Kinetic and Thermodynamic Characterization of the Bacterial Lectin FimH

Inauguraldissertation

zur

Erlangung der Würde eines Doktors der Philosophie

vorgelegt der

Philosophisch-Naturwissenschaftlichen Fakultät

der Universität Basel

von

Marleen Silbermann

aus Deutschland

Basel, 2021

Genehmigt von der Philosophisch-Naturwissenschaftlichen Fakultät
auf Antrag von

Fakultätsverantwortlicher:
Prof. em. Dr. Beat Ernst
Institut für Molekulare Pharmazie
Universität Basel

Korreferent:
Dr. Daniel Strasser
Biomarker Platform Translational Science – Drug Discovery Biology
Idorsia Pharmaceuticals Ltd., Allschwil

Basel, den 11. Dezember 2018

Prof. Dr. Martin Spiess
Dekan

Copyright waiver

Institute of Molecular Pharmacy

University of Basel, Klingelbergstrasse 50, CH-4056 Basel, Switzerland

Declaration

I hereby declare that this doctoral dissertation entitled "Kinetic and thermodynamic characterization of the bacterial lectin FimH" has been completed only with the assistance mentioned herein and that it has not been submitted for award to any other university nor to any other faculty at the University of Basel.

Marleen Silbermann, Basel, the 23st of November, 2018

“All things are difficult before they are easy”

(Dr. Thomas Fuller)

Acknowledgements

First and foremost, I would like to express my gratitude to Prof. Dr. Beat Ernst for giving me the chance to challenge myself as a scientist in his outstanding research group. I am greatly thankful to you for sharing your vast knowledge in drug discovery with me and particularly appreciated your thrilling enthusiasm and inspiring mindset. Thank you also for the time you invested to proofread my written PhD work. I wish you all the best for your future and many new exciting and refreshing projects.

I further would like to thank Dr. Daniel Strasser for accepting to be the co-referee of my thesis.

Another thanks goes to Dr. Timothy Sharpe from the Biophysics Facility for his great scientific support and assistance to overcome sophisticated biophysical issues.

Throughout my PhD time, I met many kind persons who made my years at the IMP an unforgettable experience. I thank all members of the biology group being always helpful and motivating… thank you Butrint, Philipp, Anja, and Jacqueline. A special "thank you" goes to Deniz for his instructions and help at the beginning of my PhD journey. After his company in the biology lab, Brigitte took his chair. I am very grateful for your scientific advice and friendship. Additionally, I thank Said for his time to acquaint me with protein purification/expression and to help me to navigate through cloning experiments. Many thanks also go to all the master students - Guilia, Simon, Nesrin, Kathrin, Yasmin, and Pascal - you made the biology lab a more vivid and enjoyable place. The same applies to my present lab mate Kevin.

I also thank the chemists Bea, Xiaohua, Giulio, Olly, Priska, Norbert, Blijke, Lijuan, Fan, Rachel, Hélène, Maja and Wojtek for their "chemistry help" and the colleges of "Molecular Modeling" group Christoph, Martin, Oya, Joël, Charleen, and Zhenquan for their scientific computer-based input. It was also an honor for me to have met Prof. Dr. Angelo Vedani… a very exceptional and inspiring person.

The PhD time was even more pleasant due to the skiing trips into the Alps, the "New Year`s dinners", and the cozy evenings we spent together during our "summer beer" events. Bea`s holiday tips and picture shows during the lunch break were also unique.

Furthermore, I appreciated Prof. Dr. Ricklin talent to transform the IMP group very smoothly and to create a unity of “old” and “new”. Thank you also for your ideas and help concerning SPR. I wish you and your growing group – Kevin, Clement, Richard, Christina – many scientific breakthroughs.

I owe a special thanks to my friends and family for their support, encouragement, friendship, and love. Especially, I would like to thank my dad, who always have a helping hand whenever needed, my mum for the encouraging phone calls, and my grandma for her continuous support throughout my life. I also thank my sister Laura for the healthy food she prepared for us now and then. The same also applies to Thomas, who sometimes brought me food when I was working late. The short but very colorful and adventures PhD break in France with my friend Romina as well belongs to my PhD memories. I am also grateful for Priska`s company and friendship throughout my IMP time. I enjoyed our sunny lunch breaks and the chatty coffee breaks we had with Maja and Hélène. I will never regret my decision to move to Basel also because of the city and the surrounding area, but the best thing about it is that I got to know my boyfriend Stefan.

Last of all, I want to thank Brigitte, Jonathan, Laura, and Stefan for proofreading of some of these pages.

To put it in a nutshell, I would like to thank all of you who supported me over the last years.

Abstract

One fundamental aim of drug discovery is the development of new molecular entities that have a considerably advantage over already existing therapies. Urinary tract infections (UTIs) urgently require an alternative to the conventional antibiotic therapy as resistance rates for antibiotics are increasing. The development of an anti-adhesive UTI treatment strategy with the bacterial lectin FimH as target is a promising approach to remedy such alarming tendencies. FimH is presented by uropathogenic *E. coli* (UPEC) strains on the tip of type 1 pili and mediates adhesion to mannosylated residues on the urothelium. This interaction prevents the clearance of UPECs during micturition and enables internalization of the pathogens by urothelial cells. Mannoside-derived FimH antagonists are under development and are considered as promising treatment option for UTIs. In contrast to antibiotics, FimH antagonists do not necessarily exert resistance mechanisms against drugs because they block the adhesion of bacteria to the urothelium without killing them or inhibiting their growth.

In the present thesis, FimH and its interaction with mannose-based antagonists were biophysically characterized. Additionally, new methodical approaches are introduced, which are relevant not only for a strategic development of FimH antagonists but also for drugs of other therapeutic areas. The following aspects were investigated:

- **Publication 2:** The publication "*KinITC – One method supports both thermodynamic and kinetic SARs*" (*Chemistry*, 2018, *24*(49), 13049-13057) comments on kinITC-ETC, a new method based on ITC data to reveal the kinetic fingerprint of a drug–target interaction. In this study, kinITC-ETC was independently validated for the first time. Moreover, structural properties of FimH antagonists could be correlated with kinetic parameters of FimH–antagonist interactions.

- **Manuscript 1:** The development of an off-rate screening approach is presented in the study *"Off-rate screening by surface plasmon resonance – The search for promising lead structures targeting low-affinity FimH"*. The method is subsequently applied to screen a mannose-based compound library against full-length FimH. The assay allows classification of structurally diverse FimH antagonist in order to spot chemical classes exhibiting long dissociative half-lives.

- **Publication 3:** The lectin domain is conformationally rigid and needs the pilin domain for allosteric propagation. However, the crosstalk between allosteric sites within the lectin domain takes also place in the absence of the pilin domain as demonstrated in the publication *"Conformational switch of the bacterial adhesin FimH in the absence of the regulatory domain – Engineering a minimalistic allosteric system"* (*J. Biol. Chem.*, 2018, *293*(5), 1835-1849). Mutants of the isolated lectin domain, $FimH_{LD}$ R60P and V27C/L34C, exhibited a low-affinity state and mimic full-length FimH regarding its conformational transition upon mannoside binding.

- **Publication 4:** The publication *"Target-directed dynamic combinatorial chemistry: A study on potentials and pitfalls as exemplified on a bacterial target"* (*Chemistry*, 2017, *23*, 11570-11577) illustrates a target-directed dynamic combinatorial chemistry (tdDCC) approach employing reversible acylhydrazone formation with FimH full-length as target. Optimal sample preparation and data procession are discussed in detail. Finally, the results of the tdDCC assay were subsequently compared with the affinity of library constituents by SPR.

- **Publication 5**: In the publication *"Comparison of affinity ranking by target-directed dynamic combinatorial chemistry and surface plasmon resonance"* larger FimH antagonist libraries were screened using the tdDCC method established in publication 3. The comparison of amplification rates of library substituents with respective binding affinities determined by SPR revealed a linear association. Furthermore, the hazardous acylhydrazone moiety could be replaced by various bioisosteres without changing the affinity of the parent compound.

- **Manuscript 2:** The hydrogen bond network formed between mannose derivates and the CRD of FimH is extensively elucidated in the manuscript "*High-affinity carbohydrate–lectin interaction: How nature makes it possible*". Computational methods and structural prediction in combination with binding data revealed that the hydrogen bond network forms a unified whole. The removal of only a single hydroxyl group leads to a disruption of the cooperative interplay within the network and consequently results in a dramatic loss in binding affinity.

- **Manuscript 3:** In the study "*The tyrosine gate of the bacterial adhesion FimH – An evolutionary remnant paves the way for drug discovery*", ITC measurements demonstrated the influence of the tyrosine gate on binding affinity between FimH and natural ligands. While the tyrosine gate is exploited to form optimal hydrophobic interactions with aryl aglycones of synthetic FimH antagonists in order to increase their binding affinity, the tyrosine gate has only a marginal impact on the K_D of natural ligands. In contrast to wild-type FimH, mutants that partially or completely lack the tyrosine gate exhibited a comparable binding affinity to dimannoside.

- **Publication 6:** The publication *"Improvement of aglycone π-stacking yields nanomolar to sub-nanomolar FimH antagonists"* displays that fluorination of biphenyl mannosides further improved π-π stacking with the tyrosine gate, reaching nanomolar affinities with $FimH_{FL}$ and even picomolar affinities with $FimH_{LD}$. It also could be shown that ligand binding to $FimH_{FL}$ occurs with a highly favorable enthalpic and a considerably unfavorable entropic contribution.

  Publication 7: In the publication "*Enhancing the enthalpic contribution of hydrogen bonds by solvent shielding*" microcalorimetric studies of FimH could reveal that conformational adaptions of the binding site can establish a solvent-free cavity. Shielding the solvent results in a lower dielectric environment, in which the formation of hydrogen bonds has a considerable enthalpic contribution to the binding free energy. In the case of FimH approximately -13 kJ mol^{-1} for mannoside binding.

Abbreviations

ΔC_p	change in heat capacity
ΔG	change in Gibbs free energy
$\Delta G°$	standard Gibbs free energy change
ΔH	enthalpy change
$\Delta H°$	standard enthalpy change
ΔT	temperature difference
ΔS	entropy change
$\Delta S°$	standard entropy change
ε_0	vacuum permittivity
ε_r	dielectric constant
τ	residence time
A	acceptor
ADME	absorption, distribution, metabolism, and excretion
atm	standard atmosphere (1 atm = 101.325 kPa)
AUC	analytical ultracentrifugation
BIA-MS	biomolecular interaction analysis mass spectrometry
CD	circular dichroism
CRD	carbohydrate recognition domain
D	donor
Da	dalton
DP	differential power
DsG	donor strand FimG
DsF	donor strand FimF
DSF	differential scanning fluorimetry
E	electrostatic energy
E. coli	*Escherichia coli*
EDC	1-ethyl-3-(3-dimethylaminopropyl)-carbodiimide
ESI	electrospray ionization
ETC	equilibration time curve
Fc	Fragment, crystallizable
$FimH_{LD}$	FimH lectin domain

$FimH_{PD}$	FimH pilin domain
FDA	Food and Drug Administration
GCPR	G-protein-coupled receptor
H–S	enthalpy–entropy
HBD	hydrogen bond donor
HBA	hydrogen bond acceptor
High	high-affinity state
HM	*n*-heptyl α-D- mannopyranoside
HSQC	heteronuclear single quantum correlation
HTS	high-throughput screening
IBC	intracellular bacterial communities
IFC	integrated μ-fluidic cartridge
ITC	isothermal titration calorimetry
Ig	immunoglobulin
IUPAC	International Union of Pure and Applied Chemistry
K_D	equilibrium dissociation constant
K_{eq}	equilibrium constant
KinITC	kinetic ITC
kinITC-ETC	kinetic ITC - Equilibration Time Curve
k_{on}	association rate constant
k_{off}	dissociation rate constant
L	ligand
log P	partition coefficient
Low	low-affinity state
MALDI-MS	matrix-assisted laser desorption/ionization mass spectrometry
Medium	medium-affinity
mRNA	messenger ribonucleic acid
MCK	multi-cycle kinetics
MW	molecular weight
MS	mass spectrometry
MST	microscale thermophoresis
MW	molecular weight
n	binding stoichiometry
NaOH	sodium hydroxide

NHS	N-hydroxysuccinimide
NMR	nuclear magnetic resonance
P20	polysorbate 20
PCR	polymerase chain reaction
PDB	Protein Data Bank
PID	proportional-integral-derivative
q_1	quantity of charge on object 1
q_2	quantity of charge on object 2
QIR	quiescent intracellular reservoirs
r	distance
R	gas constant
R&D	research and development
RI	refractive index
Ro5	Lipinski`s 'rule of five'
RU	response units
rUTI	recurrent UTI
SAR	structure–activity relationship
SCK	single-cycle kinetics
SEC-MS	size-exclusion chromatography mass spectroscopy
SPR	surface plasmon resonance
TIR	total internal reflection
T_M	melting temperature
TSA	thermal shift analysis
uHTS	ultra-high-throughput screening
UPEC	uropathogenic *Escherichia coli*
UPIa	uroplakin Ia
UTI	urinary tract infection
Y	analyte
Z	ligand

Table of Contents

I. Introduction

Section I
Introduction

1. Drug Discovery

1.1 The drug development process

When drug discovery was still in its infancy, drugs were discovered by a serendipity or identified based on ancient traditions or on their phenotypic effects.[1,2] Nowadays, *de novo* drug discovery mostly starts with the identification and validation of promising therapeutic targets. Target-based drug discovery has its origin in 1894 when Emil Fischer proposed his lock-and-key concept to explain enzyme specificity. In the early 1990s, the implementation of high-throughput screening (HTS) in research and development (R&D) revolutionized the search for active compounds. Henceforth, combinatorial chemistry libraries were subjected to automated screening against target proteins.[3] Because in general the overall screening attrition rate is usually extremely high, on average, one million compounds have to be screened for one licensed drug.[4] If the three-dimensional structure of the ligand binding site is available, virtual screening is an alternative approach to find promising drug-like molecules by docking ligands into the target binding site.[5] Primary screening hits are confirmed and characterized in secondary assays afterwards. In parallel, structure-activity relationship (SAR) analyses are performed to further optimize target selectivity and potency by chemical modifications. Before entering the clinical phase, candidate molecules have to survive preclinical *in vitro* and *in vivo* ADME (absorption, distribution, metabolism and excretion) and toxicology studies. Once a selected candidate enters clinical development the probability to reach the market is below 10% (Figure 1).[6,7] A failure at this stage has the highest financial consequence, especially in phase III confirmatory efficacy and safety trials, which recruit the largest number of participants and are often logistically complex.[8]

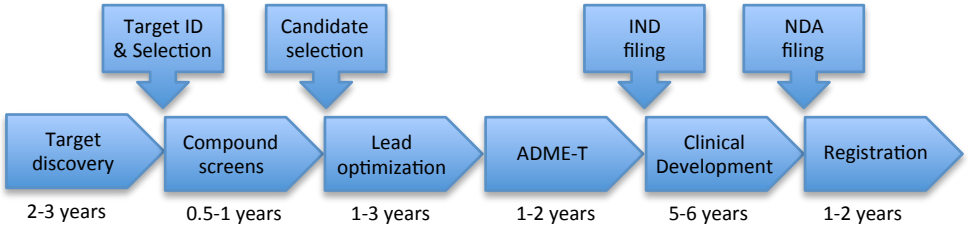


Figure 1: *De novo* drug discovery and development. The overall process usually amounts between 10-17 years. The figure is adapted from Reference [9]. ADME-T; absorption, distribution, metabolism, excretion, and toxicology; IND: Investigational New Drug; NDA; New Drug Application.

A better assessment of pharmacological and toxicological properties earlier in the drug development process enables an earlier rejection of inadequate compounds and thus an improvement in cost efficiency.

