM E-Selectin<sub>LEC2</sub> and incubated for 10 mins at room temperature. The protein concentration was determined by HPLC-UV against a BSA standard.<sup>(4,5)</sup> Except the measurements with varying laser power, all experiments were independently performed three times. Datapoints were normalized using the bound and unbound borders achieved by NanoTemper Analysis 1.2.205 software (NanoTemper Technologies GmbH, Munich, Germany) and analyzed/illustrated with GraphPad Prism 5.0 (GraphPad Software, La Jolla, CA, USA). The measurements were globally fitted using equation 1 for single site binding.<sup>(6)</sup>

$$[PL] = \frac{(C_P + C_L + K_D) - \sqrt{(C_P + C_L + K_D)^2 - 4C_P C_L}}{2C_P} \quad (\text{eq. 1})$$

where  $[PL]$  is the protein-ligand complex concentration and  $K_D$  is the dissociation constant.  $C_P$  represents the total concentration of protein and  $C_L$  the total concentration of ligand.

**Isothermal Titration Calorimetry.** Measurements were performed at 25° C with a reference power of 10  $\mu\text{cal}/\text{sec.}$ , a stirring speed of 307 rpm, high feedback, a spacing time of 600 sec., and a filter period of 2 sec. Preceding the measurements, E-selectin was dialyzed against assay buffer. Ligands were dissolved in the same buffer. For the measurement of compound **2** in the presence of dimethylsulfoxide, 10% v/v DMSO was added to both, protein and ligand. 5 – 15  $\mu\text{L}$  ligand solution were injected into the sample cell (1.4523 mL) containing the protein. The protein concentration was determined by HPLC-UV against a BSA standard.<sup>(4,5)</sup> Ligand and protein concentrations of all experiments are reported in Table S2. The first injection was always excluded from data analysis. Baseline adjustment and peak integration were carried out using Origin 7.0 as described by the manufacturer (OriginLab, Northampton, MA, USA). A global three-parameter data fitting was processed to determine  $N$  (stoichiometry),  $K_A$  (association constant) and  $\Delta H^\circ$  (change in standard enthalpy) using SEDPHAT software.<sup>(7)</sup> Previously published data from Binder *et al.* (2012)<sup>(8)</sup> (compounds **1-3** binding to E-selectin<sub>LEC6</sub>-IgG<sub>FC</sub>) were reanalyzed in a similar manner, although for **1** and **2** the stoichiometry had to be fixed to 1 due to low  $c$ -values. This allowed a reliable determination of  $K_A$  and  $\Delta H^\circ$ .<sup>(9,10)</sup> Thermodynamic parameters were calculated from equation 2:

$$\Delta G^\circ = \Delta H^\circ - T\Delta S^\circ = RT\ln K_D = -RT\ln K_A \quad (\text{eq. 2})$$

where  $\Delta G^\circ$  is the change in standard free energy of binding and  $\Delta S^\circ$  is the change in standard entropy,  $T$  is the absolute temperature, and  $R$  is the universal gas constant (8.314 J/mol K).

## References

1. Schwizer, D., *et al.* (2012). Pre-organization of the core structure of E-selectin antagonists. *Chem. Eur. J.*, 18(5): 1342-1351.
2. Jahnke, W., Kolb, H.C., J., B.M.J., Magnani, J.L., & Ernst, B. (1997). Comparison of the Bioactive Conformations of Sialyl LewisX and a Potent Sialyl LewisX Mimic. *Angew. Chem. Int. Ed.*, 36(23): 2603-2607.
3. Preston, R.C., *et al.* (2015). E-selectin ligand complexes adopt an extended high-affinity conformation. *J. Mol. Cell. Biol.*: mjev046.
4. Mesch, A., S., *et al.* (2010). Kinetic and thermodynamic properties of MAG antagonists. *Carbohydr. Res.*, 345(10): 1348-1359.
5. Bitsch, F., *et al.* (2003). Identification of natural ligands of retinoic acid receptor-related orphan receptor alpha ligand-binding domain expressed in Sf9 cells - a mass spectrometry approach. *Anal. Biochem.*, 323(1): 139-149.
6. Cooper, A. & Royal Society of Chemistry (Great Britain) (2004). *Biophysical Chemistry*. Royal Society of Chemistry, Cambridge. pp. 109-110.
7. Houtman, J.C., *et al.* (2007). Studying multisite binary and ternary protein interactions by global analysis of isothermal titration calorimetry data in SEDPHAT: application to adaptor protein complexes in cell signaling. *Protein Sci.*, 16(1): 30-42.
8. Binder, F.P., Lemme, K., Preston, R.C., & Ernst, B. (2012). Sialyl Lewis<sup>x</sup>: A “Pre-Organized Water Oligomer”? *Angew. Chem. Int. Ed.*, 51(29): 7327-7331.
9. Tellinghuisen, J. (2008). Isothermal titration calorimetry at very low  $c$ . *Anal. Biochem.*, 373(2): 395-397.
10. Turnbull, W.B. & Daranas, A.H. (2003). On the value of  $c$ : can low affinity systems be studied by isothermal titration calorimetry? *J. Am. Chem. Soc.*, 125(48): 14859-14866.

## Manuscript 8

### Fragment Screening Towards an Orally Available E-selectin Antagonist Using Microscale Thermophoresis

Tobias Mühlethaler,<sup>†</sup> Pascal Zihlmann,<sup>†</sup> Christoph P. Sager,<sup>†</sup> Brigitte Fiege,<sup>†</sup>  
Timothy Sharpe<sup>§</sup> and Beat Ernst<sup>†\*</sup>

<sup>†</sup>University of Basel, Institute of Molecular Pharmacy,  
Klingelbergstr. 50, 4056 Basel, Switzerland

<sup>§</sup> University of Basel, Biophysics Facility,  
Klingelbergstr. 70, 4056 Basel, Switzerland

\*Corresponding author.

Tel.: 0041 (0)61 267 15 51; Fax: 0041 (0)61 267 15 52;  
E-mail: beat.ernst@unibas.ch

#### Contributions of Pascal Zihlmann

- Project planning
- Master thesis supervision



**Abstract**

Microscale thermophoresis (MST) is a time-efficient method to determine affinities consuming only small amounts of ligand and receptor with a constantly growing field of applications. In the present work, we introduce MST as a primary screening method for the identification of fragments. The assay performance was demonstrated with E-selectin, a lectin that plays an important role in the early stage of inflammation. The screening was performed competitively against a reference ligand, a derivative of the natural carbohydrate binding motif. This allowed screening in cocktails and restricting the detected binding fragments to the ones binding in the natural binding site. The assay format was validated by the competitive displacement of a compound representing a sub-structure of the natural binding motif, which is known to bind with a weak affinity. By deconvoluting the active cocktails several hits could be identified. Using nuclear magnetic resonance (NMR) as an orthogonal method, the validation of fragments binding in the binding site could be achieved, confirming the reliability of the assay for a primary screening of fragment libraries.

### Introduction

E-selectin is a C-type lectin that plays a major role in the initial phase of inflammation.<sup>(1)</sup> Upon an inflammatory stimulus, macrophages activate vascular endothelial cells via proinflammatory cytokines. The endothelial cells subsequently express E-selectin and present it on their luminal cell surface.<sup>(2)</sup> Once presented, E-selectin mediates the initial contact between the vascular endothelium and leukocytes at the site of inflammation.<sup>(3)</sup> This first weak interaction slows down the leukocytes enabling stronger interactions between members of the immunoglobulin superfamily and integrins.<sup>(4)</sup> Ultimately, the tethering, rolling, and firm adhesion results in the extravasation of the leukocytes. Besides its physiological function, E-selectin is also involved in the pathology of several diseases. Even though E-selectin is not the causative molecule, the cell-adhesive properties of E-selectin are a promoting factor in a variety of diseases. These are cardiovascular diseases like atherosclerosis,<sup>(5)</sup> myocardial infarction,<sup>(6)</sup> reperfusion injury,<sup>(7)</sup> and hypertension.<sup>(8)</sup> Furthermore, E-selectin plays a role in immune diseases such as asthma bronchiale<sup>(9)</sup> or rheumatoid arthritis<sup>(10)</sup> due to facilitated neutrophil extravasation. Finally, E-selectin is also associated with sickle cell adhesion<sup>(11)</sup> and metastasis of breast and colon cancer.<sup>(12)</sup> To treat these types of diseases, E-selectin is a valuable drug target.<sup>(13)</sup>

With GMI-1070 (Rivipansel), the first intravenously applied E-selectin antagonist recently entered clinical trial phase 3 for the treatment of acute sickle cell crisis.<sup>(14)</sup> The lead structure for the development was the natural ligand of E-selectin, the tetrasaccharide sialyl Lewis<sup>x</sup> (sLe<sup>x</sup>).<sup>(15,16)</sup> As for all lectins, E-selectin binds its natural carbohydrate ligand with high specificity, which is achieved by the formation of a complex network of hydrogen bonds and salt bridges. A total of eleven interactions are formed between E-selectin and sLe<sup>x</sup> as determined from a recently published crystal structure (PDB code 4CSY).<sup>(17)</sup> However, the lead optimization is highly challenging as the removal of a single hydroxyl or carboxyl group leads to a dramatic loss in affinity due to their direct involvement in binding. The oral availability, on the other hand, is decreased by these groups due to lowered membrane permeability.<sup>(18)</sup> As most of the E-selectin related diseases are chronic or recurrent in nature requesting frequent applications, oral availability is a prerequisite for a therapeutic success. A possible

approach to this challenge is a non-carbohydrate replacement of the sLe<sup>x</sup> mimetic based on a fragment-based approach.

A fragment-based approach starts with a library of small drug-like fragments. Screening of fragments for binding to a specific target usually yields hits with affinities in the millimolar range. In a second step, hits are further characterized regarding their mode of binding. Different hits are either chemically linked together,<sup>(19)</sup> grown,<sup>(20)</sup> or serve for fragment-assisted drug design.<sup>(21)</sup> Since 2013, a single drug emerged from a fragment-based approach and eleven drug candidates have surpassed clinical phase 1.<sup>(22)</sup> A challenging step of a fragment-based approach is the initial screen. Due to their small size, solubility limits, and generally weak affinities, many of the standard assay methods to reliably determine  $K_D$  values are unsuccessful because they reach their detection limits. A method for robust, sensitive, and rapid  $K_D$  determination for weakly binding and small ligands is therefore highly desirable.

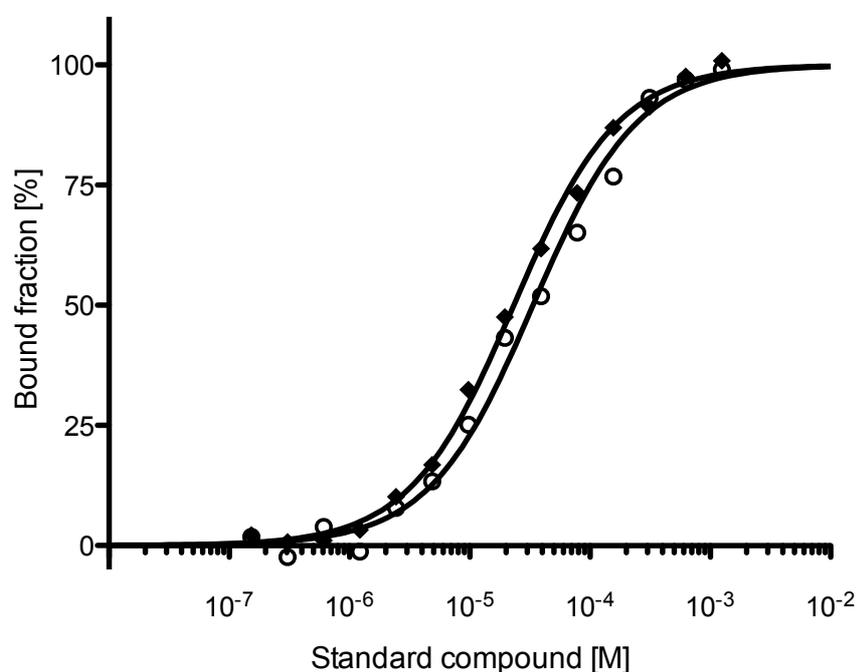
The recently introduced assay technique termed microscale thermophoresis (MST) works with small sample volumes and allows for the determination of weak affinities without the need for immobilization of either interaction partner. MST is based on the observation that different molecules exhibit individual thermophoretic properties, *i.e.* the mobility of molecules along temperature gradients. Thermophoresis is sensitive to changes in size, charge, and the solvation shell of a molecule.<sup>(23)</sup> Upon complex formation of ligand and protein, the thermophoretic parameters usually differ from the unbound state of the protein. Therefore a change of the thermophoresis of a protein is a function of its bound state. In recent studies, MST has been demonstrated to measure high-affinity ligands with low molecular weight<sup>(24)</sup> and it has furthermore been used to confirm hits found in a primary screen with other methods.<sup>(25)</sup>

In this study, we present an MST-based method for screening fragments for E-selectin binding. To the best of our knowledge, no MST-based fragment screening method has been published to date. Our method is based on competitive binding to an sLe<sup>x</sup> mimetic and provides fast screening with low sample consumption to detect fragments in the millimolar affinity range. As a result, we could successfully identify four site-specifically binding fragments.

## Results

**Assay Validation.** The principle of the applied assay format is to pre-incubate E-selectin with a titrated reference compound and a constant fragment concentration. The reference compound has a known binding affinity and binding mode analog to PDB entry 4C16. A fragment binding to the binding site competes with the reference compound, observable in a shift of the binding curve.

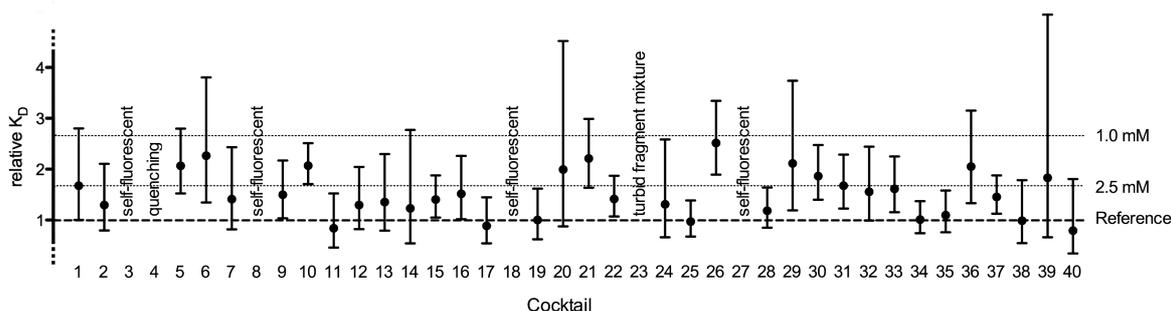
For the development of the competitive assay, we selected 1-methyl  $\alpha$ -L-fucopyranoside (MeFuc), a weakly binding fragment of both the reference compound as well as sLe<sup>x</sup>. With a direct MST experiment, we determined the affinity of MeFuc to be  $50.9 \pm 3.4$  mM. Indeed, the competing effect of 15 mM MeFuc decreased the binding affinity of the reference compound from 22.6  $\mu$ M to 33.5  $\mu$ M (Figure 1). Using equation 2, this shift corresponds to the affinity of 31.1 mM for MeFuc. Therefore, the competitive assay is in good agreement with the direct titration.



**Figure 1.** Shift of the normalized binding curve of E-selectin when binding the reference compound alone (●) and when binding the reference compound competing against MeFuc (○).

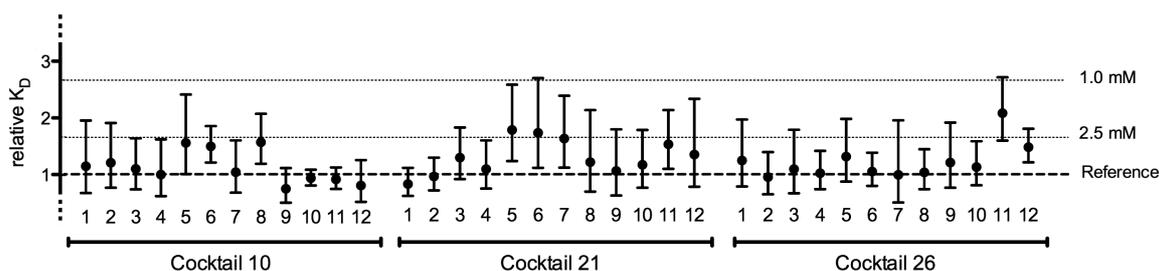
Considering the maximal DMSO-*d6* tolerance of E-selectin of 20% (v/v) and the fact that the fragments from the Maybridge library are prepared as 100 mM stock solutions, we decided to use cocktails of 12 fragments, each at a concentration of 1.66 mM. This allows the detection of fragments binding with an affinity in the low millimolar range.

**Fragment Screening with MST.** In an initial screen, we tested 480 fragments in 40 cocktails (Figure 2). To ensure comparability, we prepared a stock dilution series of the reference compound to preincubate with E-selectin before the cocktails (1 - 40) were added. In cocktails 3, 8, 18, and 27 one or more fragments were present with a self-fluorescence that shifted the baseline over the detection limit of the device. Furthermore, cocktail 4 contained a quenching substance and cocktail 23 contained turbid residues that prevented the cocktail from being properly measured.



**Figure 2.** Overview of the competitive binding experiments of the reference compound competing against fragment cocktails (1 - 40) for binding to E-selectin. The affinity shifts are given relative to the reference compound. The dashed line represents the affinity of the reference compound in the absence of competitor. The dotted lines (1 mM, 2.5 mM) indicate the theoretical binding affinity of a fragment when only one fragment in the cocktail accounted for the  $K_D$  shift of the reference compound.

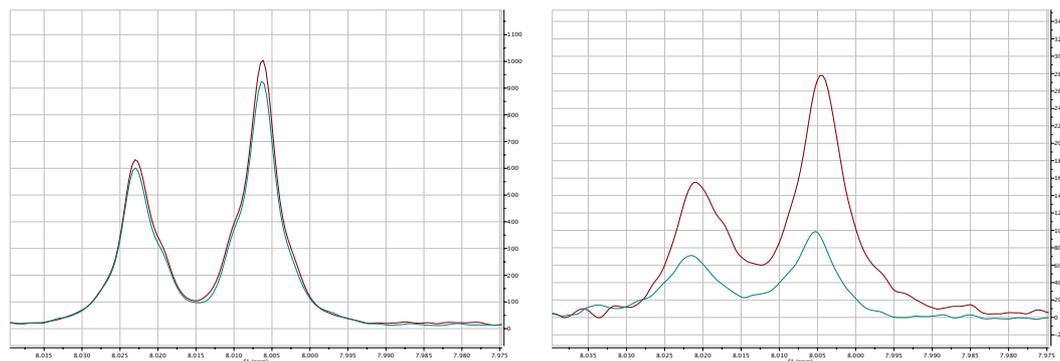
Out of the 40 cocktails the lower confidence interval border was shifted over the reference value of the reference compound in a total of 14 cocktails (5, 6, 9, 10, 15, 21, 22, 26, 29, 30, 31, 33, 36, 37). As we are interested only in the best binding fragments we selected the best three cocktails (10, 21, 26) considering  $K_D$  shift and the according confidence interval for more in depth-analysis. Therefore, we measured each fragment of these three cocktails individually applying the same assay procedures. Figure 3 depicts the results of the resolved cocktails 10, 21, and 26. In all three, multiple fragments introduced  $K_D$  shift. With 10-6, 10-8, 21-5, 21-6, 21-7, 26-11, and 26-12, a total of seven possible hit fragments were identified.



**Figure 3.** Overview of the competitive binding experiments of the reference compound competing with one fragment for the binding to E-selectin. The affinity shifts are given relative to the reference compound. The dashed line represents the reference affinity of the reference compound in the absence of competitor. The dotted lines visualize the theoretical shift of the binding curve in the presence of a competitor with an affinity of 1 mM and 2.5 mM, respectively.

To evaluate the possibilities of a direct measurement, we measured all hits again non-competitively at three different laser powers (50 %, 75 %, and 100 %). Except for fragment 10-1, none of the hits showed a clear change in thermophoresis that would unambiguously identify it as a binder. In three cases (fragments 21-5, 21-6, 26-11) there was an indication for binding at the highest ligand concentrations (10 - 20 mM). However, only a marginal signal change was achieved and a binding curve could not be fitted. Fragments 21-7 and 26-12 showed no change in thermophoresis at all. A special case was observed for the binding of fragment 10-8 for which the fluorescence decreased from the moment the LED was turned on.

**Hit Validation by NMR.** To validate fragments identified by the competitive MST assay with a second orthogonal method, we performed T1 $\rho$  NMR experiments, a label-free and well-established technique. Based on T1 $\rho$  NMR experiments, discriminating between a bound and an unbound fragment is possible as the relaxation of fragment protons in the close proximity of a protein is faster (Figure 4). Each fragment has an individual relaxation behavior in the absence of protein. Furthermore, the fragments bind to different sub-pockets providing a unique chemical environment for each proton affecting the relaxation in a distinct manner. This makes a cross-comparison between the fragments difficult and a ranking of the fragments affinity would not be reliable based on the T1 $\rho$  NMR results alone. However, T1 $\rho$  NMR is a valuable experiment to distinguish binders from non-binders.



**Figure 4.** Comparison of the relaxation of fragment 10-6 after 20 ms (red) and 200 ms (green) in the absence (left) and presence (right) of E-selectin in a  $T1\rho$  NMR experiment. In the presence of E-selectin fast relaxation can be observed, whereas in the absence of protein only slow relaxation occurs.

A foregoing COSY NMR confirmed the identity of all hit fragments except for fragment 21-6, which was therefore excluded from further experiments. A random fragment from each hit cocktail that did not show a change in  $K_D$  in the competitive MST assay was selected as negative control (10-1, 21-1, 26-8). For the confirmed binders, two additional  $T1\rho$  NMR experiments were conducted. First, a competitive measurement against MeFuc was performed to answer the question whether the fragment occupies the same part of the binding site as MeFuc. The aim of the second experiment, namely the subsequent addition of the reference compound to the same sample, was to confirm that the fragment is binding in the binding pocket of interest.

### Section III. E-selectin – Manuscript 8

**Table 1.**  $I_x$  depicts the percentage intensity change of the NMR signal after 200 ms compared to the intensity after 20 ms. Shown are the differences between the sole fragment (F) and the relaxation after adding protein (P), MeFuc and reference compound (Ref). Per fragment, a distinct aromatic peak was chosen for evaluation. The difference is given in percentage points.

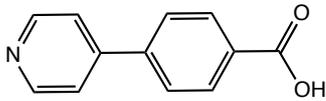
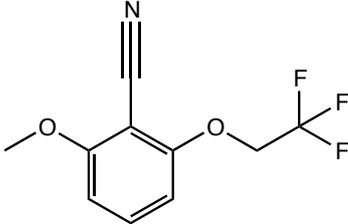
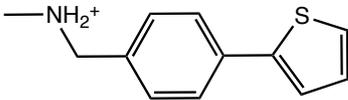
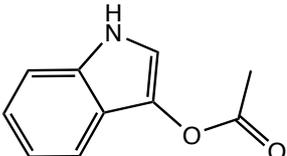
Fragment	$I_F - I_{F+P}$	$I_F - I_{F+P+MeFuc}$	$I_F - I_{F+P+MeFuc+Ref}$	Classification
10-6	23.5	24.9	2.8	competitive binder
10-8	3.4	n.d.	n.d.	non binder
21-5	26.6	20.4	15.2	competitive binder
21-7	1.8	n.d.	n.d.	non binder
26-11	18.2	11.9	8.1	competitive binder
26-12	14.6	10.5	4.5	competitive binder
10-1	1.0	n.d.	n.d.	non binder (negative control)
21-1	1.9	n.d.	n.d.	non binder (negative control)
26-8	28.37	n.d.	29.6	non-competitive binder (negative control)

n.d.: The competitive measurements were not conducted for the non-binding fragments.

A change in relaxation greater than 5 percentage points was defined as lower limit due to the signal to noise ratio. From the six hits, only four (10-6, 21-5, 26-11, and 26-12) could be confirmed as binders. The competitive measurements showed that fragment 10-6 is only competitive to the reference compound but does not occupy the binding site of MeFuc. Fragments 21-5, 26-11, and 26-12 are competitive against MeFuc and the reference compound. Because MeFuc at a 50 mM concentration does not completely saturate the binding sites, the reference compound can further increase the competitive effect. The two negative controls 10-1 and 21-1 did not bind. However, the third negative control 26-8 was a binder. However, since binding was observed after addition of the reference compound, this fragment likely does not bind in the binding site and could therefore not be detected in the competitive MST screen.

To get more precise  $K_D$  values and to test the reproducibility of the competitive MST assay we then remeasured the four confirmed binders. The values presented in Table 2 are in good agreement with the initial values depicted in Figure 3.

**Table 2.** The  $K_D$  values and structures of the hit fragments from the competitive measurements

Fragment	$K_D$ [mM]	Structure
10-6	$3.3 \pm 0.9$	
21-5	$1.8 \pm 0.5$	
26-11	$1.7 \pm 0.4$	
26-12	$2.7 \pm 0.6$	

## Discussion

Four fragments with millimolar affinities to E-selectin were identified by modifying the previously introduced MST assay.<sup>(26)</sup> Again, MST was shown to be a robust and fast method with only low consumption of ligand and protein.

In our project aimed to identify fragments binding to E-selectin, two major problems had to be solved. First, the determination of binders at a single fragment concentration, as practiced in established fragment screens with Differential Scanning Fluorimetry<sup>(27)</sup> or Surface Plasmon Resonance,<sup>(28)</sup> was not applicable with MST. For the evaluation of the MST experiment, the unbound, as well as the bound state of the protein, has to be distinguishable, which demands a dilution series. Hence, reaching the bound state for molecules with binding affinities in the low millimolar range requires high starting concentrations of the dilution series. This inevitably leads to problems regarding compound solubility and self-fluorescence of a substantial number of fragments. Second,

not all fragments are capable to significantly alter the hydration shell, a prerequisite to obtaining a measurable signal. Conformingly, except for fragment 10-6, none of the hits identified by a competitive MST screen would have been detected by a direct titration of the fragment. Furthermore, the laser power optimal to observe the thermophoretic behavior of a protein-ligand complex is not predictable. Therefore, each fragment should be measured with different laser power intensities.

Nevertheless, these problems could be solved by performing a competitive experiment with a serial dilution of the reference compound and the addition of the competing fragment at a constant concentration. The preliminary experiment with MeFuc showed that a competitive MST measurement with small molecular weight compounds causes a distinct  $K_D$  shift of the titrated reference compound. In addition, the competitive experiment provides further advantages. Most importantly it is possible to screen fragment cocktails, which drastically reduces measurement time and protein consumption. Also, the applied low fragment concentrations decreased the risk of reaching the solubility limits, diminishes the problem of self-fluorescence and reduces the fragment consumption. Furthermore, the competition with a mimetic of the natural binding epitope (reference compound) allows identifying fragments binding in the binding site of interest. In addition, stable and well-known screening conditions with an unambiguous fluorescence change of 15% allowed detecting also small shifts of the binding curves.

We used cocktails that were a compromise between the DMSO-*d6* tolerance of E-selectin, the assumed druggability representing the number of hits expected, and the desired fragment affinity. However, several weakly binding fragments added up their effect leading us to believe in an apparent large  $K_D$  shift. Therefore, the deconvolution of a cocktail might not lead to a hit with expected high affinity. Also, too strong self-fluorescence or quenching properties of a specific fragment might mask a binding fragment and not every apparent hit can be confirmed. Furthermore, an interaction of the fragment with the fluorescent label covalently attached to the protein may falsify the result. An interaction with the label of the protein or the competitor (reference compound) might also explain the false positive result obtained with fragment 21-7. In our case, a total of 13 out of 48 cocktails showed a significantly shift in  $K_D$  and thereof

seven over the 4 mM and three over the 2.5 mM mark. As lectins are generally regarded to have a low druggability,<sup>(29)</sup> this amount of hits was surprising. In future projects, we should work with smaller cocktails to decrease the number of hits in one cocktail.

The competitive NMR experiments gave a first indication for future project strategies. As fragment 10-1 is not competing the MeFuc, a linking strategy between the two fragments is at hand. However, the linking of fragments 21-5, 26-11, and 26-12 to MeFuc is not advised as they are competing for the same binding site. Here, a fragment growing approach, i.e. regarding the fragment as scaffold has to be considered. However, before any further experiments are initiated, fragments of interest should be co-crystallized to gain solid information regarding their binding mode.

In conclusion, our approach proved to be eligible to fragments with low millimolar affinities within short period of time and with low material consumption. Especially when the MST assay for high-affinity binders is already established, no changes to the existing protocol have to be made, except for adding constant amounts of fragments. This makes it an easily applicable method for an initial fast screen in a fragment-based approach.

### Experimental Section

**Reagents and Compounds.** HEPES and 1-methyl  $\alpha$ -L-fucopyranoside were purchased from Sigma-Aldrich Chemie GmbH (Steinheim, Germany). D<sub>2</sub>O, TSP-*d4*, and DMSO-*d6* were purchased from Armar Chemicals (Döttingen, Switzerland) and HEPES-*d18* was purchased from Cambridge Isotope Laboratories Inc. (Andover, MA, USA). Vivaspin<sup>®</sup> 20 centrifugal concentration tubes for membrane filtration were purchased from Sartorius-Stedim (Göttingen, Germany). Protein Labeling Kit BLUE-NHS (Amine Reactive) and Standard Treated Capillaries were purchased from Nanotemper Technologies GmbH (Munich, Germany). The Maybridge Ro3 Diversity Fragment Library Core Set was purchased from Fisher Scientific AG (Wohlen, Switzerland). SLe<sup>x</sup> methyl glycoside was obtained from Carbosynth (San Diego, CA, USA). The reference compound CGP69669A<sup>(30)</sup> was synthesized according to Kolb and Ernst.<sup>(31)</sup>

### Microscale Thermophoresis

**Protein Preparation.** The E-selectin<sub>LEC2</sub> construct was expressed and purified according to Preston *et al.*<sup>(17)</sup> The labeling procedure was performed according to Zihlmann *et al.*<sup>(26)</sup> Briefly, the E-selectin<sub>LEC2</sub> binding site was blocked and then labeled using the amine reactive protein labeling kit BLUE-NHS according to the protocol of the manufacturer. Unreacted dye was removed by dialysis against buffer A (10 mM HEPES, 150 mM NaCl, 1 mM CaCl<sub>2</sub>, pH 7.4).

**Measurements.** All experiments were carried out using a Monolith<sup>™</sup> NT.115 device (Nanotemper Technologies GmbH, Munich, Germany) at 298 K with 100% LED power, 50% laser power, laser on time of 30 s, and laser off time of 5 s with standard treated capillaries. Throughout all experiments, the E-selectin<sub>LEC2</sub> concentration was 0.1  $\mu$ M. For the competitive measurements, a stock dilution series of the reference compound was prepared. The reference compound was diluted 1:1 starting at a concentration of 5 mM with buffer A supplemented with 0.1% (v/v) Tween 20. For the cocktail measurements cocktails of 12 fragments at a concentration of 3.33 mM in buffer A with 40% (v/v) DMSO-*d6* were prepared. For the single compound measurements, the 1-methyl  $\alpha$ -L-fucopyranoside concentration was 30 mM and the fragment concentration was 3.33 mM in buffer A with 40% (v/v) DMSO-*d6*. The E-selectin<sub>LEC2</sub> was preincubated with the stock dilution series. Subsequently, the preincubated protein was mixed 1:1 with either a cocktail or single compound. For the direct measurements, three runs with the same sample at a laser power of 50%, 75%, and 100% were conducted. The fragments were diluted 1:1 with buffer A supplemented with 0.1% (v/v) Tween 20 and 40% (v/v) DMSO-*d6*. The highest concentrations were 1 M for the 1-methyl  $\alpha$ -L-fucopyranoside, 40 mM for the fragments 10-8, 21-5, 21-6, 21-7, 26-11, and 26-12, and due to solubility problems 4 mM for fragment 10-6. The dilution series were mixed 1:1 with 0.2  $\mu$ M of the E-selectin<sub>LEC2</sub> protein.

**Data Analysis.** Every measurement was conducted once except for the four hit fragments and the 1-methyl  $\alpha$ -L-fucopyranoside measurements, which were measured in triplicate. Data points were normalized using the bound and unbound borders calculated by NanoTemper Analysis 1.2.205 software (NanoTemper Technologies GmbH, Munich, Germany) and analyzed with GraphPad Prism 5.0 (GraphPad Software, La Jolla, CA, USA). To evaluate the  $K_D$  shift of the reference compound the measurements were globally fitted using equation (1) for single site binding.<sup>(32)</sup>

$$[PL] = \frac{(C_P + C_L + K_D - \sqrt{(C_P + C_L + K_D)^2 - 4C_P C_L})}{2C_P} \quad (1)$$

where [PL] is the protein-ligand complex concentration and  $K_D$  is the dissociation constant.  $C_P$  represents the total concentration of protein and  $C_L$  the total concentration of ligand. To obtain the affinity constant of the fragment, the measurements were first globally fitted using Wang formula for competitive single site binding.<sup>(33)</sup> As the same results were obtained with the simplified Schild formula<sup>(34)</sup> (equation 2) this equation was used instead.

$$p_S = \frac{\frac{C_S}{K_{D,S}}}{1 + \frac{C_S}{K_{D,S}} + \frac{C_F}{K_{D,F}}} \quad (2)$$

where  $p_S$  is the fraction of binding sites occupied by the fragment.  $K_{D,F}$  is the dissociation of the fragment,  $C_F$  is the fragment concentration,  $K_{D,S}$  is the dissociation constant of the reference compound and  $C_S$  is the concentration of the reference compound.

### Nuclear Magnetic Resonance

**Protein Preparation.** Buffer B (10 mM HEPES-*d18*, 150 mM NaCl, pH 7.4) was prepared by adjusting the pH in water with correction for deuterium effect, subsequent three times lyophilization, and resolubilization in D<sub>2</sub>O. The E-selectin<sub>LEC6</sub>-IgG<sub>Fc</sub> construct was expressed and purified according to Jahnke *et al.*<sup>(31)</sup> After purification the protein was transferred to buffer B by membrane filtration (10 kDa MWCO).