1.2 The drug-like chemical space

HTS is able to test between 10,000 to 100,000 compounds per day while ultra-high-throughput screening (uHTS) methods can even handle more than 100,000 small molecules per day.[10] These screening numbers seem negligible when compared to the drug-like chemical space, which has been estimated to be in the order of 10^{23} to 10^{60} organic molecules following **Lipinski`s 'rule of five'** (Ro5) for oral bioavailability.[11,12,13] Molecules complying with the Ro5 commonly exhibit favorable pharmacokinetic properties, like suitable solubility and permeability, which enable oral absorption and distribution.[14] The Ro5 proposes that at least two of the four following criteria have to be fulfilled:[15]

(1) number of hydrogen bond donor (HBD) atoms ≤ 5
(2) number of hydrogen bond acceptor (HBA) atoms ≤ 10
(3) partition coefficient (log P) ≤ 5
(4) molecular weight (MW) ≤ 500 Da

The partition coefficient (log P) is a logarithmic value of the distribution between water and an organic solvent. It is utilized to predict the lipophilicity/hydrophilicity of small molecules and defined as a ratio of concentrations of a unionized compound in the water and the *n*-octanol phase at equilibrium (Equation 1).[16,17]

$$log\ P = log\ \frac{[unionized\ compound]_{octanol}}{[unionized\ compound]_{water}} \qquad \text{(Equation 1)}$$

Medicinal chemists are constantly designing new structural classes of molecules to feed the drug development pipeline. Since only about 100 million drug-like compounds have been synthesized yet, a huge treasure of potential drug candidates is still available.[18]

1.3 Potential drug targets and their druggability

Most small molecule drugs currently on the market interact with protein targets to exert the desired pharmacological effect. However, for numerous protein targets the search for small molecule ligands is heavily complex or even failed. They are therefore named undruggable targets characterized by a strong hydrophilicity, a small or shallow binding shape, and/or the requirement of covalent binding.[19,20] In 2002, Hopkins and Groom introduced the concept of druggability that includes all proteins that bind Ro5 molecules with a K_D below 10 μM. At that time, the druggable genome was estimated to consist of more or less 3,000 genes of which about 600-1,500 were considered as potential drug targets.[21] However, over the time the condition of target druggability blurred and became more a contemporary assessment adapting to scientific knowledge and technical progress. Targets initially appearing to be undruggable became druggable with time, as for instance carbohydrate-binding lectins, e.g. involved in essential cellular recognition mechanisms. There original classification as undruggable targets is related to their shallow binding sites and pronounced polarity.[21] Nowadays, several promising lectin antagonists for the treatment of different disorders are in preclinical and clinical development. One of them is the pan-selectin antagonist rivipansel for the treatment of vaso-occlusive crisis in patients with sickle cell disease. It was developed in the group of Beat Ernst in collaboration with the company GlycoMimitics, licensed to Pfizer in 2011, and is currently in clinical trial phase 3.[22,23,24]

Recently, it was also assumed that not only as originally assumed 3,000 [21] but up to 10,000 [25] of the 20,000-25,000 protein-coding genes [26] in the human genome are related to diseases rendering them interesting for the pharmaceutical industry. However, this number is likely still a vast underestimation, as the number of disease-relevant proteins greatly excel the number of genes due to alternative splicing of precursor mRNAs, post-translational modifications, and the formation of heteromeric protein complexes, increasing the number of possible targets.[27]

In 2017, the overall number of FDA-approved protein targets of small molecule drugs amounted to 549.[28] Protein targets mainly belong to the families of GCPRs, kinases, ion channels, nuclear receptors, or transporters.[28,29] Further target families are proteases, epigenetic drugs, and proton pump inhibitors that are highlighted in the

following conference report of the 11th Swiss Course on Medicinal Chemistry in Leysin that took place in 2014 (**Publication 1**). The Swiss Course on Medicinal Chemistry provided an insight in the manifold aspects of medical chemistry and clarified the current status of research progress. In this context, approaches to increase the number of drug-like molecules are elucidated. Moreover, current tactical aspects and technologies in the drug development process are discussed.

References

1. Heatley, N.G., Alexander fleming. The man and the myth. *Med. Hist.*, **1984**, *28*(4), 453-455.
2. Ball, P., Synthetic organic chemistry in China: Building on an ancient tradition – an interview with Qi-Lin Zhou and Xiaoming Feng. *Natl. Sci. Rev.*, **2017**, *4*(3), 437-440.
3. Zhang, L. and Foxman, B., Molecular epidemiology of Escherichia coli mediated urinary tract infections. *Front. Biosci.*, **2003**, *8*, e235-244.
4. Carnero, A., High-throughput screening in drug discovery. *Clin. Transl. Oncol.*, **2006**, *8*(7), 482-490.
5. Cheng, T., Li, Q., Zhou, Z., Wang, Y., and Bryant, S.H., Structure-based virtual screening for drug discovery: A problem-centric review. *AAPS J.* , **2012**, *14*(1), 133-141.
6. Hartung, T., Food for thought look back in anger – what clinical studies tell us about preclinical work. *ALTEX*, **2013**, *30*(3), 275-291.
7. Arrowsmith, J., A decade of change. *Nat. Rev. Drug Discov.*, **2012**, *11*(1), 17-18.
8. Evans, S.R., Fundamentals of clinical trial design. *J. Exp. Stroke Transl. Med.*, **2010**, *3*(1), 19-27.
9. Ashburn, T.T. and Thor, K.B., Drug repositioning: Identifying and developing new uses for existing drugs. *Nat. Rev. Drug. Discov.*, **2004**, *3*(8), 673-683.
10. Mayr, L.M. and Fuerst, P., The future of high-throughput screening. *J. Biomol. Screen.*, **2008**, *13*(6), 443-448.
11. Bohacek, R.S., McMartin, C., and Guida, W.C., The art and practice of structure-based drug design: A molecular modeling perspective. *Med. Res. Rev.*, **1996**, *16*(1), 3-50.
12. Polishchuk, P.G., Madzhidov, T.I., and Varnek, A., Estimation of the size of drug-like chemical space based on GDB-17 data. *J. Comput. Aided Mol. Des.*, **2013**, *27*(8), 675-679.
13. Ertl, P., Cheminformatics analysis of organic substituents: Identification of the most common substituents, calculation of substituent properties, and automatic identification of drug-like bioisosteric groups. *J. Chem. Inform. Comput. Sci.*, **2003**, *43*(2), 374-380.
14. Reymond, J.-L. and Awale, M., Exploring chemical space for drug discovery using the chemical universe database. *ACS Chem. Neurosci.*, **2012**, *3*(9), 649-657.
15. Lipinski, C.A., Lombardo, F., Dominy, B.W., and Feeney, P.J., Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv. Drug Deliv. Rev.*, **1997**, *23*(1), 3-25.
16. Hillery, A.M., Drug delivery: The basic concepts, in *Drug delivery and targeting for pharmacists and pharmaceutical scientists*, **2001**, Taylor & Francis, 1-48.
17. Devalapally, H., Duan, Z., Seiden, M.V., and Amiji, M.M., Paclitaxel and ceramide co-administration in biodegradable polymeric nanoparticulate delivery system to overcome drug resistance in ovarian cancer. *Int. J. Cancer*, **2007**, *121*(8), 1830-1838.
18. Fink, T. and Reymond, J.L., Virtual exploration of the chemical universe up to 11 atoms of C, N, O, F: Assembly of 26.4 million structures (110.9 million stereoisomers) and analysis for new ring systems, stereochemistry, physicochemical properties, compound classes, and drug discovery. *J. Chem. Inf. Model.*, **2007**, *47*(2), 342-353.
19. Cheng, A.C., Coleman, R.G., Smyth, K.T., Cao, Q., Soulard, P., Caffrey, D.R., Salzberg, A.C., and Huang, E.S., Structure-based maximal affinity model predicts small molecule druggability. *Nat. Biotechnol.*, **2007**, *25*, 71.
20. Sakharkar, M.K., Sakharkar, K.R., and Pervaiz, S., Druggability of human disease genes. *Int. J. Biochem. Cell Biol.*, **2007**, *39*(6), 1156-1164.

21. Hopkins, A.L. and Groom, C.R., The druggable genome. *Nat. Rev. Drug Discov.*, **2002**, *1*(9), 727-730.
22. Telen, M.J., Beyond hydroxyurea: New and old drugs in the pipeline for sickle cell disease. *Blood*, **2016**, *127*(7), 810-819.
23. Ernst, B. and Magnani, J.L., From carbohydrate leads to glycomimetic drugs. *Nat. Rev. Drug Discov.*, **2009**, *8*, 661.
24. Chang, J., Patton, J.T., Sarkar, A., Ernst, B., Magnani, J.L., and Frenette, P.S., GMI-1070, a novel pan-selectin antagonist, reverses acute vascular occlusions in sickle cell mice. *Blood*, **2010**, *116*(10), 1779-1786.
25. Gonzaga-Jauregui, C., Lupski, J.R., and Gibbs, R.A., Human genome sequencing in health and disease. *Annu. Rev. Med.*, **2012**, *63*, 35-61.
26. Finishing the euchromatic sequence of the human genome. *Nature*, **2004**, *431*(7011), 931-945.
27. Kubinyi, H., Drug research: Myths, hype, and reality. *Nat. Rev. Drug Discov.*, **2003**, *2*(8), 665-668.
28. Santos, R., Ursu, O., Gaulton, A., Bento, A.P., Donadi, R.S., Bologa, C.G., Karlsson, A., Al-Lazikani, B., Hersey, A., Oprea, T.I., and Overington, J.P., A comprehensive map of molecular drug targets. *Nat. Rev. Drug Discov.*, **2017**, *16*(1), 19-34.
29. Tiikkainen, P. and Franke, L., Analysis of commercial and public bioactivity databases. *J. Chem. Inf. Model.*, **2012**, *52*(2), 319-326.

ChemPubSoc Europe
DOI: 10.1002/cmdc.201402543
CHEMMEDCHEM Conference Reports

Drug Discovery Summit: 11th Swiss Course on Medicinal Chemistry

Priska Frei,[a] Giulio Navarra,[a] Christoph P. Sager,*[a] Marleen Silbermann,[a] Norbert Varga,[a] and Eike-Christian Wamhoff[b]

Introduction

In the beautiful surroundings of the Swiss mountains and vineyards, more than 100 young scientists assembled with 33 speakers and instructors to attend the *11th Swiss Course on Medicinal Chemistry* (SCMC, October 12–17, Leysin, Switzerland).[1] The aim of the course is to attract and train young scientists working in the field of medicinal chemistry by expert speakers coming from both industry and academia, while providing a state-of-the-art scientific program focused on some of the hottest topics in life science research. The program highlighted various target families that were introduced by general overview talks, followed by well-selected case studies of recent and successful drug discovery projects that have led to development candidates or even marketed drugs. Furthermore, various tactical aspects and technologies, applied in modern drug discovery, were covered in several presentations and tutorials. Demonstrations of case studies offered detailed insight into the drug discovery and development process, all the way from target identification through the lead-finding and optimization phase into clinical trials. This report gives a comprehensive account of the topics covered at this year's conference and highlights transformative contributions and approaches.

Chemical Space

Several talks addressed strategies to access new chemotypes for drug discovery. Natural products are a particularly intriguing class of compounds, as they are evolutionarily optimized to interact with a number of biological targets. It has been shown that a wide range of biological effects might be achieved by only small changes to their structure. Inspired by this concept, Karl-Heinz Altmann (ETH Zürich, Switzerland) and Bart DeCorte (Entura, Johnson & Johnson, USA) presented approaches to exploit the chemical space of natural products for the discovery of new drugs. Other contributions focused on peptidomimetics as highly selective and active candidates. Norbert Sewald (Bielefeld University, Germany) presented strategies to overcome metabolic stability issues and highlighted their potential to target protein–protein interactions (PPI). A lecture about macrocyclic peptidomimetics (D. Obrecht, Polyphor, Switzerland) further underscored the versatility of this class of molecule (see *Target Families* below). The use of glycomimetics as inhibitors of lectins was addressed by Beat Ernst (University of Basel), who demonstrated the potential of carbohydrate-derived selectin antagonists for the treatment of vaso-occlusive crisis patients of sickle cell anemia (Figure 1).

Target Families

This year's discussion focused on six major target families of considerable pharmaceutical interest: ion channels (R. Owen, Pfizer Neusentis, UK), proteases (R. Sedrani, Novartis Pharma AG, Switzerland), kinases (G. Müller, Mercachem, Netherlands), epigenetic drug targets (P. Brennan, University of Oxford, UK), PPIs (D. Obrecht, Polyphor, Switzerland), and G-protein-coupled receptors (GPCRs; J. Mason, Heptares Therapeutics, UK). Moving from broad concepts to specific in-depth observations, a series of case studies was also presented (Figure 2). Ivacaftor, a new treatment for cystic fibrosis (see *Highlights* below), was introduced by Peter Grootenhuis (Vertex Pharmaceuticals, USA). As an important example of protease drug discovery, Sven Ruf (Sanofi-Aventis, Germany) presented the case of cathepsin A inhibitors as a novel treatment for cardiovascular diseases. Ray Finlay (AstraZeneca, UK) briefly introduced non-small-cell lung cancer and the challenges associated with resistance (kinase mutants), before he presented AZD9291 as a new mutant-selective inhibitor of the epidermal growth factor receptor (EGFR). DNA deregulation is emerging as an important causative agent for cancer and a wide variety of other diseases, which does not necessarily involve mutations in the DNA sequence. It is the epigenetic control that governs gene expression, which is regulated by post-translational modifications of histone proteins. The development of the methyltransferase inhibitor EPZ-6438 as an epigenetic drug against the EZH2 enzyme was extensively described by Richard Chesworth (Epizyme, USA). Among the several projects in Polyphor's pipeline, Daniel Obrecht presented the PPI inhibitor POL7080 as an innovative drug against *Pseudomonas* spp.

Tactical Aspects

Another thematic priority was set on strategies and approaches for efficient drug discovery. In this context, the po-

[a] *P. Frei,[+] G. Navarra,[+] C. P. Sager,[+] M. Silbermann,[+] Dr. N. Varga[+]*
University of Basel, Klingelbergstrasse 50, 4056 Basel (Switzerland)
E-mail: christoph.sager@unibas.ch
[b] *E.-C. Wamhoff[+]*
Max Planck Institute of Colloids and Interfaces
Wissenschaftspark Potsdam-Golm
Am Mühleberg 1 OT Golm, 14476 Potsdam (Germany)
[+] *These authors contributed equally to this work.*

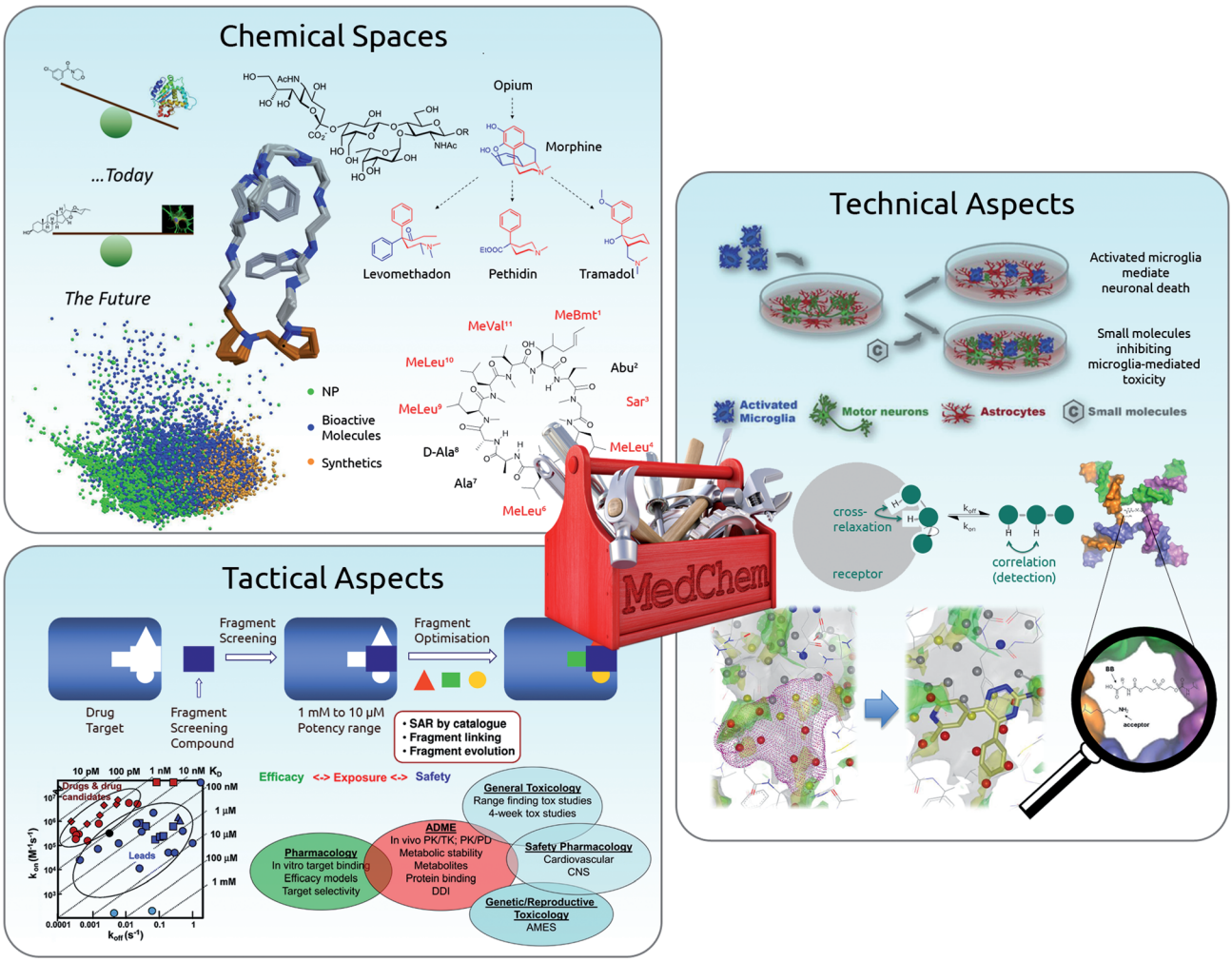


Figure 1. Schematic overview of the main groups of topics covered. Top left: chemical space (natural products, peptidomimetics, glycomimetics); right: technical aspects (phenotypic screening, NMR techniques, DNA-encoded libraries, water networks); bottom left: tactical aspects (fragment-based lead generation, binding kinetics, ADME).

tential of the bioisostere concept (N. Meanwell, Bristol-Myers Squibb, USA) was discussed as well as the influence of binding kinetics on the pharmacological properties of a candidate (D. Swinney, iRND3, USA). Additionally, the audience was introduced to ADME determinants and their crucial role during the optimization of drug candidates (C. Funk, Roche, Switzerland). Subsequently, the impact of metabolic processes on drugs was elucidated in detail (H. Kubinyi, University of Heidelberg, Germany). Polypharmacological (J.-U. Peters, Roche, Switzerland) and covalent drugs (J. Singh, Celgene Avilomics Research, USA; see *Highlights*) were illustrated by various examples as an alternative to conventional agents. Furthermore, Richard Morphy (Lilly Research Centre, UK) elaborated on the challenges and opportunities in targeting the central nervous system (CNS), which comes with specific restrictions on the molecular property space. Emphasis was also placed on tactical aspects in lead finding and optimization (A. Mortlock, AstraZeneca, UK), which was deepened further into two interactive tutorials (A. Mortlock, AstraZeneca, UK; L. van Berkom and S. Gremmen, both Mercachem, Netherlands). In this way, participants were able to apply their newly acquired knowledge and had the opportunity to discuss specific issues extensively with experts in the field. Hands-on practice was offered for fragment-based lead generation (S. Courtney and M. Mazanetz, both Evotec, UK) and patents in drug discovery (F. Schager, Actelion, Switzerland).

Technologies

In the field of screening technologies, focus was laid on recent developments in hit series generation and toward the application of novel phenotypic screening methods, which were also widely applied in the presented case studies.

The development of DNA-encoded libraries, even for hard-to-drug targets such as PPIs, was addressed by Nils Hansen (Vipergen, Denmark). He showed that the virtues of this combinatorial method are its high success rates and general applicability to soluble proteins. Dirk Ullmann (Evotec AG, Germany) commented on the importance of phenotypic screening for first-in-class small-molecule drugs over the last decade and gave an outlook on complex cellular models that can only be accessed by high-content screening. A three-dimensional in-

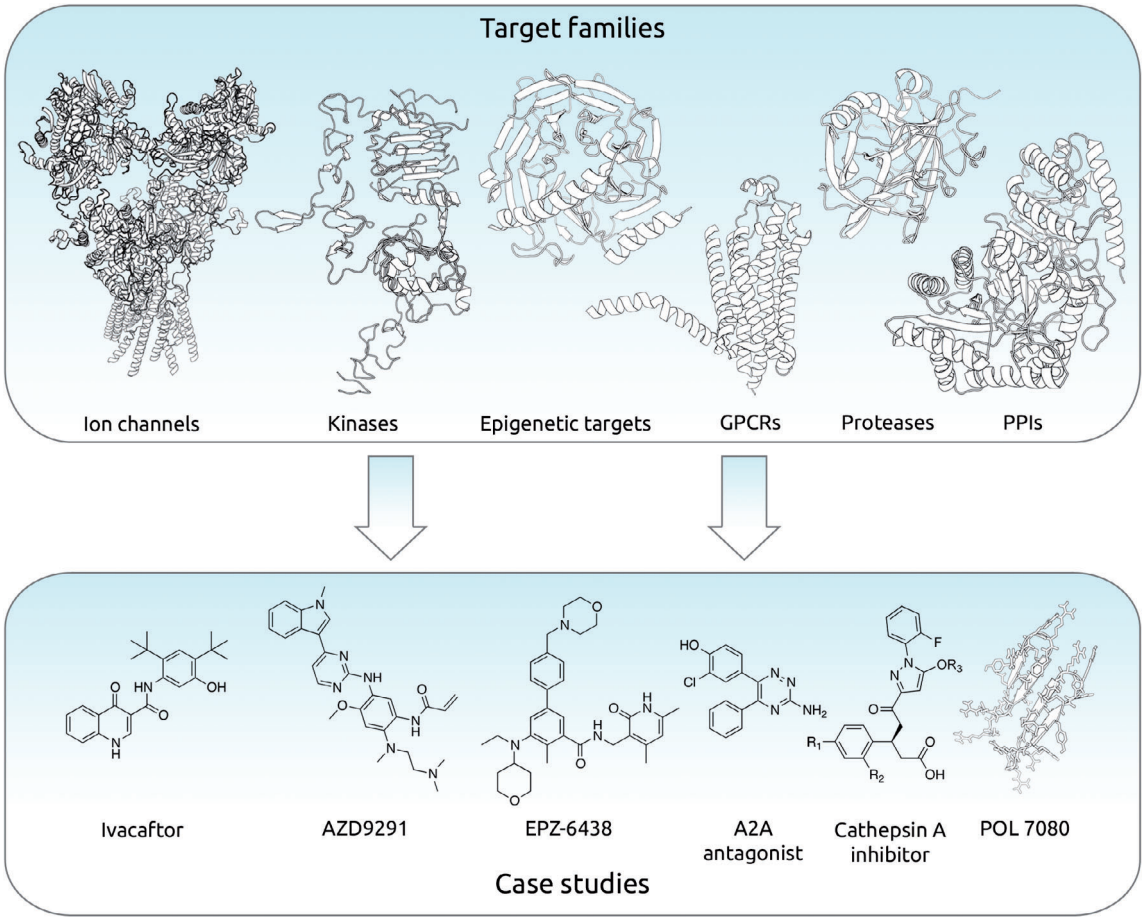


Figure 2. Examples from the target families mentioned in the lectures, and corresponding inhibitors of (left to right): sodium ion channel Na_v, EGFR, EZH2, A2A GPCR, cathepsin A, and two proteins involved in a PPI.

sight into the fundamental role of water molecules in protein–ligand interactions was given by Jonathan Mason (Heptares Therapeutics, UK). In various examples of stabilized GPCRs (StaRs®), he showed how elucidation of water molecules during the binding process improves the quality of structure-based drug design. Mark Murcko (Disruptive Biomedical, USA) presented recent advancements in force fields and free-energy perturbation (FEP; see *Highlights*). Participants could also take part in an interactive NMR tutorial lead by Christoph Rademacher (Max Planck Institute of Colloids and Interfaces, Berlin) and Anders Friberg (Bayer Healthcare Pharmaceuticals, Germany).

Highlights

A small panel of presentations was selected for more detailed accounts. This subjective selection reflects the particular scientific background and preferences of the authors.

Discovery of the CF drug ivacaftor

On the first day of the SCMC, Peter Grootenhuis (Vertex Pharmaceuticals, USA) presented the decade-spanning story of the recently approved cystic fibrosis (CF) drug, ivacaftor. CF patients suffer from decreased lung function, frequent lung infections, and pancreatic dysfunction among other symptoms. These symptoms are caused by defects in the CF transmembrane conductance regulator (CFTR) gene responsible for chloride transport. Various mechanisms of CFTR dysfunction have been proposed, and different modes of action are pursued by Vertex Pharmaceuticals accordingly. The discovery process for ivacaftor started from a high-throughput screen followed by extensive medicinal chemistry and structure–activity relationship studies to advance the initial hit into a clinical candidate. The talk emphasized the in vitro model developed in-house, which is based on cultured human bronchial epithelial cells, and was employed to monitor the activity of potential drug candidates. Ivacaftor works as a potentiator of CFTR function and was approved for the treatment of patients with gene defects leading to defective CFTR gating at the cell surface (Figure 3). At the end of his lecture, Grootenhuis left the audience with an inspiring success story demonstrating the impact of medicinal chemistry efforts on challenging targets. This stimulated many fruitful discussions in the further course of the conference.

Pushing the limits of FEP calculations

Another highlight of this year's program was Mark Murcko's (Disruptive Biomedical, USA) presentation on selected aspects

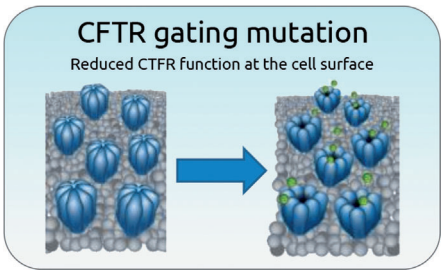


Figure 3. Closed (left) and open (right) CFTR chloride channels at the cell surface.

of computational chemistry. Starting from case studies, Murcko went on to discuss the opportunities structure-based drug design (SBDD) can offer, while also mentioning its shortcomings. He highlighted recent work aimed at overcoming some of the limitations of SBDD by advancing the free-energy perturbation (FEP) method. FEP provides a means to accurately predict affinity, selectivity, and pharmacokinetic properties based on structural data (Figure 4).

First proposed by Robert Zwanzig in 1954, the applicability of FEP to SBDD has been limited by issues with usability, accuracy, and throughput. Recent technological progress in the areas of graphic processors, sampling algorithms, and force-field parameterization has significantly pushed the boundaries in this field. Consequently, the latest developments have been able to show great predictive power for their FEP implementation in both retrospective and predictive settings. The implementation comprises a substantial set of ligand structure perturbations and is accessible for non-experts. Therefore, FEP can be envisioned to accelerate SBDD in the near future.

Resurgence of covalent drugs

Irreversible binders (i.e., covalent inhibitors) are generally considered harmful, as they can form covalent bonds with several biological molecules in vivo, thus leading to all kinds of unwanted side effects. Yet, in many cases reactive drug metabolites, rather than the original drug, are responsible for such bonds. Moreover, targeted covalent inhibitors with exceptional potency and selectivity as well as prolonged pharmacokinetic profiles have been discovered. Juswinder Singh (formerly with Celgene Avilomics Research, USA) presented an elegant technological platform for the rational design of covalent drugs. The approach uses structural data of known reversible inhibi-

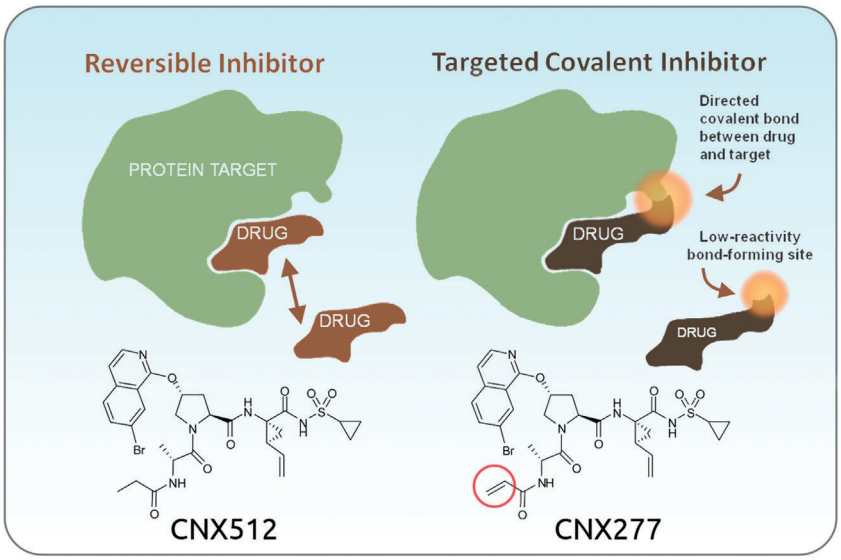


Figure 5. Schematic representation of the Avilomics approach. Structural information of the protein and its known binder (CNX512) is used to selectively introduce a reactive moiety on the latter, producing a new covalent drug (CNX277) that can form an irreversible bond to its target.

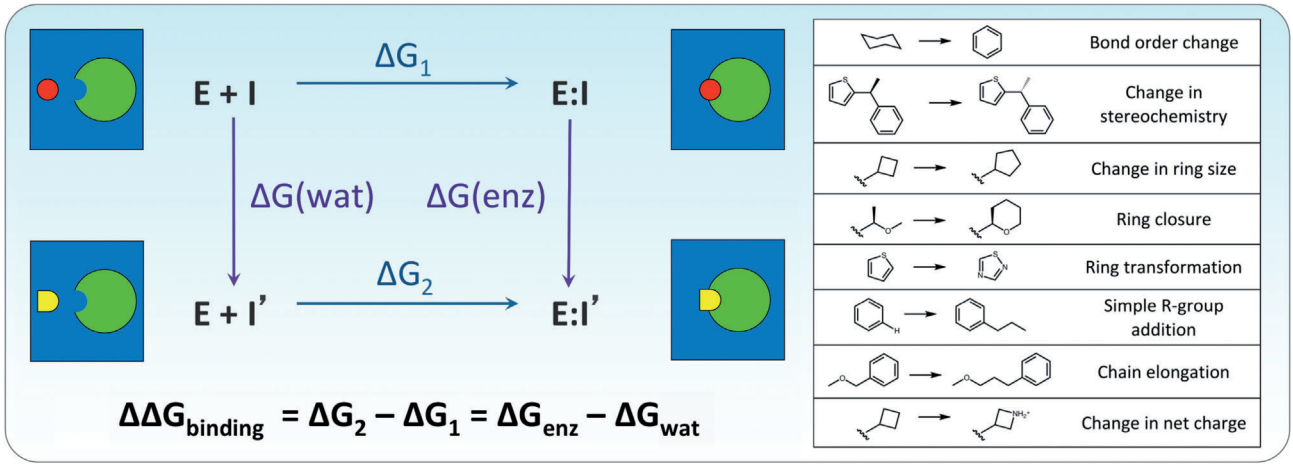


Figure 4. The principle of free-energy perturbation (FEP). Left: scheme of FEP calculations; right: examples of chemotype perturbations.

tors to install low-reactive functional groups, such as hindered Michael acceptors, on high-affinity molecules (Figure 5). Avilomics' approach has been applied to various target classes, including hepatitis C virus protease and EGFR and has produced several clinical candidates. Notably, targeted covalent inhibitors display sustained efficacy in cases of mutational resistance as observed for antiviral therapy, further validating their therapeutic potential.

Concluding Remarks

Besides the excellent talks and tutorials, the 11th SCMC provided a good opportunity for young scientists to approach the speakers and instructors during the breaks and social events. The course was filled with fruitful discussions as well as active networking between the participants. The presence of numerous trendsetters and opinion leaders in the field created a truly inspiring atmosphere and set the quality of the event at a high level. As young medicinal chemists from academia, we emphasize the uniqueness of this opportunity that allowed us to directly interact with world-class experts in biomedical research and to gain authentic insight into industrial drug discovery and development.

Outlook 2016

As we learned from the organizers, Gerhard Müller (Mercachem, Netherlands) and Beat Ernst (University of Basel), the tradition of organizing this Medicinal Chemistry Academy will be extended beyond this year's event, which marked a 20-year anniversary; the first SCMC was held in 1994 in Leysin. As this conference has developed its own distinct trademark over the last 20 years as "*the Leysin course*", it will be held again in this scenic setting in October 2016. Over the past two decades, this event has covered many technological and thematic changes in our industry at the highest possible level, and hence we can be confident that the 12th Swiss Course on Medicinal Chemistry will live up to the highest expectations. The international medicinal chemistry community can confidently look forward to learning more about the next event.

At this point, we cordially thank this year's organizing committee: first of all, Yvonne Baggerman (Mercachem, Netherlands), who did a marvelous job as the event assistant, finding a solution for each and every problem that surfaced during the conference; also Beat Ernst and Gerhard Müller, who once again put together a program based on state-of-the-art expert talks from leading trendsetters, all the while accounting for the requisite educational effort for young colleagues in the field.

[1] *11th Swiss Course on Medicinal Chemistry*, Swiss Chemical Society, Division of Medicinal Chemistry & Chemical Biology: scg.ch/leysin2014.