**T1ρ Measurements.** A sample with 1 mM fragment in D<sub>2</sub>O with 10% (v/v) DMSO-*d6* with 0.1 mM TSP-*d4* added as internal reference was prepared to evaluate the signal reduction of the unbound fragment. For the measurements combined with protein a sample of 0.5 mM fragment, 20 μM E-selectin<sub>LEC6</sub>-IgG<sub>Fc</sub>, 0.1 mM TSP-*d4*, and 1 mM CaCl<sub>2</sub> in buffer B with 10% (v/v) DMSO-*d6* was prepared. For the competitive measurements first 1-methyl α-L-fucopyranoside and second the reference compound were added to the sample protein sample to reach a final concentration of 50 mM and 5 mM, respectively. All experiments were measured on a Bruker Avance III 500 MHz NMR spectrometer equipped with a 5 mm BBO room temperature probe head at a temperature of 298 K. The measurements are based on a reference Bruker pulse sequence with excitation sculpting for suppression of residual water signals, with a relaxation delay of 10 s, an acquisition time of 2.73 s, and with 256 to 512 scans and 4 dummy scans. For analysis of the transversal relaxation in absence and presence of protein, spinlock pulses of 20 ms and 200 ms were used at a power level of 2.6 kHz. Spectra were acquired, processed and analyzed with Topspin 2.1 (Bruker, Fällanden, Switzerland).

**Data Analysis.** Data were evaluated according to Hajduk *et al.*<sup>(35)</sup> With data evaluation software Bruker Topspin v2.1 the percentaged signal reduction per proton signal of the fragment with 200 ms spinlock compared to 20 ms spinlock were obtained. Subsequently, the percentaged signal reduction of a proton signal was compared to the percentaged signal reduction in presence of protein and competitor. A difference of 5 percentage points was defined as the cut-off for binding.

## References

1. Bevilacqua, M.P. & Nelson, R.M. (1993). Selectins. *J. Clin. Invest.*, 91(2): 379-387.
2. Wyble, C.W., *et al.* (1997). TNF- $\alpha$  and IL-1 upregulate membrane-bound and soluble E-selectin through a common pathway. *J. Surg. Res.*, 73(2): 107-112.
3. Lawrence, M.B. & Springer, T.A. (1993). Neutrophils roll on E-selectin. *J. Immunol.*, 151(11): 6338-6346.
4. Lawrence, M.B. & Springer, T.A. (1991). Leukocytes roll on a selectin at physiologic flow rates: distinction from and prerequisite for adhesion through integrins. *Cell*, 65(5): 859-873.
5. Davies, M.J., *et al.* (1993). The expression of the adhesion molecules ICAM-1, VCAM-1, PECAM, and E-selectin in human atherosclerosis. *J. Pathol.*, 171(3): 223-229.
6. Pellegatta, F., *et al.* (1997). Soluble E-selectin and intercellular adhesion molecule-1 plasma levels increase during acute myocardial infarction. *J. Cardiovasc. Pharmacol.*, 30(4): 455-460.
7. Takada, M., Nadeau, K.C., Shaw, G.D., Marquette, K.A., & Tilney, N.L. (1997). The cytokine-adhesion molecule cascade in ischemia/reperfusion injury of the rat kidney. Inhibition by a soluble P-selectin ligand. *J. Clin. Invest.*, 99(11): 2682.
8. Blann, A.D., Tse, W., Maxwell, S.J., & Waite, M.A. (1994). Increased levels of the soluble adhesion molecule E-selectin in essential hypertension. *J. Hypertens.*, 12(8): 925-928.
9. Kobayashi, T., *et al.* (1994). Elevation of serum soluble intercellular adhesion Molecule-1 (sICAM-1) and sE-Selectin levels in bronchial asthma. *Clin. Exp. Immunol.*, 96(1): 110-115.
10. Koch, A.E., *et al.* (1991). Immunolocalization of endothelial and leukocyte adhesion molecules in human rheumatoid and osteoarthritic synovial tissues. *Lab. Invest.*, 64(3): 313-320.
11. Natarajan, M., Udden, M.M., & McIntire, L. (1996). Adhesion of sickle red blood cells and damage to interleukin-1 beta stimulated endothelial cells under flow in vitro. *Blood*, 87(11): 4845-4852.
12. Krause, T. & Turner, G.A. (1999). Are selectins involved in metastasis? *Clin. Exp. Metastasis*, 17(3): 183-192.
13. Ulbrich, H., Eriksson, E.E., & Lindbom, L. (2003). Leukocyte and endothelial cell adhesion molecules as targets for therapeutic interventions in inflammatory disease. *Trends Pharmacol. Sci.*, 24(12): 640-647.
14. Chang, J., *et al.* (2010). GMI-1070, a novel pan-selectin antagonist, reverses acute vascular occlusions in sickle cell mice. *Blood*, 116(10): 1779-1786.
15. Berg, E.L., Robinson, M.K., Mansson, O., Butcher, E.C., & Magnani, J.L. (1991). A carbohydrate domain common to both sialyl Le(a) and sialyl Le(X) is recognized by the endothelial cell leukocyte adhesion molecule ELAM-1. *J. Biol. Chem.*, 266(23): 14869-14872.
16. Foxall, C., *et al.* (1992). The three members of the selectin receptor family recognize a common carbohydrate epitope, the sialyl Lewis(x) oligosaccharide. *J. Cell Biol.*, 117(4): 895-902.
17. Preston, R.C., *et al.* (2015). E-selectin ligand complexes adopt an extended high-affinity conformation. *J. Mol. Cell. Biol.*: mjev046.
18. Lipinski, C.A., Lombardo, F., Dominy, B.W., & Feeney, P.J. (2001). Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings<sup>1</sup>. *Adv. Drug Del. Rev.*, 46(1-3): 3-26.
19. Jencks, W.P. (1981). On the attribution and additivity of binding energies. *Proc. Natl. Acad. Sci. U. S. A.*, 78(7): 4046-4050.
20. Hung, A.W., *et al.* (2009). Application of Fragment Growing and Fragment Linking to the Discovery of Inhibitors of Mycobacterium tuberculosis Pantothenate Synthetase. *Angew. Chem. Int. Ed.*, 48(45): 8452-8456.
21. Whittaker, M. (2009). Picking up the pieces with FBDD or FADD: invest early for future success. *Drug Discov. Today*, 14(13-14): 623-624.
22. Baker, M. (2013). Fragment-based lead discovery grows up. *Nat. Rev. Drug Discov.*, 12(1): 5-7.
23. Duhr, S. & Braun, D. (2006). Why molecules move along a temperature gradient. *Proc. Natl. Acad. Sci. U. S. A.*, 103(52): 19678-19682.
24. Seidel, S.A., *et al.* (2012). Label-free microscale thermophoresis discriminates sites and affinity of protein-ligand binding. *Angew. Chem. Int. Ed.*, 51(42): 10656-10659.
25. Pollack, S.J., *et al.* (2011). A comparative study of fragment screening methods on the p38alpha kinase: new methods, new insights. *J. Comput. Aided Mol. Des.*, 25(7): 677-687.
26. Zihlmann, P., *et al.* (2015). Microscale Thermophoresis: A powerful opportunity for carbohydrate-based drug discovery. *Manuscript in preparation*.

27. Mashalidis, E.H., Śledź, P., Lang, S., & Abell, C. (2013). A three-stage biophysical screening cascade for fragment-based drug discovery. *Nat. Protoc.*, 8(11): 2309-2324.
28. Hämmäläinen, M.D., *et al.* (2008). Label-free primary screening and affinity ranking of fragment libraries using parallel analysis of protein panels. *J. Biomol. Screen.*, 13(3): 202-209.
29. Aretz, J., Wamhoff, E.C., Hanske, J., Heymann, D., & Rademacher, C. (2014). Computational and experimental prediction of human C-type lectin receptor druggability. *Front. Immunol.*, 5: 323-335.
30. Norman, K.E., Anderson, G.P., Kolb, H.C., Ley, K., & Ernst, B. (1998). Sialyl Lewis<sup>x</sup> (sLe<sup>x</sup>) and an sLe<sup>x</sup> Mimetic, CGP69669A, Disrupt E-Selectin-Dependent Leukocyte Rolling In Vivo. *Blood*, 91(2): 475-483.
31. Jahnke, W., Kolb, H.C., J., B.M.J., Magnani, J.L., & Ernst, B. (1997). Comparison of the Bioactive Conformations of Sialyl LewisX and a Potent Sialyl LewisX Mimic. *Angew. Chem. Int. Ed.*, 36(23): 2603-2607.
32. Cooper, A. & Royal Society of Chemistry (Great Britain) (2004). *Biophysical Chemistry*. Royal Society of Chemistry, Cambridge. pp. 109-110.
33. Wang, Z.-X. (1995). An exact mathematical expression for describing competitive binding of two different ligands to a protein molecule. *FEBS Lett.*, 360(2): 111-114.
34. Colquhoun, D. (2007). Why the Schild method is better than Schild realised. *Trends Pharmacol. Sci.*, 28(12): 608-614.
35. Hajduk, P.J., Olejniczak, E.T., & Fesik, S.W. (1997). One-Dimensional Relaxation- and Diffusion-Edited NMR Methods for Screening Compounds That Bind to Macromolecules. *J. Am. Chem. Soc.*, 119(50): 12257-12261.



## Manuscript 9

# Is it a Hydrophobic Clash that Pre-Organizes Sialyl Lewis<sup>x</sup> Mimics in the Bioactive Conformation for the Interaction with E-selectin?

Norbert Varga,<sup>#</sup> Martin Smiesko,<sup>#</sup> Mario Schubert,<sup>‡</sup> Pascal Zihlmann,<sup>#</sup>  
Roland C. Preston,<sup>#</sup> Tobias Mühlethaler,<sup>#</sup> Mirko Zierke,<sup>#</sup> Roman P. Jakob,<sup>§</sup>  
Timm Maier,<sup>§</sup> Beat Ernst<sup>#\*</sup>

<sup>#</sup>University of Basel, Institute of Molecular Pharmacy,  
Klingelbergstr. 50, 4056 Basel, Switzerland

<sup>†</sup>University of Basel, Institute of Structural Biology,  
Klingelbergstr. 70, 4056 Basel, Switzerland

<sup>‡</sup>ETH Zürich, Institute of Molecular Biology and Biophysics,  
8093 Zürich, Switzerland

\*Corresponding author.

Tel.: 0041 (0)61 267 15 51; Fax: 0041 (0)61 267 15 52;  
E-mail: beat.ernst@unibas.ch

Contributions of Pascal Zihlmann:

- Aided in manuscript preparation
- MST experiments and data analysis



## Abstract

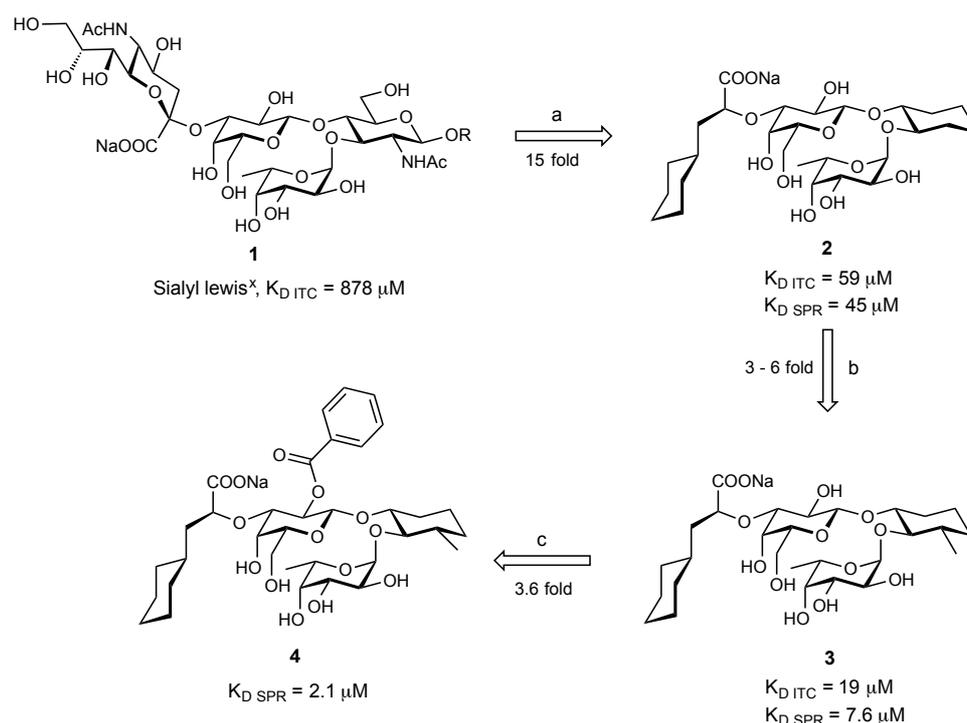
E-selectin is a cell-adhesion lectin expressed on the surface of vascular endothelial cells involved in the recruitment of leukocytes to the site of inflammation. By interacting with the tetrasaccharide sialyl Lewis<sup>x</sup> ([Neu5Ac( $\alpha$ 2-3)Gal $\beta$ (1-4)[Fuc $\alpha$ (1-3)]GlcNAc, sLe<sup>x</sup>, **1**) in the glycan of glycoproteins on leukocytes, E-selectins mediates the initial contact and enable them rolling along the endothelial vessel. This process is useful to defend the body against infections, but excessive extravasation of leukocytes can harm the tissue and is associated with a broad variety of diseases. Hence, blocking this interaction is a promising strategy to suppress an inflammation at the beginning of the cascade.

Our approach to develop a potent E-selectin antagonist is to optimize the natural ligand sLe<sup>x</sup>. We have been able to improve the affinity 15 fold by the replacement of D-Neu5Ac by (*S*)-cyclohexyl lactic acid and D-GlcNAc by (*R,R*)-cyclohexane-1,2-diol (**2**). Further improvement by a factor of 20 was achieved by methyl substitution in 3'-position of the 1,2-cyclohexyldiol (**3**) and by benzylation in 2'-position of D-Gal (**4**).

Here, we demonstrate by X-ray crystallography, NMR spectroscopy and *in silico* molecular modeling that benzylation in 2'-position of D-Gal rather stabilizes the bioactive conformation of **4** in solution than mediating new interactions with the protein. An intramolecular hydrophobic clash between the benzoate and (*S*)-cyclohexyl lactic acid was considered as cause for the improved pre-organization. Therefore we synthesized a series of antagonists to further explore the hydrophobic clash and used microscale thermophoresis to evaluate their affinities. The hydrophobic interaction between the benzoate and the (*S*)-cyclohexyl lactic acid was found support the pre-organization of the core conformation of sLe<sup>x</sup> mimics.

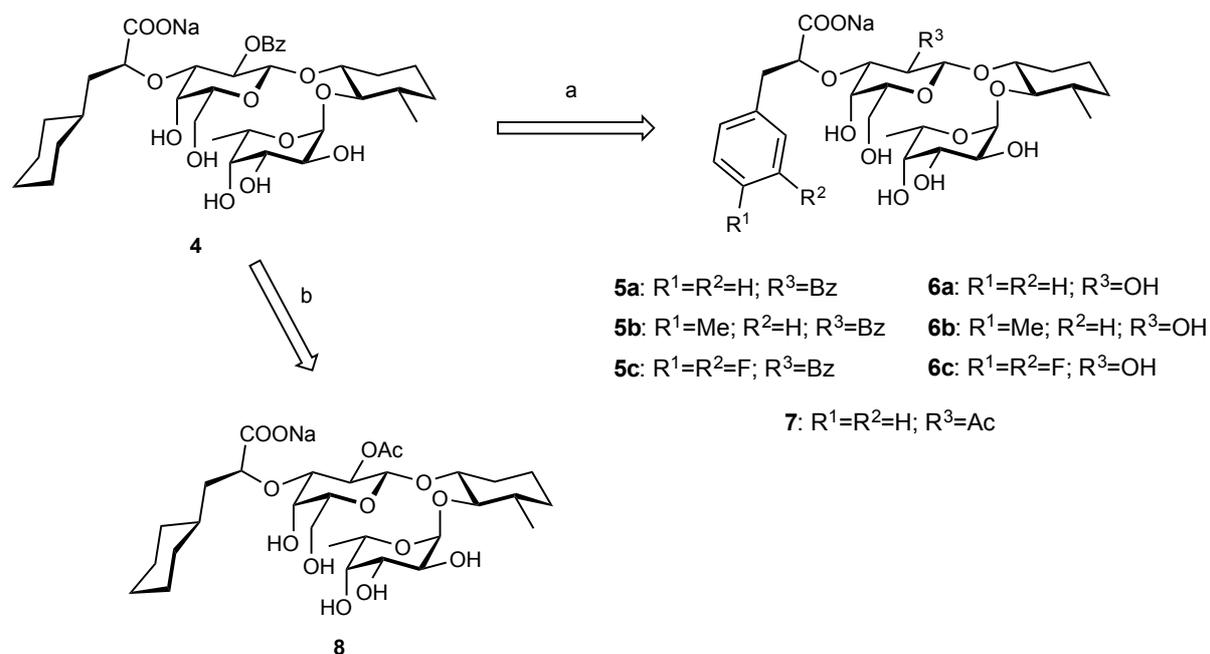
## Introduction

Lectins, such as selectins,<sup>(1-3)</sup> galectins<sup>(4,5)</sup> or siglecs<sup>(6)</sup> are subject of numerous studies as potential drug targets. Most notably, recent research has focused on selectins, a class of three cell-adhesion molecules (E-, P- and L-selectin), as they play a crucial role in the early stage of inflammation. They mediate tethering and rolling of leukocytes on endothelial cells near damaged tissue, which consequently allows for tighter integrin binding and leukocyte extravasation from blood circulation into the diseased or infected tissue.<sup>(7,8)</sup> However, exceedingly high number of leukocytes in the adjacent tissue can be destructive, leading to a range of acute and chronic diseases such as psoriasis,<sup>(9)</sup> asthma,<sup>(10)</sup> reperfusion injury<sup>(11)</sup> and rheumatoid arthritis<sup>(12)</sup>. Moreover, selectins are involved in the development of certain tumors, as they enable the migration process of tumor cells to metastasize.<sup>(13)</sup> These findings further underline the importance of selectins as a therapeutical target.



**Scheme 1.** Schematic presentation of the development of E-selectin antagonists **2-4** from sLe<sup>x</sup>, **1**. a) Replacement of D-Neu5Ac and D-GlcNAc in ligand **1** by (*S*)-cyclohexyl lactic acid and (*R,R*)-cyclohexane-1,2-diol, respectively, leading to a 15 fold increase in activity. b) An equatorial methyl group in position 3 of the d-GlcNAc mimic ((*R,R*)-cyclohexane-1,2-diol → (*1R,2R,3S*)-3-methylcyclohexane-1,2-diol) helps to stabilize the core conformation of **3**, which further improves the affinity by a factor of 3-6. c) A benzoylation of the 2-hydroxyl position of D-Gal residue additionally increases the potency of ligand **4** by a factor of 3.6.  $K_{D}$  values from isothermal titration calorimetry ( $K_{D \text{ ITC}}$ ) for compounds **1-3** are taken from Binder *et al.*<sup>(14)</sup> Steady-state  $K_{D}$  values from surface plasmon resonance ( $K_{D \text{ SPR}}$ ) data for compounds **2-4** are taken from Schwizer *et al.*<sup>(15)</sup>

Sialyl Lewis X (sLe<sup>x</sup>, **1**) is the common tetrasaccharide epitope of all natural selectin ligands. Since sLe<sup>x</sup> has only weak affinity to selectins, it is presented in multivalent forms in nature.<sup>(16,17)</sup> To develop a sLe<sup>x</sup>-based antagonist with drug like properties<sup>(1)</sup>, ligand **1** had to undergo certain structural changes. Studies employing transferred nuclear overhauser enhancement NMR spectroscopy (trNOE-NMR)<sup>(18-21)</sup>, saturation transfer difference NMR spectroscopy (STD-NMR)<sup>(22)</sup>, elucidation of the structure activity relationship (SAR)<sup>(23-26)</sup>, X-ray crystallography<sup>(27,28)</sup>, mutation studies<sup>(29)</sup>, *in-silico* molecular modeling<sup>(30)</sup> and finally isothermal titration calorimetry (ITC)<sup>(14)</sup> yielded a profound understanding of the interaction between sLe<sup>x</sup> and E-selectin on an atomic level. The main pharmacophores of **1** are the hydroxyls of the fucose moiety, the 4-OH and 6-OH of the galactose residue as well as the COOH group of the sialic acid moiety. Conformational studies revealed that the molecule contains a rigid core represented by the Galβ(1-4)[Fucα(1-3)]GlcNAc (Le<sup>x</sup>-core) trisaccharide. This rigid motif is stabilized by the exo-anomeric effect<sup>(31)</sup>, hydrophobic interactions<sup>(32)</sup>, steric effects<sup>(2,3,15)</sup> and by a non-conventional hydrogen bond<sup>(33)</sup>, consequently its conformation in solution and bound state is almost identical. Structural elucidations revealed, that the GlcNAc residue has only weak interaction with the protein and acts rather as a unique restrained linker between fucose and galactose to ensure the bioactive conformation of sLe<sup>x</sup>. The replacement of non-binding elements in molecule **1** by carefully selected hydrophobic residues yielded ligand **2**, a potent sLe<sup>x</sup> mimetic with improved drug-like properties. The  $K_D$  values determined by isothermal titration calorimetry<sup>(14)</sup> and surface plasmon resonance<sup>(15)</sup> showed a 15 fold higher activity for **2** over its natural counterpart **1**. An additional methyl group in position 3 of the 1,2-cyclohexyldiol spacer in mimic **3** stabilizes the sLe<sup>x</sup> core which leads to a 3-6 fold improvement over ligand **2**. Furthermore, benzylation in 2'-position of D-galactose increased the affinity by an additional factor of 3.6 (**3**→**4**). This effect was previously observed by Thoma *et al.*<sup>(34)</sup>, and it was presumed that the 2' benzylation has a stabilizing effect on the Le<sup>x</sup> mimic core, however, this reasoning was not further investigated.

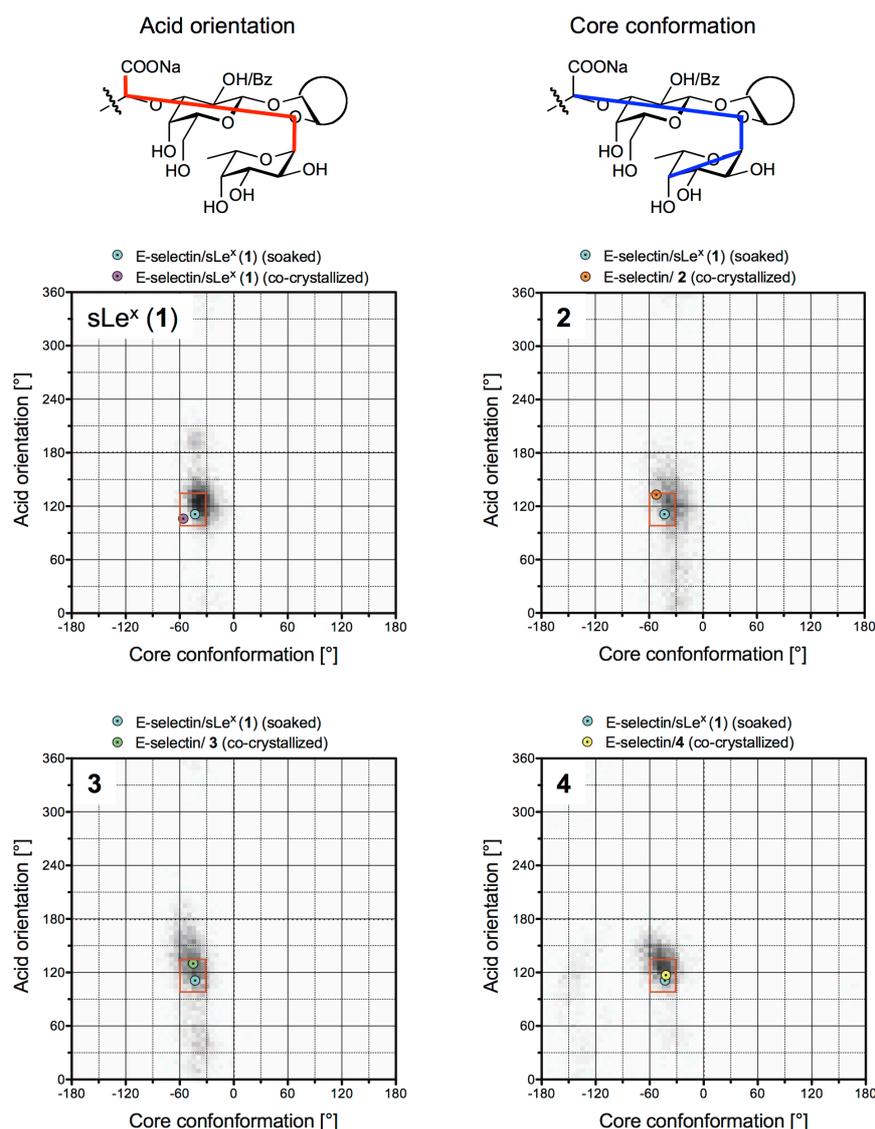


**Scheme 2.** E-selectin ligands designed and synthesized in this paper. a) the lactic acid moiety is functionalized with neutral (**5a**, **6a**), electron rich (**5b**, **6b**) and electron poor (**5c**, **6c**) phenyl rings. The hydroxyl group in position 2 of D-Gal is either un-substituted (**6a-c**) or functionalized with a benzyl group (**5a-c**) or an acetyl function (**7**) b) The benzoyl group in ligand **4** is replaced by a smaller acetyl function (**8**).

In this study, we elucidate the role of the benzoyl function on the 2' position of D-Gal in ligand **4**. By using co-crystallization X-ray experiments, NMR spectroscopy, and *in silico* molecular modeling the conformation of **4** in both solution and bound state was determined. To further investigate the influence of the substituents on C2-OH of D-Gal on the binding, a set of antagonists **5a-c**, **6a-c**, **7** and **8** was synthesized and evaluated using microscale thermophoresis for affinity ( $K_D$ ) determination. In the designed glycomimetic antagonists the lactic acid moiety is functionalized with either neutral, electron rich or electron poor phenyl rings, whereas the OH on the C2 position of D-Gal is either unprotected or functionalized with a benzoyl or an acetyl moiety. In ligand **8**, position 2 of D-Gal is protected with an acetyl group instead of the benzoyl function.

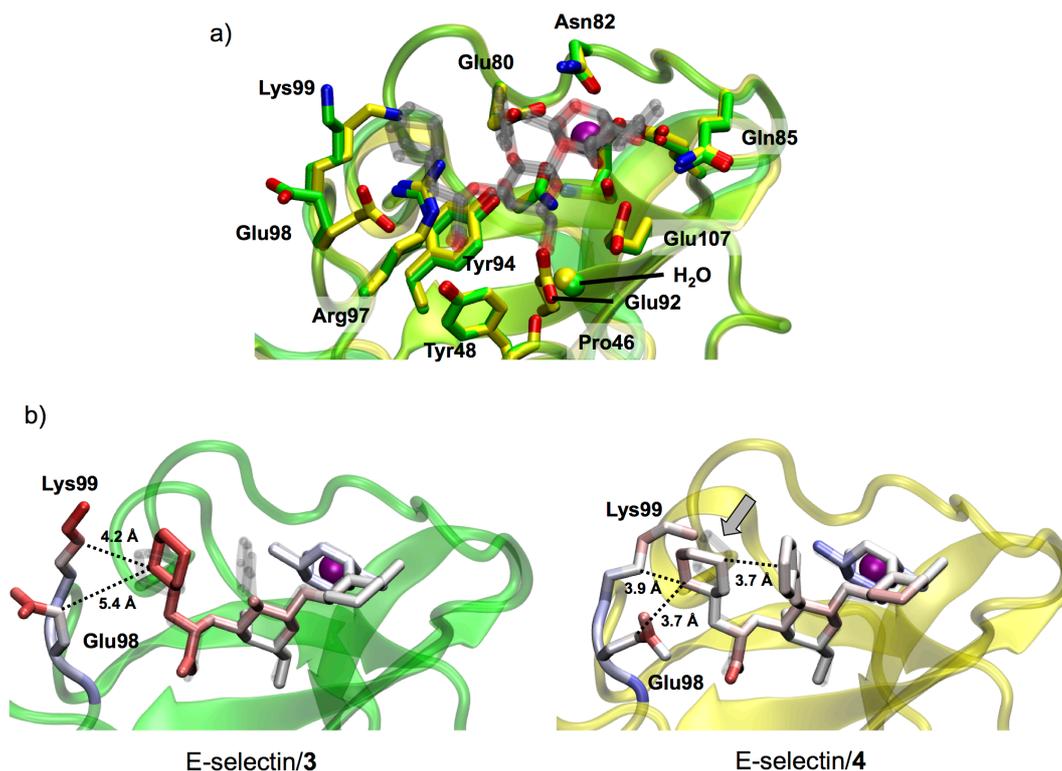
## Results & Discussion

To describe the spatial arrangement of the pharmacophores of sLe<sup>x</sup> (**1**), two dihedral angles have previously been identified.<sup>(35)</sup> The *core conformation* describes the spatial arrangement of L-Fuc and D-Gal, while the *acid orientation* describes the orientation of the carboxylic acid group relative to the Le<sup>x</sup> core (Figure 2).<sup>(36)</sup> These two torsion angles were previously determined for sLe<sup>x</sup> (**1**) in its protein-bound state by transferred nuclear overhauser-NMR (trNOE-NMR), and were later confirmed in the bound state where E-selectin crystals were soaked with sLe<sup>x</sup> (**1**).<sup>(18,27,37,38)</sup> In its bioactive conformation, the *core conformation* of sLe<sup>x</sup> (**1**) ranges between -30° and -55°, and the *acid orientation* between 105° and 130°. <sup>(20,22,39)</sup> Using molecular dynamic (MD) simulations, we calculated the two torsion angles in a solvated environment for sLe<sup>x</sup> (**1**) and for the three E-selectin antagonists (**2-4**). As shown in Figure 1, the angles for sLe<sup>x</sup> (**1**), describing the *core conformation* and the *acid orientation*, are highly populated within the experimental window. The replacement of D-GlcNAc and D-NeuNAc by smaller, hydrophobic and unsubstituted residues in compound **2**, lead to a small offset in the *core conformation*<sup>(15)</sup>, whereas the carboxylic acid gained more flexibility. The introduction of a methyl group in the 3-position of the cyclohexane-diol aided to rigidify the *core conformation* in compound **3**. The additional benzoate in the 2'-position of D-Gal in compound **4** restored the preorganization of the acid orientation to the position as determined for sLe<sup>x</sup> (**1**).



**Figure 1.** The *acid orientation* is a measurement of the carboxylic acid position relative to the Le<sup>x</sup> core. The *core conformation* describes the spatial arrangement of the core itself. These torsion angles were calculated in solution for the ligands 1–4 in absence of protein with 9.6 ns MD simulations. A darker shading correlates with higher population in the respective conformation. The red rectangle encloses experimental trNOE-NMR data of sLe<sup>x</sup> (1) bound to E-selectin.<sup>(18,37,38)</sup> The measured angles from crystallographic structures for E-selectin soaked with sLe<sup>x</sup> (1, cyan, PDB code 1G1T),<sup>(27)</sup> and the co-crystallized compounds (1, magenta; 2, orange; 3, green; 4, yellow) are given. The blue arrow indicates the small offset in core-conformation of antagonist 2.

In a co-crystallization study with sLe<sup>x</sup> (1) and antagonist 3 we recently demonstrated that the overall conformation of the protein is altered.<sup>(40)</sup> While the binding mode and the general orientation of the pharmacophores of sLe<sup>x</sup> (1) and antagonist 3 remained similar to the soaked structure, the protein undergoes significant alterations.<sup>(40)</sup> We demonstrated that small-molecule ligand interaction lead to a ligand-induced fit of the protein in solution and an overall extension of the protein, which involves rearrangements over the entire length of the lectin domain to the EGF-like domain interface.

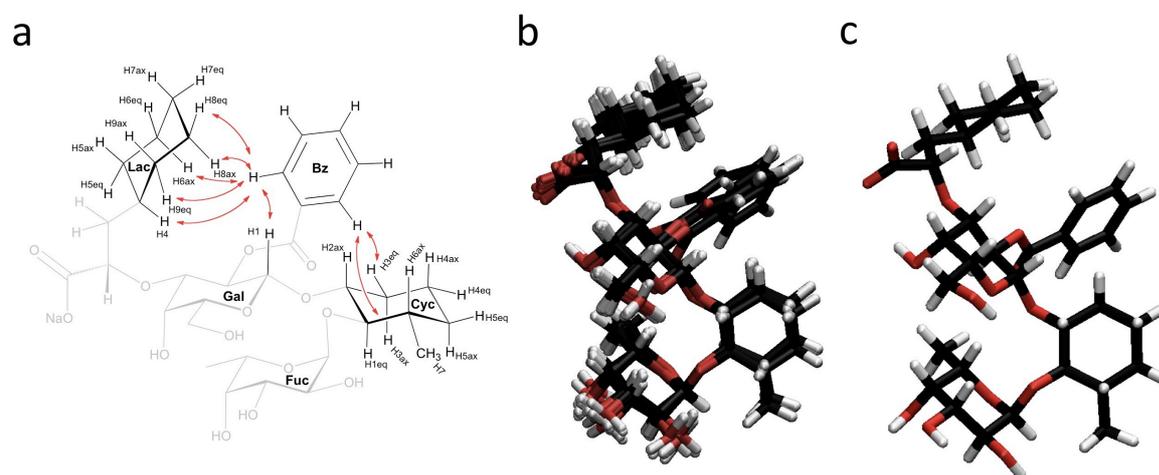


**Figure 2.** E-selectin co-crystallized with the flexible antagonists **3** (green), and the pre-organized antagonist **4** (yellow). a) Comparison of binding site residues in direct or water-mediated (Pro46) contact to the ligands (semi-transparent). All residues are in identical orientations, except for Glu89 and Lys99 side-chains. The water mediating an interaction to Pro46 is shown as green or yellow sphere. The Ca<sup>2+</sup> ion is depicted as a purple sphere. b) Comparison of crystallographic b-values in blue (below average low b-value), white (average), and red (above average). For the Glu98 and Lys99 side-chains, the b-values were normalized against their backbone atoms. For the ligands, the b-values of the (*S*)-cyclohexyl lactic acid was normalized against the modified Le<sup>x</sup> core. The arrow indicates the shift of the cyclohexane ring orientation.

For further insights into the E-selectin-ligand interactions, we herein report a co-crystal structure of compound **4** with E-selectin consisting of the four N-terminal domains (lectin domain, EGF-like domain, and first two short consensus repeats). While the binding site residues which directly interact with the L-Fuc and D-Gal moieties have identical orientations as previously observed in the co-crystal E-Selectin-ligand **3** (Figure 2a), the side-chain rotamers of Glu98 and Lys99 surrounding the (*S*)-cyclohexyl lactic acid moiety shows significant alterations. The benzoate on D-Gal in antagonist **4** can potentially create both hydrophobic and  $\sigma$ - $\pi$  interactions with the cyclohexane moiety. These intra-molecular stabilizations shifts the cyclohexane residue into a well-defined orientation compared to antagonists **3**. This re-orientation helps to accommodate the cyclohexyl unit between the side-chain residues of Glu98 and Lys99 causing an overall rigidification. Indeed, crystallographic b-factor analysis of the cyclohexyl moiety (normalized to the Le<sup>x</sup> core) reveals a significantly higher thermal motion in antagonist **3**

in comparison to antagonist **4** (Figure 2b). This suggests a stabilization of the cyclohexyl moiety in antagonist **4**, which in turn constrains the orientation of carboxylic pharmacophore and adopts a more rigid pre-organized conformation. This is reflected in the b-values for the carboxylic acid that are decreased in compound **4**.

**Compound 4 is pre-organized in the bioactive conformation in solution.** To prove our hypothesis that the acid pharmacophore of compound **4** is pre-organized in the bioactive conformation by an interaction between the benzoate and the cyclohexyl lactic acid, we solved the conformation in solution. Recently, it was demonstrated that NOESY NMR spectroscopy is an effective tool to analyze the conformation of oligosaccharides at ~277 K at 900 MHz high field in solution.<sup>(41)</sup>

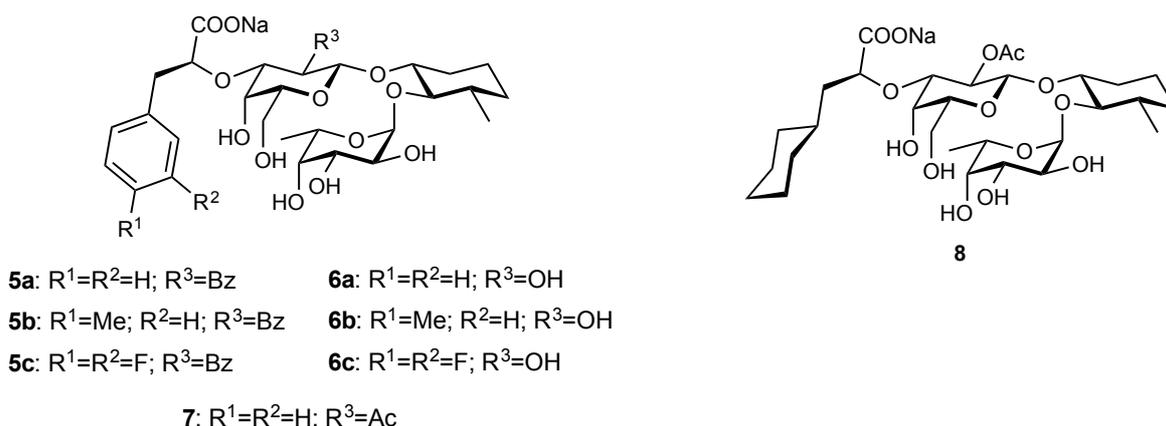


**Figure 3.** a) Schematic overview of interresidual NOEs between cyclohexane moieties and H2-benzoate (red arrows) protons; b) Ensemble of 20 lowest energy conformations of **4** at 277 K in solution, calculated and refined using experimental NOE restraints; c) Crystal structure of **4**.

The chemical shift assignment of the resonances of compound **4** was achieved by homo- and heteronuclear 2D spectra recorded in D<sub>2</sub>O (Figures S1 and S2). Using a 2D NOESY spectrum, we could extract 56 unambiguous NOE cross-peaks (36 intraresidual and 20 interresidual) between various non-exchangeable protons (Table S1 and statistics in Table S2). Specifically 10 interresidual NOEs between the benzoate and the two cyclohexane residues could be extracted (Figure 3a and Table S1). With 57 NOE derived distance restraints we were able to calculate a well-defined structural ensemble of compound **4** using Cyana<sup>(42)</sup> and subsequent refinement with AMBER 9<sup>(43)</sup> GAFF force field (Figure 3b). A comparison of this ensemble of solution conformations and the crystal structure of **4** displays an obvious similarity (Figure 3b and 3c). Furthermore, the

dihedral angles of conformations of **4** observed in solution, crystal and by molecular dynamic calculations showed values in a narrow range in the core conformation/acid orientation plot (Figure S3; statistics in Table S2). Another confirmation for the interactions between the cyclohexyl lactic acid and the benzoate are the pronounced upfield shifts of cyclohexane protons with spatial proximity to the aromatic ring in **4** compared to **3** (Figure S4). The results propose that the stabilization of the cyclohexyl lactic acid by the benzoate in **4** is established already in solution, which guarantees a pre-organization of the acid pharmacophore before the binding event.

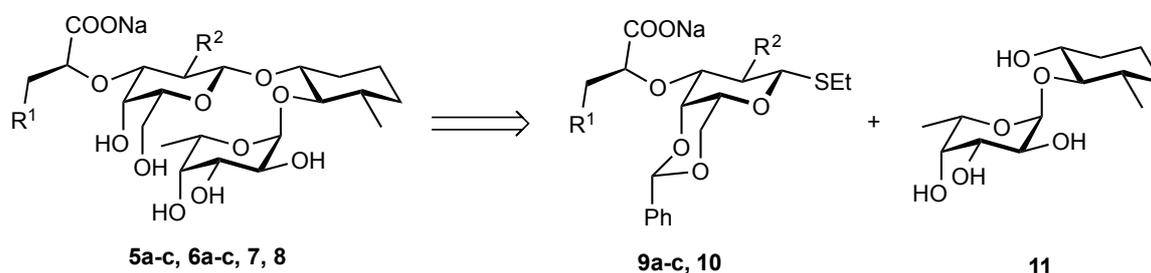
**Synthesis of the ligands.** The structural data obtained from the NMR and x-ray studies of ligand **4** suggest, that by tuning the intra-molecular interactions between the residue on position 3 of the lactic acid moiety and the substituent on position 2 of D-Gal, the pre-organization of the carboxylic pharmacophore could be improved.



**Scheme 3.** E-Selectin antagonists synthesized in this study. In order to achieve better pre-organization of the acid pharmacophore, the lactic acid moiety and the hydroxyl group in position 2 of D-Gal were functionalized with different substituents. The used moieties can potentially exhibit various intra-molecular interactions what should be reflected in the activity of the ligand.

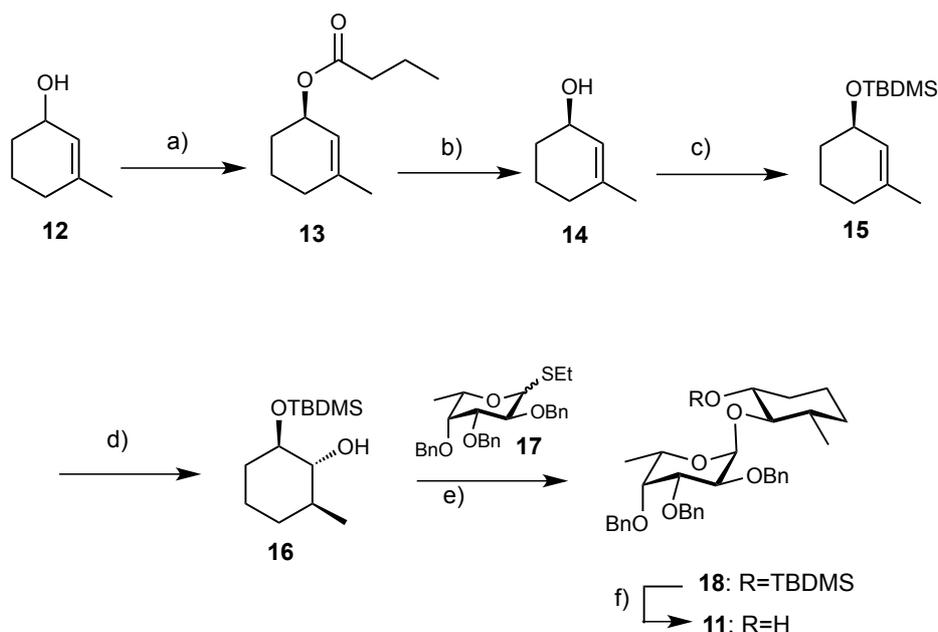
We propose a small library of E-selectin antagonists **5a-c**, in which we replace the cyclohexyl residue of lactic acid moiety in ligand **4** by phenyl groups with different substituents. The two aromatic functions could potentially create intra-molecular pi-pi or sigma-pi interactions and affect the orientation of the acid pharmacophore. For comparison, we also prepared the analogues without the benzoate on the D-Gal unit (**6a-c**). Furthermore, we replaced the benzoyl function in ligand **4** and **5a** by an acetyl group

(**8**, **7** resp.) to study the effect of smaller substituents in position 2 of D-Gal on the activity.



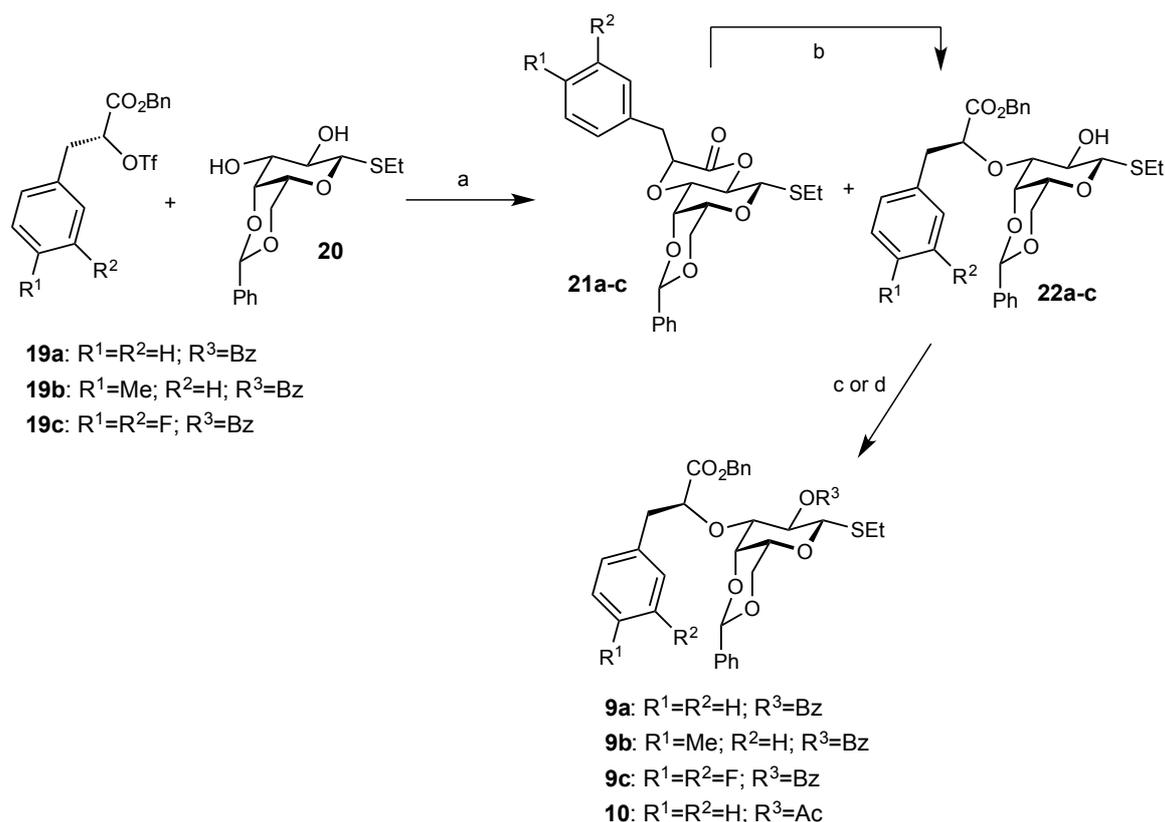
**Scheme 4.** Schematic presentation of the main building blocks (**9a-c**, **10** and **11**) used for the preparation of the desired ligands **5a-c**, **6a-c**, **7**, **8**.

The synthesis of the designed ligands starts with the preparation of building blocks **9a-c**, **10** and **11**. We recently reported the synthesis of the  $\text{Fu}\alpha(1-3)\text{GlcNAc}$  mimic **11**<sup>(15)</sup> and herein we describe a faster and more efficient approach to prepare **7** from commercially available racemic seudenol 3-methyl-2-cyclohexenol, **12**. Using optimized conditions for the butanoylation of racemic seudenol with immobilized *Candida antarctica* lipase C (Novozym 435),<sup>(44)</sup> we could isolate (*R*)-seudenolester **13**. Subsequent saponification with NaOH afforded (*R*)-seudenol **14** in 84% yield and 97.5% enantiomeric excess (*ee*). Since the protection group of the hydroxy group in **16** has to be stable under strongly basic and acidic conditions, should not hamper fucosylation by steric bulk, and finally allow cleavage under mild conditions orthogonal to benzyl protecting groups, a *tert*-butyldimethylsilyl (TBS) ether **15** was synthesized. Hydroboration followed by oxidation yielded all-*trans* **16** in 81% over two steps. Fucosylation of **16** under *in situ* anomerisation conditions<sup>(45,46)</sup> gave **18**, which was smoothly deprotected with tetrabutylammonium fluoride, affording pseudodisaccharide **11**<sup>(15)</sup> in good yield over two steps. Starting from racemic seudenol, this short sequence allowed the gram scale synthesis of **11** in 27% overall yield, requiring only four chromatographic purifications.



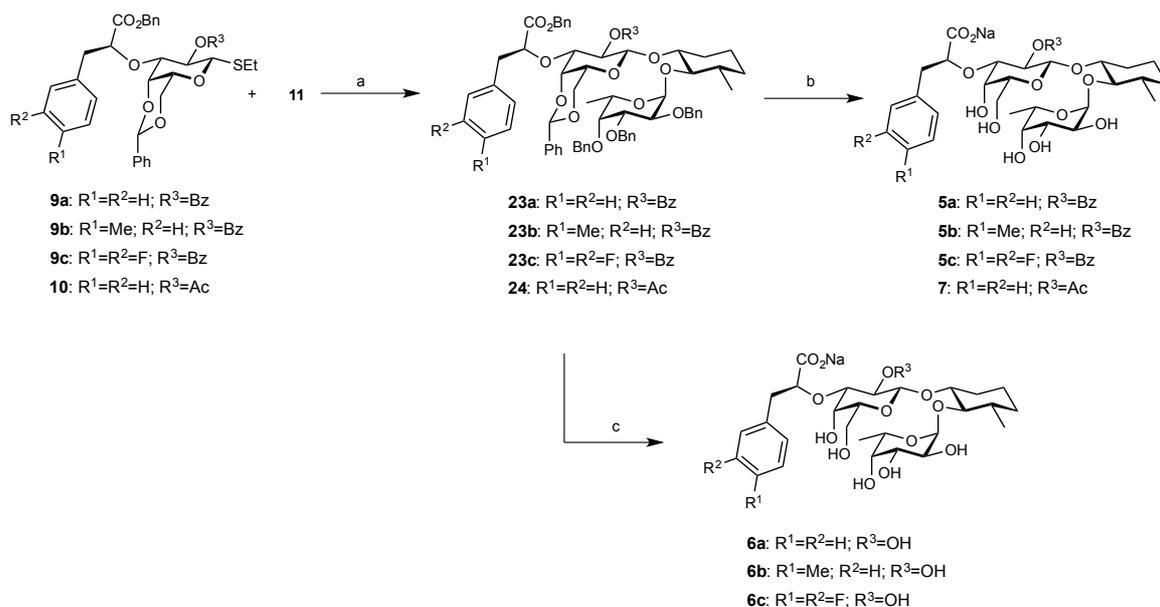
**Scheme 5.** Synthesis of disaccharide mimic **11**: a) Novozyme 435, vinylbutyrate, heptane; b) aqueous NaOH, MeOH, 84%; c) TBSCl, imidazol, DMAP, CH<sub>2</sub>Cl<sub>2</sub>; d) (i) BH<sub>3</sub>·THF, THF; (ii) H<sub>2</sub>O<sub>2</sub>, aqueous NaOH, 81% from **10**; e) CuBr<sub>2</sub>, DTBMP, TBAB, CH<sub>2</sub>Cl<sub>2</sub>, DMF, MS 4 Å, 87%; f) TBAF, THF, quant.

The preparation of building blocks **9a-c** and **10** starts with the synthesis of lactic acid derivatives **19a-c**, which can be generated from the corresponding D-amino acids (SI scheme S1 and S2). Tin mediated alkylation of position 3 of the D-galactose derivative **20**<sup>(47)</sup> with triflate **19a-c**, gave a mixture of **22a-c** and its lactonisation by-product **21a-c**, whereas **21a-c** was converted back to **22a-c** using neat benzyl alcohol and DMAP in catalytic amount. Benzoylation of the hydroxyl in **22a-c** afforded glycosyl donors **9a-c**, whereas the acetylation of **18a** resulted in donor **10**.



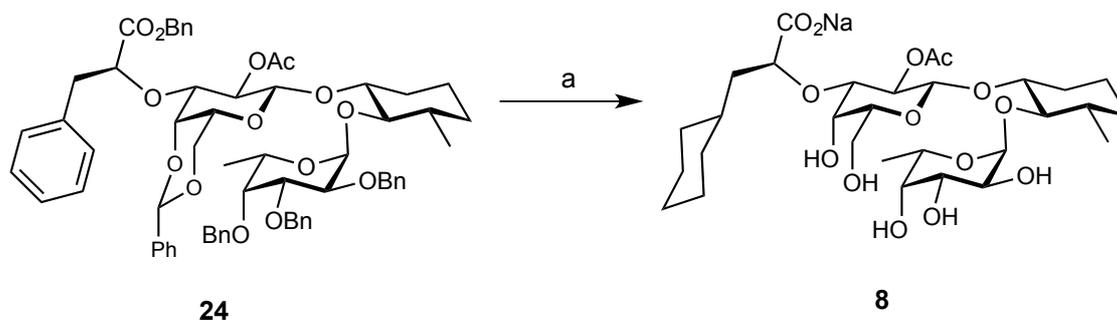
**Scheme 6.** Synthesis of galactoside derivatives **9a-c** and **10**: a) (i) Bu<sub>2</sub>SnO, MeOH; (ii) CsF, DME, (lactone - **17a**: 16%, **17b**: 25%, **17c**: 25%; benzylester - **18a**: 25%, **18b**: 0%, **18c**: 5%); b) BnOH, DMAP (cat.) (**17a**: 73%, **17b**: 80%, **17c**: 65%); c) BzCl, py, DMAP (cat.) (**8a**: 87%, **8b**: 72%, **8c**: 64%); d) Ac<sub>2</sub>O, pyridine, **9**: 78%.

In the following step disaccharide mimic **11** was glycosylated with the corresponding galactoside donor (**9a-c** and **10**) using dimethyl(methylthio)sulfonium triflate (DMTST) as promoter, what gave product **23a-c** and **24** respectively, in good yields. In order to remove the benzyl protecting groups as well as the benzylyden acetal function, compound **23a-c** was hydrogenated under heterogeneous conditions using Pd(OH)<sub>2</sub>/C as catalyst resulting in final ligands **5a-c** with a benzoyl moiety on position 2 of the galactose unit, whereas hydrogenation of **24** generated ligand **7** with an acetyl function. To obtain the fully deprotected antagonist **6a-c**, compound **23a-c** was additionally treated with LiOH after the hydrogenation step.



**Scheme 7.** Synthesis of E-selectin ligands **5a-c**, **6a-c** and **7**: a) DMTST, DCM, mol. sieves (4Å) (**19a**: 79%, **19b**: 80%, **19c**: 83%, **20**: 78%); b) H<sub>2</sub>, Pd(OH)<sub>2</sub>/C, MeOH or THF (**5a**: 42%, **5b**: 37%, **5c**: 63%, **7**: 33%) c) (i) H<sub>2</sub>, Pd(OH)<sub>2</sub>/C, MeOH; (ii) LiOH, MeOH/H<sub>2</sub>O (**6a**: 12%, **6b**: 41%, **6c**: 59%).

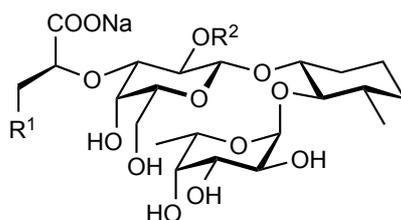
In order to generate ligand **8** with a cyclohexylactic acid residue on position 3 and an acetyl group on position 2 of the galactoside unit, molecule **24** was additionally hydrogenated in the presence of Rh/Al<sub>2</sub>O<sub>3</sub> after the removal of benzyl and benzyldine acetal protecting groups.

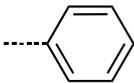
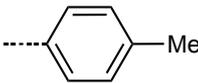
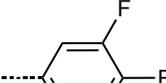


**Scheme 8.** Synthesis of sLe<sup>x</sup> mimic **8**, a) i) H<sub>2</sub>, Pd(OH)<sub>2</sub>/C, THF; ii) H<sub>2</sub>, Rh/Al<sub>2</sub>O<sub>3</sub>, H<sub>2</sub>O/dioxane/AcOH : 2/1/0.5 (10%)

**Benzoylation in 2'-position of the D-galactose residue effects affinity in lactic acid derivatives.** The prepared ligands **5a-c**, **6a-c**, **7** and **8** were tested using microscale thermophoresis (MST) along with the previously reported antagonists **3** and **4**. MST is a new method to determine the dissociation constant ( $K_D$ ) of a binding reaction that was recently applied to E-selectin.<sup>(48)</sup> The technique is based on the phenomenon that proteins move different along a temperature gradient if they undergo a change in size, charge or

hydration shell.<sup>(49)</sup> If binding a ligand changes one of these parameters the amount of bound protein can be monitored by a change in fluorescence. Measuring the thermophoresis of the protein with different amounts of ligand allows deducing the binding affinity. MST turned out to be a highly valid and reliable method to determine binding affinities of E-selectin – ligand interactions. The results are summarized in table 1:

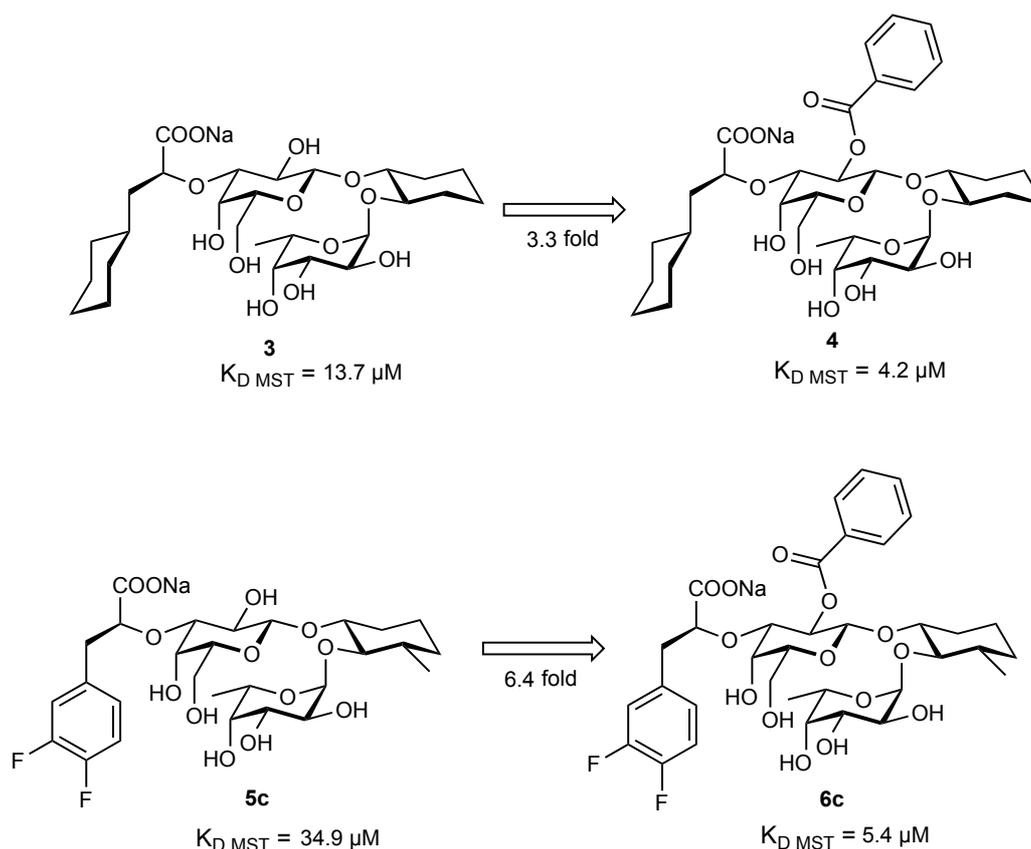


entry	R <sub>1</sub>	R <sub>2</sub>	Ligand	K <sub>D</sub>	rK <sub>D</sub>
1		H	<b>3</b>	13.69 ± 0.93	3.28
2		Bz	<b>4</b>	4.18 ± 0.47	1.00
3		Ac	<b>8</b>	4.85 ± 0.41	1.16
4		H	<b>5a</b>	38.99 ± 1.69	9.34
5		Bz	<b>6a</b>	7.81 ± 0.49	1.87
6		Ac	<b>7</b>	9.97 ± 0.54	2.39
7		H	<b>5b</b>	35.16 ± 1.66	8.42
8		Bz	<b>6b</b>	5.75 ± 0.36	1.38
9		H	<b>5c</b>	34.90 ± 1.65	8.36
10		Bz	<b>6c</b>	5.42 ± 0.37	1.30

**Table 1.** Affinity data of E-Selectin antagonists **3**, **4**, **5a-c**, **6a-c**, **7** and **8** determined by microscale thermophoresis (MST) assay.

Ligands, which contain the phenyl lactic acid derivatives and lack the benzoyl group on position 2' of D-Gal (**5a-c**) showed lower activity in comparison to the cyclohexyl lactic acid analogue **3** by a factor of 2.5. However, an additional benzoylation of ligands **6a-c** improved the potency up to a factor of 6.5 (**5c**→**6c**), whereas the difference between **3** and **4** is only 3.2 fold. This suggests, that the benzoate in position 2 of D-Gal helps to stabilize the acid orientation in a similar fashion as in ligand **4**, and sets the K<sub>D</sub> values of

**6a-c** and **4** in the same range. The substitution on the phenyl rings had only a small effect on the potency of the ligands. Interestingly, ligands with electron donating (**6b**) and electron withdrawing (**6c**) substituents on the phenyl groups exhibited higher potency in comparison to the un-functionalized derivative (**6a**).



**Scheme 9.** Substitution in position 2' of D-Gal in E-selectin ligand **3** improves the affinity by a factor of 3.2 whereas the increase in the phenyl lactic acid derivatives is up to 6.5 fold. **5c**→**6c**

An acetyl group in position 2' of D-Gal in ligands **7** and **8** had similar effect on the activity as the benzoyl function in **4** and **6a**, respectively. Activity evaluation of the prepared antagonists showed that substitution in position 2' of the galactose unit improves the activity of the ligands in the same manner as in inhibitor **4**, what further underlines the importance of the substitution in this position. This effect can be achieved by a benzoyl function as well as an acetyl moiety, a functional group with a much smaller molecular weight. Ligand **8** revealed, that no aromatic group is required to achieve the desired acid pre-organization.

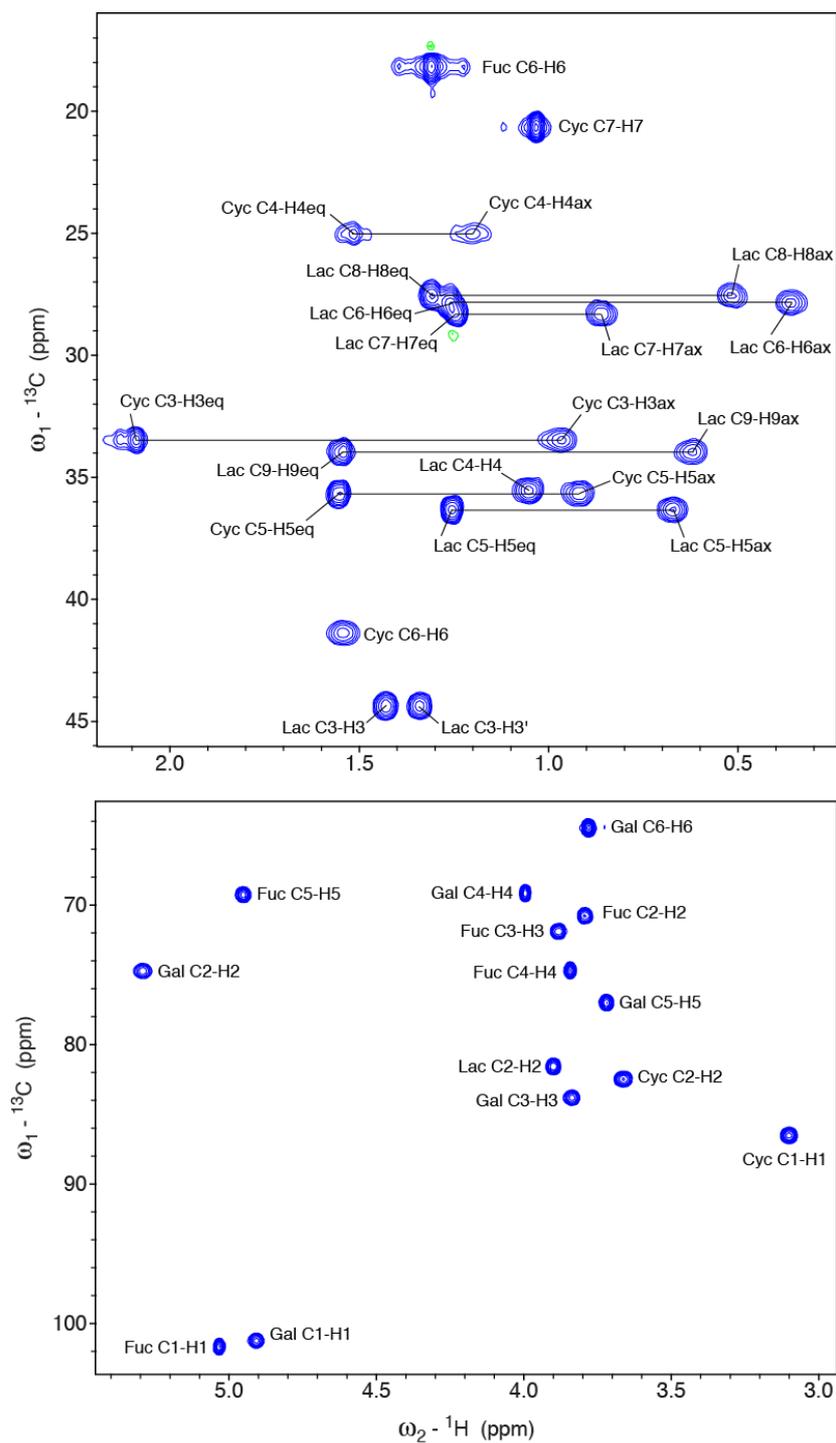
In summary, we could demonstrate with by NMR spectroscopy, X-ray crystallography and *in silico* modeling that pre-organization rather than additional interactions are responsible for the improved affinity of **4**. Using NOE derived distance restraints we were able to calculate solution structures of **4** that are highly comparable to the bound conformation in the co-crystal structure. The crystal structure furthermore confirmed that the lactic acid substituents and the benzoate in position 2' of D-Gal are pointing to the solvent and interact with each other by hydrophobic effects. An MD-simulation could show that the acid pre-organization of **4** is improved. Additionally, a b-factor analysis revealed lower thermal motion on the (*S*)-cyclohexyl lactic acid site of **4** compared to **3**, which presumably constrains the orientation of carboxylic pharmacophore. However, substituting the benzoyl by a non-aromatic acetyl group in position 2' of D-Gal (**7**, **8**) had a comparable effect, while having a much smaller molecular weight. It remains an issue of further investigation to study the impact of smaller substituents in position 2' of D-Gal.