Received: December 17, 2014
Published online on January 14, 2015

3. The Bacterial Lectin FimH

3.1 Urinary tract infection

Urinary tract infections (UTIs) are the second most often diagnosed type of infection worldwide that affects any part of the urinary tract.[1] A distinction is made between uncomplicated and complicated UTIs: Uncomplicated UTIs are predominant and typically occur in otherwise healthy individuals, while complicated UTIs are often related to functionally or structurally abnormal urinary tracts, including pregnant women, patients with kidney transplantations, catheters, or metabolic diseases.[2,3,4] Uncomplicated UTIs are characterized by episodes of acute cystitis and pyelonephritis and in approximately 80% of the cases are caused by a heterogeneous group of uropathogenic *E. coli* (UPEC) strains.[4,5,6] The infection rate is significantly higher in women than in men. Around 50-60% of women experience at least one symptomatic UTI during their lifetime.[7,8] The health condition of these affected women is further aggravated due to a high incidence of recurrent UTIs (rUTIs).[9] Besides their effects on health issues, UTIs additionally impose a great economic burden, also due to the high recurrence rate.[10] The standard therapy for UTI is a short-term administration of antibiotics to treat symptoms and to prevent the infection from ascending to the upper urinary tract, where it can progress to a life threatening pyelonephritis or urosepsis.[5] Unfortunately, frequent antimicrobial treatment is followed by an increased resistance rate of uropathogens to antibiotics. Further risk factors that additionally promote antimicrobial resistances include rUTIs, hospitalization, and renal transplantation.[11] In order to prevent the development of bacterial resistance, alternative therapies are required.[12,13]

3.2 Acute infection cycle of uropathogenic *Escherichia coli*

The gram-negative and facultative anaerobic bacterium *E. coli* primarily inhabits the gastrointestinal tract where it lives in symbiosis with its host.[14] Although UPEC strains are commensals in the gut, they exhibit a pathogenic behavior once they reach the urinary tract. In order to invade and colonize the urothelium, UPECs use a set of virulence factors such as adhesins.[15,16]

In the first step of the acute infection cycle (Figure 1) UPECs adhere to the urothelium. This step prevents the clearance of UPECs during micturition and enables internalization of the pathogens by urothelial cells.[17] It is mediated by the mannose-specific bacterial lectin FimH localized at the tip of type 1 pili. FimH binds to mannose residues of high-mannose N-linked glycans of the transmembrane glycoprotein uroplakin Ia (UPIa) that is expressed on urothelial facet cells.[17,18] Upon FimH-mediated internalization, UPECs replicate and are able to form biofilm-like intracellular bacterial communities (IBCs). IBCs are enclosed by a proteinaceous polysaccharide matrix that protects the bacteria from antibiotics and the host's innate immune response.[19,20] At the early stage of IBC formation, UPECs are rod-shaped cells in loosely organized colonies that mature into a tightly packed cluster of coccoid cells.[20] In the matured IBC, UPECs revert into their rod-like morphology or transition into long filaments. Both subpopulations flux from infected cells into the bladder lumen and may re-infect adjacent differentiated superficial facet cells.[21] An UPEC infection can also lead to an activation of the apoptotic machinery resulting in cell exfoliation. This leads to a clearance of adherent and intracellular bacteria, but also enables the bacteria to establish quiescent intracellular reservoirs (QIRs) in underlying cell layers.[22]

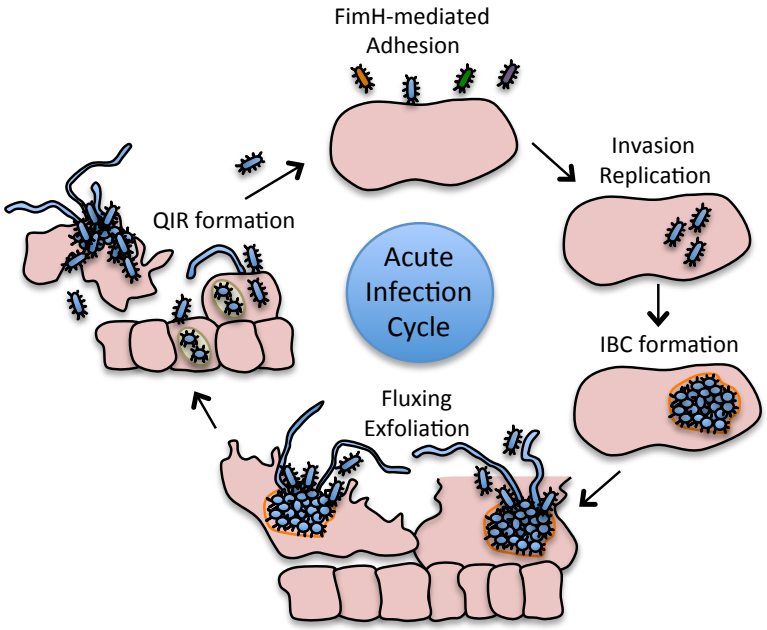


Figure 1: Acute infection cycle of uropathogenic *E. coli* (UPEC). Adhesion and internalization is achieved by FimH on the tip of type I pili interacting with oligomannosides presented on urothelial cells. UPECs replicate and form intracellular bacterial communities (IBCs). After IBCs are matured, the infected urothelial cell may exfoliate and UPECs flux into the bladder infecting other cells or building quiescent intracellular reservoirs (QIRs). The figure is modified from References [23,24].

Upon epithelial turnover, latent bacteria can provoke relapse.[25] Besides FimH-mediated adhesion and host defense avoidance mechanisms, toxins, and iron acquisition systems are further virulence factors of UPECs.[26]

3.3 Structure of type I pili and FimH

A single *E. coli* cell expresses between 100-500 peritrichously arranged type 1 pili, which enable the attachment to UPIa (Figure 2-A & 2-B). Type 1 pili are hetero-oligomeric mannose-binding fibers of about 7 nm in diameter, up to 2 µm in length,[27,28] and are assembled via the chaperone-usher pathway.[29] They are made of homologous proteins encoded by a set of *fim* genes located in the *fim* operon.[30] The pilus is divided into a fibrillar tip element (FimF, FimG, and FimH) and a rod-shaped main element (FimA subunits) (Figure 2-B). All proteins possess an incomplete, immunoglobulin (Ig)-like fold and interact via donor strand complementation, i.e. an N-terminal donor strand of each protein completes the *β*-sandwich structure of the following subunit.[31] The helical pilus rod is formed by 500 to 3000 copies of FimA and is anchored to the assembly platform FimD in the outer *E. coli* membrane. The fibrillar tip is composed of a single or multiple copies of FimF and FimG, which are connected to the distal end of the pilus rod, and a single copy of FimH at the pilus tip.[32]

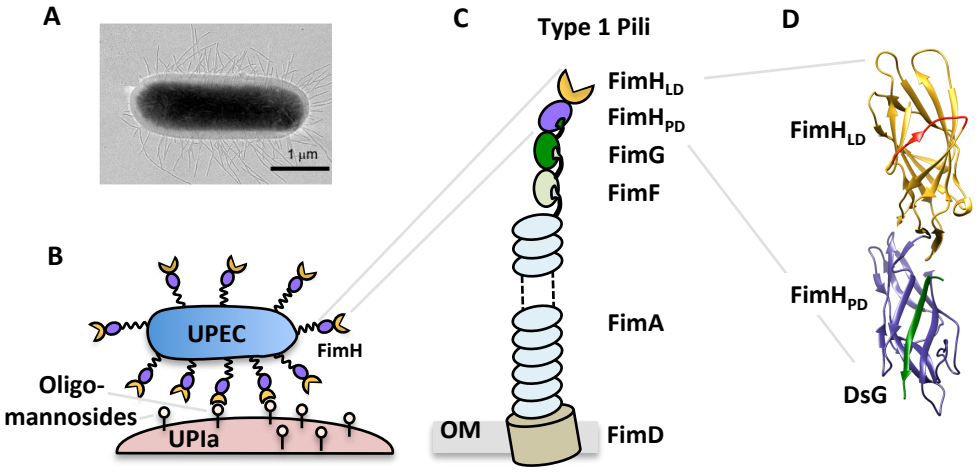


Figure 2: Structural elements of UPECs. A) Electron micrograph of a type I-fimbriated uropathogenic *E. coli* (UPEC) cell.[33] B) Schematic representation (modified from[34,35]) of an UPEC binding to oligomannosylated UPIa. C) Illustration of a type I pilus consisting of a helical rod (FimA subunits) and a fibrillar tip (FimF, FimG, and FimH). D) Crystal structure of full-length FimH consisting of a lectin domain ($FimH_{LD}$) and a pilin domain ($FimH_{PD}$), which is complemented with the N-terminal donor peptide of FimG (PDB ID: 4XOD).[36]

FimH is a 29 kDa protein consisting of 279 amino acids consisting of two domains; the N-terminal FimH lectin domain (Fim_{LD}, residues 1-156) containing the mannose-binding site, and the C-terminal FimH pilin domain ($FimH_{PD}$, residues 159-279) connected by a three-amino acid linker (Figure 2-C).[31] $FimH_{PD}$ anchors the protein to FimG whereas $FimH_{LD}$ acts as a mannose-specific carbohydrate recognition domain (CRD). Both domains predominately comprise β-sheets forming β-sandwich folds. Isolated FimH is unstable and requires a 14-amino acid donor strand from FimG for stabilization (ADVTITVNGKVVAK).[31,36] In experimental studies, a 15-amino acid synthetic peptide donor strand derived from FimG with an additional arginine at the C-terminus to improve solubility (DsG) is used to stabilize recombinant FimH (FimH*DsG or full-length FimH, $FimH_{FL}$).[36,37,38] FimH is also stable when isolated in complex with the donor strand-donating chaperone FimC.[39] In its isolated form, the lectin domain $FimH_{LD}$ can be stably expressed as well.[40]

3.4 Conformations and binding behavior of FimH

$FimH_{FL}$ exhibits a sophisticated allosteric mechanism to modulate its affinity for mannosides by conformational regulation. Recent work of Sauer *et al.* demonstrated that FimH*DsG mainly adopts three distinct conformational states: the low-affinity state (unbound $FimH_{FL}$, **Low**), the medium-affinity state (mannoside-bound $FimH_{FL}$ under static conditions, **Medium**), and the high-affinity state (mannoside-bound $FimH_{FL}$ under shear force, **High**) (Figure 3-A).[36] Through the conformational flexibility of $FimH_{FL}$, UPECs are able to regulate their binding affinity to mannose as an evolutionary adaption. A weak interaction with host cells is advantageous for bacterial motility to colonize the urinary tract for nutrient acquisition whereas strong adherence avoids clearance from the bladder by urination.[38]

The crystal structure of a ligand-unbound fimbrial tip (PDB: 3JWN) revealed a compressed FimH lectin domain that harbors a shallow binding site.[41] The same structure was observed for ligand-free $FimH_{FL}$ (Figure 3-A, **Low**).[36] In this conformation, $FimH_{LD}$ and $FimH_{PD}$ form a hook-shaped structure.[41,42] The lectin domain interacts with the pilin domain with three loop segments: the swing loop (residues 27-33), the linker loop (residues 154-160), and the insertion loop (residues 112-118).

Upon ligand binding and under static conditions, the clamp loop moves towards the ligand resulting in a deep and well-defined binding pocket (Figure 3-A, **Medium**). Apart from the change in the CRD, the overall conformation of the medium-affinity state is very similar to the unbound low-affinity conformation, maintaining the interaction between the two domains. Dissociation of $FimH_{LD}$ from UPIa is promoted via dynamic allostery by $FimH_{PD}$, which functions as a negative allosteric regulator.[36] In the medium-affinity state, the binding affinity of $FimH_{FL}$ to mannose is low. In the absence of urine flow, fast binding and release from UPIa enables fast reversible attachment and bacterial expansion in the urinary tract.[36,41,43] However, under shear stress, the binding affinity for mannose is dramatically increased due to the catch bond behavior of FimH, first discovered in 2002 (Figure 3-B).[44] In general, many adhesive proteins form catch bonds that are induced by mechanical force.[45,46] In case of FimH, the separation of $FimH_{LD}$ and $FimH_{PD}$ is induced by tensil forces during micturition. Nowadays, it is known that the consequent disruption of the allosteric interplay between the domains results in an up to 300-fold increase of the interaction strength of FimH with UPIa (Figure 3-A, **High**).[42]

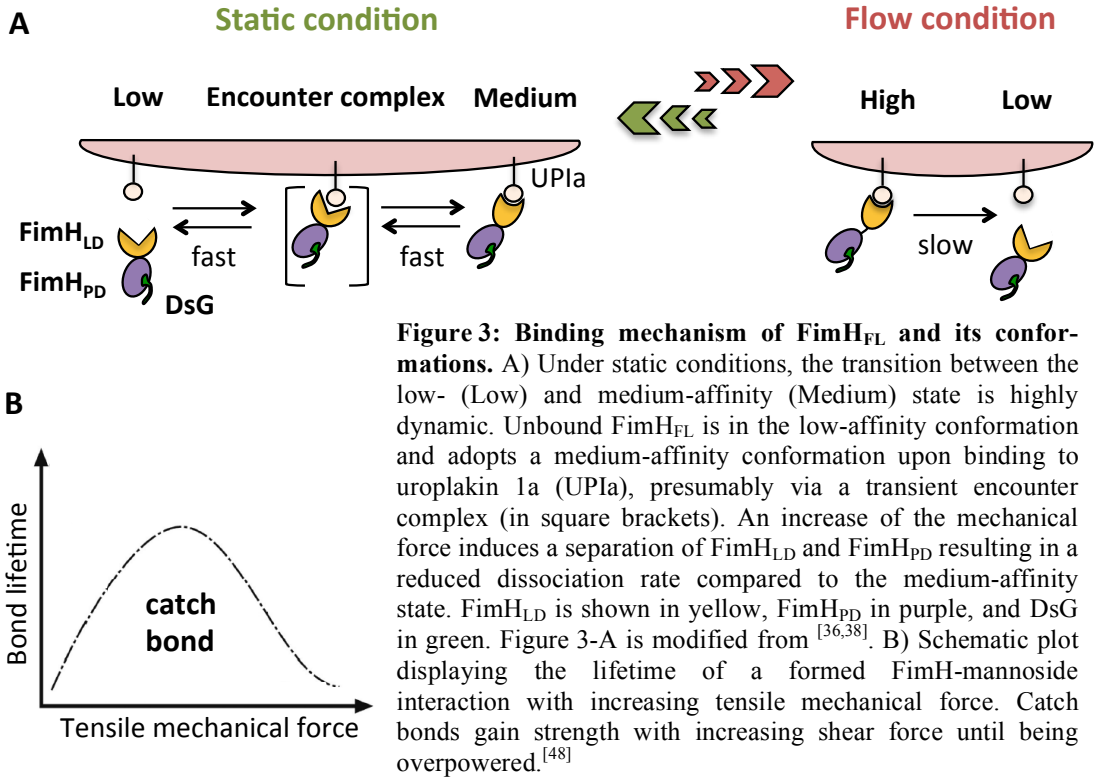


Figure 3: Binding mechanism of $FimH_{FL}$ and its conformations. A) Under static conditions, the transition between the low- (Low) and medium-affinity (Medium) state is highly dynamic. Unbound $FimH_{FL}$ is in the low-affinity conformation and adopts a medium-affinity conformation upon binding to uroplakin 1a (UPIa), presumably via a transient encounter complex (in square brackets). An increase of the mechanical force induces a separation of $FimH_{LD}$ and $FimH_{PD}$ resulting in a reduced dissociation rate compared to the medium-affinity state. $FimH_{LD}$ is shown in yellow, $FimH_{PD}$ in purple, and DsG in green. Figure 3-A is modified from [36,38]. B) Schematic plot displaying the lifetime of a formed FimH-mannoside interaction with increasing tensile mechanical force. Catch bonds gain strength with increasing shear force until being overpowered.[48]

Nevertheless, in our study "*Conformational switch of the bacterial adhesin FimH in the absence of the regulatory domain – Engineering a minimalistic allosteric system*"

(**publication 3)** we could demonstrate that in absence of the pilin domain, the crosstalk of allosteric sites within the lectin domain is also present in $FimH_{LD}$ variants harboring specific site mutations. In the flow-induced conformation, the shape of the binding pocket is similar to the medium-affinity state, though the overall shape of the lectin domain is narrower and more elongated in comparison to the low- and medium-affinity states.[41,47,48] Flow experiments with bacterial cells have shown that the FimH high-affinity state is nearly incapable to initiate bacterial adhesion at high shear stresses while FimH in the low-affinity state exhibits shear-enhanced attachment to urothelial cells and is thus fundamental for adhesion under flow conditions.[49]

Crystal structures are available for both the apoprotein in the low-affinity conformation (PDB: 4XOD) and the ligand-bound medium-affinity conformation (PDB: 4XOE). A high-affinity conformation of $FimH_{FL}$ can be experimentally induced in certain $FimH_{FL}$ crystals, in which the peptide DsG is replaced with the non-native donor strand of FimF (DsF). DsF stabilizes the pilin domain less well facilitating the domain separation even under crystal growing conditions (Figure 4).[36]

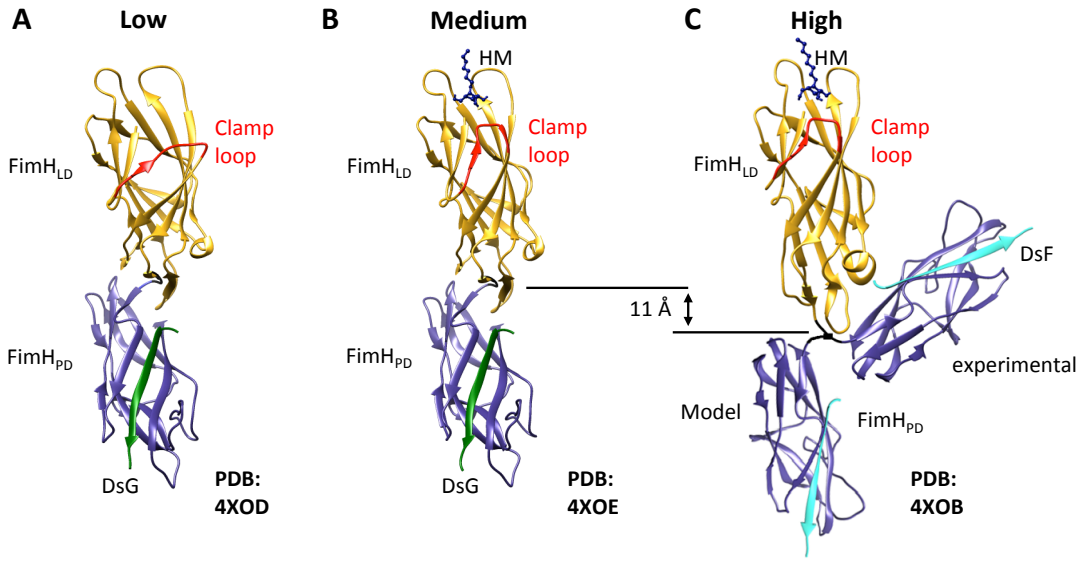


Figure 4: Crystal structures of FimH conformational states. A) Unbound $FimH_{FL}$ is present in the low-affinity conformation (PDB: 4XOD) and consists of the lectin domain ($FimH_{LD}$, golden) and the DsG-stabilized (green) pilin domain ($FimH_{PD}$, purple) interacting with each other. The lectin domain has a compressed shape with an open clamp loop (red) forming a shallow binding site. B) Binding of a ligand (*n*-heptyl *α*-D-mannopyranoside, blue) under static conditions leads to the medium-affinity state (PDB: 4XOE) with a closed clamp loop building a well-defined binding pocket. C) Under shear stress, both domains separate leading to an elongation of $FimH_{LD}$ by 11 Å in the high-affinity conformation. Domain separation of $FimH_{FL}$ under crystallization conditions was achieved by replacing DsG with the donor strand of FimF (DsF, cyan) (PDB: 4XOB). The crystal structures were adapted from Reference [36].