## References

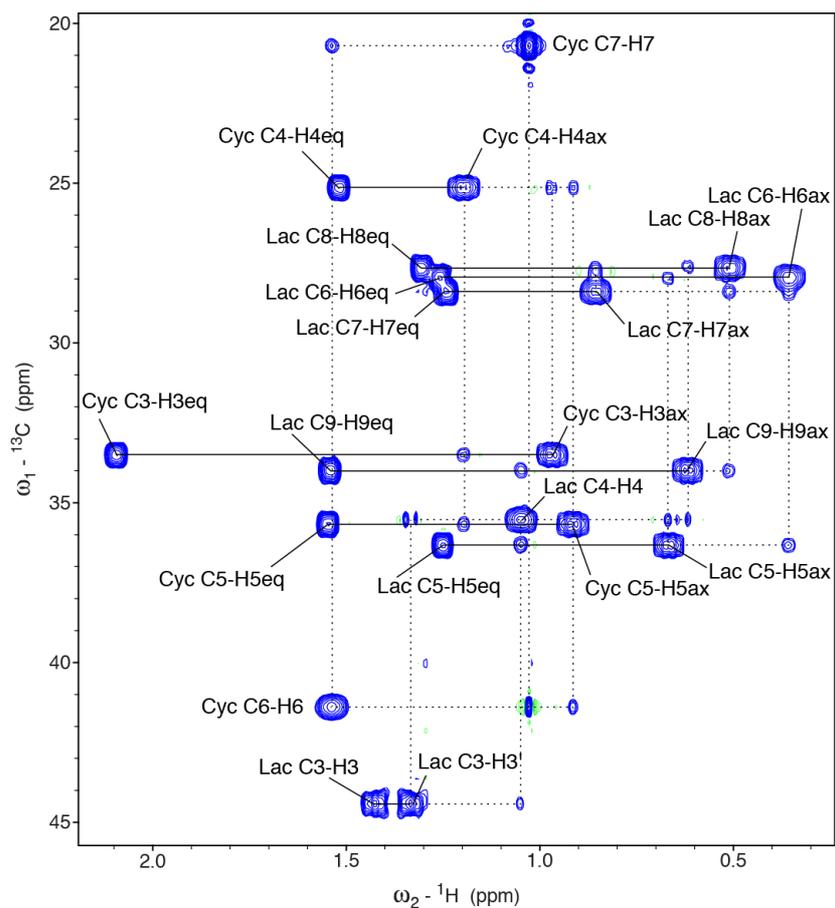
1. Ernst, B. & Magnani, J.L. (2009). From carbohydrate leads to glycomimetic drugs. *Nat. Rev. Drug Discov.*, 8(8): 661-677.
2. Kaila, N. & Thomas, B.E. (2002). Design and synthesis of sialyl Lewis<sup>x</sup> mimics as E- and P-selectin inhibitors. *Med. Res. Rev.*, 22(6): 566-601.
3. Simanek, E.E., McGarvey, G.J., Jablonowski, J.A., & Wong, C.H. (1998). Selectin-Carbohydrate Interactions: From Natural Ligands to Designed Mimics. *Chem. Rev.*, 98(2): 833-862.
4. Oberg, C.T., Leffler, H., & Nilsson, U.J. (2011). Inhibition of galectins with small molecules. *Chimia (Aarau)*, 65(1-2): 18-23.
5. Ingrassia, L., et al. (2006). Anti-galectin compounds as potential anti-cancer drugs. *Curr. Med. Chem.*, 13(29): 3513-3527.
6. Crocker, P.R., Paulson, J.C., & Varki, A. (2007). Siglecs and their roles in the immune system. *Nat. Rev. Immunol.*, 7(4): 255-266.
7. Kansas, G.S. (1996). Selectins and their ligands: current concepts and controversies. *Blood*, 88(9): 3259-3287.
8. Cines, D.B., et al. (1998). Endothelial cells in physiology and in the pathophysiology of vascular disorders. *Blood*, 91(10): 3527-3561.
9. Bock, D., Philipp, S., & Wolff, G. (2006). Therapeutic potential of selectin antagonists in psoriasis. *Expert Opin. Investig. Drugs*, 15(8): 963-979.
10. Romano, S.J. (2005). Selectin antagonists : therapeutic potential in asthma and COPD. *Treat. Respir. Med.*, 4(2): 85-94.
11. Lefer, D.J., Flynn, D.M., Phillips, M., Ratcliffe, M., & Buda, A.J. (1994). A novel sialyl Lewis<sup>x</sup> analog attenuates neutrophil accumulation and myocardial necrosis after ischemia and reperfusion. *Circulation*, 90(5): 2390-2401.
12. Sfrikakis, P.P. & Mavrikakis, M. (1999). Adhesion and lymphocyte costimulatory molecules in systemic rheumatic diseases. *Clin. Rheumatol.*, 18(4): 317-327.
13. Witz, I.P. (2006). The involvement of selectins and their ligands in tumor-progression. *Immunol. Lett.*, 104(1-2): 89-93.
14. Binder, F.P., Lemme, K., Preston, R.C., & Ernst, B. (2012). Sialyl Lewis<sup>x</sup>: A “Pre-Organized Water Oligomer”? *Angew. Chem. Int. Ed.*, 51(29): 7327-7331.
15. Schwizer, D., et al. (2012). Pre-organization of the core structure of E-selectin antagonists. *Chem. Eur. J.*, 18(5): 1342-1351.
16. Spiess, M. (1990). The asialoglycoprotein receptor: a model for endocytic transport receptors. *Biochemistry*, 29(43): 10009-10018.
17. Mammen, M., Choi, S., & Whitesides, G. (1998). Polyvalent interactions in biological systems: Implications for design and use of multivalent ligands and inhibitors. *Angew. Chem. Int. Ed.*, 37(20): 2755-2794.
18. Harris, R., et al. (1999). Stable-isotope-assisted NMR studies on C-13-enriched sialyl Lewis<sup>x</sup> in solution and bound to E-selectin. *J. Am. Chem. Soc.*, 121(11): 2546-2551.
19. Poppe, L., Brown, G., Philo, J., Nikrad, P., & Shah, B. (1997). Conformation of sLe<sup>x</sup> tetrasaccharide, free in solution and bound to E-, P-, and L-selectin. *J. Am. Chem. Soc.*, 119(7): 1727-1736.
20. Scheffler, K., et al. (1997). Application of homonuclear 3D NMR experiments and 1D analogs to study the conformation of sialyl Lewis<sup>x</sup> bound to E-selectin. *J. Biomol. NMR*, 9(4): 423-436.
21. Peters, T., et al. (1995). Determination of the Bioactive Conformation of the Carbohydrate Ligand in the E-Selectin/Sialyl Lewis<sup>x</sup> Complex. *Angew. Chem. Int. Ed.*, 34(17): 1841-1844.
22. Rinnbauer, M., et al. (2003). Epitope mapping of sialyl Lewis<sup>x</sup> bound to E-selectin using saturation transfer difference NMR experiments. *Glycobiology*, 13(6): 435-443.
23. Tyrrell, D., et al. (1991). Structural requirements for the carbohydrate ligand of E-selectin. *Proc. Natl. Acad. Sci. U. S. A.*, 88(22): 10372-10376.
24. Brandley, B.K., et al. (1993). Structure-function studies on selectin carbohydrate ligands. Modifications to fucose, sialic acid and sulphate as a sialic acid replacement. *Glycobiology*, 3(6): 633-641.
25. Ramphal, J.Y., et al. (1994). Structure-activity relationships of sialyl Lewis<sup>x</sup>-containing oligosaccharides. 1. Effect of modifications of the fucose moiety. *J. Med. Chem.*, 37(21): 3459-3463.
26. Stahl, W., Sprengard, U., Kretschmar, G., & Kunz, H. (1994). Synthesis of Deoxy Sialyl Lewis<sup>x</sup> Analogues, Potential Selectin Antagonists. *Angew. Chem. Int. Ed.*, 33(20): 2096-2098.

27. Somers, W.S., Tang, J., Shaw, G.D., & Camphausen, R.T. (2000). Insights into the molecular basis of leukocyte tethering and rolling revealed by structures of P- and E-selectin bound to SLe<sup>x</sup> and PSGL-1. *Cell*, 103(3): 467-479.
28. Graves, B.J., *et al.* (1994). Insight into E-selectin/ligand interaction from the crystal structure and mutagenesis of the lec/EGF domains. *Nature*, 367(6463): 532-538.
29. Erbe, D.V., *et al.* (1992). Identification of an E-selectin region critical for carbohydrate recognition and cell adhesion. *J. Cell Biol.*, 119(1): 215-227.
30. Ishida, T. (2010). Computational modeling of carbohydrate-recognition process in E-selectin complex: structural mapping of sialyl Lewis X onto ab initio QM/MM free energy surface. *J. Phys. Chem. B*, 114(11): 3950-3964.
31. Wolfe, S., Myung-Hwan, W., & Mitchell, D.J. (1979). On the magnitudes and origins of the “anomeric effects”, “exo-anomeric effects”, “reverse anomeric effects”, and C-X and C-Y bond lengths in XCH<sub>2</sub>YH molecules. *Carbohydr. Res.*, 69(1): 1-26.
32. Titz, A., *et al.* (2012). Conformational Constraints: Nature Does It Best with Sialyl Lewis<sup>x</sup>. *Eur. J. Org. Chem.*, (28): 5534-5539.
33. Zierke, M., *et al.* (2013). Stabilization of Branched Oligosaccharides: Lewis<sup>x</sup> Benefits from a Nonconventional C–H···O Hydrogen Bond. *J. Am. Chem. Soc.*, 135(36): 13464-13472.
34. Thoma, G., Banteli, R., Jahnke, W., Magnani, J.L., & Patton, J.T. (2001). A Readily Available, Highly Potent E-Selectin Antagonist. *Angew. Chem. Int. Ed. Engl.*, 40(19): 3644-3647.
35. Lis, H. & Sharon, N. (1998). Lectins: Carbohydrate-specific proteins that mediate cellular recognition. *Chem. Rev.*, 98(2): 637-674.
36. Kolb, H.C. & Ernst, B. (1997). Development of tools for the design of selectin antagonists. *Chem. Eur. J.*, 3(10): 1571-1578.
37. Cooke, R.M., Hale, R.S., Lister, S.G., Shah, G., & Weir, M.P. (1994). The Conformation of the Sialyl Lewis X Ligand Changes Upon Binding to E-Selectin. *Biochemistry*, 33(35): 10591-10596.
38. Hensley, P., *et al.* (1994). The Soluble Form of E-Selectin Is an Asymmetric Monomer - Expression, Purification, and Characterization of the Recombinant Protein. *J. Biol. Chem.*, 269(39): 23949-23958.
39. Scheffler, K., *et al.* (1995). Determination of the bioactive conformation of the carbohydrate ligand in the E-selectin/sialyl Lewis<sup>x</sup> complex. *Angew. Chem. Int. Ed.*, 34(17): 1841-1844.
40. Preston, R.C., *et al.* (2015). E-selectin ligand complexes adopt an extended high-affinity conformation. *J. Mol. Cell. Biol.*: mjev046.
41. Aeschbacher, T. (2013). *Assisted and automated RNA resonance assignment using chemical shift statistics and the finding of a widespread C-H···O hydrogen-bond stabilized glycoepitope architecture*. Dissertation, ETH, Zürich.
42. Herrmann, T., Güntert, P., & Wüthrich, K. (2002). Protein NMR structure determination with automated NOE-identification in the NOESY spectra using the new software ATNOS. *J. Biomol. NMR*, 24(3): 171-189.
43. Case, D., *et al.* (2005). The Amber biomolecular simulation programs. *J. Comput. Chem.*, 26(16): 1668-1688.
44. ter Halle, R., *et al.* (2004). Development of a Practical Multikilogram Production of (R)-Seudenol by Enzymatic Resolution. *Org. Process Res. Dev.*, 8(2): 283-286.
45. Lemieux, R.U. & Morgan, A.R. (1965). Mechanisms for the Reactions of the Anomeric Tetra-O-Acetyl-D-Glucopyranosyl Bromides with Pyridine and 2-Pyridone. *Can. J. Chem.*, 43(8): 2214-2221.
46. Lemieux, R.U., Hendriks, K.B., Stick, R.V., & James, K. (1975). Halide ion catalyzed glycosidation reactions. Syntheses of .alpha.-linked disaccharides. *J. Am. Chem. Soc.*, 97(14): 4056-4062.
47. Gridley, J.J., Hacking, A.J., Osborn, H.M.I., & Spackman, D.G. (1998). Regioselective lipase-catalysed acylation of 4,6-O-benzylidene- $\alpha$ - and  $\beta$ -d-pyranoside derivatives displaying a range of anomeric substituents. *Tetrahedron*, 54(49): 14925-14946.
48. Zihlmann, P., *et al.* (2015). Microscale Thermophoresis: A powerful opportunity for carbohydrate-based drug discovery. *Manuscript in preparation*.
49. Jerabek-Willemsen, M., Wienken, C., Braun, D., Baaske, P., & Duhr, S. (2011). Molecular Interaction Studies Using Microscale Thermophoresis. *Assay Drug Dev. Technol.*, 9(4): 342-353.

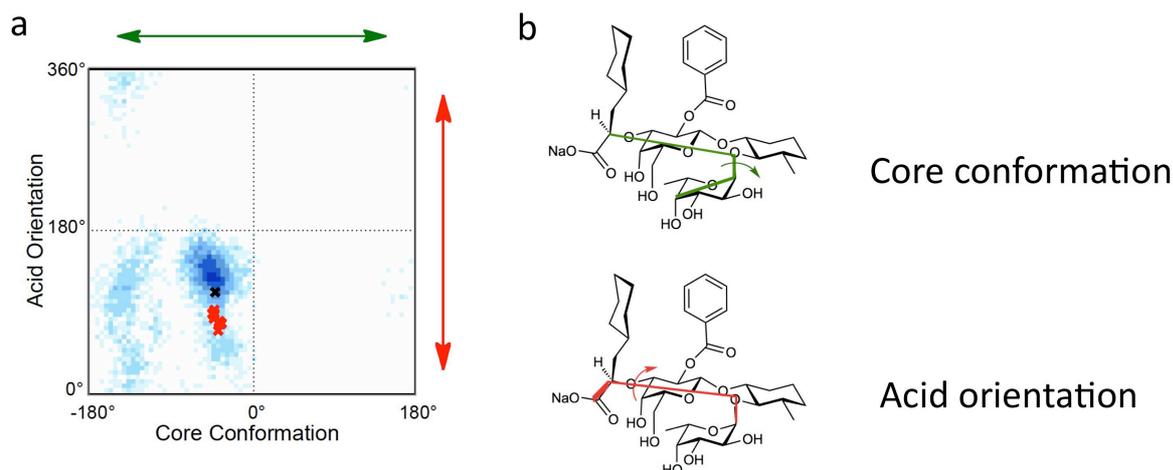
## Supporting Information



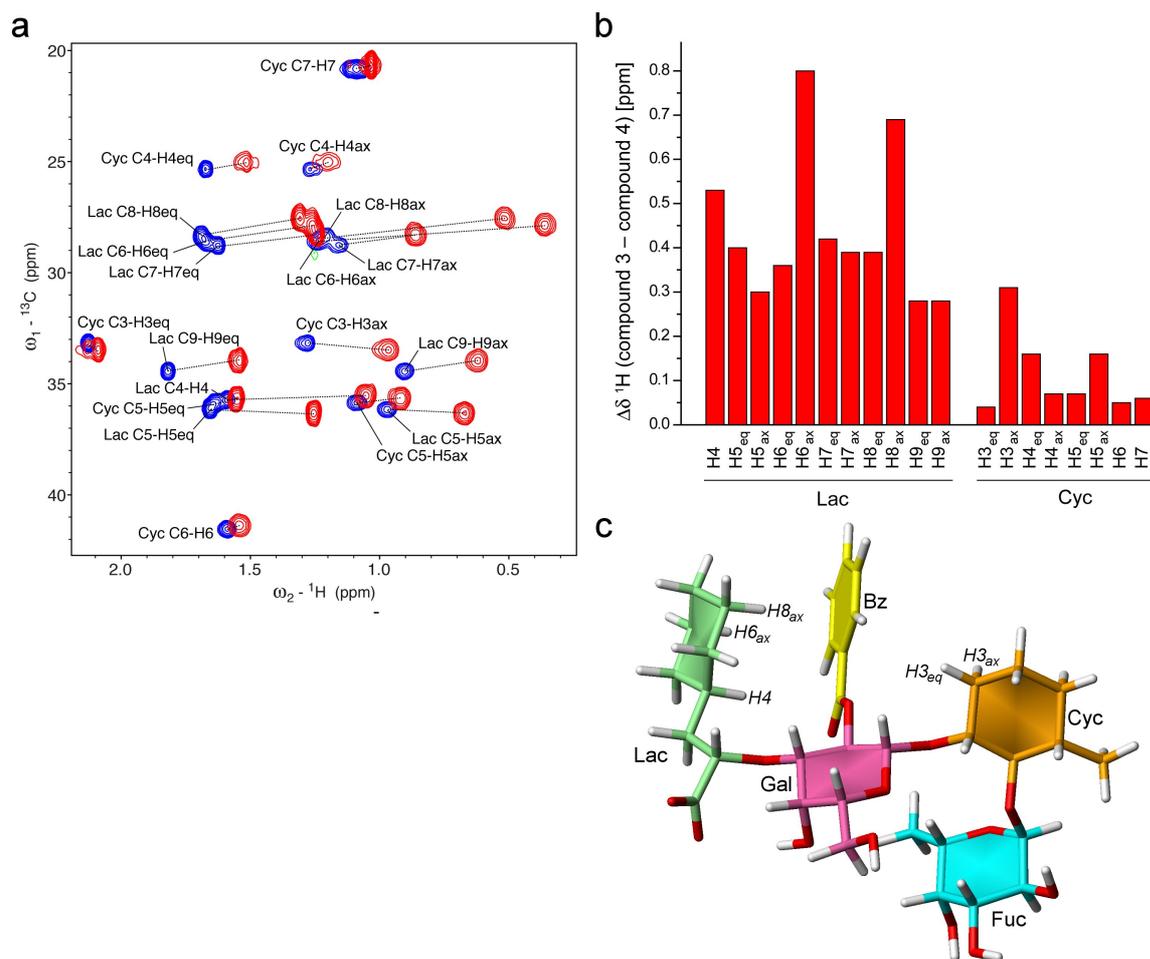
**Figure S1.**  $^{13}\text{C}$ - $^1\text{H}$  HSQC of **4** recorded at 750 MHz and 275 K with assignments. The aliphatic region is shown on top, the carbohydrate region at the bottom.



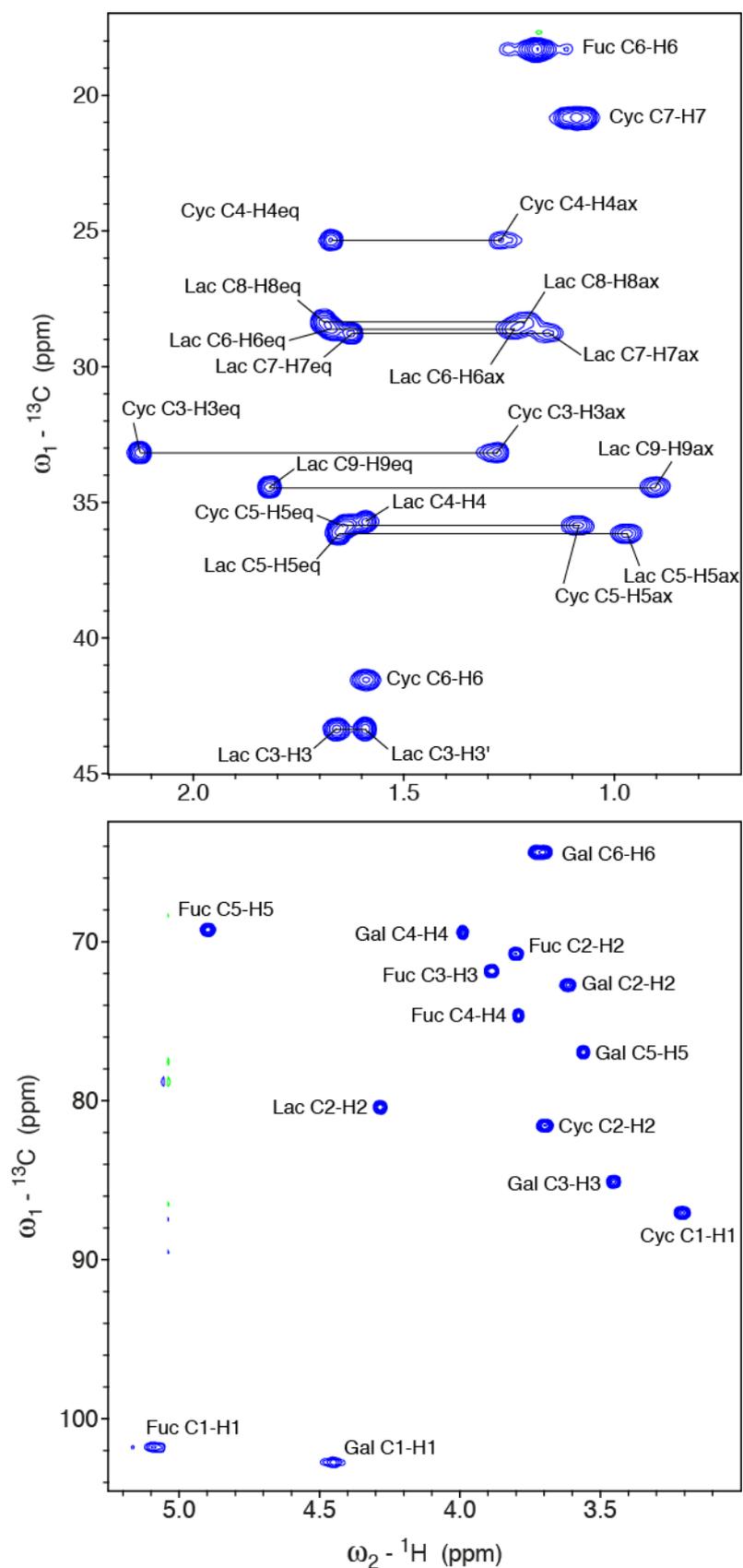
**Figure S2.** Chemical shift assignment of cyclohexane resonances using a  $^{13}\text{C}$ - $^1\text{H}$  HMQC-COSY. The small unassigned signals originating from two-bond scalar couplings  $^2J_{\text{CH}}$  are ideally suited to assign the cyclohexane resonances by following the dotted lines. Resonances from axial protons were distinguished from equatorial ones by their strength of  $^3J_{\text{HH}}$  scalar coupling constants and NOE signal intensities (data not shown). As is typical the axial resonances are more upfield than the equatorial ones.



**Figure S3.** a) Acid orientation / core conformation blot of GMI1077: crystal structure of  $\text{sLe}^x$  (black cross), (Lit: Somers *et al.*  $\text{sLe}^x$  crystal) molecular dynamics simulations (blue) and the ensemble of 20 solution conformations determined by NMR at 277 K at 900 MHz high field (red crosses). b) Schematic overview of core conformation (= relative orientation of D-galactose and L-fucose) and acid orientation (= tilting angle of the carboxylic acid relative to the core).



**Figure S4.** Confirmation of the three-dimensional solution structure by chemical shift deviations caused by the benzoyl group. a) Overlay of HSQC spectra of compound **4** (red) and compound **3** (blue). The region with signals of the methyl-cyclohexane diol mimic and the (*S*)-cyclohexane lactic acid are shown. Corresponding signals are connected by dotted lines. The benzoyl group causes a large variety of upfield chemical shift changes, which are typically caused by ring current effects on protons perpendicular to the aromatic ring plane. b) Histogram of the chemical shift deviations of cyclohexane protons caused by the benzoyl group. c) Representative of the obtained structural ensemble by solution NMR spectroscopy. Relevant protons in the proximity of the aromatic ring are labeled. The closest protons perpendicular to the aromatic plane are expected to experience the largest ring current effect and indeed they show the largest chemical shift perturbation.



**Figure S5.**  $^{13}\text{C}$ - $^1\text{H}$  HSQC of 3 recorded at 750 MHz and 275 K with assignments. The aliphatic region is shown on top, the carbohydrate region at the bottom.

**Table S1.** Intra and interresidual NOEs of **4** at 275K with the corresponding calculated  $^1\text{H}$ - $^1\text{H}$  distances.

proton pair	S/N of NOEs cross peaks	Corresponding $^1\text{H}$ - $^1\text{H}$ distance <sup>b</sup>	proton pair	S/N of NOEs cross peaks	Corresponding $^1\text{H}$ - $^1\text{H}$ distance <sup>b</sup>
Intra			Intra		
Lac H2-H3ax	476	2.6	Cyc H3eq-H5ax	167	3.1
Lac H2-H3eq	943	2.3	Cyc H3ax-H4eq	848	2.4
Lac H2-H4eq	187	3.0	Cyc H5ax-H7	303 <sup>f</sup>	2.8
Lac H3ax-H4eq	441	2.6	Fuc H1-H2	798	2.4
Lac H4-H6ax	268 <sup>a</sup>	2.9	Fuc H1-H3	113 <sup>a</sup>	3.3
Lac H4-H8ax	231 <sup>a</sup>	2.9	Fuc H3-H5	962	2.3
Lac H3ax-H5ax	312	2.8	Fuc H4-H5	1494	2.1
Lac H3eq-H5ax	407	2.7	Inter		
Lac H5ax-H6ax	341	2.7	Lac H9eq-Gal H3	223	2.9
Lac H7ax-H8eq	413 <sup>a</sup>	2.7	Lac H4-Bz H2	88 <sup>a,c</sup>	3.4
Lac H2-H9eq	1012	2.3	Lac H6ax-Bz H2	56 <sup>a,c</sup>	3.7
Lac H2-H9ax	261	2.9	Lac H8eq-Bz H2	84 <sup>a,c</sup>	3.5
Lac H3ax-H9ax	102	3.4	Lac H8ax-Bz H2	114 <sup>a,c</sup>	3.3
Lac H7ax-H9ax	335	2.7	Lac H9eq-Bz H2	80 <sup>a,c</sup>	3.5
Gal H1-H5	1271	2.2	Gal H1-Cyc H2	1221	2.2
Gal H1-H2	277	2.8	Gal H1-Cyc H3a	969	2.3
Gal H1-H3	541	2.5	Gal H1-Cyc H3b	167	3.1
Gal H2-H3	198 <sup>a</sup>	3.0	Gal H2-Fuc H5	387	2.7
Gal H4-H5	1568 <sup>a</sup>	2.1	Gal H2-Fuc H6	241 <sup>f</sup>	2.9
Cyc H1-H2	290	2.8	Gal H6-Fuc H6	62 <sup>e,f</sup>	3.6
Cyc H1-H3eq	146	3.2	Gal H1-Bz H2	57 <sup>a,c</sup>	3.7
Cyc H1-H3ax	712	2.4	Bz H2-Cyc H3eq	157 <sup>e</sup>	3.1
Cyc H1-H5ax	612	2.5	Bz H2-Cyc H3ax	123 <sup>a,c</sup>	3.2
Cyc H1-H7	424 <sup>f</sup>	2.6	Fuc H1-Cyc H1	1576	2.1
Cyc H2-H3	929	2.3	Fuc H1-Cyc H6	173	3.1
Cyc H2-H4ax	625 <sup>a</sup>	2.5	Fuc H1-Cyc H7	777 <sup>f</sup>	2.4
Cyc H3eq-H3ax	4648 <sup>d</sup>	1.77	Fuc H5-Cyc H1	173 <sup>a</sup>	3.1
Cyc H3eq-H4eq	894	2.3	Fuc H5-Cyc H2	140	3.2
Cyc H3eq-H4ax	742	2.4			

<sup>a</sup> Only one cross-peak was used because of artifacts or overlap.

<sup>b</sup> The  $^1\text{H}$ - $^1\text{H}$  distances were calculated from experimentally obtained NOE intensities using the H3a-H3b cross-peak of Cyc as a reference with a distance of 1.77 Å assuming a  $r^{-6}$  dependence of the NOE intensities.

<sup>d</sup> Reference restraints for the  $^{13}\text{C}$ -filtered-filtered NOESY.

<sup>e,f</sup> Signal to noise ratios from cross-peaks involving methyl- or methylene protons with coinciding frequencies were divided by 3 or 2, respectively

**Table S2.** NMR structure determination statistics of **4** in solution.

	Compound <b>4</b>
NMR distance and dihedral restraints	
Total NOE restraints	56
Intra-residue	36
Inter-residue	20
Sequential ( $ i-j  = 1$ )	10
Nonsequential ( $ i-j  > 1$ )	10
Structure statistics *	
Violations (mean and s.d.)	
Number of distance constraint violations $> 0.1$ Å	0±0
Max. distance constraint violation [Å]	0.10±0.01
Deviations from idealized geometry	
Bond lengths [Å]	0.0125±0.0002
Bond angles [°]	1.54±0.03
Heavy atom RMSD to mean [Å]	0.79±0.30
Glycosidic linkage phi / psi angles **	
Fuca(1,1)CycMe	-68.9±0.8/-100.5±0.4 (crystal: -78.0/-101.4)
Galb(1,2)CycMe	-75.1±4.8/151.3±2.3 (crystal: -99.8/142.4)
Acid Orientation	82.0±6.4 (crystal: 120.5)
Core	-37.8±3.6 (crystal: -41.6)

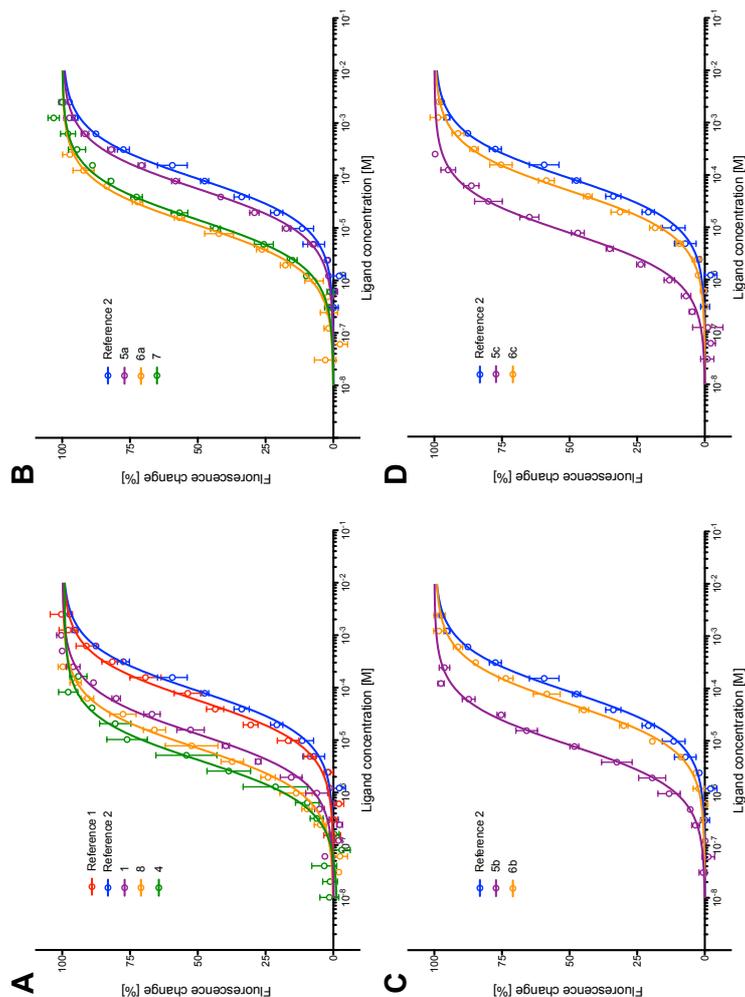
\* for an ensemble of 20 refined structures

\*\* phi is defined as  $O_5-C_1-O_x-C'_x$  and psi as  $C_1-O_x-C'_x-C'_{x-1}$ 

† extracted by XtalView

**Table S3.** Best-fit values, errors and normalized values of antagonists binding to E-selectin measured by MST. For each compound a global analysis was performed over a total of three independent measurements using GraphPad Prism software version 5.0. Batch variations were normalized according to a measurement with a reference substance (GCP69669A)<sup>(1)</sup> to an absolute  $K_D$  which was determined by isothermal titration calorimetry.<sup>(2,3)</sup>

	Reference 1	Compound 3	Compound 4
<b>Best-fit values</b>			
$Y_{off}$	-0.005346	-0.05184	-1.324
$Y_{max}$	99.83	99.92	99.10
$pK_D$	4.243	4.890	5.406
$K_D$ [ $\mu$ M]	57.11	12.88	3.928
<b>Std. Error</b>			
$pK_D$	0.03236	0.02848	0.04650
$K_D$ [ $\mu$ M]	53.04 to 61.57	12.06 to 13.76	3.528 to 4.370
<b>95% Conf. Int.</b>			
$pK_D$	4.178 to 4.309	4.832 to 4.948	5.312 to 5.500
$K_D$ [ $\mu$ M]	49.13 to 66.39	11.28 to 14.71	3.165 to 4.876
<b>Goodness of Fit</b>			
R square	0.9898	0.9923	0.9796
Sum of Squares	656.5	471.0	1369
<b>Constraints</b>			
Conc. Protein	2.500e-007	2.000e-007	2.000e-007
Points analyzed	44	41	45
<b>Normalize</b>			
$pK_D$	4.217	4.864	5.379
$K_D$ [ $\mu$ M]	<b>60.70</b>	<b>13.69</b>	<b>4.175</b>
Std. Error [ $\mu$ M]	56.34 to 65.40	12.82 to 14.62	3.751 to 4.647
95% Conf. [ $\mu$ M]	52.22 to 70.56	11.99 to 15.63	3.364 to 5.183



**Figure S6.** Binding curves of antagonists binding to E-selectin by MST. The error bars represent the standard deviation over a total of three independent measurements. The color of a curve indicates the substitution in 2'-position of D-Gal, where a hydroxyl group is orange, a benzoate is purple and an acetyl group is green. Furthermore, the cyclohexyl on the lactic acid moiety (A) was substituted by either a neutral (B), electron rich (C), or electron poor (D) phenyl ring. The binding curves in red and blue represent the standard reference measurements from two different protein batches.

**Table S4.** Best-fit values, errors and normalized values of antagonists binding to E-selectin measured by MST. For each compound a global analysis was performed over a total of three independent measurements using GraphPad Prism software version 5.0. Batch variations were normalized according to a measurement with a reference substance (GCP69669A)<sup>(1)</sup> to an absolute  $K_D$  which was determined by isothermal titration calorimetry.<sup>(2,3)</sup>

	Reference 2 (GCP69669A)							
	Compound 8	Compound 5a	Compound 6a	Compound 7	Compound 5b	Compound 6b	Compound 5c	Compound 6c
<b>Best-fit values</b>								
$Y_{off}$	-0.01076	0.002130	0.01265	0.04739	0.05827	-0.001900	-0.02106	-0.008091
$Y_{max}$	99.97	100.0	99.96	99.99	99.56	99.97	99.95	99.99
$pK_D$	4.062	4.255	4.953	4.847	4.300	5.086	4.303	5.112
$K_D$ [ $\mu$ M]	86.62	55.64	11.15	14.23	50.17	8.204	49.77	7.730
<b>Std. Error</b>								
$pK_D$	0.02640	0.01843	0.02643	0.02282	0.02005	0.02608	0.02040	0.02841
$K_D$ [ $\mu$ M]								
<b>95% Conf. Int.</b>								
$pK_D$	4.009 to 4.116	4.217 to 4.292	4.899 to 5.006	4.800 to 4.893	4.259 to 4.340	5.033 to 5.139	4.262 to 4.344	5.054 to 5.169
$K_D$ [ $\mu$ M]	76.60 to 97.95	51.06 to 60.63	9.861 to 12.61	12.80 to 15.83	45.69 to 55.09	7.266 to 9.264	45.26 to 54.73	6.772 to 8.824
<b>Goodness of Fit</b>								
R square	0.9934	0.9968	0.9935	0.9951	0.9964	0.9937	0.9961	0.9925
Sum of Squares	378.1	195.4	355.5	299.4	215.9	366.0	241.6	438.5
<b>Constraints</b>								
Conc. Protein	1.000e-007	2.000e-007						
Points analyzed	42	42	42	42	41	42	42	42
<b>Normalize</b>								
$pK_D$	4.217	4.409	5.107	5.001	4.454	5.240	4.457	5.266
$K_D$ [ $\mu$ M]	<b>60.70</b>	<b>38.99</b>	<b>7.81</b>	<b>9.97</b>	<b>35.16</b>	<b>5.75</b>	<b>34.9</b>	<b>5.417</b>
Std. Error [ $\mu$ M]	57.12 to 64.50	37.37 to 40.68	7.352 to 8.304	9.461 to 10.51	33.57 to 36.82	5.414 to 6.105	33.28 to 36.55	5.074 to 5.783
95% Conf. [ $\mu$ M]	53.68 to 68.64	35.78 to 42.49	6.910 to 8.837	8.970 to 11.09	32.02 to 38.61	5.092 to 6.492	31.72 to 38.35	4.746 to 6.184

## Experimental Section

### NMR spectroscopy for three-dimensional structure determination

Unless indicated otherwise spectra were measured at a 750 MHz Bruker Avance III spectrometer equipped with a TXI triple-resonance probe at 275 K. For complete chemical shift assignment of compounds **3** and **4** they were dissolved in D<sub>2</sub>O at concentrations of 6.6 mM **4** and 7.6 mM **3**, respectively. Assignment was achieved using the following 2D spectra: <sup>13</sup>C-<sup>1</sup>H HSQC, <sup>13</sup>C-<sup>1</sup>H HMBC, <sup>13</sup>C-<sup>1</sup>H HMQC-COSY<sup>(4)</sup> and TOCSY <sup>1</sup>H-<sup>1</sup>H TOCSY (mixing times 13 ms and 120 ms). The <sup>13</sup>C-<sup>1</sup>H HMQC-COSY spectrum was especially useful for assigning the cyclohexane resonances, see Supplementary Figure S2. Complete chemical shift assignments for compounds **3** and **4** can be found in Supplementary Figures S1 and S5. A 2D NOESY spectrum of compound **3** was recorded at 900 MHz with a mixing time of 150ms, 96 transients and 2,048×530 points. Spectra were processed using Topspin 2.1 and analyzed by Sparky<sup>(5)</sup>. All spectra were referenced to DSS according to Markley *et al.*<sup>(6)</sup>

### Structure Calculation and Refinement

Initial structures were calculated using CYANA 3.0<sup>(7)</sup>. Signal to noise (S/N) ratios of all NOE signals were extracted using the program Sparky<sup>(5)</sup> and converted to distances using the  $r^{-6}$  dependence and the Cyc H3a–H3b cross-peaks (1.77 Å) as reference. S/N ratios of signals involving CH<sub>2</sub> and CH<sub>3</sub> groups with coinciding <sup>1</sup>H resonances were divided by a factor of 2 or 3, respectively. Upper limit restraints with an additional tolerance of 0.5 Å were applied. Out of 200 structures, the 30 structures with the lowest target function were further refined in AMBER 9<sup>(8)</sup> applying the general AMBER force field (GAFF)<sup>(9)</sup>. A generalized Born model<sup>(10)</sup> was used to mimic solvent. Initial charge parameters were generated by Antechamber for the GAFF force field.

### Microscale Thermophoresis

MST experiments were carried out at 25° C with 100% LED power, 10-100% laser power (standard condition 50% laser power), a laser on time of 30 sec, and a laser off time of 5 sec using standard treated capillaries. Ligands were dissolved in assay buffer supplemented with 0.05% (v/v) Tween 20 at a concentration of 10 mM (reference substance, **5a-c** and **7**), 4 mM (**3**) 1 mM (**6a-c** and **8**) and 0.66 mM (**4**). Titration series were generated diluting the ligand stocks 15 times 1:1 with assay buffer. E-Selectin was purified and labeled according to Zihlmann *et al.* 2014.<sup>(2)</sup> The protein concentration was determined by HPLC-UV against a BSA standard.<sup>(11,12)</sup> The dilution series of ligand was 1:1 mixed with a solution of 0.1 to 0.25 μM E-Selectin and incubated for 10 mins at room temperature. All experiments were independently performed three times. Datapoints were normalized using the bound and unbound borders achieved by NanoTemper Analysis 1.2.205 software (NanoTemper Technologies GmbH, Munich, Germany) and analyzed/illustrated with GraphPad Prism 5.0 (GraphPad Software, La Jolla, CA, USA). The measurements were globally fitted using equation 1 for single site binding.<sup>(13)</sup>

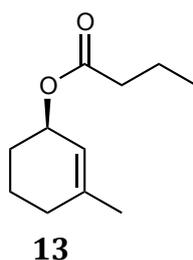
$$[PL] = \frac{(C_P + C_L + K_D - \sqrt{(C_P + C_L + K_D)^2 - 4C_P C_L})}{2C_P} \quad (\text{eq. 1})$$

where  $[PL]$  is the protein-ligand complex concentration and  $K_D$  is the dissociation constant.  $C_P$  represents the total concentration of protein and  $C_L$  the total concentration of ligand. Batch variations were normalized according to a measurement with a reference substance (GCP69669A)<sup>(1)</sup> to an absolute  $K_D$  which was determined by isothermal titration calorimetry.<sup>(2,3)</sup>

### Synthesis

General Methods: NMR spectra were recorded on a Bruker Avance DMX-500 (500 MHz) spectrometer. Assignment of  $^1\text{H}$  and  $^{13}\text{C}$  NMR spectra was achieved using 2D methods (COSY, HSQC). Chemical shifts are expressed in ppm using residual TMS,  $\text{CHCl}_3$ ,  $\text{CHD}_2\text{OD}$  and HDO as references. For complex molecules, the following prefixes for substructures are used: Cy (cyclohexyl), Fuc (fucose), Gal (galactose) and Lac (lactate). Optical rotations were measured on a Perkin-Elmer Polarimeter 341. Electron spray ionization mass spectra (ESI-MS) were obtained on a Waters micromass ZQ. HRMS analysis was carried out using a Bruker Daltonics micrOTOF spectrometer equipped with a TOF hexapole detector. Reactions were monitored by TLC using glass plates coated with silica gel 60 F254 (Merck) and visualized by using UV light and/or by heating to  $140^\circ\text{C}$  for 5 min with aq.  $\text{KMnO}_4$  solution or a molybdate solution (a 0.02 M solution of ammonium cerium sulfate dihydrate and ammonium molybdate tetrahydrate in aq. 10%  $\text{H}_2\text{SO}_4$ ). Column chromatography was performed on a CombiFlash Companion (Teledyne-ISCO, Inc.) using RediSep<sup>®</sup> normal phase disposable flash columns (silica gel). Reversed phase chromatography was performed on LiChroprep<sup>®</sup> RP-18 (Merck, 40-63  $\mu\text{m}$ ). Hydrogenation reactions were performed in a shaking apparatus (Parr Instruments Company, Moline, Illinois, USA) in 250 mL or 500 mL bottles. Methanol (MeOH), dichloromethane (DCM), dimethoxyethane (DME) were dried by filtration through  $\text{Al}_2\text{O}_3$  (Fluka, type 5016 A basic) and stored on activated molecular sieves. Dioxane, were dried by distillation from sodium/benzophenone. DMF was dried by distillation from calcium hydride.

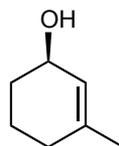
### (*R*)-3-Methylcyclohex-2-en-1-yl butyrate (**13**)



Immobilized Novozyme 435 (222 mg, 444 U, EC 232-619-9) was added to a solution of **12** (10.0 g, 89 mmol) and vinyl butyrate (22.6 mL, 20.3 g, 178 mmol) in heptane (90 mL). The mixture was stirred at  $23^\circ\text{C}$  and 200 rpm. After 2 h 25 min the mixture was filtered and volatiles were evaporated at  $60^\circ\text{C}$  and 10 mbar to give 12 g of a clear oil. Column chromatography on silica ( $\text{CH}_2\text{Cl}_2$ ) yielded pure **13** (7.50 g, 41 mmol, 46%).  $[\alpha]_D^{22} +168.7$  (*c* 9.28,  $\text{CHCl}_3$ );  $^1\text{H}$  NMR (500.1 MHz,  $\text{CDCl}_3$ ):  $\delta = 5.44$  (m, 1H, H-2), 5.23 (m, 1H, H-1), 2.24 (t,  $J = 7.4$  Hz, 2H,  $\text{COCH}_2\text{CH}_2\text{CH}_3$ ), 2.02-1.84 (m, 2H, H-4, H-4'), 1.81-1.56 (m, 9H, H-5, H-5', H-6, H-6',  $-\text{CH}_3$ ,  $\text{COCH}_2\text{CH}_2\text{CH}_3$ ), 0.92 (t,  $J = 7.4$  Hz, 3H,

COCH<sub>2</sub>CH<sub>2</sub>CH<sub>3</sub>); <sup>13</sup>C NMR (125.8 MHz, CDCl<sub>3</sub>): δ 173.5 (COCH<sub>2</sub>CH<sub>2</sub>CH<sub>3</sub>), 141.0 (C-3), 120.2 (C-2), 68.6 (C-1), 36.7 (COCH<sub>2</sub>CH<sub>2</sub>CH<sub>3</sub>), 30.0 (C-4), 28.1 (C-6), 23.8 (-CH<sub>3</sub>), 19.1, 18.7 (2C, C-5, COCH<sub>2</sub>CH<sub>2</sub>CH<sub>3</sub>), 13.7 (COCH<sub>2</sub>CH<sub>2</sub>CH<sub>3</sub>); MS (ESI) *m/z*: Calcd for C<sub>11</sub>H<sub>18</sub>NaO<sub>2</sub><sup>+</sup> [M+Na]<sup>+</sup>: 205.12; found: 204.83; elemental analysis calcd (%) for C<sub>11</sub>H<sub>18</sub>O<sub>2</sub> (182.26): C 72.49, H 9.95; found: C 72.87, H 9.65.

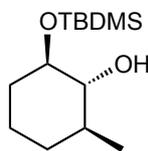
**(R)-seudenol (14)**<sup>(14)</sup>



**14**

A solution of NaOH in H<sub>2</sub>O (10.3 mL, 4N) was slowly added to a solution of seudenol butyrate **13** (3.50 g, 19 mmol) in MeOH (30 mL) at 0°C and stirred at 0°C for 5 h. The mixture was diluted with H<sub>2</sub>O (25 mL) and extracted with CH<sub>2</sub>Cl<sub>2</sub> (3 x 20 mL). The combined organic layers were washed with brine (25 mL) and dried over Na<sub>2</sub>SO<sub>4</sub>. Filtration and evaporation of volatiles (200 mbar, 40°C) gave spectroscopically pure (*R*)-seudenol (**14**) (1.81 g, 16.0 mmol, 84%) as a clear oil, which was directly used in the next step. [α]<sub>D</sub> +91.7 (*c* 0.74, CHCl<sub>3</sub>); HPLC: 97.5% ee, 98% purity; NMR data were in accordance with literature.<sup>(14)</sup>

**(1R,2R,3S)-1-[(*tert*-butyldimethylsilyl)oxy]-3-methylcyclohexan-2-ol (16)**

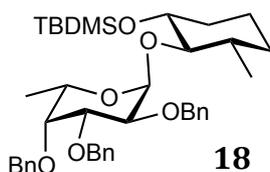


**16**

Imidazol (4.40 g, 65 mmol) was added to a solution of (*R*)-seudenol **14** (3.50 g, 31 mmol), DMAP (cat.), and TBSCl (7.31 g, 48 mmol) in anhydrous CH<sub>2</sub>Cl<sub>2</sub> (65 mL) at r.t. under argon. After stirring for 15 h, the reaction mixture was quenched with satd. aq. NaHCO<sub>3</sub> (50 mL) and extracted with CH<sub>2</sub>Cl<sub>2</sub> (20 mL). The organic layer was washed with aq. HCl (20 mL, 0.01 N), satd. aq. NaHCO<sub>3</sub> (20 mL), and brine (20 mL) and dried over Na<sub>2</sub>SO<sub>4</sub>. Filtration and evaporation of volatiles (200 mbar, 40°C) gave the TBS ether **15** as clear oil. A solution of BH<sub>3</sub>·THF (60 mL, 1M in THF) was slowly added to a solution of the crude TBS ether (**15**) in anhydrous THF (60 mL) under argon at 0°C. After stirring for 2 h at rt, the reaction mixture was cooled to 0°C again and aq. NaOH (180 mL, 3N) followed by aq. H<sub>2</sub>O<sub>2</sub> (180 mL, 30%) were slowly added *via* dropping funnel (**CAUTION**: strong gas development). The mixture was stirred at 0°C for 1 h, subsequently acidified to pH 3 by slow addition of 10% aq. HCl *via* dropping funnel (**CAUTION**: strong gas development) and extracted with DCM (2 x 300 mL). The extracts were dried over Na<sub>2</sub>SO<sub>4</sub>, filtered, concentrated (100 mbar, 40°C) and purified by column chromatography (petroleum ether/Et<sub>2</sub>O 98.5/1.5) to yield pure **16** (6.20 g, 25

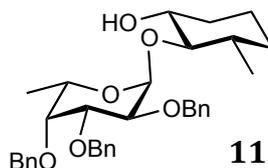
mmol, 81%) as clear oil.  $[\alpha]_D^{22} - 13.7$  ( $c$  3.14,  $\text{CHCl}_3$ );  $^1\text{H NMR}$  (500.1 MHz,  $\text{CDCl}_3$ ):  $\delta = 3.34$  (m, 1H, H-1), 2.92 (dd,  $J = 8.5, 10.0$  Hz, 1H, H-2), 2.47 (s, 1H, OH), 1.81 (m, 1H, H-6), 1.63-1.56 (m, 2H, H-4, H-5), 1.41 (m, 1H, H-3), 1.34-1.99 (m, 2H, H-5', H-6'), 1.04-0.92 (m, 4H, H-4', Me), 0.91-0.83 (m, 9H,  $\text{SiC}(\text{CH}_3)_3$ ), 0.07 (s, 3H,  $\text{SiCH}_3$ ), 0.05 (s, 3H,  $\text{SiCH}_3$ );  $^{13}\text{C NMR}$  (125.8 MHz,  $\text{CDCl}_3$ ):  $\delta = 81.0$  (C-2), 77.0 (C-1), 37.0 (C-3), 33.4, 33.9 (2C, C-4, C-6), 25.9 (3C,  $\text{SiC}(\text{CH}_3)_3$ ), 23.6 (C-5), 18.5 ( $-\text{CH}_3$ ), 18.1 ( $\text{SiC}(\text{CH}_3)_3$ ), -3.9, -4.6 ( $\text{SiCH}_3$ ); HR-MS (ESI)  $m/z$ : calcd for  $\text{C}_{13}\text{H}_{28}\text{NaO}_2\text{Si}^+$   $[\text{M}+\text{Na}]^+$ : 267.1751; found: 267.1752.

**[(1R,2R,3S)-1-(*tert*-Butyldimethylsilyloxy)-3-methyl-cyclohex-2-yl] 2,3,4-tri-O-benzyl-6-deoxy- $\alpha$ -L-galactopyranoside (**18**)**

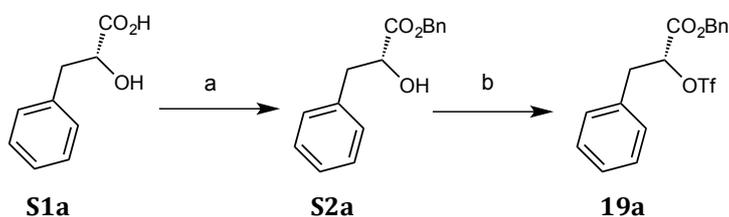


Ethylthio fucoside **17**<sup>(15)</sup> (3.90 g, 8.15 mmol) and TBAB (4.00 g, 12.4 mmol) were dried at high vacuum overnight. Powdered activated molecular sieves 4 Å (5.0 g), compound **16** (1.00 g, 4.09 mmol), 2,6-di-*tert*-butyl-4-methylpyridine (2.50 g, 12.2 mmol), anhydrous DCM (35 ml) and DMF (5 ml) were added and the mixture was stirred for 4 h at rt under argon.  $\text{CuBr}_2$  (2.70 g, 12.1 mmol), dried under high vacuum overnight at 70°C, was added and the resulting dark mixture was stirred at rt under argon. After completion of the reaction (17 h), the solution was filtered through a pad of celite and the filtrate was washed with a solution of satd. aq.  $\text{NH}_4\text{Cl}$  and aqueous  $\text{NH}_3$  (9/1 (v/v), 2 x 200 mL) and brine (100 mL). The aqueous layers were extracted with DCM (2 x 200 mL) and the combined organic layers were dried ( $\text{Na}_2\text{SO}_4$ ) and concentrated. Column chromatography on silica (petroleum ether/EtOAc 98/2 to 97/3) gave the pseudodisaccharide **18** as clear oil (2.34 g, 3.54 mmol, 87%).  $[\alpha]_D - 53.7$  ( $c$  2.1,  $\text{CHCl}_3$ );  $^1\text{H NMR}$  (500.1 MHz,  $\text{CDCl}_3$ ):  $\delta = 7.47$ -7.27 (m, 15H, 3  $\text{C}_6\text{H}_5$ ), 5.16 (d,  $^3J = 3.4$  Hz, 1H, Fuc H-1), 5.03 (A of AB,  $^2J = 11.6$  Hz, 1H,  $\text{CH}_2\text{Ph}$ ), 4.89, 4.85, 4.78, 4.76 (4d,  $^2J = 11.8$  Hz, 4H,  $\text{CH}_2\text{Ph}$ ), 4.70 (B of AB,  $^2J = 11.6$  Hz, 1H,  $\text{CH}_2\text{Ph}$ ), 4.26 (q,  $^3J = 6.4$  Hz, 1H, Fuc H-5), 4.10 (dd,  $^3J = 3.4, 10.2$  Hz, 1H, Fuc H-2), 4.05 (dd,  $^3J = 2.6, 10.2$  Hz, 1H, Fuc H-3), 3.75 (m, 1H, H-1), 3.70 (m, 1H, Fuc H-4), 3.36 (t,  $^3J = 6.4$  Hz, H-2), 1.88-1.77 (m, 2H, H-3, H-6a), 1.76-1.68 (m, 2H, H-4a, H-5a), 1.43 (m, 1H, H-6b), 1.34-1.11 (m, 8H, Fuc-H6,  $-\text{CH}_3$ , H-4b, H-5b), 0.93 (s, 9H,  $\text{SiC}(\text{CH}_3)_3$ ), 0.09 (s, 3H,  $\text{SiCH}_3$ ), 0.06 (s, 3H,  $\text{SiCH}_3$ );  $^{13}\text{C NMR}$  (125.8 MHz,  $\text{CDCl}_3$ ):  $\delta$  139.1, 138.9, 138.8, 128.4, 128.3, 128.2, 127.6, 127.5 (18C, 3  $\text{C}_6\text{H}_5$ ), 96.8 (Fuc C-1), 81.5 (C-2), 79.3 (Fuc C-3), 78.2 (Fuc C-4), 76.7 (Fuc C-2), 74.9 ( $\text{CH}_2\text{Ph}$ ), 73.6 ( $\text{CH}_2\text{Ph}$ ), 73.3 (C-1), 73.0 ( $\text{CH}_2\text{Ph}$ ), 66.4 (Fuc C-5), 35.6 (C-3), 33.2 (C-6), 31.1 (C-5), 26.1 (3C,  $\text{SiC}(\text{CH}_3)_3$ ), 19.8 (C-4), 18.9 (Fuc C-6), 18.2 ( $\text{SiC}(\text{CH}_3)_3$ ), 17.0 ( $\text{CH}_3$ ), -3.9, -5.0 (2C,  $\text{SiCH}_3$ ); HR-MS (ESI)  $m/z$ : Calcd for  $\text{C}_{40}\text{H}_{56}\text{NaO}_6\text{Si}^+$   $[\text{M}+\text{Na}]^+$ : 683.3738; found: 683.3740.

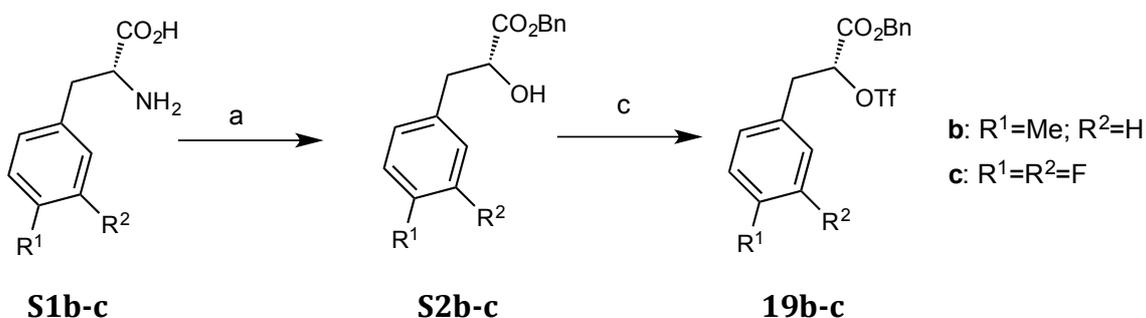
[(1R,2R,3S)-1-Hydroxy-3-methyl-cyclohex-2-yl] 2,3,4-tri-O-benzyl-6-deoxy- $\alpha$ -L-galactopyranoside (**11**)<sup>(16)</sup>



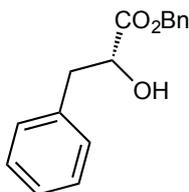
Compound **18** (2.10 g, 3.18 mmol) was dissolved in a solution of TBAF in THF (20 mL, 1M) and stirred for 24 h at rt. The solution was diluted with DCM (50 mL) and washed with H<sub>2</sub>O (100 mL). The aq. layer was extracted with DCM (2 x 50 mL) and the combined organic layers were dried (Na<sub>2</sub>SO<sub>4</sub>) and concentrated. Column chromatography on silica (petroleum ether/EtOAc 80/20) gave **11** as white solid (1.74 g, 3.18 mmol, quant.); [ $\alpha$ ]<sub>D</sub> - 42.0 (*c* 0.45, CHCl<sub>3</sub>); NMR data were in accordance with literature.<sup>(16)</sup>



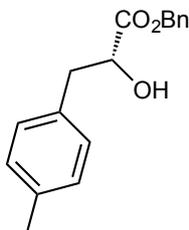
**Scheme S1:** Synthesis of triflate derivative **19a** a) BnOH, TsOH, benzene (88%); b) Tf<sub>2</sub>O, lutidine, DCM (98%).



**Scheme S2.** Synthesis of triflate derivatives **19b-c** a) (i) NaNO<sub>2</sub>, H<sub>2</sub>SO<sub>4</sub>, H<sub>2</sub>O (ii) BnOH, TsOH, benzene (**b**: 57%, **c**: 50%); b) Tf<sub>2</sub>O, lutidine, DCM (**b**: 97%, **c**: 95%).

**(R)-Benzyl 2-hydroxy-3-phenylpropanoate (S2a)<sup>(17)</sup>****S2a**

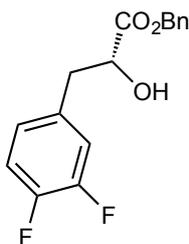
To a solution of 2-hydroxy-3-phenyl-2-(*R*)-propanoic acid (**S1a**) (1.94 g, 11.6 mmol) in benzene (25 mL) benzyl alcohol (1.62 mL, 15.6 mmol) and *p*-toluenesulfonic acid (200 mg, 1.2 mmol) were added. The solution was refluxed at 80°C until completion of the reaction (3 h, TLC: DCM/MeOH, 19:1 and petroleum ether/EtOAc, 4:1). The reaction mixture was allowed to cool down to rt, diluted with EtOAc (50 mL) and washed with satd. aq. NaHCO<sub>3</sub> (20 mL), water (20 mL) and brine (10 mL). The organic layer was dried over Na<sub>2</sub>SO<sub>4</sub>, filtered and solvent was removed *in vacuo*. The crude product was purified by flash chromatography (petroleum ether/EtOAc, gradient 0-25%) to afford (**S2a**)<sup>(17)</sup> as colorless wax (2.62 g, 10.2 mmol, 88%). *R*<sub>f</sub> = (petroleum ether/EtOAc, 4:1) 0.50; [α]<sub>D</sub><sup>22</sup> +53.5 (*c* 0.69, CHCl<sub>3</sub>); <sup>1</sup>H NMR (500.1 MHz, CDCl<sub>3</sub>): δ = 7.54 - 7.06 (m, 10H, Ar-H), 5.18 (s, 2H, Ar-CH<sub>2</sub>), 4.50 (br s, 1H, H-2), 3.22 - 3.06 (m, 1H, H-3), 3.09 - 2.90 (m, 1H, H-3'); <sup>13</sup>C NMR (125.8 MHz, CDCl<sub>3</sub>): δ = 174.0 (C-1), 136.1, 135.0, 129.5, 128.7, 128.7, 128.6, 128.4 (8C, Ar-C); 71.3 (C-2); 67.5 (Ar-CH<sub>2</sub>); 40.5 (C-3); ESI-MS: *m/z*: Calcd for C<sub>16</sub>H<sub>16</sub>NaO<sub>3</sub> [M+Na]<sup>+</sup>: 279.3, found: 278.9.

**(R)-Benzyl 2-hydroxy-3-(4-methylphenyl)-propanoate (S2b)****S2b**

A solution of sodium nitrite (11.7 g, 0.17 mol) in water (40 mL) was added dropwise to a stirred and ice-cooled solution of *p*-methyl-D-phenylalanine (**S1b**) (5.10 g, 28.4 mmol) in 1 M H<sub>2</sub>SO<sub>4</sub> (50 mL) over 1 h, and the resulting mixture was stirred for 24 h at rt. The reaction mixture was extracted with Et<sub>2</sub>O (3 x 150 mL), the combined organic phases were washed with brine (30 mL), dried over Na<sub>2</sub>SO<sub>4</sub>, filtered and concentrated *in vacuo*. The crude product (4.85 g) was dissolved in benzene (30 mL) and the resulting solution was treated with benzyl alcohol (3.63 mL, 35.0 mmol, 1.3 eq) and *p*-toluenesulfonic acid (463 mg, 2.70 mmol, 0.1 eq). The solution was refluxed until completion of the reaction (5 h, TLC: DCM/MeOH, 19:1 and petroleum ether/EtOAc, 4:1). The reaction mixture was allowed to cool down to rt, diluted with EtOAc (150 mL) and washed with satd. aq. NaHCO<sub>3</sub> (30 mL), water (30 mL) and brine (20 mL). The organic layer was dried over Na<sub>2</sub>SO<sub>4</sub>, filtered and concentrated *in vacuo*. The crude product was purified by flash

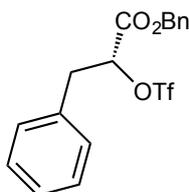
chromatography (petroleum ether/EtOAc, gradient 0-25%) to afford (**S2b**) as colorless wax (4.2 g, 15.16 mmol, 53% over two steps).  $R_f$  = (petroleum ether/EtOAc, 4:1) 0.52;  $[\alpha]_D^{22}$  -46.5 ( $c$  0.63,  $\text{CHCl}_3$ );  $^1\text{H NMR}$  (500.1 MHz,  $\text{CDCl}_3$ ):  $\delta$  = 7.40 - 7.30 (m, 5H, Ar-H), 7.05 (dd,  $J$  = 8.1 Hz, 4H, Ar-H); 5.15 - 5.02 (m, 2H, Ar- $\text{CH}_2$ ), 4.47 (dd,  $J$  = 4.7, 6.4 Hz, 1H, H-2), 3.09 (dd,  $J$  = 4.7, 13.9 Hz, 1H, H-3), 2.95 (dd,  $J$  = 6.4, 13.9 Hz, 1H, H-3'), 2.31 (s, 3H, Ar-Me);  $^{13}\text{C NMR}$  (125.8 MHz,  $\text{CDCl}_3$ ):  $\delta$  = 174.2 (C-1), 136.6, 135.2, 133.0, 129.5, 129.3, 128.8, 128.7 (8C, Ar-C), 71.5 (C-2), 67.5 (Ar- $\text{CH}_2$ ), 40.2 (C-3), 21.2 (Ar-Me); ESI-MS:  $m/z$ : Calcd for  $\text{C}_{17}\text{H}_{18}\text{NaO}_3$   $[\text{M}+\text{Na}]^+$ : 293.3, found: 292.9.

**(R)-Benzyl 2-hydroxy-3-(3,4-difluorophenyl)-propanoate (S2c)**

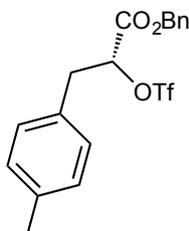


**S2c**

A solution of sodium nitrite (2.05 g, 39.8 mmol) in water (10 mL) was added dropwise to a stirred and ice-cooled solution of 3,4-difluoro-D-phenylalanine (**S1c**) (1.00 g, 4.97 mmol) in 1 M  $\text{H}_2\text{SO}_4$  (10 mL) over 30 min, and the resulting mixture was stirred for 2 d at rt. The reaction mixture was extracted with  $\text{Et}_2\text{O}$  (3 x 50 mL), the combined organic phases were washed with brine (20 mL), dried over  $\text{Na}_2\text{SO}_4$ , filtered and concentrated *in vacuo*. The crude product (790 mg) was dissolved in benzene (10 mL) and the resulting solution was treated with benzyl alcohol (0.55 mL, 5.08 mmol) and *p*-toluenesulfonic acid (62 mg, 0.39 mmol). The solution was refluxed at 80°C until completion of the reaction (4 h, TLC: DCM/MeOH, 19:1 and petroleum ether/EtOAc, 4:1). The reaction mixture was allowed to cool down to rt, diluted with EtOAc (50 mL) and washed with satd. aq.  $\text{NaHCO}_3$  solution (20 mL), water (20 mL) and brine (10 mL). The organic layer was dried over  $\text{Na}_2\text{SO}_4$ , filtered and concentrated *in vacuo*. The crude product was purified by flash chromatography (Tol/EtOAc, gradient 0-10%) to afford (**S2c**) as colorless wax (720 mg, 2.46 mmol, 50% over two steps).  $R_f$  = (petroleum ether/EtOAc, 4:1) 0.38;  $[\alpha]_D^{22}$  +61.9 ( $c$  0.76,  $\text{CHCl}_3$ );  $^1\text{H NMR}$  (500.1 MHz,  $\text{CDCl}_3$ ):  $\delta$  = 7.43 - 7.21 (m, 6H, Ar-H), 7.03 - 6.91 (m, 2H, Ar-H), 6.87 - 6.77 (m, 1H, Ar-H), 5.26 - 5.11 (m, 2H, Ar- $\text{CH}_2$ ), 4.45 (q,  $J$  = 5.3 Hz, 1H, H-2), 3.07 (dd,  $J$  = 4.0, 14.1 Hz, 1H, H-3), 2.91 (dd,  $J$  = 6.2, 14.1 Hz, 1H, H-3');  $^{13}\text{C NMR}$  (125.8 MHz,  $\text{CDCl}_3$ ):  $\delta$  = 173.7 (C-1), 151.3 - 148.2 (m, 2C, ArF-C), 134.8 (Ar-C) 133.1 (m, Ar-C), 128.9, 128.8 (2C, Ar-C), 125.5 (m, Ar-C), 118.4 (d,  $J$  = 17.1 Hz, Ar-C), 116.9 (d,  $J$  = 17.1 Hz, Ar-C), 70.8 (C-2); 67.7 (Ar- $\text{CH}_2$ ); 39.3 (C-3). ESI-MS:  $m/z$ : Calcd for  $\text{C}_{16}\text{H}_{14}\text{F}_2\text{NaO}_3$   $[\text{M}+\text{Na}]^+$ : 315.3, found: 314.6.

**(R)-Benzyl 3-phenyl-2-(trifluoromethyl)sulfonyloxy-propanoate (19a)****19a**

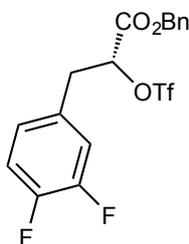
To a stirred solution of (**S2a**) (2.44 g, 8.35 mmol) and lutidine (1.16 mL, 10.8 mmol) in dry DCM (30 mL) Tf<sub>2</sub>O (1.68 mL, 10.0 mmol) was added dropwise at -60°C under argon. The reaction was stirred at -60°C for 20 min, then at -5°C for 40 min. TLC (petroleum ether/EtOAc, 9:1) indicated completion of the reaction and the mixture was quenched by addition of water (10 mL) at 0°C. The organic phase was separated and the water phase was extracted with DCM (50 mL). The combined organic phases were washed with water (20 mL), dried over Na<sub>2</sub>SO<sub>4</sub>, filtered and concentrated *in vacuo*. The crude product was purified by flash chromatography using a short silica pad (DCM/petroleum ether, 1:1) to afford (**19a**) as pale orange oil. (3.17 g, 8.17 mmol, 98%). R<sub>f</sub> = (petroleum ether/EtOAc, 9:1) 0.50; [α]<sub>D</sub><sup>22</sup> +29.5 (*c* 1.33, CHCl<sub>3</sub>); <sup>1</sup>H NMR (500.1 MHz, CDCl<sub>3</sub>): δ = 7.41 - 7.12 (m, 10H, Ar-H), 5.30 - 5.20 (m, 1H, H-2), 5.24 (s, 2H, Ar-CH<sub>2</sub>), 3.34 (dd, *J* = 3.6, 14.6 Hz, 1H, H-3), 3.21 (dd, *J* = 8.5, 14.6 Hz, 1H, H-3'); <sup>13</sup>C NMR (126 MHz, CDCl<sub>3</sub>): δ = 166.5 (C-1), 134.2, 133.2, 129.5, 128.9, 128.8, 128.7, 128.9, 127.8 (8C, Ar-C), 83.8 (C-2), 68.4 (Ar-CH<sub>2</sub>), 38.2 (C-3).