The high-affinity conformation has also been observed in a crystal structure of FimH in complex with pilus assembly chaperone FimC, where FimC is wedged between $FimH_{LD}$ and $FimH_{PD}$ (PDB: 1QUN).[31] Crystal structures of the isolated $FimH_{LD}$, both in the free (PDB: 4AUU) and ligand-bound (PDB: 4AV4) state, always show the extended high-affinity conformation.[50] In contrast to $FimH_{LD}$ within full-length FimH, the isolated lectin domain does not undergo significant conformational changes upon ligand binding.[51] Binding affinities of $FimH_{FL}$ for mannose-based antagonists are reduced by a factor of about 100 compared to isolated $FimH_{LD}$.[38] In **manuscript 5** *"The Different Affinity Stages of FimH and Their Thermodynamics of Binding"* a difference of two orders of magnitude in affinity was also discovered between the low-affinity and the high-affinity state. Isothermal titration calorimetry (ITC; method described in chapter I.5) demonstrated that ligand binding to $FimH_{FL}$ occurs with a highly favorable enthalpic and a considerably unfavorable entropic contribution. A considerable influence on the thermodynamics can be attributed to the dielectric constant of the binding pocket.

3.5 Molecular insight into the FimH binding site

In comparison to other lectin–monosaccharide interactions, the CRD of FimH has not only a remarkable binding specificity but also an unusually high-affinity with a K_D value of 2.3 μM for its ligand mannose.[52] This is due to a well-defined and charged binding pocket that envelops the sugar monomer and forms an extended network of 12 direct and indirect hydrogen bonds with the pyranose ring (Figure 5-A). Apart from the α-anomeric position, all hydroxyl groups of mannose form direct hydrogen bonds with the protein (Asp47, Asp54, Asn135, Asp140, and Phe1). Indirect hydrogen bonds with Gln133 and Gly14 are established via a water molecule, which is well-resolved in crystal structures of the high-affinity conformation. In total, the nine direct and the three indirect hydrogen bonds are formed with an appropriate hydrogen-bonding geometry.[52,53] Any modification on the hydroxyl groups of mannose considerably diminishes its binding affinity to FimH.[52,53,54]

The entrance of the mannose-binding site consists of a hydrophobic ridge (Ile13, Phe1, and Phe142) and the tyrosine gate motif (Tyr48, Ile52, and Tyr137) at the opposite side (Figure 5-B).[39,52] The X-ray crystal structure of the isolated $FimH_{LD}$

with oligomannose-3 (PDB: 2VCO) provided evidence that the tyrosine gate is involved in hydrophobic interactions with oligomannosides.[55] Thereupon, many studies focused on the development of mannoside derivatives with hydrophobic aglycones for anti-adhesive therapy of UTI. Favorable stacking and van der Waals interactions of the aglycone with aromatic side chains of the tyrosine gate lead to continuously improved affinities of synthetic FimH antagonists.[50,56,57] However, in contrast to synthetic FimH ligands, our study "*The tyrosine gate of the bacterial adhesion FimH – An evolutionary remnant paves the way for drug discovery*" (**manuscript 4**) on natural ligands of FimH has shown that a partial or complete removal of the tyrosine gate had no impact on the binding affinity of dimannosides in comparison with wild-type FimH. Furthermore, the binding affinity of oligomannoside–FimH complexes was not significantly affected by different glycosidic linkage types or a varying number of mannose residues.

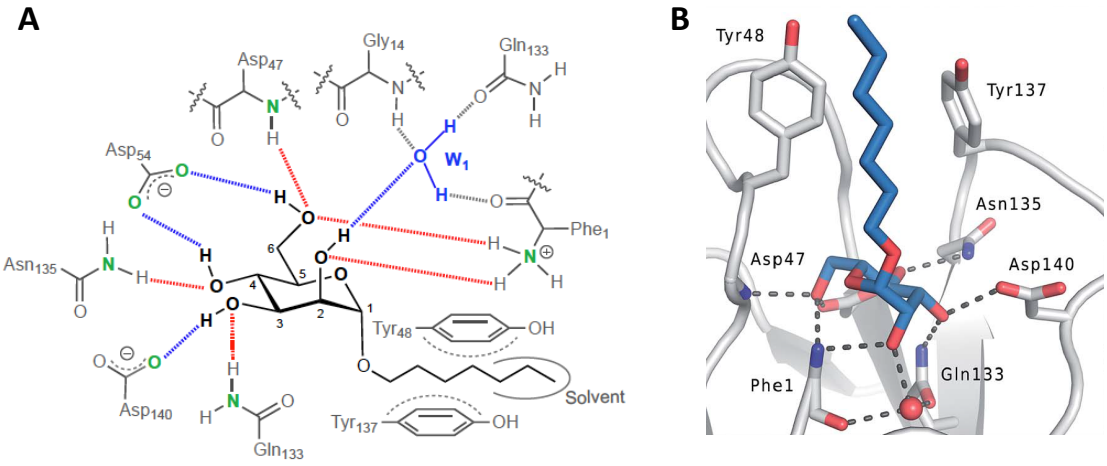


Figure 5: The carbohydrate recognition domain of the isolated FimH lectin domain in complex with *n*-heptyl *α*-D- mannopyranoside (HM). A) Schematic illustration of the hydrogen bond network with its ligand-accepted (red) and -donated (blue) hydrogen bonds. The structural water W1 is highlighted in blue. For clarity reasons, Gln133 is displayed twice. The figure is adapted from **manuscript 3**. B) Co-crystal structure of the $FimH_{LD}$-HM complex (PDB: 4BUQ)[58] displaying the hydrogen bond network (dashed lines) and the tyrosine gate.

The binding pocket is discussed in **manuscript 3** "*High-affinity carbohydrate-lectin interaction: How nature makes it possible*". In this study, the hydrogen bond network between FimH and mannose is elucidated at high detail by a systematic analysis of mannose derivatives by several biophysical methods. Deoxy- and deoxy-halogeno derivates of HM displayed a dramatic loss in binding affinity by the removal of only a single hydroxyl group. The reason for this is the disruption of the cooperative interplay of all hydrogen bonds in the network.

3.6 FimH antagonists

To date, antibiotics are still the first line treatment for UTIs,[59] although they may induce selection pressure on UPECs.[60] As a result, antibiotic-resistant organisms might emerge more often and are an expanding health problem worldwide.[61] In order to remedy such alarming tendencies, an alternative to the conventional antibiotic therapy is the development of an anti-adhesive UTI treatment strategy with FimH as promising target. In contrast to antibiotics, FimH antagonists do not necessarily exert resistance mechanisms against drugs because they block the adhesion of bacteria to the urothelium without killing them or inhibiting their growth.[62,63]

The development of high-affinity FimH antagonists with a therapeutic profile is still an ongoing challenge. As early as in the late 1960s, researchers discovered that type 1 pili enable *E. coli* to adhere to several cell types including erythrocytes and epithelial cells.[64] A few years later, it was observed that D-mannose and methyl *α*-D-mannopyranoside (**1**) prevented the attachment of type 1-pili-presenting *E. coli* strains to epithelial cells, and the type 1-pili-mediated agglutination of mannan-containing yeast.[65,66] In the 1980s, linear and branched oligosaccharides were studied and over the years several multivalent antagonists with nanomolar affinities were developed.[67,68,69] However, due to physicochemical and pharmacokinetic drawbacks, these large molecules are not suited for oral application, a prerequisite for the treatment of UTIs.

For the identification of orally available FimH antagonists, the development of monovalent ligands is the most promising approach. The first monovalent binders with reasonably high affinities to $FimH_{LD}$ were mannosides bearing an aromatic aglycone. One of them, *p*-nitrophenyl *α*-D-mannopyranoside (**2a**), inhibited the adherence of *E. coli* to epithelial cells 70 times more effective than methyl *α*-D-mannopyranoside. The introduction of a chloro substituent (**2b**) in *ortho* position further improved the inhibition by a factor of 7.[70] From 1999 on, crystal structures were used to rationally design high-affinity antagonists. Only since the first crystal structure of $FimH_{LD}$ in complex with a mannoside ligand was solved, detailed information was available regarding the tyrosine gate as a potential contact area for antagonists.[52] Long-chain alkyl aglycones turned out to provide hydrophobic contacts with the tyrosine gate allowing a drastic affinity increase. Among the studied

alkyl mannosides, *n*-heptyl *α*-D-mannopyranoside (HM; **3**) was the best $FimH_{LD}$ antagonist with a K_D of 5 nM in surface plasmon resonance assays (SPR; method described in chapter I.4).[52] The interaction of anti-adhesives with the tyrosine gate could be further enhanced by *π*-*π* stacking of several aryl aglycones with Tyr48. Substitution of the phenyl ring in *para* position resulted in squaric acid monoamide derivatives (**4**),[71,72] indolinyl phenyl mannosides (**5**),[73] and biphenyl *α*-D-mannopyranosides with distinct ring substitutions (**6a** and **6b**).[63,74,75] All previously mentioned FimH antagonists of this subchapter are illustrated in Figure 6.

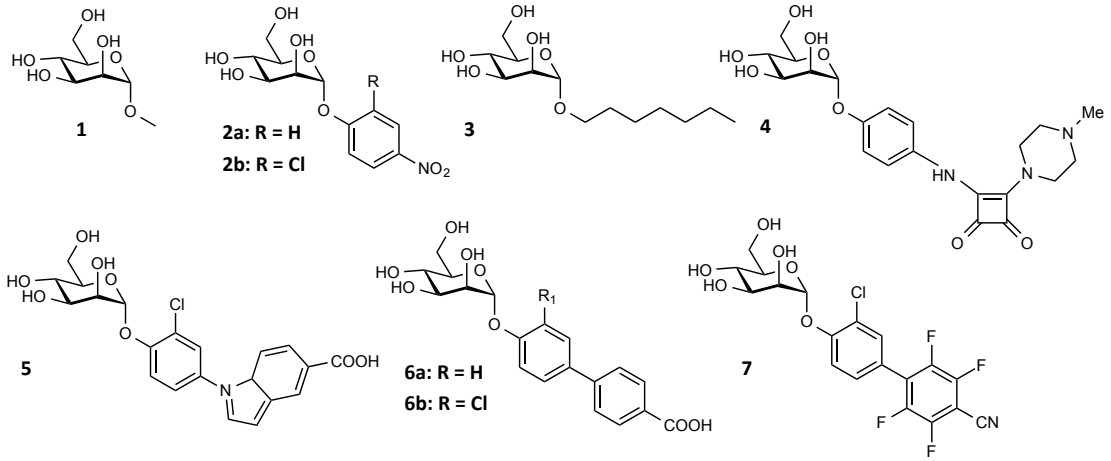


Figure 6: Chemical structures of selected FimH antagonists. Monovalent alkyl (**1** and **3**) and aryl (**2a**, **2b** and **4**-**7**) *α*-D-mannopyranosides that represent milestones in the development of high-affinity FimH antagonists.

Although a therapeutically effective FimH antagonist should as well exhibit a high affinity for FimH in the low-affinity conformation, previous drug development efforts mostly based on the high-affinity conformation by using isolated $FimH_{LD}$ as target protein. Only recently, $FimH_{FL}$ was included in the search of an appropriate FimH antagonist for a successful anti-adhesive UTI therapy. SPR and ITC measurements have shown that binding affinities of mannose-based antagonists are about 100 times lower for $FimH_{FL}$ compared to $FimH_{LD}$.[38,76] This two orders of magnitude difference in affinity was also confirmed in **manuscript 5** *"The Different Affinity Stages of FimH and Their Thermodynamics of Binding"*. In this study, it was demonstrated that fluorination of biphenyl mannosides further improved *π*-*π* stacking with the tyrosine gate, reaching nanomolar affinities with $FimH_{FL}$ and even picomolar affinities with $FimH_{LD}$. Perfluorination of the terminal phenyl ring in combination with a cyano group in *para* position resulted in a K_D of 349 pM for $FimH_{LD}$ and a K_D of 22.1 nM

for $FimH_{FL}$, currently representing the highest published affinity for a $FimH_{FL}$ antagonist.

To develop successful FimH anti-adhesives, medicinal chemists not only focus on optimized affinities and long residence times. Target selectivity as an indicator for toxicity and unwanted side effects is also an essential component of drug development and should be carefully investigated early on.[77] FimH antagonists based on α-D-mannopyranose may produce adverse side effects due to non-specific binding to mannose receptors of the human host system. As a promising result, competitive binding assays demonstrated a 10^5 lower affinity of antagonists **3**, **4**, **5**, **6a**, and **6b** to eight human mannose receptors compared to the isolated lectin domain, likely due to less optimal hydrophobic interactions and the fundamental design of human mannose receptors for multivalent interactions.[78] Furthermore, physicochemical and pharmacokinetic properties including solubility, lipophilicity, and membrane permeability have to be optimized to ensure oral bioavailability. Prodrug approaches can be taken into consideration to balance between solubility and lipophilicity for effective dosage and membrane permeation. For instance, insufficient solubility of FimH antagonists, which otherwise show high passive permeability, has been overcome by a phosphate prodrug strategy.[79] In another study, a series of ester prodrugs of biphenyl α-D-mannopyranosides with linear short-chain acyl promoieties at the C-6 position showed not only good permeability but also adequate solubility for intestinal absorption.[80]

3.7 Kinetics of FimH

Typically, carbohydrate–lectin interactions display fast binding kinetics that complicate the development of successful carbohydrate-based drugs.[81] Of particular interest is the dissociation rate constant (k_{off}), because it determines the half-life ($t_{1/2} = \ln 2/k_{off}$) of a drug–target complex, i.e. the time a target is occupied by a drug. Since target occupancy is directly related to *in vivo* drug efficacy considering pharmacokinetic effects such as metabolism and elimination from the body, an extended $t_{1/2}$ is very often required to obtain a therapeutically potent antagonist.[82,83] In the case of E-, L- and P-selectins, SPR experiments unveiled fast off-rates that resulted in half-lives of less than a second.[84,85,86] Other carbohydrate–lectin

interactions showed similar kinetic properties with $t_{1/2}$ values in the second range. The longest half-lives were observed for concanavalin A[87] and mannose-binding protein[88] that exhibited half-lives of 1.2 min and 4.4 min in complex with a D-mannose derivate and D-Man_{16}-BSA, respectively. All the more surprising was the discovery of $FimH_{LD}$–antagonist complexes exhibiting half-lives of above 3.6 h.[81] In contrast to the other investigated lectins, which possess shallow and water-accessible CRDs,[89,90,91] the deep and polar binding pocket of $FimH_{LD}$ an extended hydrogen bond network can be established being responsible for slow dissociation rates and thus for the observed long complex half-lives.

Yakovenko *et al*. performed SPR measurements of isolated fimbriae to investigate the kinetic difference between the low-affinity state in wild-type K12 FimH and the engineered FocH variant[76] that is locked in the high-affinity state.[49] In this study, the low-affinity conformation exhibited a 23-fold faster association rate and a 600-fold higher dissociation rate compared to the high-affinity conformation (Figure 7).