**(R)-Benzyl 3-(4-methylphenyl)-2-(trifluoromethyl)sulfonyloxy-propanoate (19b)****19b**

To a stirred solution of (**S2b**) (4.10 g, 15.2 mmol) and lutidine (2.29 mL, 19.7 mmol) in dry DCM (30 mL) a solution of Tf<sub>2</sub>O in dry DCM (1 M, 18.2 mL, 18.2 mmol) was added at -80°C under argon. The reaction was allowed to warm up to -20°C (2 h) and stirred at that temperature for an additional 1 h. TLC (petroleum ether/EtOAc, 9:1) indicated completion of the reaction and the mixture was quenched by addition of water (10 mL) at 0°C. The organic phase was separated and the water phase was extracted with DCM (2 x 50 mL). The combined organic phases were washed with water (20 mL), dried over Na<sub>2</sub>SO<sub>4</sub>, filtered and concentrated *in vacuo*. The crude product was purified by flash chromatography using a short silica pad (DCM/petroleum ether, 1:1) to afford (**19b**) as pale grey oil (5.90 g, 14.7 mmol, 97%). R<sub>f</sub> = (petroleum ether/EtOAc, 9:1) 0.45; [α]<sub>D</sub><sup>22</sup> -26.1 (*c* 1.22, CHCl<sub>3</sub>); <sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>): δ = 7.39 - 7.27 (m, 5H, Ar-H), 7.09 (d, *J* = 7.9 Hz, 2H, Ar-H), 7.04 (d, *J* = 7.9 Hz, 2H, Ar-H), 5.26 (dd, *J* = 4.5, 8.1 Hz, 1H,

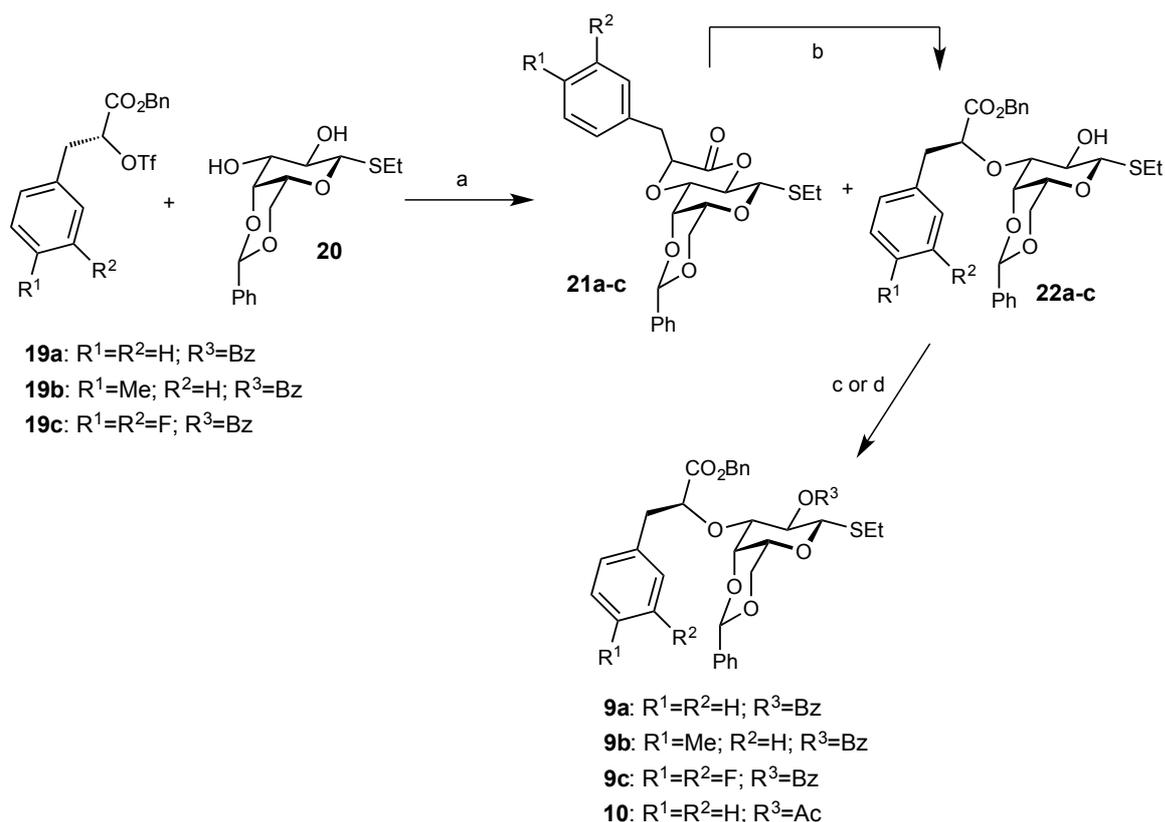
H-2), 5.23 (s, 2H, Ar-CH<sub>2</sub>), 3.29 (dd, *J* = 4.5, 14.6 Hz, 1H, H-3), 3.18 (dd, *J* = 8.1, 14.6 Hz, 1H, H-3'), 2.32 (s, 3H, Ar-Me); <sup>13</sup>C NMR (126 MHz, CDCl<sub>3</sub>): δ = 166.6 (C-1), 137.7, 134.4, 130.1, 129.6, 129.5, 129.0, 128.8, 128.8 (8C, Ar-C), 84.0 (C-2), 68.5 (Ar-CH<sub>2</sub>), 38.0 (C-3), 21.2 (Ar-Me).

**(*R*)-Benzyl 3-(3,4-difluorophenyl)-2-(trifluoromethyl)sulfonyloxy-propanoate (19c)**



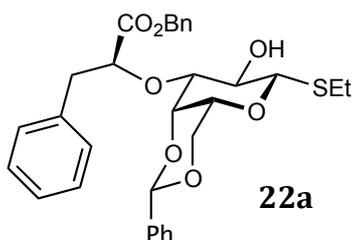
**19c**

To a stirred solution of (**S2c**) (670 mg, 2.29 mmol) and lutidine (0.35 mL, 2.97 mmol) in dry DCM (10 mL) a solution of Tf<sub>2</sub>O in dry DCM (1 M, 2.75 mL, 2.75 mmol) was added at -80°C under argon. The reaction was allowed to warm up to -20°C (1 h) and stirred at this temperature for an additional 1 h. TLC (petroleum ether/EtOAc, 9:1) indicated completion of the reaction and the mixture was quenched by addition of water (10 mL) at 0°C. The organic phase was separated and the water phase was extracted with DCM (2 x 30 mL). The combined organic phases were washed with water (10 mL), dried over Na<sub>2</sub>SO<sub>4</sub>, filtered and concentrated *in vacuo*. The crude product was purified by flash chromatography using a short silica pad (DCM/petroleum ether, 1:1) to afford (**19c**) as colorless oil (900 mg, 2.12 mmol, 93%). *R*<sub>f</sub> = (petroleum ether/EtOAc, 9:1) 0.40; [α]<sub>D</sub><sup>22</sup> +31.1 (*c* 0.62, CHCl<sub>3</sub>); <sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>): δ = 7.41 - 7.22 (m, 5H, Ar-H), 7.09 - 7.00 (m, 1H, Ar-H), 7.00 - 6.92 (m, 1H, Ar-H), 6.89 - 6.83 (m, 1H, Ar-H), 5.32 - 5.14 (m, 3H, H-2, Ar-CH<sub>2</sub>), 3.28 (dd, *J* = 3.6, 14.8 Hz, 1H, H-3), 3.18 (dd, *J* = 7.9, 14.8 Hz, 1H, H-3'); <sup>13</sup>C NMR (126 MHz, CDCl<sub>3</sub>): δ = 166.0 (C-1); 151.5 - 148.8 (m, 2C, ArF-C), 134.1 (Ar-C) 130.0 (m, Ar-C), 129.1, 128.8 (2C, Ar-C), 125.7 (m, Ar-C), 118.6 (d, *J* = 17.6 Hz, Ar-C), 117.6 (d, *J* = 17.4 Hz, Ar-C), 82.9 (C-2), 68.6 (Ar-CH<sub>2</sub>), 37.3 (C-3).



**Scheme S3.** Synthesis of galatoside derivatives **9a-c**: a) (i) Bu<sub>2</sub>SnO, MeOH; (ii) CsF, DME, (lactone - **21a**: 16%, **21b**: 25%, **21c**: 25%; benzylester - **22a**: 25%, **22b**: 0%, **22c**: 5%); b) BnOH, DMAP (cat.) (**21a**: 73%, **21b**: 80%, **21c**: 65%); c) BzCl, py, DMAP (cat.) (**9a**: 87%, **9b**: 72%, **9c**: 64%,); d) Ac<sub>2</sub>O, pyridine, **10**: 78%.

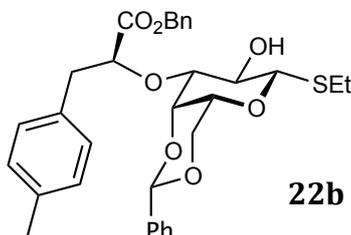
### Ethyl 4,6-*O*-benzylidene-3-*O*-[(*S*)-1-(benzyloxy)-1-oxo-3-phenylpropan-2-yl]-1-thio-β-*D*-galactopyranoside (**22a**)



Thiogalactoside (**20**) (482 mg, 1.544 mmol, 1 eq) and *n*-Bu<sub>2</sub>SnO (575 mg, 2.31 mmol) were dissolved in dry MeOH (15 mL) and refluxed at 65°C under argon for 4 h. The solvent was removed under reduced pressure and the crude product was dried in high vacuum for 16 h. The residue was dissolved in dry DME (10 mL) under argon and a solution of triflate (**19a**) (2 g, 5.14 mmol) in dry DME (10 mL) was added followed by the addition of CsF (700 mg, 4.62 mmol, dried in high vacuum at 100°C for 16 h). The turbid solution was stirred for 5 h after which TLC (petroleum ether/EtOAc, 9:1, DCM/MeOH, 19:1) indicated still presence of starting material (**20**) but no triflate (**19a**) so another portion of (**19a**) (1 g, 2.57 mmol) was added and the mixture was stirred for additional 16 h. A solution of KF (10%) in aq. KH<sub>2</sub>PO<sub>4</sub> (1 N, 10 mL) was added, and after stirring for 1 h, the reaction was extracted with DCM (3 x 70 mL). The combined organic phases were washed with brine (20 mL) dried over Na<sub>2</sub>SO<sub>4</sub> and the solvents were

removed *in vacuo*. The crude product was purified by flash chromatography (petroleum ether/EtOAc, gradient 0-50%) to afford a 1:1.6 mixture of **21a** (lactone) and **22a** (open) (330 mg, 41%), which was used in the following reaction without further purification. ESI-MS: *m/z*: Calcd for (**21a**) (lactone) C<sub>24</sub>H<sub>26</sub>NaO<sub>6</sub>S [M+Na]<sup>+</sup>: 465.5, found: 465.2; Calcd for (**22a**) (open) C<sub>31</sub>H<sub>34</sub>NaO<sub>7</sub>S [M+Na]<sup>+</sup>: 573.7, found: 573.1. To a solution of the mixture (330 mg) obtained in the previous reaction, benzyl alcohol (2mL) and DMAP (cat.) were added. The reaction was heated up to 60°C for 4 h, then the reaction mixture was purified by flash chromatography (petroleum ether/EtOAc, gradient 0-50%) to afford pure (**22a**) as white foam (260 mg, 0.47 mmol, 30% starting from **20**). R<sub>f</sub> = (petroleum ether/EtOAc, 7:3) 0.33; [α]<sub>D</sub><sup>22</sup> -44.9 (*c* 0.31, CHCl<sub>3</sub>); <sup>1</sup>H NMR (500.1 MHz, CDCl<sub>3</sub>): δ = 7.52 - 7.47 (m, 2H, Ar-H), 7.41 - 7.31 (m, 5H, Ar-H), 7.28 - 7.20 (m, 8H, Ar-H), 5.34 (s, 1H, Ph-CH), 5.10 (d, *J* = 12.1 Hz, 1H, Ar-CH<sub>2</sub>), 5.04 (d, *J* = 12.1 Hz, 1H, Ar-CH<sub>2</sub>), 4.62 (dd, *J* = 4.8, 8.4 Hz, 1H, Lac-H2), 4.26 - 4.16 (m, 3H, Gal-H1, Gal-H4, Gal-H6), 3.88 - 3.82 (m, 2H, Gal-H2, Gal-H6'), 3.37 (dd, *J* = 3.4, 9.2 Hz, 1H, Gal-H3), 3.29 (br s, 1H, Gal-H5), 3.14 - 3.03 (m, 2H, Lac-H3), 2.80 - 2.62 (m, 2H, SCH<sub>2</sub>CH<sub>3</sub>), 2.09 (d, *J* = 1.9 Hz, 1H, OH), 1.28 (t, *J* = 7.5 Hz, 3H, SCH<sub>2</sub>CH<sub>3</sub>); <sup>13</sup>C NMR (125.8 MHz, CDCl<sub>3</sub>): δ = 172.3 (Lac-C1), 138.0, 137.0, 135.4, 129.6, 128.9, 128.6, 128.5, 128.5, 128.5, 128.3, 128.1, 126.9, 126.4 (12C, Ar-C), 100.9 (Ph-CH), 85.2 (Gal-C1), 81.8 (Gal-C3), 80.4 (Lac-C2), 75.2 (Gal-C4), 70.2 (Gal-C5), 69.2 (Gal-C6), 68.7 (Gal-C2), 66.8 (Ph-CH<sub>2</sub>), 39.2 (Lac-C3), 23.0 (SCH<sub>2</sub>CH<sub>3</sub>), 21.1 (Ar-Me), 15.2 (SCH<sub>2</sub>CH<sub>3</sub>); ESI-MS: *m/z*: Calcd for C<sub>31</sub>H<sub>34</sub>NaO<sub>7</sub>S [M+Na]<sup>+</sup>: 573.7, found: 573.1.

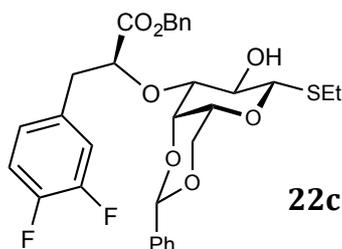
**Ethyl 4,6-O-benzylidene-3-O-[(R)-1-(benzyloxy)-1-oxo-3-(4-methylphenyl)propan-2-yl]-1-thio-β-D-galactopyranoside (22b)**



Thiogalactoside (**20**) (50 mg, 0.16 mmol) and *n*-Bu<sub>2</sub>SnO (60 mg, 0.24 mmol) were dissolved in dry MeOH (3 mL) and refluxed at 65°C under argon for 4 h. The solvent was removed under reduced pressure and the crude product was dried in high vacuum for 16 h. The residue was dissolved in dry DME (3 mL) under argon and a solution of triflate (**19b**) (193 mg, 0.48 mmol) in dry DME (2 mL) was added followed by the addition of CsF (73 mg, 0.48 mmol, dried in high vacuum at 100°C for 16 h). The turbid solution was stirred for 3 h after which TLC (petroleum ether/EtOAc, 9:1, DCM/MeOH, 19:1) indicated still presence of starting material (**20**) but no triflate (**19b**) so another portion of (**19b**) (128 mg, 0.32 mmol) was added and the mixture was stirred for additional 3 h. A solution of KF (10%) in aq. KH<sub>2</sub>PO<sub>4</sub> (1 N, 10 mL) was added, and after stirring for 1 h, the reaction was extracted with DCM (3 x 20 mL). The combined organic phases were washed with brine (10 mL) dried over Na<sub>2</sub>SO<sub>4</sub> and the solvents were removed *in vacuo*. The crude product was purified by flash chromatography (petroleum ether/EtOAc, gradient 0-40%) to afford (**21b**) as colorless wax (22.6 mg). ESI-MS: *m/z*: Calcd for C<sub>25</sub>H<sub>28</sub>NaO<sub>6</sub>S [M+Na]<sup>+</sup>: 479.5, found: 479.1. To a solution of **21b** (22.6 mg, 0.050 mmol) in benzyl alcohol (0.5 mL) DMAP (cat.) was added. The reaction was

heated to 60°C for 4 h, then the reaction mixture was purified by flash chromatography (petroleum ether/EtOAc, gradient 0-50%) to afford (**22b**) as colorless wax (18.2 mg, 0.032 mmol, 20% after two steps).  $R_f$  = (petroleum ether/EtOAc, 7:3) 0.30;  $[\alpha]_D^{22}$  -79.3 ( $c$  0.25, CHCl<sub>3</sub>); <sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>):  $\delta$  = 7.53 - 7.45 (m, 2H, Ar-H), 7.41 - 7.30 (m, 6H, Ar-H), 7.12 (d,  $J$  = 7.4 Hz, 2H, Ar-H), 7.04 (d,  $J$  = 7.4 Hz, 2H, Ar-H), 5.34 (s, 1H, Ph-CH), 5.09 (d,  $J$  = 12.1 Hz, 1H, Ar-CH<sub>2</sub>), 5.03 (d,  $J$  = 12.1 Hz, 1H, Ar-CH<sub>2</sub>), 4.66 - 4.54 (m, 1H, Lac-H2), 4.26 - 4.20 (m, 2H, Gal-H1, Gal-H6), 4.18 (s, 1H, Gal-H4), 3.91 - 3.82 (m, 2H, Gal-H2, Gal-H6'), 3.32 (d,  $J$  = 9.2 Hz, 1H, Gal-H3), 3.29 (s 1H, Gal-H5), 3.10 - 2.98 (m, 2H, Lac-H3), 2.81 - 2.71 (m, 2H, SCH<sub>2</sub>CH<sub>3</sub>), 2.30 (s, 3H, Ar-Me), 1.28 (t,  $J$  = 7.5 Hz, 3H, SCH<sub>2</sub>CH<sub>3</sub>); <sup>13</sup>C NMR (125.8 MHz, CDCl<sub>3</sub>):  $\delta$  = 172.4 (Lac-C1), 138.3, 136.3, 135.4, 133.8, 129.5, 129.1, 128.9, 128.6, 128.5, 128.5, 128.1, 126.4 (12C, Ar-C), 100.3 (Ph-CH), 85.1 (Gal-C1), 81.9 (Gal-C3), 80.6 (Lac-C2), 75.1 (Gal-C4), 70.2 (Gal-C5), 69.2 (Gal-C6), 68.6 (Gal-C2), 66.8 (Ph-CH<sub>2</sub>), 38.8 (Lac-C3), 23.0 (SCH<sub>2</sub>CH<sub>3</sub>), 21.1 (Ar-Me), 15.2 (SCH<sub>2</sub>CH<sub>3</sub>); ESI-MS:  $m/z$ : Calcd for C<sub>32</sub>H<sub>36</sub>NaO<sub>7</sub>S [M+Na]<sup>+</sup>: 587.7, found: 587.2.

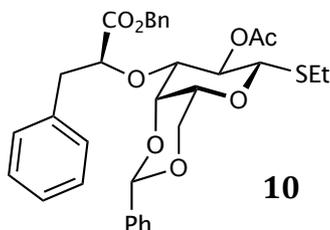
**Ethyl 4,6-O-benzylidene-3-O-[(R)-1-(benzyloxy)-1-oxo-3-(3,4-difluorophenyl)propan-2-yl]-1-thio- $\beta$ -D-galactopyranoside (**22c**)**



Thiogalactoside derivative (**20**) (208 mg, 0.66 mmol) and *n*-Bu<sub>2</sub>SnO (247 mg, 1.00 mmol) were dissolved in dry MeOH (10 mL) and refluxed at 65°C under argon for 4 h. The solvent was removed under reduced pressure and the crude product was dried in high vacuum for 16 h. This residue was dissolved in dry DME (4 mL) under argon and a solution of triflate (**19c**) (533 mg, 1.33 mmol) in dry DME (6 mL) was added followed by the addition of CsF (302 mg, 2.00 mmol, dried in high vacuum at 100°C for 16 h). The turbid solution was stirred for 3 h after which TLC (petroleum ether/EtOAc, 9:1, DCM/MeOH, 19:1) indicated still presence of starting material (**20**) but no triflate (**19c**) so another portion of (**19c**) (266 mg, 0.66 mmol) was added and the mixture was stirred for additional 16 h. A solution of KF (10%) in aq. KH<sub>2</sub>PO<sub>4</sub> (1 N, 10 mL) was added, and after stirring for 1 h, the reaction was extracted with DCM (3 x 50 mL). The combined organic phases were washed with brine (10 mL), dried over Na<sub>2</sub>SO<sub>4</sub> and the solvents were removed *in vacuo*. The crude product was purified by flash chromatography (petroleum ether/EtOAc, gradient 0-50%) to afford a 5:1 mixture of **21c** (lactone) and **22c** (benzylester), which was used in the following reaction without further purification. (94 mg). ESI-MS:  $m/z$ : Calcd for **21c** (lactone) C<sub>24</sub>H<sub>24</sub>F<sub>2</sub>NaO<sub>6</sub>S [M+Na]<sup>+</sup>: 501.5, found: 501.1, found: Calcd for **22c** (open) C<sub>31</sub>H<sub>32</sub>F<sub>2</sub>NaO<sub>7</sub>S [M+Na]<sup>+</sup>: 609.6, found: 609.1. To a solution of **21c** (90.0 mg, 0.188 mmol) in benzyl alcohol (1 mL) DMAP (cat.) was added. The reaction was heated to 60°C for 3 h, then the reaction mixture was purified by flash chromatography (petroleum ether/EtOAc, gradient 0-50%) to afford (**22c**) as colorless wax (88 mg, 0.15 mmol, 22% after two steps).  $R_f$  = (petroleum ether/EtOAc, 3:2) 0.32;

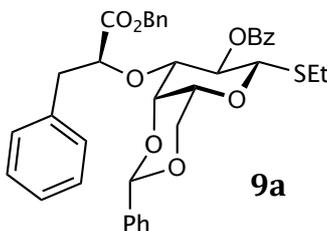
$[\alpha]_D^{22}$  -35.8 (*c* 0.35,  $\text{CHCl}_3$ );  $^1\text{H NMR}$  (500.1 MHz,  $\text{CDCl}_3$ ):  $\delta$  = 7.52 - 7.47 (m, 2H, Ar-H), 7.40 - 7.32 (m, 6H, Ar-H), 7.31 - 7.24 (m, 2H, Ar-H), 7.12 - 7.05 (m, 1H, Ar-H), 6.92 - 6.82 (m, 2H, Ar-H), 5.42 (s, 1H, Ph-CH), 5.11 (d,  $J$  = 12.0 Hz, 1H, Ar-CH<sub>2</sub>), 5.05 (d,  $J$  = 12.0 Hz, 1H, Ar-CH<sub>2</sub>), 4.75 (t,  $J$  = 5.7 Hz, 1H, Lac-H2), 4.29 - 4.18 (m, 3H, Gal-H1, Gal-H4, Gal-H6a), 3.93 - 3.87 (m, 2H, Gal-H2, Gal-H6b), 3.47 (d,  $J$  = 9.3 Hz, 1H, Gal-H3), 3.34 (s 1H, Gal-H5), 3.10 - 2.95 (m, 2H, Lac-H3), 2.82 - 2.64 (m, 2H, SCH<sub>2</sub>CH<sub>3</sub>), 2.31 (br s, 1H, OH), 1.30 (t,  $J$  = 7.4 Hz, 3H, SCH<sub>2</sub>CH<sub>3</sub>);  $^{13}\text{C NMR}$  (125.8 MHz,  $\text{CDCl}_3$ ):  $\delta$  = 172.0 (Lac-C1), 153.5 - 148.1 (m, 2C, ArF-C), 130.0, 135.3, 128.9, 128.7, 128.6, 128.1 (6C, Ar-C), 126.3 (Ar-C), 125.5 (m, Ar-C), 118.8 (d,  $J$  = 17.7 Hz, Ar-C), 116.6 (d,  $J$  = 17.0 Hz, Ar-C), 100.9 (Ph-CH), 85.7 (Gal-C1), 80.5 (Gal-C3), 79.3 (Lac-C2), 75.6 (Gal-C4), 70.4 (Gal-C5), 69.2 (Gal-C6), 69.1 (Gal-C2), 66.8 (Ph-CH<sub>2</sub>), 38.2 (Lac-C3), 23.0 (SCH<sub>2</sub>CH<sub>3</sub>), 15.4 (SCH<sub>2</sub>CH<sub>3</sub>); ESI-MS: *m/z*: Calcd for C<sub>31</sub>H<sub>32</sub>F<sub>2</sub>NaO<sub>7</sub>S [M+Na]<sup>+</sup>: 609.6, found: 609.11.

**Ethyl 2-*O*-acetyl-4,6-*O*-benzylidene-3-*O*-[(*R*)-1-(benzyloxy)-1-oxo-3-phenylpropan-2-yl]-1-thio- $\beta$ -D-galactopyranoside (**9a**)**



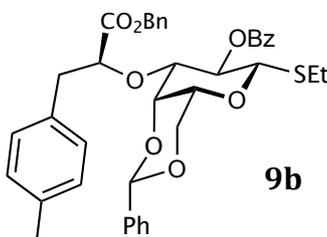
To a solution of (**22a**) (60 mg, 0.109 mmol) in pyridine (1 mL) Ac<sub>2</sub>O (1 ml) was added. The reaction was stirred at rt for 18 h. The solvent was removed *in vacuo*, and the crude product was purified by flash chromatography (petroleum ether/EtOAc, gradient 0-60%) to afford (**10**) as white solid (60 mg, 0.087 mmol, 80%). *R<sub>f</sub>* = (petroleum ether/EtOAc, 7:3) 0.40;  $[\alpha]_D^{22}$  -14.1 (*c* 1.0,  $\text{CHCl}_3$ );  $^1\text{H NMR}$  (500.1 MHz,  $\text{CDCl}_3$ ):  $\delta$  = 7.51 (dd,  $J$  = 7.7, 1.5 Hz, 2H, Ar-H), 7.42 - 7.33 (m, 6H, Ar-H), 7.31 - 7.14 (m, 7H, Ar-H), 5.36 (m, 2H, Gal-H2, Ph-CH), 5.07 (q,  $J$  = 12.0 Hz, 2H, Ar-CH<sub>2</sub>), 4.35 (dd,  $J$  = 5.0, 7.2 Hz, 1H, Lac-H2), 4.28 - 4.20 (m, 3H, Gal-H1, Gal-H4, Gal-H6), 3.85 (dd,  $J$  = 1.4, 12.3 Hz, Gal-H6'), 3.50 (dd,  $J$  = 3.3, 9.7 Hz, 1H, Gal-H3), 3.28 (s, 1H, Gal-H5), 3.05 - 2.96 (m, 2H, Lac-H3), 2.86 - 2.78 (m, 1H, SCH<sub>2</sub>CH<sub>3</sub>), 2.70 - 2.62 (m, 1H, SCH<sub>2</sub>'CH<sub>3</sub>), 1.78 (s, 3H, CO-CH<sub>3</sub>), 1.22 (t,  $J$  = 7.5 Hz, 3H, SCH<sub>2</sub>CH<sub>3</sub>);  $^{13}\text{C NMR}$  (126 MHz,  $\text{CDCl}_3$ ):  $\delta$  = 172.1 (Lac-C1), 169.1 (Ac-CO), 137.9, 136.0, 135.3, 129.6, 128.9, 128.7, 128.6, 128.6, 128.3, 128.1, 126.7, 126.4 (12C, Ar-C), 100.9 (Ph-CH), 82.8 (Gal-C1), 81.3 (Lac-C2), 79.6 (Gal-C3), 75.3 (Gal-C4), 70.2 (Gal-C5), 69.1 (Gal-C6), 68.7 (Gal-C2), 66.8 (Ph-CH<sub>2</sub>), 39.3 (Lac-C3), 22.5 (SCH<sub>2</sub>CH<sub>3</sub>), 20.7 (CO-CH<sub>3</sub>), 14.6 (SCH<sub>2</sub>CH<sub>3</sub>); ESI-MS: *m/z*: Calcd for C<sub>33</sub>H<sub>36</sub>NaO<sub>8</sub>S [M+Na]<sup>+</sup>: 615.7, found: 615.2.

**Ethyl 2-*O*-benzoyl-4,6-*O*-benzylidene-3-*O*-[(*R*)-1-(benzyloxy)-1-oxo-3-phenylpropan-2-yl]-1-thio- $\beta$ -D-galactopyranoside (**9a**)**



To a solution of (**22a**) (250 mg, 0.45 mmol, 1 eq) in pyridine (2 mL) DMAP (6 mg, 0.05 mmol, 0.1 eq) and benzoyl chloride (0.16 mL, 1.36 mmol, 3 eq) were added at 0°C. The reaction was stirred at rt for 3h after which TLC (petroleum ether/EtOAc, 3:2) indicated completion of the reaction. The solvent was removed *in vacuo*, the crude residue was taken up in DCM (100 mL), washed with 0.5 M aq. HCl (20 mL), satd aq. NaHCO<sub>3</sub> (20 mL) and brine (20 mL), dried over Na<sub>2</sub>SO<sub>4</sub> and concentrated *in vacuo*. The crude product was purified by flash chromatography (petroleum ether/EtOAc, gradient 0-50%) to afford (**9a**) as white solid (260 mg, 0.40 mmol, 87%).  $R_f$  = (petroleum ether/EtOAc, 7:3) 0.45;  $[\alpha]_D^{22}$  -22.5 (*c* 2.0, CHCl<sub>3</sub>); <sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>):  $\delta$  = 7.95 (d, *J* = 7.3 Hz, 2H, Ar-H), 7.62 - 7.51 (m, 3H, Ar-H), 7.46 - 7.28 (m, 10H, Ar-H), 7.02 (t, *J* = 7.3 Hz, 1H, Ar-H), 6.98 - 6.90 (m, 2H, Ar-H), 7.02 (d, *J* = 7.2 Hz, 2H, Ar-H), 5.67 (t, *J* = 9.7 Hz, 1H, Gal-H2), 5.39 (s, 1H, Ph-CH), 5.05 - 4.95 (m, 2H, Ar-CH<sub>2</sub>), 4.44 (d, *J* = 9.7 Hz, 1H, Gal-H1), 4.34 (t, *J* = 6.2 Hz, 1H, Lac-H2), 4.30 - 4.24 (m, 2H, Gal-H6a, Gal-H4), 3.88 (d, *J* = 12.2 Hz, Gal-H6b), 3.68 (d, *J* = 7.6 Hz, 1H, Gal-H3), 3.34 (s, 1H, Gal-H5), 2.92 - 2.82 (m, 3H, Lac-H3, SCH<sub>2</sub>CH<sub>3</sub>), 2.77 - 2.66 (m, 1H, SCH<sub>2</sub>CH<sub>3</sub>), 1.22 (t, *J* = 7.5 Hz, 3H, SCH<sub>2</sub>CH<sub>3</sub>); <sup>13</sup>C NMR (126 MHz, CDCl<sub>3</sub>):  $\delta$  = 172.0 (Lac-C1), 164.8 (Bz-CO), 137.9, 136.0, 135.3, 133.1, 129.9, 129.8, 129.4, 128.9, 128.7, 128.6, 128.4, 128.2, 128.1, 126.5, 126.4 (16C, Ar-C), 100.9 (Ph-CH), 82.9 (Gal-C1), 81.3 (Lac-C2), 79.9 (Gal-C3), 75.3 (Gal-C4), 70.3 (Gal-C5), 69.6 (Gal-C2), 69.2 (Gal-C6), 66.7 (Ph-CH<sub>2</sub>), 39.2 (Lac-C3), 22.7 (SCH<sub>2</sub>CH<sub>3</sub>), 14.8 (SCH<sub>2</sub>CH<sub>3</sub>); ESI-MS: *m/z*: Calcd for C<sub>38</sub>H<sub>38</sub>O<sub>8</sub>S [M+Na]<sup>+</sup>: 577.8, found: 577.2.

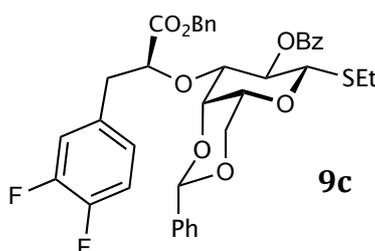
**Ethyl 2-*O*-benzoyl-4,6-*O*-benzylidene-3-*O*-[(*R*)-1-(benzyloxy)-1-oxo-3-(4-methylphenyl)-propan-2-yl]-1-thio- $\beta$ -D-galactopyranoside (**9b**)**



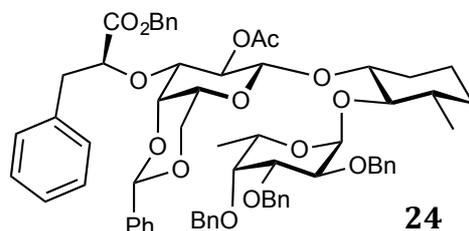
To a solution of (**22b**) (195 mg, 0.34 mmol) in pyridine (2 mL) DMAP (4 mg, 0.03 mmol) and benzoyl chloride (0.12 mL, 1.04 mmol) were added at 0°C. The reaction was stirred at rt for 16 h after which TLC (petroleum ether/EtOAc, 3:2) indicated completion of the reaction. The solvent was removed *in vacuo*, the crude residue was taken up in DCM (100 mL), washed with 0.5 M aq. HCl (20 mL), satd aq. NaHCO<sub>3</sub> (20 mL) and

brine (20 mL), dried over Na<sub>2</sub>SO<sub>4</sub> and concentrated *in vacuo*. The crude product was purified by flash chromatography (petroleum ether/EtOAc, gradient 0-50%) to afford (**9b**) as white solid (166 mg, 0.25 mmol, 72%). R<sub>f</sub> = (petroleum ether/EtOAc, 7:3) 0.48; [α]<sub>D</sub><sup>22</sup> -26.3 (c 0.45, CHCl<sub>3</sub>); <sup>1</sup>H NMR (500.1 MHz, CDCl<sub>3</sub>): δ = 7.94 (d, *J* = 7.7 Hz, 2H, Ar-H), 7.64 - 7.50 (m, 3H, Ar-H), 7.49 - 7.29 (m, 8H, Ar-H), 7.25 - 7.17 (m, 2H, Ar-H), 6.81 - 6.68 (m, 4H, Ar-H), 5.67 (t, *J* = 9.7 Hz, 1H, Gal-H2), 5.39 (s, 1H, Ph-CH), 5.01 (s, 2H, Ar-CH<sub>2</sub>), 4.43 (d, *J* = 9.6 Hz, 1H, Gal-H1), 4.35 - 4.23 (m, 3H, Gal-H6, Lac-H2, Gal-H4), 3.88 (d, *J* = 12.2 Hz, Gal-H6'), 3.67 (d, *J* = 9.6 Hz, 1H, Gal-H3), 3.35 (s 1H, Gal-H5), 2.93 - 2.66 (m, 4H, Lac-H3, SCH<sub>2</sub>CH<sub>3</sub>), 2.18 (s, 3H, Ar-Me), 1.23 (t, *J* = 7.7 Hz, 3H, SCH<sub>2</sub>CH<sub>3</sub>); <sup>13</sup>C NMR (125.8 MHz, CDCl<sub>3</sub>): δ = 172.1 (Lac-C1), 164.8 (Bz-CO), 137.9, 135.8, 135.3, 133.0, 132.9, 129.9, 129.8, 129.3, 128.9, 128.8, 128.6, 128.6, 128.5, 128.3, 128.2, 126.5 (16C, Ar-C), 101.0 (Ph-CH), 82.9 (Gal-C1), 81.7 (Lac-C2), 79.9 (Gal-C3), 75.3 (Gal-C4), 70.3 (Gal-C5), 69.5 (Gal-C2), 69.2 (Gal-C6), 66.7 (Ph-CH<sub>2</sub>), 38.8 (Lac-C3), 22.7 (SCH<sub>2</sub>CH<sub>3</sub>), 21.0 (Ar-Me), 14.7 (SCH<sub>2</sub>CH<sub>3</sub>); ESI-MS: *m/z*: Calcd for C<sub>39</sub>H<sub>40</sub>NaO<sub>8</sub>S [M+Na]<sup>+</sup>: 691.8, found: 691.2.