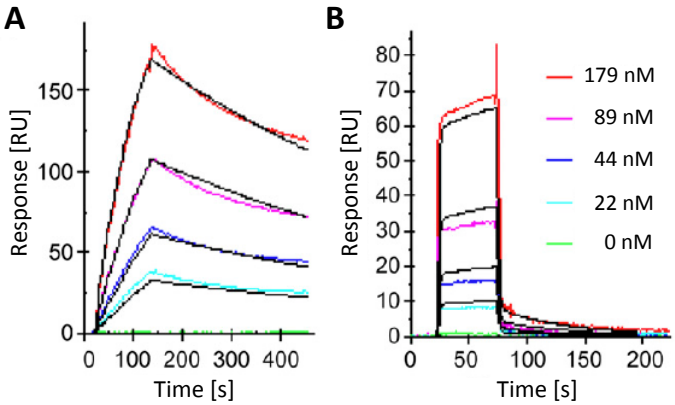


Figure 7: Kinetic fingerprints of fimbriae presenting FimH in the high- and low-affinity conformations. Binding behavior of fimbriae from the strains FocH (high-affinity conformation, A) and wild type K12 (low-affinity conformation, B) to immobilized man-BSA. Measurements were performed with the indicated fimbriae concentrations using a Biacore 2000. Sensorgrams were fitted with a one-state (FocH, A) and a two-state model (K12, B), respectively. Results were taken from Yakovenko *et al.*, 2015.[49]

Fast association and dissociation rates of HM (**3**) binding to FimH in the low affinity conformation were also demonstrated by Sauer *et al*. for the full-length FimH construct stabilized by DsG. A k_{on} of $5.0*10^6$ $M^{-1}s^{-1}$ and a k_{off} of 22 s^{-1} were measured, resulting in a complex half-life of 32 ms.[36] Apart from these studies, kinetic information about the low-affinity state is still rare, despite the fact that a therapeutically successful FimH antagonist should exhibit a reasonable dissociation

half-life. A stable antagonist–$FimH_{FL}$ complex should prevent the adhesion of UPEC to the bladder epithelium under static conditions and consequently promotes bacterial clearance from the urinary tract. In order to find chemical structures with optimized half-lives, we designed a SPR method to screen a compound library against low-affinity $FimH_{FL}$. Assay development and screening of structurally diverse FimH antagonist classes is highlighted in **manuscript 1** *"Off-rate screening by surface plasmon resonance - The search for promising lead structures targeting low-affinity FimH"*.

Considering the emerging importance of binding kinetics in drug discovery, efforts have also been made to develop new methods for determination of association and dissociation rates. **Publication 2** "*KinITC – One method supports both thermo-dynamic and kinetic SARs*" elaborates on kinetic ITC based upon the determination of an "Equilibration Time Curve" (kinITC-ETC), a novel technique that extracts the kinetic fingerprint of a drug–target complex from ITC data. In this context, structural properties of FimH antagonists could be correlated with kinetic parameters of $FimH_{LD}$–antagonist interactions. In case of the isolated FimH lectin domain, it was shown that the rate constant k_{off} is mainly influenced by formation of the hydrogen bond network, whereas electrostatic interactions and conformational restrictions primarily affect the rate constant k_{on}.

References

1. Harding, G.K. and Ronald, A.R., The management of urinary infections: What have we learned in the past decade? *Int. J. Antimicrob. Agents*, **1994**, *4*(2), 83-88.
2. Lichtenberger, P. and Hooton, T.M., Complicated urinary tract infections. *Curr. Infect. Dis. Rep.*, **2008**, *10*(6), 499-504.
3. Nicolle, L.E., Urinary tract pathogens in complicated infection and in elderly individuals. *J. Infect. Dis.*, **2001**, *183 Suppl 1*, S5-8.
4. Ronald, A., The etiology of urinary tract infection: Traditional and emerging pathogens. *Am. J. Med.*, **2002**, *113 Suppl 1*, S14-19.
5. Hooton, T.M. and Stamm, W.E., Diagnosis and treatment of uncomplicated urinary tract infection. *Infect. Dis. Clin. North Am.*, **1997**, *11*(3), 551-581.
6. Zhang, L. and Foxman, B., Molecular epidemiology of Escherichia coli mediated urinary tract infections. *Front. Biosci.*, **2003**, *8*, e235-244.
7. Foxman, B., Epidemiology of urinary tract infections: Incidence, morbidity, and economic costs. *Am. J. Med.*, **2002**, *113 Suppl 1A*, S5-13.
8. Leigh, D., Urinary tract infections, in *Topley and Wilson's principles of bacteriology, virology and immunity*, **1990**, Butler and Tanler Ltd., Frome and London.
9. Ikaheimo, R., Siitonen, A., Heiskanen, T., Karkkainen, U., Kuosmanen, P., Lipponen, P., and Makela, P.H., Recurrence of urinary tract infection in a primary care setting: Analysis of a 1-year follow-up of 179 women. *Clin. Infect. Dis.*, **1996**, *22*(1), 91-99.

10. Kostakioti, M., Hultgren, S.J., and Hadjifrangiskou, M., Molecular blueprint of uropathogenic Escherichia coli virulence provides clues toward the development of anti-virulence therapeutics. *Virulence*, **2012**, *3*(7), 592-594.
11. Bischoff, S., Walter, T., Gerigk, M., Ebert, M., and Vogelmann, R., Empiric antibiotic therapy in urinary tract infection in patients with risk factors for antibiotic resistance in a German emergency department. *BMC Infect. Dis.*, **2018**, *18*, 56.
12. Foxman, B. and Buxton, M., Alternative approaches to conventional treatment of acute uncomplicated urinary tract infection in women. *Curr. Infect. Dis. Rep.*, **2013**, *15*(2), 124-129.
13. Wagenlehner, F., Wullt, B., Ballarini, S., Zingg, D., and Naber, K.G., Social and economic burden of recurrent urinary tract infections and quality of life: A patient web-based study (GESPRIT). *Expert. Rev. Pharmacoecon. Outcomes Res.*, **2018**, *18*(1), 107-117.
14. Moreno, E., Andreu, A., Perez, T., Sabate, M., Johnson, J.R., and Prats, G., Relationship between Escherichia coli strains causing urinary tract infection in women and the dominant faecal flora of the same hosts. *Epidemiol. Infect.*, **2006**, *134*(5), 1015-1023.
15. Bien, J., Sokolova, O., and Bozko, P., Role of uropathogenic Escherichia coli virulence factors in development of urinary tract infection and kidney damage. *Int. J. Nephrol.*, **2012**, *2012*, 681473.
16. Spaulding, C.N., Klein, R.D., Ruer, S., Kau, A.L., Schreiber, H.L., Cusumano, Z.T., Dodson, K.W., Pinkner, J.S., Fremont, D.H., Janetka, J.W., Remaut, H., Gordon, J.I., and Hultgren, S.J., Selective depletion of uropathogenic E. coli from the gut by a FimH antagonist. *Nature*, **2017**, *546*, 528.
17. Mulvey, M.A., Adhesion and entry of uropathogenic Escherichia coli. *Cell. Microbiol.*, **2002**, *4*(5), 257-271.
18. Zhou, G., Mo, W.J., Sebbel, P., Min, G., Neubert, T.A., Glockshuber, R., Wu, X.R., Sun, T.T., and Kong, X.P., Uroplakin ia is the urothelial receptor for uropathogenic Escherichia coli: Evidence from in vitro FimH binding. *J. Cell Sci.*, **2001**, *114*(Pt 22), 4095-4103.
19. Anderson, G.G., Palermo, J.J., Schilling, J.D., Roth, R., Heuser, J., and Hultgren, S.J., Intracellular bacterial biofilm-like pods in urinary tract infections. *Science*, **2003**, *301*(5629), 105-107.
20. Justice, S.S., Hung, C., Theriot, J.A., Fletcher, D.A., Anderson, G.G., Footer, M.J., and Hultgren, S.J., Differentiation and developmental pathways of uropathogenic Escherichia coli in urinary tract pathogenesis. *Proc. Natl. Acad. Sci. U. S. A.*, **2004**, *101*(5), 1333-1338.
21. Hunstad, D.A. and Justice, S.S., Intracellular lifestyles and immune evasion strategies of uropathogenic Escherichia coli. *Annu. Rev. Microbiol.*, **2010**, *64*, 203-221.
22. Mulvey, M.A., Lopez-Boado, Y.S., Wilson, C.L., Roth, R., Parks, W.C., Heuser, J., and Hultgren, S.J., Induction and evasion of host defenses by type 1-piliated uropathogenic Escherichia coli. *Science*, **1998**, *282*(5393), 1494-1497.
23. Cegelski, L., Marshall, G.R., Eldridge, G.R., and Hultgren, S.J., The biology and future prospects of antivirulence therapies. *Nat. Rev. Microbiol.*, **2008**, *6*(1), 17-27.
24. King, J.E. and Roberts, I.S., Bacterial surfaces: Front lines in host-pathogen interaction. *Adv. Exp. Med. Biol.*, **2016**, *915*, 129-156.
25. Mysorekar, I.U. and Hultgren, S.J., Mechanisms of uropathogenic Escherichia coli persistence and eradication from the urinary tract. *PNAS*, **2006**, *103*(38), 14170-14175.
26. Santo, E., Macedo, C., and Marin, J.M., Virulence factors of uropathogenic Escherichia coli from a university hospital in Ribeirao Preto, Sao Paulo, Brazil. *Rev. Inst. Med. Trop. Sao Paulo*, **2006**, *48*(4), 185-188.
27. Hahn, E., Wild, P., Hermanns, U., Sebbel, P., Glockshuber, R., Haner, M., Taschner, N., Burkhard, P., Aebi, U., and Muller, S.A., Exploring the 3d molecular architecture of Escherichia coli type 1 pili. *J. Mol. Biol.*, **2002**, *323*(5), 845-857.
28. Klemm, P., Hancock, V., and Schembri, M.A., Fimbrial adhesins from extraintestinal Escherichia coli. *Environ. Microbiol. Rep.*, **2010**, *2*(5), 628-640.
29. Sauer, F.G., Remaut, H., Hultgren, S.J., and Waksman, G., Fiber assembly by the chaperone-usher pathway. *Biochim. Biophys. Acta*, **2004**, *1694*(1-3), 259-267.
30. Abraham, J.M., Freitag, C.S., Clements, J.R., and Eisenstein, B.I., An invertible element of DNA controls phase variation of type 1 fimbriae of Escherichia coli. *PNAS*, **1985**, *82*(17), 5724-5727.
31. Choudhury, D., Thompson, A., Stojanoff, V., Langermann, S., Pinkner, J., Hultgren, S.J., and Knight, S.D., X-ray structure of the FimC-FimH chaperone-adhesin complex from uropathogenic Escherichia coli. *Science*, **1999**, *285*(5430), 1061-1066.

32. Jones, C.H., Pinkner, J.S., Roth, R., Heuser, J., Nicholes, A.V., Abraham, S.N., and Hultgren, S.J., FimH adhesin of type 1 pili is assembled into a fibrillar tip structure in the Enterobacteriaceae. *PNAS*, **1995**, *92*(6), 2081-2085.
33. Capitani, G., Eidam, O., Glockshuber, R., and Grütter, M.G., Structural and functional insights into the assembly of type 1 pili from Escherichia coli. *Microbes Infect.*, **2006**, *8*(8), 2284-2290.
34. Mydock-McGrane, L.K., Hannan, T.J., and Janetka, J.W., Rational design strategies for FimH antagonists: New drugs on the horizon for urinary tract infection and Crohn's disease. *Expert. Opin. Drug Discov.*, **2017**, *12*(7), 711-731.
35. Puorger, C., Vetsch, M., Wider, G., and Glockshuber, R., Structure, folding and stability of fima, the main structural subunit of type 1 pili from uropathogenic Escherichia coli strains. *J. Mol. Biol.*, **2011**, *412*(3), 520-535.
36. Sauer, M.M., Jakob, R.P., Eras, J., Baday, S., Eriş, D., Navarra, G., Bernèche, S., Ernst, B., Maier, T., and Glockshuber, R., Catch-bond mechanism of the bacterial adhesin FimH. *Nat. Commun.*, **2016**, *7*, 10738.
37. Frei, P., Pang, L., Silbermann, M., Eris, D., Muhlethaler, T., Schwardt, O., and Ernst, B., Target-directed dynamic combinatorial chemistry: A study on potentials and pitfalls as exemplified on a bacterial target. *Chemistry*, **2017**, *23*(48), 11570-11577.
38. Mayer, K., Eris, D., Schwardt, O., Sager, C.P., Rabbani, S., Kleeb, S., and Ernst, B., Urinary tract infection: Which conformation of the bacterial lectin FimH is therapeutically relevant? *J. Med. Chem.*, **2017**, *60*(13), 5646-5662.
39. Hung, C.S., Bouckaert, J., Hung, D., Pinkner, J., Widberg, C., DeFusco, A., Auguste, C.G., Strouse, R., Langermann, S., Waksman, G., and Hultgren, S.J., Structural basis of tropism of Escherichia coli to the bladder during urinary tract infection. *Mol. Microbiol.*, **2002**, *44*(4), 903-915.
40. Bouckaert, J., Mackenzie, J., de Paz, J.L., Chipwaza, B., Choudhury, D., Zavialov, A., Mannerstedt, K., Anderson, J., Piérard, D., Wyns, L., Seeberger, P.H., Oscarson, S., De Greve, H., and Knight, S.D., The affinity of the FimH fimbrial adhesin is receptor-driven and quasi-independent of Escherichia coli pathotypes. *Mol. Microbiol.*, **2006**, *61*(6), 1556-1568.
41. 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., and Thomas, W.E., Structural basis for mechanical force regulation of the adhesin FimH via finger trap-like beta sheet twisting. *Cell*, **2010**, *141*(4), 645-655.
42. Aprikian, P., Interlandi, G., Kidd, B.A., Le Trong, I., Tchesnokova, V., Yakovenko, O., Whitfield, M.J., Bullitt, E., Stenkamp, R.E., Thomas, W.E., and Sokurenko, E.V., The bacterial fimbrial tip acts as a mechanical force sensor. *PLoS Biol.*, **2011**, *9*(5), e1000617.
43. Nilsson, L.M., Thomas, W.E., Trintchina, E., Vogel, V., and Sokurenko, E.V., Catch bond-mediated adhesion without a shear threshold: Trimannose versus monomannose interactions with the FimH adhesin of escherichia coli. *J. Biol. Chem.*, **2006**, *281*(24), 16656-16663.
44. Thomas, W.E., Trintchina, E., Forero, M., Vogel, V., and Sokurenko, E.V., Bacterial adhesion to target cells enhanced by shear force. *Cell*, **2002**, *109*(7), 913-923.
45. Marshall, B.T., Long, M., Piper, J.W., Yago, T., McEver, R.P., and Zhu, C., Direct observation of catch bonds involving cell-adhesion molecules. *Nature*, **2003**, *423*(6936), 190-193.
46. Sarangapani, K.K., Yago, T., Klopocki, A.G., Lawrence, M.B., Fieger, C.B., Rosen, S.D., McEver, R.P., and Zhu, C., Low force decelerates L-selectin dissociation from P-selectin glycoprotein ligand-1 and endoglycan. *J. Biol. Chem.*, **2004**, *279*(3), 2291-2298.
47. Yakovenko, O., Sharma, S., Forero, M., Tchesnokova, V., Aprikian, P., Kidd, B., Mach, A., Vogel, V., Sokurenko, E., and Thomas, W.E., FimH forms catch bonds that are enhanced by mechanical force due to allosteric regulation. *J. Biol. Chem.*, **2008**, *283*(17), 11596-11605.
48. Sun, L., Cheng, Q.H., Gao, H.J., and Zhang, Y.W., Effect of loading conditions on the dissociation behaviour of catch bond clusters. *J. Royal Soc. Interface*, **2012**, *9*(70), 928-937.
49. Yakovenko, O., Tchesnokova, V., Sokurenko, E.V., and Thomas, W.E., Inactive conformation enhances binding function in physiological conditions. *PNAS*, **2015**, *112*(32), 9884-9889.
50. Wellens, A., Lahmann, M., Touaibia, M., Vaucher, J., Oscarson, S., Roy, R., Remaut, H., and Bouckaert, J., The tyrosine gate as a potential entropic lever in the receptor-binding site of the bacterial adhesin FimH. *Biochemistry*, **2012**, *51*(24), 4790-4799.
51. Fiege, B., Rabbani, S., Preston, R.C., Jakob, R.P., Zihlmann, P., Schwardt, O., Jiang, X., Maier, T., and Ernst, B., The tyrosine gate of the bacterial lectin FimH: A conformational

analysis by NMR spectroscopy and X-ray crystallography. *Chembiochem*, **2015**, *16*(8), 1235-1246.

52. Bouckaert, J., Berglund, J., Schembri, M., De Genst, E., Cools, L., Wuhrer, M., Hung, C.S., Pinkner, J., Slattegard, R., Zavialov, A., Choudhury, D., Langermann, S., Hultgren, S.J., Wyns, L., Klemm, P., Oscarson, S., Knight, S.D., and De Greve, H., Receptor binding studies disclose a novel class of high-affinity inhibitors of the Escherichia coli FimH adhesin. *Mol. Microbiol.*, **2005**, *55*(2), 441-455.
53. Zihlmann, P.J., X.; Sager, C. P.; Fiege, B.; Jakob, R. P.; Siegrist, S.; Zalewski, A.; Rabbani, S.; Eriş, D.; Silbermann, M.; Pang, L.; Mühlethaler, T.; Sharpe, T.; Maier, T.; Ernst, B., High-affinity carbohydrate-lectin interaction: How nature makes it possible. *unpublished results*.
54. Old, D.C., Inhibition of the interaction between fimbrial haemagglutinins and erythrocytes by D-mannose and other carbohydrates. *J. Gen. Microbiol.*, **1972**, *71*(1), 149-157.
55. Wellens, A., Garofalo, C., Nguyen, H., Van Gerven, N., Slattegard, R., Hernalsteens, J.P., Wyns, L., Oscarson, S., De Greve, H., Hultgren, S., and 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*(4), e2040.
56. Han, Z., Pinkner, J.S., Ford, B., Obermann, R., Nolan, W., Wildman, S.A., Hobbs, D., Ellenberger, T., Cusumano, C.K., Hultgren, S.J., and Janetka, J.W., Structure-based drug design and optimization of mannoside bacterial FimH antagonists. *J. Med. Chem.*, **2010**, *53*(12), 4779-4792.
57. Kleeb, S., Pang, L., Mayer, K., Eris, D., Sigl, A., Preston, R.C., Zihlmann, P., Sharpe, T., Jakob, R.P., Abgottspon, D., Hutter, A.S., Scharenberg, M., Jiang, X., Navarra, G., Rabbani, S., Smiesko, M., Lüdin, N., Bezençon, J., Schwardt, O., Maier, T., and Ernst, B., FimH antagonists: Bioisosteres to improve the in vitro and in vivo PK/PD profile. *J. Med. Chem.*, **2015**, *58*(5), 2221-2239.
58. Sager, C.P., Fiege, B., Zihlmann, P., Vannam, R., Rabbani, S., Jakob, R.P., Preston, R.C., Zalewski, A., Maier, T., Peczuh, M.W., and Ernst, B., The price of flexibility - a case study on septanoses as pyranose mimetics. *Chem. Sci.*, **2018**, *9*(3), 646-654.
59. Fihn, S.D., Clinical practice. Acute uncomplicated urinary tract infection in women. *N. Engl. J. Med.*, **2003**, *349*(3), 259-266.
60. Sanchez, G.V., Master, R.N., Karlowsky, J.A., and 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*(4), 2181-2183.
61. Ventola, C.L., The antibiotic resistance crisis: Part 1: Causes and threats. *Pharm. Ther.*, **2015**, *40*(4), 277-283.
62. Cusumano, C.K., Pinkner, J.S., Han, Z., Greene, S.E., Ford, B.A., Crowley, J.R., Henderson, J.P., Janetka, J.W., and Hultgren, S.J., Treatment and prevention of urinary tract infection with orally active FimH inhibitors. *Sci. Transl. Med.*, **2011**, *3*(109), 109ra115.
63. 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., and Ernst, B., FimH antagonists for the oral treatment of urinary tract infections: From design and synthesis to in vitro and in vivo evaluation. *J. Med. Chem.*, **2010**, *53*(24), 8627-8641.
64. Duguid, J.P., The function of bacterial fimbriae. *Arch. Immunol. Ther. Exp. (Warsz.)*, **1968**, *16*(2), 173-188.
65. Ofek, I., Mirelman, D., and Sharon, N., Adherence of Escherichia coli to human mucosal cells mediated by mannose receptors. *Nature*, **1977**, *265*, 623.
66. Sharon, N., Bacterial lectins, cell‐cell recognition and infectious disease. *FEBS Lett.*, **1987**, *217*(2), 145-157.
67. Hartmann, M. and Lindhorst Thisbe, K., The bacterial lectin FimH, a target for drug discovery: Carbohydrate inhibitors of type 1 fimbriae-mediated bacterial adhesion. *Eur. J. Org. Chem.*, **2011**, *2011*(20-21), 3583-3609.
68. Imberty, A., Chabre, Y.M., and Roy, R., Glycomimetics and glycodendrimers as high affinity microbial anti-adhesins. *Chemistry*, **2008**, *14*(25), 7490-7499.
69. Pieters, R.J., Maximising multivalency effects in protein–carbohydrate interactions. *Org. Biomol. Chem.*, **2009**, *7*(10), 2013-2025.
70. Firon, N., Ashkenazi, S., Mirelman, D., Ofek, I., and 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*(2), 472-476.

71. Rabbani, S., Jiang, X., Schwardt, O., and Ernst, B., Expression of the carbohydrate recognition domain of FimH and development of a competitive binding assay. *Anal. Biochem.*, **2010**, *407*(2), 188-195.
72. Sperling, O., Fuchs, A., and 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*(21), 3913-3922.
73. Jiang, X., Abgottspon, D., Kleeb, S., Rabbani, S., Scharenberg, M., Wittwer, M., Haug, M., Schwardt, O., and Ernst, B., Antiadhesion therapy for urinary tract infections: A balanced PK/PD profile proved to be key for success. *J. Med. Chem.*, **2012**, *55*(10), 4700-4713.
74. Han, Z., Pinkner, J.S., Ford, B., Chorell, E., Crowley, J.M., Cusumano, C.K., Campbell, S., Henderson, J.P., Hultgren, S.J., and Janetka, J.W., Lead optimization studies on FimH antagonists: Discovery of potent and orally bioavailable ortho-substituted biphenyl mannosides. *J. Med. Chem.*, **2012**, *55*(8), 3945-3959.
75. Pang, L., Kleeb, S., Lemme, K., Rabbani, S., Scharenberg, M., Zalewski, A., Schadler, F., Schwardt, O., and Ernst, B., FimH antagonists: Structure-activity and structure-property relationships for biphenyl alpha-D-mannopyranosides. *ChemMedChem*, **2012**, *7*(8), 1404-1422.
76. Aprikian, P., Tchesnokova, V., Kidd, B., Yakovenko, O., Yarov-Yarovoy, V., Trinchina, E., Vogel, V., Thomas, W., and Sokurenko, E., Interdomain interaction in the FimH adhesin of Escherichia coli regulates the affinity to mannose. *J. Biol. Chem.*, **2007**, *282*(32), 23437-23446.
77. Huggins, D.J., Sherman, W., and Tidor, B., Rational approaches to improving selectivity in drug design. *J. Med. Chem.*, **2012**, *55*(4), 1424-1444.
78. Scharenberg, M., Schwardt, O., Rabbani, S., and Ernst, B., Target selectivity of FimH antagonists. *J. Med. Chem.*, **2012**, *55*(22), 9810-9816.
79. Kleeb, S., Jiang, X., Frei, P., Sigl, A., Bezençon, J., Bamberger, K., Schwardt, O., and Ernst, B., FimH antagonists: Phosphate prodrugs improve oral bioavailability. *J. Med. Chem.*, **2016**, *59*(7), 3163-3182.
80. Schönemann, W., Kleeb, S., Dätwyler, P., Schwardt, O., and Ernst, B., Prodruggability of carbohydrates – oral FimH antagonists. *Can. J. Chem.*, **2016**, *94*(11), 909-919.
81. Scharenberg, M., Jiang, X., Pang, L., Navarra, G., Rabbani, S., Binder, F., Schwardt, O., and Ernst, B., Kinetic properties of carbohydrate–lectin interactions: FimH antagonists. *ChemMedChem*, **2014**, *9*(1), 78-83.
82. Tummino, P.J. and Copeland, R.A., Residence time of receptor–ligand complexes and its effect on biological function. *Biochemistry*, **2008**, *47*(20), 5481-5492.
83. Lu, H. and Tonge, P.J., Drug-target residence time: Critical information for lead optimization. *Curr. Opin. Chem. Biol.*, **2010**, *14*(4), 467-474.
84. Wild, M.K., Huang, M.C., Schulze-Horsel, U., van der Merwe, P.A., and Vestweber, D., Affinity, kinetics, and thermodynamics of E-selectin binding to E-selectin ligand-1. *J. Biol. Chem.*, **2001**, *276*(34), 31602-31612.
85. Nicholson, M.W., Barclay, A.N., Singer, M.S., Rosen, S.D., and van der Merwe, P.A., Affinity and kinetic analysis of L-selectin (cd62l) binding to glycosylation-dependent cell-adhesion molecule-1. *J. Biol. Chem.*, **1998**, *273*(2), 763-770.
86. Mehta, P., Cummings, R.D., and McEver, R.P., Affinity and kinetic analysis of P-selectin binding to P-selectin glycoprotein ligand-1. *J. Biol. Chem.*, **1998**, *273*(49), 32506-32513.
87. Murthy, B.N., Sinha, S., Surolia, A., Indi, S.S., and Jayaraman, N., SPR and ITC determination of the kinetics and the thermodynamics of bivalent versus monovalent sugar ligand–lectin interactions. *Glycoconj. J.*, **2008**, *25*(4), 313-321.
88. Terada, T., Nishikawa, M., Yamashita, F., and Hashida, M., Analysis of the molecular interaction between mannosylated proteins and serum mannan-binding lectins. *Int. J. Pharm.*, **2006**, *316*(1), 117-123.
89. Hardman, K.D. and Ainsworth, C.F., Structure of the concanavalin A–methyl α-D-mannopyranoside complex at 6-Å resolution. *Biochemistry*, **1976**, *15*(5), 1120-1128.
90. Somers, W.S., Tang, J., Shaw, G.D., and Camphausen, R.T., Insights into the molecular basis of leukocyte tethering and rolling revealed by structures of P- and E-Selectin bound to sLex and PSGL-1. *Cell*, **2000**, *103*(3), 467-479.
91. Ng, K.K., Drickamer, K., and Weis, W.I., Structural analysis of monosaccharide recognition by rat liver mannose-binding protein. *J. Biol. Chem.*, **1996**, *271*(2), 663-674.

4. Biophysical Methods in Drug Discovery

Modern drug discovery heavily relies on biophysical techniques and quite a number of these have matured into key technologies for the drug development process. In the present studies, a number of them have been applied.

For structure elucidation and structure-based rational drug design, X-ray crystallography is usually the method of choice. Based on X-ray diffraction patterns three-dimensional structures of macromolecules are unveiled at an atomic level.[1,2] Up until the beginning of 2018, more than 120.000 structures of apoproteins and protein–ligand complexes in the Protein Data Bank (PDB) were determined using X-ray crystallography.[3]

Additional 12.000 structures were solved by nuclear magnetic resonance (NMR),[3] a versatile method to study protein structure, dynamics, and protein–ligand interactions including the identification of binding partners, determination of the binding mode, and affinity measurements.[4] In ligand-observed NMR experiments, the resonances of small molecules are observed in the presence of a target protein to screen for ligand binding and to further characterize the complex in terms of affinity and kinetics. Isotopic labeling of the interaction partners is not necessary. The pharmaceutical industry not only applies ligand-based NMR as a primary screening tool but also for the validation of hits from various screening techniques and for the optimization of leads.[5] Protein-based NMR measurements are often indispensable for structure-based drug discovery that demands information of the binding pocket. However, these techniques usually require isotopic labeling of the target protein. The single most-widely used protein-based NMR approach for analyses of drug-target interactions is 2D ^{1}H–^{15}N heteronuclear single quantum correlation (HSQC) of ^{15}N-labelled protein samples.[6]

The secondary structural elements of proteins, such as helices, sheets, turns, and coils can be investigated by circular dichroism (CD) spectroscopy. CD determines the protein conformation in solution by detecting the difference in absorption of left- and right-circularly polarized light over a range of wavelengths.[7] CD can be combined with stopped-flow instruments to measure protein-folding kinetics or ligand-binding

kinetics if the binding event involves a change in chirality of the protein.[8] The stopped-flow device enables a simultaneous pressure-injection of solution in different syringes into a mixing chamber. Thereupon, the flow is then promptly stopped and the reaction is measured at different time points.[9] Stopped-flow kinetic measurements can also be performed in combination with a UV/Vis or fluorescence spectrometer if a change in absorbance or fluorescence takes place during the binding event.[10]

The gold standard to obtain kinetic parameters of bimolecular interactions is surface plasmon resonance (SPR). In addition, kinetic structure–activity relationship (SAR) studies can also be performed with the recently introduced analytical tool kinetic ITC (kinITC) that extracts kinetic information from isothermal titration calorimetry (ITC) data. In **publication 2** "KinITC – One method supports both thermodynamic and kinetic SARs" kinITC was independently validated for the first time and applied to deduce the kinetic parameters from existing ITC data. ITC is the standard technique to obtain thermodynamic parameters of a binding event. Both methods, SPR and ITC (including kinITC), are described in detail in Section 4 and 5, respectively.

Mass spectroscopy (MS) is a powerful tool for the determination of concentration, purity, and molecular mass of molecules, but also for the identification of unknown samples, and revelation of protein structure and dynamics.[11,12] For protein–ligand or protein–protein interaction studies, size-exclusion chromatography in combination with MS (SEC-MS) is a suitable technique.[13] An MS analysis requires the conversion of the sample molecule into gaseous ions, e.g. by using electrospray ionization (ESI). Subsequently, the ionized fragments are assessed and recorded by their mass to charge ratios (m/z).[14] Frequently, gas or liquid chromatography is used as pre-MS separation technique especially for analysis of complex samples.[15] MS can also be combined with SPR devices for ligand-fishing experiments. A target immobilized on a SPR chip captures unknown ligands, which can be eluted and analyzed by ESI-MS or identified by matrix-assisted laser desorption/ionization mass spectroscopy (MALDI-MS) directly from the chip.[16]

Ligands with low affinity can be identified by differential scanning fluorimetry (DSF), also known as thermal shift analysis (TSA). DSF determines the melting temperature (T_m) of a protein by monitoring its thermal denaturation in the presence

of a fluorescent dye that binds to exposed hydrophobic residues.[17,18] Binding of a ligand typically stabilizes but sometimes also destabilizes the protein resulting in a T_m shift in each case. DSF is a robust and inexpensive method that makes use of a real-time quantitative PCR cycler, and therefore is often applied for primary fragment-based library screenings.[19,20]

With the rather new technique microscale thermophoresis (MST) biomolecular interactions with pico- to millimolar affinities can be studied based on the movement of fluorescent molecules in a temperature gradient.[21,22] The fluorescence is either an intrinsic property of the molecule of interest or can be generated by attaching a fluorescent protein or dye.[23] The method is highly sensitive and detects changes in the hydration shell of biomolecules as well as size- and charge-alterations, which have an impact on the thermophoretic mobility of proteins.[24] Thus, the affinity of protein–protein as well as of protein–small molecule interactions can be determined.[22]

The multifaceted method analytical ultracentrifugation (AUC) quantitatively analyzes the influence of a gravitational field generated by a centrifuge on macromolecules in solution. The technique allows the study of biomolecules over a broad concentration range and tolerates numerous solvent additives.[25] The two main AUC application methods are sedimentation velocity and sedimentation equilibrium. Sedimentation velocity experiments assess the movement rate of macromolecules in a strong centrifugal field. Based on the sedimentation rate the shape and the size of macromolecules can be ascertained. In sedimentation equilibrium experiments a comparatively lower centrifugal force acts on the protein in such a way that sedimentation and diffusion are exactly balanced resulting in a concentration distribution at equilibrium. Studying the sedimentation equilibrium of macro-molecules in a centrifugal field provides information about molecular masses, binding stoichiometry, and affinities as well as the protein quaternary structure.[26,27]

In Table 1 the main application areas of the previously mentioned biophysical methods are summarized. In addition, advantages and limitations are listed.