**Ethyl 2-*O*-benzoyl-4,6-*O*-benzylidene-3-*O*-[(*R*)-1-(benzyloxy)-1-oxo-3-(3,4-difluorophenyl)propan-2-yl]-1-thio-β-D-galactopyranoside (**9c**)**

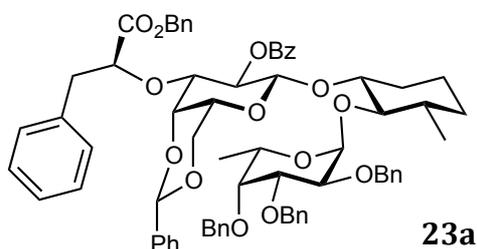


To a solution of (**22c**) (73 mg, 0.2 mmol) in pyridine (2 mL) DMAP (1.5 mg, 0.01 mmol) and benzoyl chloride (0.045 mL, 0.37 mmol) were added at 0°C. The reaction was stirred at rt for 16 h after which TLC (petroleum ether/EtOAc, 3:2) indicated completion of the reaction. The solvent was removed *in vacuo*, the crude residue was taken up in DCM (100 mL), washed with 0.5 M aq. HCl (20 mL), sat. NaHCO<sub>3</sub> (20 mL) and brine (20 mL), dried over Na<sub>2</sub>SO<sub>4</sub> and concentrated *in vacuo*. The crude product was purified by flash chromatography (petroleum ether/EtOAc, gradient 0-50%) to afford (**9c**) as white solid (55 mg, 0.08 mmol, 64%). R<sub>f</sub> = (petroleum ether/EtOAc, 7:3) 0.40; [α]<sub>D</sub><sup>22</sup> -35.3 (c 0.28, CHCl<sub>3</sub>); <sup>1</sup>H NMR (500.1 MHz, CDCl<sub>3</sub>): δ = 7.92 (d, *J* = 7.7 Hz, 2H, Ar-H), 7.65 - 7.51 (m, 3H, Ar-H), 7.50 - 7.32 (m, 8H, Ar-H), 7.32 - 7.22 (m, 2H, Ar-H), 6.77 - 6.53 (m, 3H, Ar-H), 5.65 (t, *J* = 9.6 Hz, 1H, Gal-H2), 5.42 (s, 1H, Ph-CH), 5.13 (d, *J* = 11.9 Hz, 1H, Ar-CH<sub>2</sub>), 5.02 (d, *J* = 11.9 Hz, 1H, Ar-CH<sub>2</sub>), 4.43 (d, *J* = 9.7 Hz, 1H, Gal-H1), 4.38 - 4.23 (m, 3H, Lac-H2, Gal-H6a, Gal-H4), 3.89 (d, *J* = 12.3 Hz, Gal-H6b), 3.69 (d, *J* = 9.5 Hz, 1H, Gal-H3), 3.36 (s, 1H, Gal-H5), 2.93 - 2.66 (m, 4H, Lac-H3, SCH<sub>2</sub>CH<sub>3</sub>), 1.22 (t, *J* = 7.5 Hz, 3H, SCH<sub>2</sub>CH<sub>3</sub>); <sup>13</sup>C NMR (125.8 MHz, CDCl<sub>3</sub>): δ = 171.6 (Lac-C1), 164.7 (Bz-CO), 160.8 - 140.9 (m, 2C, Ar-C), 137.9, 135.1, 133.7, 133.3 (4C, Ar-C), 133.0 (m, Ar-C), 130.2, 129.7, 129.0, 128.8, 128.8, 128.8, 128.5, 128.5, 128.2, 126.4 (10C, Ar-C), 125.7 - 125.4 (m, Ar-C), 118.4 (d, *J* = 17.2 Hz, Ar-C), 116.6 (d, *J* = 17.0 Hz, Ar-C), 101.0 (Ph-CH), 82.9 (Gal-C1), 80.8 (Lac-C2), 79.8 (Gal-C3), 75.5 (Gal-C4), 70.3 (Gal-C5), 69.6 (Gal-C2), 69.1 (Gal-C6), 67.0 (Ph-CH<sub>2</sub>), 38.2 (Lac-C3), 22.7 (SCH<sub>2</sub>CH<sub>3</sub>), 14.7 (SCH<sub>2</sub>CH<sub>3</sub>); ESI-MS: *m/z*: Calcd for C<sub>38</sub>H<sub>36</sub>F<sub>2</sub>NaO<sub>8</sub>S [M+Na]<sup>+</sup>: 713.7, found: 713.1.

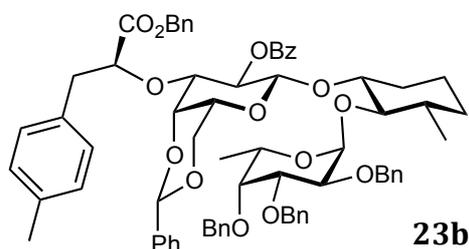
**(1*R*,2*R*,3*S*)-2-[(2,3,4-Tri-*O*-benzyl-6-deoxy- $\alpha$ -*L*-galactopyranosyl)oxy]-3-methylcyclo-hex-1-yl 2-*O*-benzoyl-4,6-*O*-benzylidene-3-*O*-[(*R*)-1-(benzyloxy)-1-oxo-3-phenylpropan-2-yl]- $\beta$ -*D*-galactopyranoside (**24**)**

A solution of donor (**10**) (45 mg, 0.076 mmol) and acceptor (**11**) (62 mg, 0.113 mmol) in dry DCM (5 mL) was added *via* syringe to activated 4Å molecular sieves (0.5 g) under argon. A suspension of dimethyl(methylthio)sulfonium triflate (DMTST, 97 mg, 0.395 mmol) and activated 4Å molecular sieves (0.5 g) in DCM (5 mL) was prepared in a second flask under argon. Both suspensions were stirred at rt for 4 h, then the DMTST suspension was added *via* syringe to the other suspension. The reaction was stopped after 18 h, filtered through celite and the celite was washed with DCM (20 mL). The filtrate was successively washed with satd. aq. NaHCO<sub>3</sub> (25 mL) and water (25 mL), dried over Na<sub>2</sub>SO<sub>4</sub>, filtered and concentrated. The crude product was purified by flash chromatography (petroleum ether/EtOAc, gradient 0-50%) to afford (**24**) as white solid (64 mg, 0.059 mmol, 78%).  $R_f$  = (petroleum ether/EtOAc, 3:2) 0.30;  $[\alpha]_D^{22}$  -66.0 (*c* 0.48, CHCl<sub>3</sub>); <sup>1</sup>H NMR (500.1 MHz, CDCl<sub>3</sub>):  $\delta$  = 7.60 (d, *J* = 7.4 Hz, 2H, Ar-H), 7.36 - 7.13 (m, 28H, Ar-H), 5.39 (s, 1H, Ph-CH), 5.25 (dd, *J* = 8.2, 10.0 Hz, 1H, Gal-H2), 5.07 - 4.99 (m, 2H, Ph-CH<sub>2</sub>), 4.94 (d, *J* = 3.1 Hz, 1H, Fuc-H1), 4.88 (q, *J* = 6.5 Hz, 1H, Fuc-H5), 4.80 (d, *J* = 7.4 Hz, 1H, Ph-CH<sub>2</sub>), 4.70 (d, *J* = 11.7 Hz, 1H, Ph-CH<sub>2</sub>), 4.59 (q, *J* = 11.5 Hz, 2H, Ph-CH<sub>2</sub>), 4.37 (t, *J* = 6.0 Hz, 1H, Lac-H2), 4.29 (d, *J* = 8.0 Hz, 1H, Gal-H1), 4.25 (d, *J* = 11.5 Hz, 1H, Gal-H6), 4.22 - 4.17 (m, 2H, Gal-H4, Ph-CH<sub>2</sub>), 3.96 - 3.86 (m, 3H, Gal-H6', Fuc-H2, Fuc-H3), 3.60 (d, *J* = 11.4 Hz, 1H, Ph-CH<sub>2</sub>), 3.52 (dd, *J* = 3.3, 10.1 Hz, 1H, Gal-H3), 3.49 - 3.41 (m, 2H, MeCy-H1), 3.23 (s, 1H, Fuc-H4), 3.21 - 3.14 (m, 2H, MeCy-H2, Gal-H5), 3.07 - 3.04 (m, 2H, Lac-H3), 1.87 - 1.85 (m, 1H, MeCy-H3), 1.85 (s, 3H, CO-CH<sub>3</sub>), 1.63 - 1.50 (m, 3H, MeCy-H6, MeCy-H4), 1.16 (d, *J* = 6.4 Hz, 3H, Fuc-H6), 1.05 (d, *J* = 6.4 Hz, 3H, MeCy-Me), 1.16 - 0.81 (m, 3H, MeCy-H4', MeCy-H5); <sup>13</sup>C NMR (125.8 MHz, CDCl<sub>3</sub>):  $\delta$  = 171.9 (Lac-C1), 168.6 (Ac-CO), 139.8, 139.6, 138.7, 138.1, 136.5, 135.2, 129.6, 128.6, 128.6, 128.6, 128.5, 128.3, 128.1, 128.0, 127.8, 127.5, 127.4, 127.3, 127.0, 126.7, 126.7, 125.9 (Ar-C), 99.9 (Gal-C1), 99.5 (Ph-CH), 98.2 (Fuc-C1), 81.3 (MeCy-C2), 81.0 (Lac-C2), 80.7 (MeCy-C1), 79.7 (Fuc-C3), 79.0 (Fuc-C4), 78.6 (Gal-C3), 75.6 (Fuc-C3), 74.9, 74.8 (2C, Gal-C4, Ph-CH<sub>2</sub>), 74.4 (Ph-CH<sub>2</sub>), 71.1 (Ph-CH<sub>2</sub>), 70.9 (Gal-C2), 69.1 (Gal-C6), 66.8 (Ph-CH<sub>2</sub>), 66.3, 66.2 (2C, Fuc-C5, Gal-C5), 39.5 (MeCy-C6), 39.3 (Lac-C3), 33.6 (MeCy-C4), 31.3 (MeCy-C3), 23.3 (MeCy-C5), 20.8 (CO-CH<sub>3</sub>), 18.7 (MeCy-Me), 16.0 (Fuc-C6); ESI-MS: *m/z*: Calcd for C<sub>65</sub>H<sub>72</sub>NaO<sub>14</sub> [M+Na]<sup>+</sup>: 1100.27, found: 1099.43.

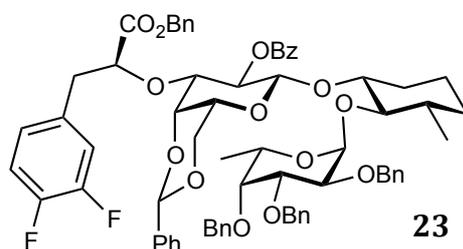
**(1*R*,2*R*,3*S*)-2-[(2,3,4-Tri-*O*-benzyl-6-deoxy- $\alpha$ -*L*-galactopyranosyl)oxy]-3-methylcyclo-hex-1-yl 2-*O*-benzoyl-4,6-*O*-benzylidene-3-*O*-[(*R*)-1-(benzyloxy)-1-oxo-3-phenylpropan-2-yl]- $\beta$ -*D*-galactopyranoside (23a)**



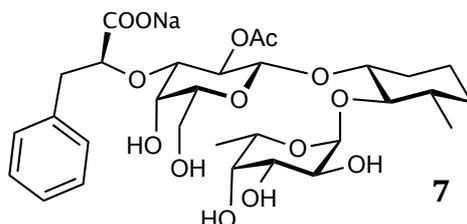
A solution of donor (**9a**) (101 mg, 0.159 mmol) and acceptor (**11**) (101 mg, 0.185 mmol) in dry DCM (5 mL) was added *via* syringe to activated 4Å molecular sieves (0.5 g) under argon. A suspension of dimethyl(methylthio)sulfonium triflate (DMTST, 119 mg, 0.462 mmol) and activated 4Å molecular sieves (0.5 g) in DCM (5 mL) was prepared in a second flask under argon. Both suspensions were stirred at rt for 4 h, then the DMTST suspension was added *via* syringe to the other suspension. The reaction was stopped after 18 h, filtered through celite and the celite was washed with DCM (20 mL). The filtrate was successively washed with satd. aq. NaHCO<sub>3</sub> (25 mL) and water (25 mL), dried over Na<sub>2</sub>SO<sub>4</sub>, filtered and concentrated. The crude product was purified by flash chromatography (petroleum ether/EtOAc, gradient 0-50%) to afford (**23a**) as white solid (138 mg, 0.12 mmol, 79%).  $R_f$  = (petroleum ether/EtOAc, 3:2) 0.35;  $[\alpha]_D^{22}$  -78.6 (*c* 0.39, CHCl<sub>3</sub>); <sup>1</sup>H NMR (500.1 MHz, CDCl<sub>3</sub>):  $\delta$  = 7.99 (d, *J* = 7.7 Hz, 2H, Ar-H), 7.63 (d, *J* = 7.5 Hz, 2H, Ar-H), 7.59 (t, *J* = 7.2 Hz, 1H, Ar-H), 7.45 (t, *J* = 7.4 Hz, 2H, Ar-H), 7.35 - 7.14 (m, 23H, Ar-H), 7.07 - 6.93 (m, 5H, Ar-H), 5.57 (t, *J* = 8.8 Hz, 1H, Gal-H2), 5.42 (s, 1H, Ph-CH), 5.02 - 4.85 (m, 5H, 2 Ph-CH<sub>2</sub>, Fuc-H1), 4.80 (d, *J* = 11.7 Hz, 1H, Ph-CH<sub>2</sub>), 4.69 (d, *J* = 11.7 Hz, 1H, Ph-CH<sub>2</sub>), 4.63 (d, *J* = 11.5 Hz, 1H, Ph-CH<sub>2</sub>), 4.58 (d, *J* = 11.3 Hz, 1H, Ph-CH<sub>2</sub>), 4.51 (d, *J* = 7.9 Hz, 1H, Gal-H1), 4.36 - 4.16 (m, 4H, Lac-H2, Gal-H6, Gal-H4, Fuc-H2), 4.00 - 3.87 (m, 3H, Gal-H6', Fuc-H3, Fuc-H5), 3.71 - 3.57 (m, 2H, Gal-H3, Fuc-H4), 3.51 (t, *J* = 9.5 Hz, 1H, MeCy-H2), 3.33 - 3.22 (m, 2H, Gal-H5, Fuc-H4), 3.13 (t, *J* = 9.5 Hz, 1H, MeCy-H1), 2.98 - 2.85 (m, 2H, Lac-H3), 1.85 - 1.77 (m, 1H, MeCy-H3), 1.58 - 1.41 (m, 3H, MeCy-H6, MeCy-H4), 1.25 (d, *J* = 6.3 Hz, 3H, Fuc-H6), 1.01 (d, *J* = 6.0 Hz, 3H, MeCy-Me), 1.16 - 0.81 (m, 3H, MeCy-H4', MeCy-H5); <sup>13</sup>C NMR (125.8 MHz, CDCl<sub>3</sub>):  $\delta$  = 171.8 (Lac-C1), 164.5 (Bz-CO), 139.8, 139.6, 138.8, 138.2, 136.0, 135.2, 133.0, 130.1, 129.7, 129.5, 128.7, 128.6, 128.6, 128.5, 128.5, 128.4, 128.1, 128.0, 128.0, 127.8, 127.5, 127.4, 127.3, 127.0, 126.8, 126.5, 126.0 (28C, Ar-C), 99.7 (Gal-C1), 99.5 (Ph-CH), 98.3 (Fuc-C1), 81.3 (MeCy-C1), 81.1 (Lac-C2), 80.7 (MeCy-C2), 79.7 (Fuc-C3), 79.1 (Fuc-C4), 89.9 (Gal-C3), 75.7 (Fuc-C5), 74.9 (2C, Gal-C4, Fuc-C4), 74.5 (Ph-CH<sub>2</sub>), 71.6 (Gal-C2), 71.2 (Ph-CH<sub>2</sub>), 69.2 (Gal-C6), 66.7 (Ph-CH<sub>2</sub>), 66.4 (Ph-CH<sub>2</sub>), 66.3 (Gal-C5), 39.5 (MeCy-C6), 39.2 (Lac-C3), 33.5 (MeCy-C4), 31.1 (MeCy-C3), 23.3 (MeCy-C5), 18.7 (MeCy-Me), 16.2 (Fuc-C6); ESI-MS: *m/z*: Calcd for C<sub>70</sub>H<sub>74</sub>NaO<sub>14</sub>[M+Na]<sup>+</sup>: 1162.3, found: 1161.4.

**(1*R*,2*R*,3*S*)-2-[(2,3,4-Tri-*O*-benzyl-6-deoxy- $\alpha$ -L-galactopyranosyl)oxy]-3-methylcyclohex-1-yl 2-*O*-benzoyl-4,6-*O*-benzylidene-3-*O*-[(*R*)-1-(benzyloxy)-1-oxo-3-(4-methylphenyl)-propan-2-yl]- $\beta$ -D-galactopyranoside (**23b**)**

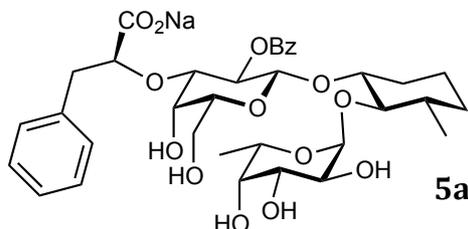
A solution of donor (**9b**) (30 mg, 0.044 mmol, 1 eq) and acceptor (**11**) (30 mg, 0.058 mmol) in dry DCM (1.5 mL) was added *via* syringe to activated 4Å molecular sieves (0.2 g) under argon. A suspension of dimethyl(methylthio)sulfonium triflate (DMTST, 35 mg, 0.134 mmol) and activated 4Å molecular sieves (0.2 g) in DCM (1.5 mL) was prepared in a second flask under argon. Both suspensions were stirred at rt for 4 h, then the DMTST suspension was added *via* syringe to the other suspension. The reaction was stopped after 18 h, filtered through celite and the celite was washed with DCM (10 mL). The filtrate was successively washed with satd. aq. NaHCO<sub>3</sub> (15 mL) and water (15 mL), dried over Na<sub>2</sub>SO<sub>4</sub>, filtered and concentrated. The crude product was purified by flash chromatography (petroleum ether/EtOAc, gradient 0-50%) to afford (**23b**) as white solid (41 mg, 0.035 mmol, 80%).  $R_f$  = (petroleum ether/EtOAc, 3:2) 0.33;  $[\alpha]_D^{22}$  -79.3 (*c* 0.25, CHCl<sub>3</sub>); <sup>1</sup>H NMR (500.1 MHz, CDCl<sub>3</sub>):  $\delta$  = 7.98 (d, *J* = 7.7 Hz, 2H, Ar-H), 7.64 (d, *J* = 7.4 Hz, 2H, Ar-H), 7.59 (t, *J* = 7.2 Hz, 1H, Ar-H), 7.39 - 7.12 (m, 27H, Ar-H), 6.87 (d, *J* = 7.5 Hz, 2H, Ar-H), 6.26 (d, *J* = 7.5 Hz, 2H, Ar-H), 5.56 (t, *J* = 9.6 Hz, 1H, Gal-H2), 5.43 (s, 1H, Ph-CH), 5.04 - 4.90 (m, 5H, 2 Ph-CH<sub>2</sub>, Fuc-H1), 4.80 (d, *J* = 11.7 Hz, 1H, Ph-CH<sub>2</sub>), 4.68 (d, *J* = 11.7 Hz, 1H, Ph-CH<sub>2</sub>), 4.63 (d, *J* = 11.3 Hz, 1H, Ph-CH<sub>2</sub>), 4.58 (d, *J* = 11.3 Hz, 1H, Ph-CH<sub>2</sub>), 4.51 (d, *J* = 9.7 Hz, 1H, Gal-H1), 4.30 - 4.15 (m, 4H, Lac-H2, Gal-H6, Gal-H4, Fuc-H2), 4.00 - 3.86 (m, 3H, Gal-H6', Fuc-H3, Fuc-H5), 3.71 - 3.57 (m, 2H, Gal-H3, Fuc-H4), 3.55 (t, *J* = 9.5 Hz, 1H, MeCy-H2), 3.30 - 3.21 (m, 2H, Gal-H5, Fuc-H4), 3.13 (t, *J* = 9.5 Hz, 1H, MeCy-H1), 2.93 - 2.85 (m, 2H, Lac-H3), 2.18 (s, 3H, Ar-Me), 1.86 - 1.74 (m, 1H, MeCy-H3), 1.55 - 1.40 (m, 3H, MeCy-H6, MeCy-H4), 1.25 (d, *J* = 6.0 Hz, 3H, Fuc-H6), 1.01 (d, *J* = 6.1 Hz, 3H, MeCy-Me), 1.16 - 0.81 (m, 3H, MeCy-H4', MeCy-H5); <sup>13</sup>C NMR (125.8 MHz, CDCl<sub>3</sub>):  $\delta$  = 172.0 (Lac-C1), 164.5 (Bz-CO), 139.8, 139.6, 138.8, 138.2, 135.9, 135.3, 132.9, 132.9, 130.1, 129.7, 129.4, 128.8, 128.7, 128.6, 128.6, 128.5, 128.5, 128.4, 128.1, 128.0, 128.0, 127.8, 127.6, 127.5, 127.3, 127.0, 126.8, 126.0 (28C, Ar-C), 99.7 (Gal-C1), 99.5 (Ph-CH), 98.3 (Fuc-C1), 81.4 (Lac-C2, MeCy-C1), 80.7 (MeCy-C2), 79.7 (Fuc-C3), 79.1 (Fuc-C4), 79.0 (Gal-C3), 75.7 (Fuc-C5), 74.9 (2C, Gal-C4, Fuc-C4), 74.5 (Ph-CH<sub>2</sub>), 71.6 (Gal-C2), 71.2 (Ph-CH<sub>2</sub>), 69.2 (Gal-C6), 66.7 (Ph-CH<sub>2</sub>), 66.4 (Ph-CH<sub>2</sub>), 66.3 (Gal-C5), 39.5 (MeCy-C6), 38.8 (Lac-C3), 33.5 (MeCy-C4), 31.1 (MeCy-C3), 23.3 (MeCy-C5), 18.7 (MeCy-Me), 16.2 (Fuc-C6); ESI-MS: *m/z*: Calcd for C<sub>71</sub>H<sub>76</sub>NaO<sub>14</sub> [M+Na]<sup>+</sup>: 1176.3, found: 1175.6.

**(1*R*,2*R*,3*S*)-2-[(2,3,4-Tri-*O*-benzyl-6-deoxy- $\alpha$ -L-galactopyranosyl)oxy]-3-methylcyclohex-1-yl 2-*O*-benzoyl-4,6-*O*-benzylidene-3-*O*-[(*R*)-1-(benzyloxy)-1-oxo-3-(3,4-difluorophenyl)-propan-2-yl]- $\beta$ -D-galactopyranoside (23c)**

A solution of donor (**9c**) (53 mg, 0.076 mmol) and acceptor (**11**) (50 mg, 0.092 mmol) in dry DCM (1.5 mL) was added *via* syringe to activated 4Å molecular sieves (0.25 g) under argon. A suspension of dimethyl(methylthio)sulfonium triflate (DMTST, 60 mg, 0.230 mmol) and activated 4Å molecular sieves (0.25 g) in DCM (1.5 mL) was prepared in a second flask under argon. Both suspensions were stirred at rt for 4 h, then the DMTST suspension was added *via* syringe to the other suspension. The reaction was stopped after 18 h, filtered through celite and the celite was washed with DCM (10 mL). The filtrate was successively washed with satd. aq. NaHCO<sub>3</sub> (20 mL) and water (20 mL), dried over Na<sub>2</sub>SO<sub>4</sub>, filtered and concentrated. The crude product was purified by flash chromatography (petroleum ether/EtOAc, gradient 0-50%) to afford (**23c**) as colorless wax (41 mg, 0.035 mmol, 80%). *R*<sub>f</sub> = (petroleum ether/EtOAc, 3:2) 0.28; [ $\alpha$ ]<sub>D</sub><sup>22</sup> -84.6 (*c* 0.60, CHCl<sub>3</sub>); <sup>1</sup>H NMR (500.1 MHz, CDCl<sub>3</sub>):  $\delta$  = 7.96 (d, *J* = 7.5 Hz, 2H, Ar-H), 7.66 - 7.55 (m, 3H, Ar-H), 7.45 (t, *J* = 7.6 Hz, 2H, Ar-H), 7.40 - 7.14 (m, 23H, Ar-H), 6.89 - 6.81 (m, 1H, Ar-H), 6.74 - 6.60 (m, 2H, Ar-H), 5.56 (t, *J* = 9.0 Hz, 1H, Gal-H2), 5.47 (s, 1H, Ph-CH), 5.09 (d, *J* = 12.0 Hz, 1H, Ph-CH<sub>2</sub>), 4.98 (d, *J* = 12.0 Hz, 1H, Ph-CH<sub>2</sub>), 4.95 - 4.90 (m, 2H, Fuc-H1, Fuc-H2), 4.80 (d, *J* = 11.7 Hz, 1H, Ph-CH<sub>2</sub>), 4.69 (d, *J* = 11.7 Hz, 1H, Ph-CH<sub>2</sub>), 4.63 (d, *J* = 11.5 Hz, 1H, Ph-CH<sub>2</sub>), 4.58 (d, *J* = 11.3 Hz, 1H, Ph-CH<sub>2</sub>), 4.51 (d, *J* = 7.9 Hz, 1H, Gal-H1), 4.34 - 4.24 (m, 3H, Lac-H2, Gal-H6, Gal-H4), 4.20 (d, *J* = 11.3 Hz, 1H, Ph-CH<sub>2</sub>), 4.00 - 3.88 (m, 3H, Gal-H6', Fuc-H2, Fuc-H3), 3.68 (d, *J* = 8.2 Hz, 1H, Gal-H3), 3.63 (d, *J* = 11.3 Hz, 1H, Ph-CH<sub>2</sub>), 3.57 - 3.46 (m, 1H, MeCy-H2), 3.33 - 3.23 (m, 2H, Gal-H5, Fuc-H4), 3.13 (t, *J* = 9.5 Hz, 1H, MeCy-H1), 2.94 - 2.80 (m, 2H, Lac-H3), 1.85 - 1.77 (m, 1H, MeCy-H3), 1.56 - 1.40 (m, 3H, MeCy-H6, MeCy-H4), 1.25 (d, *J* = 6.3 Hz, 3H, Fuc-H6), 1.01 (d, *J* = 6.4 Hz, 3H, MeCy-Me), 1.16 - 0.81 (m, 3H, MeCy-H4', MeCy-H5); <sup>13</sup>C NMR (126 MHz, CDCl<sub>3</sub>):  $\delta$  = 171.5 (Lac-C1), 164.4 (Bz-CO), 150.9 - 145.1 (m, 2C, ArF-C), 139.8, 139.6, 138.8, 138.1, 135.0, 133.1, 132.9, 130.1, 129.8, 129.6, 129.0, 128.8, 128.7, 128.6, 128.5, 128.2, 128.1, 128.1, 128.0, 127.8, 127.6, 127.4, 127.3, 127.0, 126.8, 125.9, 125.5 (26C, Ar-C), 118.5 (d, *J* = 17.3 Hz, Ar-C), 116.9 (d, *J* = 17.3 Hz, Ar-C), 99.6, 99.6 (2C, Gal-C1, Ph-CH), 98.3 (Fuc-C1), 81.3 (MeCy-C1), 80.7 (Lac-C2), 80.5 (MeCy-C2), 79.7 (Fuc-C3), 79.1 (Fuc-C4), 79.9 (Gal-C3), 75.7 (Fuc-C2), 75.1 (Gal-C4), 74.9 (Ph-CH<sub>2</sub>), 74.5 (Ph-CH<sub>2</sub>), 71.8 (Gal-C2), 71.2 (Ph-CH<sub>2</sub>), 69.2 (Gal-C6), 66.7 (Ph-CH<sub>2</sub>), 66.4, 66.3 (2C, Gal-C5, Fuc-C5), 39.5 (MeCy-C6), 38.2 (Lac-C3), 33.5 (MeCy-C4), 31.1 (MeCy-C3), 23.3 (MeCy-C5), 18.7 (MeCy-Me), 16.2 (Fuc-C6); ESI-MS: *m/z*: Calcd for C<sub>70</sub>H<sub>72</sub>F<sub>2</sub>NaO<sub>14</sub> [M+Na]<sup>+</sup>: 1198.3, found: 1198.9.

**(1R,2R,3S)-2-[( $\alpha$ -L-Fucopyranosyl)oxy]-3-methyl-cyclohex-1-yl 2-O-acetyl-3-O-[sodium (*S*)-1-carboxy-2-phenylethyl]- $\beta$ -D-galactopyranoside (7)**

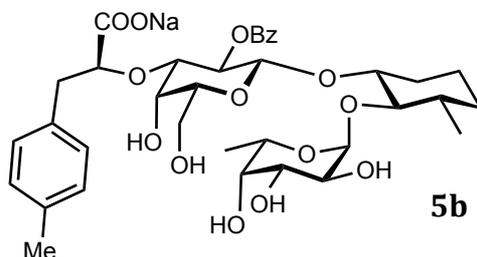
A suspension of (**24**) (15 mg, 0.014 mmol) and Pd(OH)<sub>2</sub>/C (50 mg, 10% Pd) in THF (2 mL) was hydrogenated (2 bar H<sub>2</sub>) at rt. After 1 h TLC (silica: petroleum ether/EtOAc, 3:2; C-18: H<sub>2</sub>O/MeOH, 1:1) indicated completion of the reaction. The reaction mixture was filtered and concentrated *in vacuo*. The crude product was purified by reversed phase column chromatography (C-18, H<sub>2</sub>O/MeOH, gradient 0-100%) followed by a Dowex 50 (Na<sup>+</sup> form) ion exchange column, a Sephadex G15 column, and lyophilization from water to yield the product (**7**) as white foam (3 mg, 0.0046 mmol, 33%). R<sub>f</sub> = (C-18, H<sub>2</sub>O/MeOH, 1:1) 0.55; [α]<sub>D</sub><sup>22</sup> -97.6 (c 0.35, MeOH); <sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>): δ = 7.42 - 7.36 (m, 2H, Ar-H), 7.34 - 7.28 (m, 3H, Ar-H), 5.06 (d, *J* = 4.0 Hz, 1H, Fuc-H1), 4.92 - 4.84 (m, 2H, Fuc-H5, Gal-H2), 4.52 (d, *J* = 8.1 Hz, 1H, Gal-H1), 4.16 (dd, *J* = 3.3, 9.8 Hz, 1H, Lac-H2), 3.99 (d, *J* = 2.6 Hz, 1H, Gal-H4), 3.89 (dd, *J* = 3.3, 10.5 Hz, 1H, Fuc-H3), 3.84 - 3.77 (m, 2H, Fuc-H2, Fuc-H4), 3.74 (d, *J* = 5.9 Hz, 2H, Gal-H6), 3.61 - 3.54 (m, 2H, Gal-H5, MeCy-H1), 3.51 (dd, *J* = 3.1, 9.8 Hz, 1H, Gal-H3), 3.18 - 3.01 (m, 2H, MeCy-H2, Lac-H3), 2.84 (dd, *J* = 9.8, 14.3 Hz, 1H, Lac-H3'), 2.04 - 1.92 (m, 1H, MeCy-H6), 1.83 (s, 3H, CO-CH<sub>3</sub>), 1.65 - 1.51 (m, 3H, MeCy-H3, MeCy-H4a, MeCy-H5), 1.21 (d, *J* = 6.6 Hz, 3H, Fuc-H6), 1.17 - 1.09 (m, 1H, MeCy-H5'), 1.07 (d, *J* = 6.4 Hz, 3H, MeCy-Me), 1.09 - 0.97 (m, 2H, MeCy-H4', MeCy-H6'); <sup>13</sup>C NMR (126 MHz, D<sub>2</sub>O): δ = 180.9 (Lac-C1), 173.4 (Me-CO), 138.3, 129.0, 128.6, 126.6 (4C, Ar-C), 98.7, 98.6 (2C, Gal-C1, Fuc-C1), 83.5 (MeCy-C2), 82.1 (Lac-C2), 80.6 (Gal-C3), 79.8 (MeCy-C1), 74.1 (Gal-C5), 72.0 (Fuc-C4), 71.6 (Gal-C2), 69.3 (Fuc-C3), 68.2 (Fuc-C2), 66.5 (Fuc-C5), 66.4 (Gal-C4), 61.7 (Gal-C6), 39.6 (Lac-C3), 38.6 (MeCy-C3), 33.0 (MeCy-C4), 30.7 (MeCy-C6), 22.4 (MeCy-C5), 20.6 (CO-CH<sub>3</sub>), 18.1 (MeCy-C6), 15.3 (Fuc-C6); HR-MS: *m/z*: Calcd for C<sub>30</sub>H<sub>43</sub>Na<sub>2</sub>O<sub>14</sub> [M+Na]<sup>+</sup>: 673.2448, found: 673.2451.

**(1R,2R,3S)-2-[( $\alpha$ -L-Fucopyranosyl)oxy]-3-methyl-cyclohex-1-yl 2-O-benzoyl-3-O-[sodium (*S*)-1-carboxy-2-phenylethyl]- $\beta$ -D-galactopyranoside (5a)**

A suspension of (**23a**) (76 mg, 0.068 mmol) and Pd(OH)<sub>2</sub>/C (50 mg, 10% Pd) in MeOH (5 mL) was hydrogenated (2 bar H<sub>2</sub>) at rt. After 1 h TLC (silica: petroleum ether/EtOAc, 3:2; C-18: H<sub>2</sub>O/MeOH, 1:1) indicated completion of the reaction. The reaction mixture was filtered through celite and concentrated *in vacuo*. The crude product was purified by

reversed phase column chromatography (C-18, H<sub>2</sub>O/MeOH, gradient 0-100%) followed by a Dowex 50 (Na<sup>+</sup> form) ion exchange column, a Sephadex G15 column, and lyophilization from water to yield the product (**5a**) as white foam (20.3 mg, 0.028 mmol, 42%).  $R_f$  = (C-18, H<sub>2</sub>O/MeOH, 1:1) 0.48;  $[\alpha]_D^{22}$  -97.6 (*c* 0.35, MeOH); <sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>):  $\delta$  = 7.92 (d, *J* = 8.1 Hz, 2H, Ar-H), 7.76 (t, *J* = 7.5 Hz, 1H, Ar-H), 7.57 (t, *J* = 7.9 Hz, 2H, Ar-H), 7.04 - 7.96 (m, 2H, Ar-H), 6.87 - 6.75 (m, 2H, Ar-H), 5.19 (m, 1H, Gal-H2), 5.04 (d, *J* = 4.0 Hz, 1H, Fuc-H1), 4.90 (q, *J* = 6.6 Hz, 1H, Fuc-H5), 4.70 (d, *J* = 8.1 Hz, 1H, Gal-H1), 4.12 (dd, *J* = 3.2, 9.7 Hz, 1H, Lac-H2), 4.04 (d, *J* = 2.5 Hz, 1H, Gal-H4), 3.91 - 3.75 (m, 5H, Gal-H6, Fuc-H2, Fuc-H3, Fuc-H4), 3.69 - 3.62 (m, 2H, Gal-H3, Gal-H5), 3.57 (m, 1H, MeCy-H1), 3.06 (t, *J* = 9.4 Hz, 1H, MeCy-H2), 3.00 (dd, *J* = 3.1, 14.3 Hz, 1H, Lac-H3), 2.74 (dd, *J* = 9.7, 14.3 Hz, 1H, Lac-H3'), 1.95 (s, 1H, Ar-Me), 1.97 - 1.88 (m, 1H, MeCy-H6), 1.55 - 1.39 (m, 3H, MeCy-H3, MeCy-H4a, MeCy-H5), 1.30 (d, *J* = 8.1 Hz, 3H, Fuc-H6), 1.14 (m, 1H, MeCy-H5'), 1.01 (d, *J* = 6.4 Hz, 3H, MeCy-Me), 1.82 - 0.76 (m, 2H, MeCy-H4', MeCy-H6'); <sup>13</sup>C NMR (126 MHz, D<sub>2</sub>O):  $\delta$  = 181.1 (Lac-C1), 167.8 (Bz-CO), 137.9, 134.0, 129.9, 129.0, 128.7, 128.6, 128.1, 126.2 (8C, Ar-C), 98.7, 98.7 (2C, Gal-C1, Fuc-C1), 83.5 (MeCy-C2), 82.4 (Lac-C2), 80.9 (MeCy-C2), 79.8 (Gal-C3), 74.2 (MeCy-C1), 72.1, 72.0 (2C, Gal-C5, Gal-C2), 69.3 (Fuc-C4), 68.2 (Fuc-C2), 66.5 (Fuc-C5), 66.4 (Gal-C4), 61.7 (Gal-C6), 39.6 (Lac-C3), 38.5 (MeCy-C3), 32.9 (MeCy-C4), 30.8 (MeCy-C6), 22.3 (MeCy-C5), 20.0 (Ar-Me), 18.0 (MeCy-C6), 15.4 (Fuc-C6); HR-MS: *m/z*: Calcd for C<sub>35</sub>H<sub>45</sub>Na<sub>2</sub>O<sub>14</sub>[M+Na]<sup>+</sup>: 735.2605, found: 735.2607.

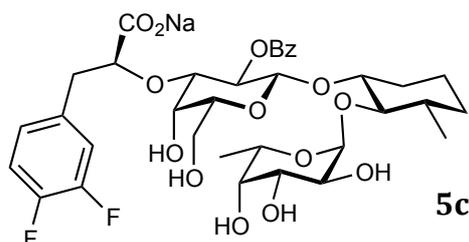
**(1*R*,2*R*,3*S*)-2-[( $\alpha$ -L-Fucopyranosyl)oxy]-3-methyl-cyclohex-1-yl 2-*O*-benzoyl-3-*O*-[sodium (*S*)-1-carboxy-2-(4-methylphenyl)-ethyl]- $\beta$ -D-galactopyranoside (**5b**)**



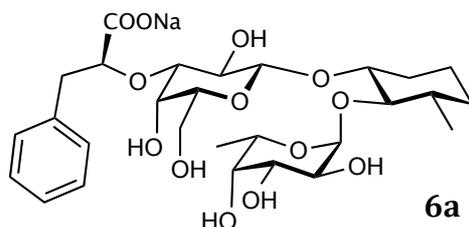
A suspension of (**23b**) (40 mg, 0.035 mmol) and Pd(OH)<sub>2</sub>/C (50 mg, 10% Pd) in MeOH (5 mL) was hydrogenated (2 bar H<sub>2</sub>) at rt. After 3 h TLC (silica: petroleum ether/EtOAc, 3:2; C-18: H<sub>2</sub>O/MeOH, 1:1) indicated completion of the reaction. The reaction mixture was filtered through celite and concentrated *in vacuo*. The crude product was purified by reversed phase column chromatography (C-18, H<sub>2</sub>O/MeOH, gradient 0-100%) followed by a Dowex 50 (Na<sup>+</sup> form) ion exchange column, a Sephadex G15 column, and lyophilization from water to yield the product (**5b**) as colorless foam (9.4 mg, 0.013 mmol, 37%).  $R_f$  = (C-18, H<sub>2</sub>O/MeOH, 1:1) 0.42;  $[\alpha]_D^{22}$  -64.5 (*c* 0.23, MeOH); <sup>1</sup>H NMR (500.1 MHz, CDCl<sub>3</sub>):  $\delta$  = 7.92 (d, *J* = 7.9 Hz, 2H, Ar-H), 7.77 (t, *J* = 7.2 Hz, 1H, Ar-H), 7.59 (t, *J* = 7.7 Hz, 2H, Ar-H), 6.92 (d, *J* = 7.5 Hz, 2H, Ar-H), 6.59 (d, *J* = 7.5 Hz, 2H, Ar-H), 5.19 (t, *J* = 8.7 Hz, 1H, Gal-H2), 5.03 (d, *J* = 2.8 Hz, 1H, Fuc-H1), 4.90 (dd, *J* = 6.5, 13.0 Hz, 1H, Fuc-H5), 4.71 (d, *J* = 8.1 Hz, 1H, Gal-H1), 4.08 (d, *J* = 10.4 Hz, 1H, Lac-H2), 4.03 (s, 1H, Gal-H4), 3.91 - 3.73 (m, 5H, Gal-H6, Fuc-H2, Fuc-H3, Fuc-H4), 3.69 - 3.60 (m, 2H, Gal-H3, Gal-H5), 3.57 (t, *J* = 13.2 Hz, 1H, MeCy-H1), 3.08 (t, *J* = 9.3 Hz, 1H, MeCy-H2), 2.96 (d, *J* = 13.7 Hz, 1H, Lac-H3a), 2.69 (m, 1H, Lac-H3b), 1.95 (s, 1H, Ar-Me), 1.93 (m, 1H, MeCy-H6), 1.58 - 1.40 (m, 3H, MeCy-H3, MeCy-H4,

MeCy-H5), 1.30 (d,  $J = 6.5$  Hz, 3H, Fuc-H6), 1.14 (m, 1H, MeCy-H5'), 1.01 (d,  $J = 6.4$  Hz, 3H, MeCy-Me), 1.93 - 0.76 (m, 2H, MeCy-H4', MeCy-H6');  $^{13}\text{C}$  NMR (125.8 MHz,  $\text{D}_2\text{O}$ ):  $\delta = 181.2$  (Lac-C1), 167.8 (Bz-CO), 136.0, 134.9, 134.0, 129.7, 129.1, 128.8, 128.8, 128.6 (8C, Ar-C), 98.7, 98.7 (2C, Gal-C1, Fuc-C1), 83.6 (MeCy-C2), 82.6 (Lac-C2), 80.5 (MeCy-C2), 80.9 (Gal-C3), 79.7 (MeCy-C1), 74.2 (Gal-C5), 72.1 (2C, Gal-C2, Fuc-C3), 69.3 (Fuc-C4), 68.2 (Fuc-C2), 66.5 (Fuc-C5), 66.4 (Gal-C4), 61.7 (Gal-C6), 39.2 (Lac-C3), 38.5 (MeCy-C3), 32.9 (MeCy-C4), 30.8 (MeCy-C6), 22.3 (MeCy-C5), 20.0 (MeCy-Me), 18.0 (MeCy-C6), 15.0 (Fuc-C6); HR-MS:  $m/z$ : Calcd for  $\text{C}_{36}\text{H}_{47}\text{Na}_2\text{O}_{14}[\text{M}+\text{Na}]^+$ : 749.2761, found: 749.2760.

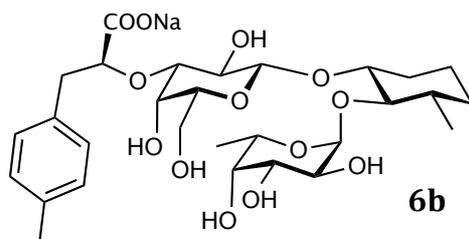
**(1*R*,2*R*,3*S*)-2-[( $\alpha$ -L-Fucopyranosyl)oxy]-3-methyl-cyclohex-1-yl 2-*O*-benzoyl-3-*O*-[sodium (*S*)-1-carboxy-2-(3,4-difluorophenyl)-ethyl]- $\beta$ -D-galactopyranoside (**5c**)**



A suspension of (**23c**) (48 mg, 0.040 mmol) and  $\text{Pd}(\text{OH})_2/\text{C}$  (50 mg, 10% Pd) in MeOH (5 mL) was hydrogenated (2 bar  $\text{H}_2$ ) at rt. After 1 h TLC (silica: petroleum ether/EtOAc, 3:2; C-18:  $\text{H}_2\text{O}/\text{MeOH}$ , 1:1) indicated completion of the reaction. The reaction mixture was filtered through celite and concentrated *in vacuo*. The crude product was purified by reversed phase column chromatography (C-18,  $\text{H}_2\text{O}/\text{MeOH}$ , gradient 0-100%) followed by a Dowex 50 ( $\text{Na}^+$  form) ion exchange column, a Sephadex G15 column, and lyophilization from water to yield the product (**5c**) as colorless foam (18.8 mg, 0.025 mmol, 63%).  $R_f = (\text{C-18}, \text{H}_2\text{O}/\text{MeOH}, 1:1) 0.45$ ;  $[\alpha]_D^{22} -80.7$  ( $c$  0.24, MeOH);  $^1\text{H}$  NMR (500.1 MHz,  $\text{CDCl}_3$ ):  $\delta = 7.86$  (d,  $J = 7.7$  Hz, 2H, Ar-H), 7.72 (t,  $J = 7.2$  Hz, 1H, Ar-H), 7.54 (t,  $J = 7.7$  Hz, 2H, Ar-H), 6.95 (m, 1H, Ar-H), 6.77 (m, 1H, Ar-H), 6.58 (m, 1H, Ar-H), 5.16 (t,  $J = 8.8$  Hz, 1H, Gal-H2), 5.03 (br s, 1H, Fuc-H1), 4.89 (dd,  $J = 6.6, 13.0$  Hz, 1H, Fuc-H5), 4.70 (d,  $J = 8.1$  Hz, 1H, Gal-H1), 4.16 (d,  $J = 10.8$  Hz, 1H, Lac-H2), 4.08 (br s, 1H, Gal-H4), 3.90 - 3.65 (m, 7H, Gal-H6, Gal-H5, Fuc-H2, Fuc-H3, Fuc-H4, Gal-H3), 3.56 (t,  $J = 11.6$  Hz, 1H, MeCy-H1), 3.05 (t,  $J = 9.3$  Hz, 1H, MeCy-H2), 2.97 (d,  $J = 14.1$  Hz, 1H, Lac-H3), 2.72 (m, 1H, Lac-H3b), 1.90 (m, 1H, MeCy-H6), 1.57 - 1.34 (m, 3H, MeCy-H3, MeCy-H4, MeCy-H5), 1.30 (d,  $J = 6.2$  Hz, 3H, Fuc-H6), 1.14 (m, 1H, MeCy-H5'), 1.01 (d,  $J = 6.2$  Hz, 3H, MeCy-Me), 1.96 - 0.70 (m, 2H, MeCy-H4', MeCy-H6');  $^{13}\text{C}$  NMR (125.8 MHz,  $\text{D}_2\text{O}$ ):  $\delta = 180.7$  (Lac-C1), 167.4 (Bz-CO), 150.2 - 146.4 (m, 2C, ArF-C), 135.1, 134.0, 129.5, 128.6, 128.5, 124.8 (8C, Ar-C), 117.1 (d,  $J = 17.2$  Hz, Ar-C), 116.6 (d,  $J = 17.3$  Hz, Ar-C), 98.8, 98.7 (2C, Gal-C1, Fuc-C1), 83.6 (MeCy-C2), 82.2 (Lac-C2), 80.9 (MeCy-C2), 80.9 (Gal-C3), 79.8 (MeCy-C1), 74.3 (Gal-C5), 72.1, 72.0 (2C, Gal-C2, Fuc-C3), 69.3 (Fuc-C4), 68.2 (Fuc-C2), 66.6, 66.5 (Fuc-C5, Gal-C4), 61.7 (Gal-C6), 38.7 (Lac-C3), 38.5 (MeCy-C3), 32.9 (MeCy-C4), 30.7 (MeCy-C6), 22.3 (MeCy-C5), 18.0 (MeCy-Me), 15.4 (Fuc-C6); HR-MS:  $m/z$ : Calcd for  $\text{C}_{35}\text{H}_{43}\text{F}_2\text{NaO}_{14}[\text{M}+\text{Na}]^+$ : 771.2416, found: 771.2415.

**(1R,2R,3S)-2-[( $\alpha$ -L-Fucopyranosyl)oxy]-3-methyl-cyclohex-1-yl 3-O-[sodium (*S*)-1-carboxy-2-phenylethyl]- $\beta$ -D-galactopyranoside (**6a**)**

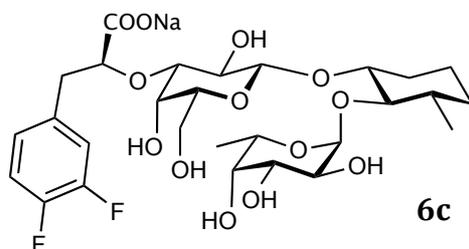
A suspension of (**23a**) (15 mg, 0.013 mmol) and Pd(OH)<sub>2</sub>/C (50 mg, 10% Pd) in MeOH (5 mL) was hydrogenated (1 bar H<sub>2</sub>) at rt. After 2 h TLC (silica: petroleum ether/EtOAc, 3:2; C-18: H<sub>2</sub>O/MeOH, 1:1) indicated completion of the reaction. The reaction mixture was filtered through celite and concentrated *in vacuo*. The residue was dissolved in MeOH/H<sub>2</sub>O (4:1, 1 mL) and treated with LiOH (16 mg, 0.65 mmol) for 7 days. The mixture was neutralized with Dowex 50X8 (H<sup>+</sup> form), filtered through a Dowex 50 (Na<sup>+</sup> form) ion exchange column and concentrated. The residue was purified by reversed phase column chromatography (C-18, H<sub>2</sub>O/MeOH, gradient 0-100%) followed by a Sephadex G15 column, and lyophilization from water to yield the product (**6a**) as white foam (1.5 mg, 0.0024 mmol, 18%). *R*<sub>f</sub> = (C-18, H<sub>2</sub>O/MeOH, 1:1) 0.25; [α]<sub>D</sub><sup>22</sup> -76.6 (*c* 0.26, MeOH); <sup>1</sup>H NMR (500.1 MHz, CDCl<sub>3</sub>): δ = 7.43 – 7.36 (m, 4H, Ar-H), 7.35 – 7.30 (m, 1H, Ar-H), 5.10 (d, *J* = 4.0 Hz, 1H, Fuc-H1), 4.86 – 4.80 (m, 1H, Fuc-H5), 4.40 (d, *J* = 8.0 Hz, 1H, Gal-H1), 4.18 (dd, *J* = 4.7, 8.7 Hz, 1H, Lac-H2), 3.92 (d, *J* = 2.8 Hz, 1H, Gal-H4), 3.89 (dd, *J* = 3.2, 10.6 Hz, 1H, Fuc-H3), 3.87 - 3.78 (m, 2H, Fuc-H2, Fuc-H4), 3.76 - 3.65 (m, 3H, MeCy-H1, Gal-H6), 3.58 – 3.51 (m, 2H, Gal-H2, Gal-H5), 3.30 (dd, *J* = 3.2, 9.5 Hz, 1H, Gal-H3), 3.22 (t, *J* = 9.6 Hz, 1H, MeCy-H2), 3.15 (dd, *J* = 4.6, 14.0 Hz, 1H, Lac-H3), 2.98 (dd, *J* = 8.8, 14.0 Hz, 1H, Lac-H3'), 2.12 - 2.05 (m, 1H, MeCy-H6), 1.70 - 1.53 (m, 3H, MeCy-H3, MeCy-H5), 1.37 - 1.21 (m, 2H, MeCy-H6', MeCy-H4), 1.18 (d, *J* = 6.6 Hz, 3H, Fuc-H6), 1.01 (d, *J* = 6.3 Hz, 3H, MeCy-Me), 1.14 – 1.04 (m, 1H, MeCy-H4'); <sup>13</sup>C NMR (125.8 MHz, D<sub>2</sub>O): δ = 181.0 (Lac-C1), 138.2, 129.4, 128.6, 126.7 (4C, Ar-C), 99.4 (Gal-C1), 98.9 (Fuc-C1), 84.1 (MeCy-C2), 82.7 (Gal-C3), 82.0 (Lac-C2), 78.6 (MeCy-C1), 74.2 (Gal-C2), 72.0 (Fuc-C2), 69.8 (Gal-C5), 69.3 (Fuc-C3), 68.2 (Fuc-C4), 66.5 (Fuc-C5), 66.2 (Gal-C4), 61.6 (Gal-C6), 39.3 (Lac-C3), 38.7 (MeCy-C3), 33.1 (MeCy-C4), 30.2 (MeCy-C6), 22.6 (MeCy-C5), 18.1 (MeCy-Me), 15.5 (Fuc-C6); HR-MS: *m/z*: Calcd for C<sub>28</sub>H<sub>42</sub>NaO<sub>13</sub><sup>+</sup> [M+H]<sup>+</sup>: 609.2518, found: 609.2521.

**(1R,2R,3S)-2-[( $\alpha$ -L-Fucopyranosyl)oxy]-3-methyl-cyclohex-1-yl 3-O-[sodium (*S*)-1-carboxy-2-(4-methylphenyl)-ethyl]- $\beta$ -D-galactopyranoside (**6b**)**

A suspension of (**23b**) (40 mg, 0.035 mmol) and Pd(OH)<sub>2</sub>/C (50 mg, 10% Pd) in MeOH (5 mL) was hydrogenated (2 bar H<sub>2</sub>) at rt. After 2 h TLC (silica: petroleum ether/EtOAc,

3:2; C-18: H<sub>2</sub>O/MeOH, 1:1) indicated completion of the reaction. The reaction mixture was filtered through celite and concentrated *in vacuo*. The residue was dissolved in MeOH/H<sub>2</sub>O (4:1, 1 mL) and treated with LiOH (41 mg, 1.73 mmol) for 32 h. The mixture was neutralized with Dowex 50X8 (H<sup>+</sup> form), filtered through a Dowex 50 (Na<sup>+</sup> form) ion exchange column and concentrated. The residue was purified by reversed phase column chromatography (C-18, H<sub>2</sub>O/MeOH, gradient 0-100%) followed by a Dowex 50 (Na<sup>+</sup> form) ion exchange column, a Sephadex G15 column, and lyophilization from water to yield the product (**6b**) as colorless foam (8.9 mg, 0.014 mmol, 41%).  $R_f$  = (C-18, H<sub>2</sub>O/MeOH, 1:1) 0.31;  $[\alpha]_D^{22}$  -76.6 (*c* 0.26, MeOH); <sup>1</sup>H NMR (500.1 MHz, CDCl<sub>3</sub>):  $\delta$  = 7.28 (d, *J* = 8.0 Hz, 2H, Ar-H), 7.23 (d, *J* = 8.0 Hz, 2H, Ar-H), 5.10 (d, *J* = 4.0 Hz, 1H, Fuc-H1), 4.86 – 4.80 (m, 1H, Fuc-H5), 4.39 (d, *J* = 8.0 Hz, 1H, Gal-H1), 4.15 (dd, *J* = 4.6, 8.9 Hz, 1H, Lac-H2), 3.92 – 3.86 (m, 2H, Gal-H4, Fuc-H3), 3.83 – 3.78 (m, 2H, Fuc-H2, Fuc-H4), 3.74 – 3.65 (m, 3H, MeCy-H1, Gal-H6), 3.56 – 3.49 (m, 2H, Gal-H2, Gal-H5), 3.28 (dd, *J* = 3.2, 9.3 Hz, 1H, Gal-H3), 3.22 (t, *J* = 9.6 Hz, 1H, MeCy-H2), 3.12 (dd, *J* = 4.5, 14.0 Hz, 1H, Lac-H3), 2.93 (dd, *J* = 8.9, 14.0 Hz, 1H, Lac-H3'), 2.33 (s, 3H, Ar-CH<sub>3</sub>), 2.13 – 2.05 (m, 1H, MeCy-H6), 1.70 – 1.53 (m, 3H, MeCy-H3, MeCy-H5), 1.30 – 1.21 (m, 2H, MeCy-H6', MeCy-H4), 1.18 (d, *J* = 6.6 Hz, 3H, Fuc-H6), 1.14 – 1.04 (m, 1H, MeCy-H4'), 1.01 (d, *J* = 6.3 Hz, 3H, MeCy-Me); <sup>13</sup>C NMR (125.8 MHz, D<sub>2</sub>O):  $\delta$  = 180.6 (Lac-C1), 136.7, 135.1, 129.4, 129.1 (4C, Ar-C), 99.4 (Gal-C1), 98.9 (Fuc-C1), 84.1 (MeCy-C2), 82.7 (Gal-C3), 82.1 (Lac-C2), 78.6 (MeCy-C1), 74.2 (Gal-C2), 72.0 (Fuc-C2), 69.8 (Gal-C5), 69.3 (Fuc-C3), 68.2 (Fuc-C4), 66.5 (Fuc-C5), 66.2 (Gal-C4), 61.6 (Gal-C6), 38.9 (Lac-C3), 38.8 (MeCy-C3), 33.2 (MeCy-C4), 30.2 (MeCy-C6), 22.6 (MeCy-C5), 20.1 (Ar-CH<sub>3</sub>), 18.1 (MeCy-Me), 15.5 (Fuc-C6); HR-MS: *m/z*: Calcd for C<sub>35</sub>H<sub>43</sub>F<sub>2</sub>Na<sub>2</sub>O<sub>14</sub>[M+Na]<sup>+</sup>: 645.2499, found: 645.2496.

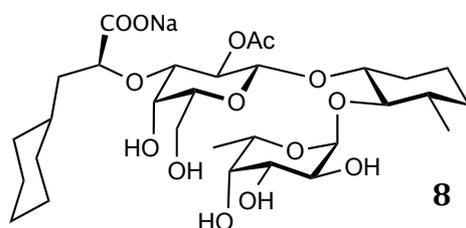
**(1R,2R,3S)-2-[( $\alpha$ -L-Fucopyranosyl)oxy]-3-methyl-cyclohex-1-yl 3-O-[sodium (S)-1-carboxy-2-(3,4-difluorophenyl)-ethyl]- $\beta$ -D-galactopyranoside (**6c**)**



A suspension of (**23c**) (12 mg, 0.010 mmol) and Pd(OH)<sub>2</sub>/C (50 mg, 10% Pd) in MeOH (5 mL) was hydrogenated (2 bar H<sub>2</sub>) at rt. After 3 h TLC (silica: petroleum ether/EtOAc, 3:2; C-18: H<sub>2</sub>O/MeOH, 1:1) indicated completion of the reaction. The reaction mixture was filtered through celite and concentrated *in vacuo*. The residue was dissolved in MeOH/H<sub>2</sub>O (4:1, 1 mL) and treated with LiOH (12 mg, 0.51 mmol) for 7 days. The mixture was neutralized with Dowex 50X8 (H<sup>+</sup> form), filtered through a Dowex 50 (Na<sup>+</sup> form) ion exchange column and concentrated. The residue was purified by reversed phase column chromatography (C-18, H<sub>2</sub>O/MeOH, gradient 0-100%) followed by a Sephadex G15 column, and lyophilization from water to yield the product (**6c**) as white foam (3.5 mg, 0.0054 mmol, 59%).  $R_f$  = (C-18, H<sub>2</sub>O/MeOH, 1:1) 0.25;  $[\alpha]_D^{22}$  -76.6 (*c* 0.26, MeOH); <sup>1</sup>H NMR (500.1 MHz, CDCl<sub>3</sub>):  $\delta$  = 7.33 – 7.25 (m, 1H, Ar-H), 7.24 – 7.15 (m, 1H, Ar-H), 7.15 – 7.03 (m, 1H, Ar-H), 5.11 (d, *J* = 4.0 Hz, 1H, Fuc-H1), 4.90 – 4.80 (m, 1H, Fuc-H5), 4.43 (d, *J* = 8.0 Hz, 1H, Gal-H1), 4.15 (dd, *J* = 4.8, 8.2 Hz, 1H, Lac-

H2), 3.93 (d,  $J = 2.9$  Hz, 1H, Gal-H4), 3.90 (dd,  $J = 3.2, 10.6$  Hz, 1H, Fuc-H3), 3.87 - 3.78 (m, 2H, Fuc-H2, Fuc-H4), 3.76 - 3.66 (m, 3H, MeCy-H1, Gal-H6), 3.59 - 3.53 (m, 2H, Gal-H2, Gal-H5), 3.33 (dd,  $J = 3.1, 9.5$  Hz, 1H, Gal-H3), 3.22 (t,  $J = 9.6$  Hz, 1H, MeCy-H2), 3.09 (dd,  $J = 4.6, 14.1$  Hz, 1H, Lac-H3), 2.97 (dd,  $J = 8.2, 14.1$  Hz, 1H, Lac-H3'), 2.15 - 2.08 (m, 1H, MeCy-H6), 1.72 - 1.54 (m, 3H, MeCy-H3, MeCy-H5), 1.33 - 1.22 (m, 2H, MeCy-H6', MeCy-H4), 1.19 (d,  $J = 6.6$  Hz, 3H, Fuc-H6), 1.10 (d,  $J = 6.3$  Hz, 3H, MeCy-Me), 1.14 - 1.04 (m, 1H, MeCy-H4');  $^{13}\text{C}$  NMR (125.8 MHz,  $\text{D}_2\text{O}$ ):  $\delta = 180.2$  (Lac-C1), 135.1 (m, Ar-C), 125.7 (m, Ar-C), 118.3 (d,  $J = 17.0$  Hz, Ar-C), 116.8 (d,  $J = 16.7$  Hz, Ar-C), 99.7 (Gal-C1), 98.9 (Fuc-C1), 84.2 (MeCy-C2), 82.4 (Gal-C3), 81.6 (Lac-C2), 78.7 (MeCy-C1), 74.1 (Gal-C2), 72.0 (Fuc-C2), 69.8 (Gal-C5), 69.3 (Fuc-C3), 68.2 (Fuc-C4), 66.5 (Fuc-C5), 66.2 (Gal-C4), 61.7 (Gal-C6), 38.7 (Lac-C3), 38.4 (MeCy-C3), 33.2 (MeCy-C4), 30.3 (MeCy-C6), 22.6 (MeCy-C5), 18.1 (MeCy-Me), 15.5 (Fuc-C6); HR-MS:  $m/z$ : Calcd for  $\text{C}_{28}\text{H}_{39}\text{F}_2\text{Na}_2\text{O}_{13}^+ [\text{M}+\text{H}]^+$ : 667.2149, found: 667.2150.

**(1R,2R,3S)-2-[( $\alpha$ -L-Fucopyranosyl)oxy]-3-methyl-cyclohex-1-yl 3-O-[sodium (S)-1-carboxy-2-(4-methylphenyl)-ethyl]- $\beta$ -D-galactopyranoside (**8**)**



A suspension of (**24**) (20 mg, 0.018 mmol) and  $\text{Pd}(\text{OH})_2/\text{C}$  (50 mg, 10% Pd) in THF (5 mL) was hydrogenated (2 bar  $\text{H}_2$ ) at rt. After 2 h TLC (silica: petroleum ether/EtOAc, 3:2; C-18:  $\text{H}_2\text{O}/\text{MeOH}$ , 1:1) indicated completion of the reaction. The reaction mixture was filtered through celite and concentrated *in vacuo*. The residue was dissolved in dioxane/ $\text{H}_2\text{O}/\text{AcOH}$  (4:2:1, 3.5 mL), treated with  $\text{Rh}/\text{Al}_2\text{O}_3$  (50 mg, 5% Rh) for 24 h. The reaction mixture was filtered through celite and concentrated *in vacuo*. The residue was purified by reversed phase column chromatography (C-18,  $\text{H}_2\text{O}/\text{MeOH}$ ) followed by a flash chromatography (silica, DCM/MeOH gradient 10-40%, MeOH containing 5%  $\text{H}_2\text{O}$ ). The product was treated with Dowex 50 ( $\text{Na}^+$  form) ion exchange resin, then filtered through a Sephadex G15 column and lyophilization from water to yield the product (**8**) as white foam (2.4 mg, 0.004 mmol, 20%).  $R_f = (\text{C-18}, \text{H}_2\text{O}/\text{MeOH}, 1:1)$  0.55;  $[\alpha]_{\text{D}}^{22} -76.6$  (c 0.26, MeOH);  $^1\text{H}$  NMR (500.1 MHz,  $\text{CDCl}_3$ ):  $\delta = 5.09$  (d,  $J = 4.0$  Hz, 1H, Fuc-H1), 5.02 (dd,  $J = 8.4, 9.5$  Hz, 1H, Gal-H2), 4.91 (q,  $J = 6.4$  Hz, 1H, Fuc-H5), 4.68 (d,  $J = 8.1$  Hz, 1H, Gal-H1), 3.97 (d,  $J = 2.7$  Hz, 1H, Gal-H4), 3.94 - 3.87 (m, 2H, Lac-H2, Fuc-H3), 3.86 - 3.79 (m, 2H, Fuc-H2, Fuc-H4), 3.78 - 3.74 (m, 2H, Gal-H6), 3.69 - 3.61 (m, 3H, MeCy-H1, Gal-H5, Gal-H3), 3.19 (t,  $J = 9.5$  Hz, 1H, MeCy-H2), 2.24 (s, 3H, Ar- $\text{CH}_3$ ), 2.17 - 2.04 (m, 1H, MeCy), 1.82 - 1.46 (m, 9H, Cy, MeCy, Lac-H3), 1.40 - 1.15 (m, 6H, Cy, MeCy), 1.25 (d,  $J = 6.2$  Hz, 3H, Fuc-H6), 1.13 - 1.02 (m, 1H, Cy or MeCy), 1.09 (d,  $J = 6.4$  Hz, 3H, MeCy-Me), 1.02 - 0.85 (m, 2H, Cy, MeCy);  $^{13}\text{C}$  NMR (125.8 MHz,  $\text{D}_2\text{O}$ ):  $\delta = 182.5$  (Lac-C1), 173.5 ( $\text{CH}_3\text{-CO}$ ), 98.8 (Fuc-C1), 98.5 (Gal-C1), 83.6 (MeCy-C2), 80.8 (Gal-C3), 79.8 (MeCy-C1), 79.1 (Lac-C2), 74.2 (Gal-C5), 72.0 (Fuc-C4, Gal-C2), 69.3 (Fuc-C3), 68.2 (Fuc-C2), 66.5 (Fuc-C5), 66.4 (Gal-C4), 61.7 (Gal-C6), 41.5 (Lac-C3), 38.7, 33.7, 33.4, 33.1, 31.6, 30.7, 26.0, 26.0, 25.7, 22.5 (Cy, MeCy), 20.9 (Ar- $\text{CH}_3$ ), 18.1 (MeCy-Me), 15.4 (Fuc-C6); HR-MS:  $m/z$ : Calcd for  $\text{C}_{30}\text{H}_{49}\text{Na}_2\text{O}_{14} [\text{M}+\text{Na}]^+$ : 679.2918, found: 679.2919

## References

1. Norman, K.E., Anderson, G.P., Kolb, H.C., Ley, K., & Ernst, B. (1998). Sialyl Lewis<sup>x</sup> (sLe<sup>x</sup>) and an sLe<sup>x</sup> Mimetic, CGP69669A, Disrupt E-Selectin-Dependent Leukocyte Rolling In Vivo. *Blood*, 91(2): 475-483.
2. Zihlmann, P., *et al.* (2015). Microscale Thermophoresis: A powerful opportunity for carbohydrate-based drug discovery. *Manuscript in preparation*.
3. Binder, F.P., Lemme, K., Preston, R.C., & Ernst, B. (2012). Sialyl Lewis<sup>x</sup>: A “Pre-Organized Water Oligomer”? *Angew. Chem. Int. Ed.*, 51(29): 7327-7331.
4. Fesik, S.W., Gampe, R.T., & Zuiderweg, E.R.P. (1989). Heteronuclear three-dimensional NMR spectroscopy. Natural abundance carbon-13 chemical shift editing of 1H-1H COSY spectra. *J. Am. Chem. Soc.*, 111(2): 770-772.
5. SPARKY 3. University of California, San Francisco, 2004.
6. Markley, J., *et al.* (1998). Recommendations for the presentation of NMR structures of proteins and nucleic acids. *Pure Appl. Chem.*, 70(1): 117-142.
7. Guntert, P., Mumenthaler, C., & Wuthrich, K. (1997). Torsion angle dynamics for NMR structure calculation with the new program DYANA. *J. Mol. Biol.*, 273(1): 283-298.
8. Case, D., *et al.* (2005). The Amber biomolecular simulation programs. *J. Comput. Chem.*, 26(16): 1668-1688.
9. Wang, J., Wolf, R.M., Caldwell, J.W., Kollman, P.A., & Case, D.A. (2004). Development and testing of a general amber force field. *J. Comput. Chem.*, 25(9): 1157-1174.
10. Bashford, D. & Case, D. (2000). Generalized born models of macromolecular solvation effects. *Annu. Rev. Phys. Chem.*, 51: 129-152.
11. Mesch, A., S., *et al.* (2010). Kinetic and thermodynamic properties of MAG antagonists. *Carbohydr. Res.*, 345(10): 1348-1359.
12. Bitsch, F., *et al.* (2003). Identification of natural ligands of retinoic acid receptor-related orphan receptor alpha ligand-binding domain expressed in Sf9 cells - a mass spectrometry approach. *Anal. Biochem.*, 323(1): 139-149.
13. Cooper, A. & Royal Society of Chemistry (Great Britain) (2004). *Biophysical Chemistry*. Royal Society of Chemistry, Cambridge. pp. 109-110.
14. ter Halle, R., *et al.* (2004). Development of a Practical Multikilogram Production of (R)-Seudenol by Enzymatic Resolution. *Org. Process Res. Dev.*, 8(2): 283-286.
15. Lönn, H. (1985). Synthesis of a tri- and a hepta-saccharide which contain  $\alpha$ -l-fucopyranosyl groups and are part of the complex type of carbohydrate moiety of glycoproteins. *Carbohydr. Res.*, 139(0): 105-113.
16. Schwizer, D., *et al.* (2012). Pre-organization of the core structure of E-selectin antagonists. *Chem. Eur. J.*, 18(5): 1342-1351.
17. Xu, Y., *et al.* (2005). Total synthesis of hirsutellide A. *Tetrahedron Lett.*, 46(25): 4377-4379.