Method	Main applications in drug discovery	Selection of strengths (+) and weaknesses (-)	Refs.
X-ray crystal-lography	3D-structure at atomic resolution	+ Highly accurate structure determination + Suitable for large complexes − Need of high-quality homogeneous crystals − Examination of static structures − No quantitative affinity information	[1]
Protein-observed NMR	3D-structure at atomic resolution Dynamics Affinity	+ Measurement in solution + No crystal packing artifacts + Determination of affinity and dynamics − Large consumption of isotopically labeled protein − High expenditure of time − Suitable for small proteins (MW < 40 kDa)	[28,29]
Ligand-observed NMR	Interaction studies	+ Does not require isotopic labeling + Suitable for high-throughput studies + Can detect low affinity interactions − Limited structural information	[5]
MS	Molecular weight Structure Interaction studies	+ Highly sensitive and label-free + Low protein consumption − Desalting of sample required − Buffer limitations − No detergents tolerated − No 3D information − Difficulty to identify stereoisomers	[30,31,32]
CD	2D-structure determination Stability	+ Fast measurements and data acquisition + Non-destructive method − Careful sample preparation (exact protein concentration) − Protein sequence information is needed − No residue-specific information − Interference with solvent absorption possible	[33,34]
Stopped-flow	Kinetics	+ In solution measurements + Can examine slow and relatively fast kinetics − Dead time of instrument − Need of optical signals to monitor changes − Interpretation of data	[35,36]
SPR	Kinetics Affinity	+ Label-free real-time detection + Low sample consumption + High sensitivity − Immobilization of one interaction partner − High stability of immobilized sample required − System prone to artifacts	[37,38,39]
ITC	Thermodynamic Affinity Stoichiometry	+ Direct determination of thermodynamic parameters + Both interaction partners are free in solution + Recently applicable for kinetic analysis (see 5.5) − High protein consumption	[40]
MST	Affinity	+ Both interaction partners are free in solution + Low sample consumption − If no intrinsic fluorescence, labeling required	[22,23,41]
AUC	Size and shape Molecular weight Stoichiometry Affinity	+Versatile and non-destructive method + In solution measurement + No solvent and pH limitations − Sophisticated measurements and data analysis	[42]
TSA	Stability	+ *High-throughput* screening + Low protein consumption − Requirement of a fluorescent dye − Unsuitable for hydrophobic proteins − Artifacts due to fluorescence quenching & aggregation − No quantitative affinity information	[17,43]

Table 1. Important biophysical techniques in drug discovery.

References

1. Smyth, M.S. and Martin, J.H.J., X-ray crystallography. *Mol. Pathol.*, **2000**, *53*(1), 8-14.
2. Blundell, T.L. and Patel, S., High-throughput X-ray crystallography for drug discovery. *Curr. Opin. Pharmacol.*, **2004**, *4*(5), 490-496.
3. Smart, O.S., Horsky, V., Gore, S., Svobodova Varekova, R., Bendova, V., Kleywegt, G.J., and Velankar, S., Worldwide Protein Data Bank validation information: Usage and trends. *Acta Cryst. Section D*, **2018**, *74*(3), 237-244.
4. Marion, D., An introduction to biological NMR spectroscopy. *Mol. Cell. Proteomics*, **2013**, *12*(11), 3006-3025.
5. Fernandez, C. and Jahnke, W., New approaches for NMR screening in drug discovery. *Drug Discov. Today Technol.*, **2004**, *1*(3), 277-283.
6. Bodenhausen, G. and Ruben, D.J., Natural abundance nitrogen-15 NMR by enhanced heteronuclear spectroscopy. *Chem. Phys. Lett.*, **1980**, *69*(1), 185-189.
7. Martin, S.R. and Schilstra, M.J., Circular dichroism and its application to the study of biomolecules, in *Methods in cell biology*, **2008**, Elsevier, 263-293.
8. Braun, C.S., Fisher, M.T., Tomalia, D.A., Koe, G.S., Koe, J.G., and Middaugh, C.R., A stopped-flow kinetic study of the assembly of nonviral gene delivery complexes. *Biophys. J.*, **2005**, *88*(6), 4146-4158.
9. Biro, F.N., Zhai, J., Doucette, C.W., and Hingorani, M.M., Application of stopped-flow kinetics methods to investigate the mechanism of action of a DNA repair protein. *J .Vis. Exp.*, **2010**, (37), 1874.
10. Sauer, M.M., Jakob, R.P., Eras, J., Baday, S., Eriş, D., Navarra, G., Bernèche, S., Ernst, B., Maier, T., and Glockshuber, R., Catch-bond mechanism of the bacterial adhesin FimH. *Nat. Commun.*, **2016**, *7*, 10738.
11. Di Girolamo, F., Lante, I., Muraca, M., and Putignani, L., The role of mass spectrometry in the "omics" era. *Curr. Org. Chem.*, **2013**, *17*(23), 2891-2905.
12. Artigues, A., Nadeau, O.W., Rimmer, M.A., Villar, M.T., Du, X., Fenton, A.W., and Carlson, G.M., Protein structural analysis via mass spectrometry-based proteomics. *Adv. Exp. Med. Biol.*, **2016**, *919*, 397-431.
13. Annis, D.A., Nickbarg, E., Yang, X., Ziebell, M.R., and Whitehurst, C.E., Affinity selection-mass spectrometry screening techniques for small molecule drug discovery. *Curr. Opin. Chem. Biol.*, **2007**, *11*(5), 518-526.
14. Kuprowski, M.C., Boys, B.L., and Konermann, L., Analysis of protein mixtures by electrospray mass spectrometry: Effects of conformation and desolvation behavior on the signal intensities of hemoglobin subunits. *J. Am. Soc. Mass Spectrom.*, **2007**, *18*(7), 1279-1285.
15. Bianchi, F., Riboni, N., Termopoli, V., Mendez, L., Medina, I., Ilag, L., Cappiello, A., and Careri, M., MS-based analytical techniques: Advances in spray-based methods and EI-LC-MS applications. *J. Anal. Methods Chem.*, **2018**, *2018*, 24.
16. Nedelkov, D. and Nelson, R.W., Surface plasmon resonance mass spectrometry: Recent progress and outlooks. *Trends Biotechnol.*, **2003**, *21*(7), 301-305.
17. Niesen, F.H., Berglund, H., and Vedadi, M., The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability. *Nat. Protoc.*, **2007**, *2*(9), 2212-2221.
18. Ladbury, J.E., Klebe, G., and Freire, E., Adding calorimetric data to decision making in lead discovery: A hot tip. *Nat. Rev. Drug Discov.*, **2010**, *9*(1), 23-27.
19. Silvestre, H.L., Blundell, T.L., Abell, C., and Ciulli, A., Integrated biophysical approach to fragment screening and validation for fragment-based lead discovery. *PNAS*, **2013**, *110*(32), 12984-12989.
20. Kranz, J.K. and Schalk-Hihi, C., Protein thermal shifts to identify low molecular weight fragments. *Methods Enzymol.*, **2011**, *493*, 277-298.
21. Entzian, C. and Schubert, T., Mapping the binding site of an aptamer on ATP using microscale thermophoresis. *J. Vis. Exp.*, **2017**, (119).
22. Jerabek-Willemsen, M., André, T., Wanner, R., Roth, H.M., Duhr, S., Baaske, P., and Breitsprecher, D., Microscale thermophoresis: Interaction analysis and beyond. *J. Mol. Struct.*, **2014**, *1077*, 101-113.
23. Seidel, S.A., Dijkman, P.M., Lea, W.A., van den Bogaart, G., Jerabek-Willemsen, M., Lazic, A., Joseph, J.S., Srinivasan, P., Baaske, P., Simeonov, A., Katritch, I., Melo, F.A., Ladbury, J.E., Schreiber, G., Watts, A., Braun, D., and Duhr, S., Microscale thermophoresis quantifies

biomolecular interactions under previously challenging conditions. *Methods*, **2013**, *59*(3), 301-315.
24. Entzian, C. and Schubert, T., Studying small molecule-aptamer interactions using microscale thermophoresis (MST). *Methods*, **2016**, *97*, 27-34.
25. Cole, J.L., Lary, J.W., Moody, T., and Laue, T.M., Analytical ultracentrifugation: Sedimentation velocity and sedimentation equilibrium. *Methods Cell Biol.*, **2008**, *84*, 143-179.
26. Howlett, G.J., Minton, A.P., and Rivas, G., Analytical ultracentrifugation for the study of protein association and assembly. *Curr. Opin. Chem. Biol.*, **2006**, *10*(5), 430-436.
27. Ebel, C., Moller J.V., le Maire, M., Analytical ultracentrifugation: Membrane protein assemblies in the presence of detergent, in *Biophysical analysis of membrane proteins: Investigating structure and function*, **2008**, Wiley, 91-120.
28. Fielding, L., Rutherford, S., and Fletcher, D., Determination of protein–ligand binding affinity by NMR: Observations from serum albumin model systems. *Magn. Reson. Chem.*, **2005**, *43*(6), 463-470.
29. Rosato, A., Tejero, R., and Montelione, G.T., Quality assessment of protein NMR structures. *Curr. Opin. Struct. Biol.*, **2013**, *23*(5), 715-724.
30. Hofstadler, S.A. and Sannes-Lowery, K.A., Applications of ESI-MS in drug discovery: Interrogation of noncovalent complexes. *Nat. Rev. Drug Discov.*, **2006**, *5*(7), 585-595.
31. Vivat Hannah, V., Atmanene, C., Zeyer, D., Van Dorsselaer, A., and Sanglier-Cianferani, S., Native MS: An 'ESI' way to support structure- and fragment-based drug discovery. *Future Med. Chem.*, **2010**, *2*(1), 35-50.
32. Renaud, J.P., Neumann, T., Van Hijfte, L., Fragment-based drug discovery, in *Small molecule medicinal chemistry: Strategies and technologies*, **2016**, John Wiley & Sons, 221-249.
33. Greenfield, N.J., Using circular dichroism spectra to estimate protein secondary structure. *Nat. Protoc.*, **2006**, *1*(6), 2876-2890.
34. Kelly, S.M., Price N.C., Sample preparation and good practice in circular dichroism spectroscopy, in *Modern techniques for circular dichroism and synchontron radiation circular dichroism spectroscopy*, **2009**, IOS Press, 91-107.
35. Gilbert, S.P. and Mackey, A.T., Kinetics: A tool to study molecular motors. *Methods*, **2000**, *22*(4), 337-354.
36. Zheng, X., Bi, C., Li, Z., Podariu, M., and Hage, D.S., Analytical methods for kinetic studies of biological interactions: A review. *J. Pharm. Biomed. Anal.*, **2015**, *113*, 163-180.
37. Day, Y.S.N., Baird, C.L., Rich, R.L., and Myszka, D.G., Direct comparison of binding equilibrium, thermodynamic, and rate constants determined by surface- and solution-based biophysical methods. *Protein Sci.*, **2002**, *11*(5), 1017-1025.
38. Rich, R.L. and Myszka, D.G., Survey of the year 2007 commercial optical biosensor literature. *J. Mol. Recognit.*, **2008**, *21*(6), 355-400.
39. Nguyen, H.H., Park, J., Kang, S., and Kim, M., Surface plasmon resonance: A versatile technique for biosensor applications. *Sensors (Basel)*, **2015**, *15*(5), 10481-10510.
40. Fulroth, B., Kaushik V.K., Mesleh M.F., Modern biophysical methods for screening and drug discovery, in *High-troughput screening methods: Evolution and refinement*, **2017**, The Royal Society of Chemistry, 117-161.
41. Seidel, S.A., Wienken, C.J., Geissler, S., Jerabek-Willemsen, M., Duhr, S., Reiter, A., Trauner, D., Braun, D., and Baaske, P., Label-free microscale thermophoresis discriminates sites and affinity of protein–ligand binding. *Angew. Chem. Int. Ed. Engl.*, **2012**, *51*(42), 10656-10659.
42. den Engelsman, J., Garidel, P., Smulders, R., Koll, H., Smith, B., Bassarab, S., Seidl, A., Hainzl, O., and Jiskoot, W., Strategies for the assessment of protein aggregates in pharmaceutical biotech product development. *Pharm. Res.*, **2011**, *28*(4), 920-933.
43. Lo, M.C., Aulabaugh, A., Jin, G., Cowling, R., Bard, J., Malamas, M., and Ellestad, G., Evaluation of fluorescence-based thermal shift assays for hit identification in drug discovery. *Anal. Biochem.*, **2004**, *332*(1), 153-159.

5. Surface Plasmon Resonance (SPR)

5.1 SPR – A short historical overview

Surface plasmon resonance (SPR) is a optical sensing technique that involves a collective charge-density oscillation at a metal-dielectric interface.[1] In 1902, Wood initially observed the effect of SPR while allowing polarized light to be incident on a metallic diffraction grating. He noticed some puzzling dark bands in the reflected light but did not associate this optical phenomenon with surface plasmons.[2] It was the work of Otto,[3] Kretschmann, and Raether [4] that completely described SPR in 1968. They also established the instrumental basis - the Otto configuration and the Kretschmann-Raether configuration - for the investigation of surface plasmons. In 1983, Liedberg and coworkers applied an SPR-based biosensor for the first time to detect antibody–antigen interactions [1] and seven years later Pharmacia Biosensor AB (today GE Healthcare Life Science) launched the first SPR instrument (BIACORE 1000).[5,6] Thereupon, SPR rapidly developed to become a powerful label-free method for real-time monitoring of biomolecular interactions in liquids.[7] For many years, Biacore was undisputed market leader for SPR devices. Nowadays, other suppliers (e.g. Bio-Rad, ForteBio, Horbia Scientific, Reichert Analytical Instruments) also brought SPR instruments on the market.[8]

5.2 Principles of SPR

SPR is based on total internal reflection (TIR) a physical effect that occurs in non-absorbing media when light travels from a medium with a higher refractive index (e.g. glass, n_1) to a medium with a lower refractive index (e.g. air, buffer, n_2). A light beam that hits a half circular prism with an incident light angle of 0 ($\Theta = 0$) exits the flat prism site straight into the air. Increasing Θ gives rise to refracted beams until the critical angle of incidence is reached. This occurs when the refracted beam travels between the two media. Above the critical incidence angle the beam fully reflects within the prism without any light passing through (Figure 1-A).[9] In case of TIR, an electromagnetic field emerges in the medium of lower refractive index without a loss in net energy of the reflected beam. This phenomenon is called an evanescent field wave that has the same wavelength as the incident light. The amplitude of the wave exponentially decreases with increasing distance from the media boundary.[10,11]

Nowadays, the most prevalent operation mode in SPR devices, including Biacore instruments, is still the Kretschmann-Raether configuration. In this configuration, the reflection site of a prism is coated with a non-magnetic metal like gold. Under TIR, the *p*-polarized component of the evanescent field wave (the electrical field component lying in the plane of incidence) penetrates the thin metal layer and SPR can occur.[11] In the resonance event, free electrons at the gold surface completely absorb the light photon energy and convert it into collective oscillating metal electrons (also referred to as surface plasmons, surface plasmon polaritons or surface plasma waves), a form of electromagnetic energy.[12,13,14] This requires that the wave vector of the photons equals the wave vector of surface plasmons that is dependent on the refractive index of the metal and the adjacent lower refractive index medium.[15] Surface plasmons propagate on the surface of the metal layer and create an enhanced evanescent wave. Simultaneously a sharp decrease in reflection intensity is observed (Figure 1-B, right).[16,17,18] The angle at which the greatest loss in intensity of the reflected light emerges is defined as the SPR angle (Θ = SPR angle).[19] The shape and location of the SPR dip can be used to convey information about the sensor surface. In comparison to the evanescent wave field generated by TIR at a non-coating interface, the SPR-generated evanescent wave field is enhanced. It is still *p*-polarized and exponentially decays out from the gold (Figure 1-B, left).[20,21]

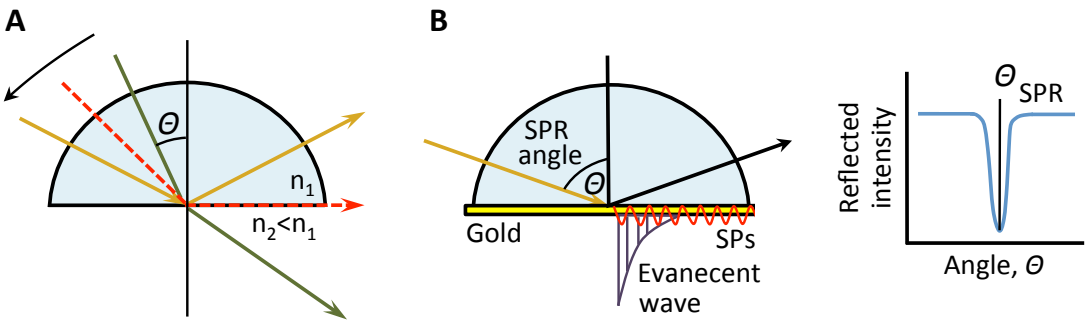


Figure 1: Principles of SPR. A) A light beam that travels from a dense to a less dense medium is refracted (green line). When the refracted light travels along the boundary of the two media, the corresponding incidental angle is referred to as the critical angle (dashed red line). An incident angle greater than the critical angle results in total internal reflection (TIR, yellow line). B) Prism coated with a gold layer (Kretschmann-Raether configuration) visualizing surface plasmon resonance (left); SPs, surface plasmons. An intensity decrease of reflected light is observed at the SPR angle (right). Pictures modified from References [11,20].

The resonance event is dependent on the incidence angle, the wavelength of light and the refractive index close to the gold surface. SPR can be received by variation of the incident light angle and the refractive index of the medium adjacent to the gold layer

while keeping the metal and its thickness constant. Consequently, changes in the adjacent medium can be monitored by adjusting the incident light angle until the occurrence of SPR.[22,23]

5.3 Biacore T200

The Biacore (Bia standing for biomelecular / biospecific interaction analysis) technology exploits SPR for the detection of differences of the refractive index caused by a change in mass concentration close to a surface of a sensor chip.[24] The key components of all Biacore instruments are a SPR optical detection unit, a sensor chip, and a microfluidic handling system. All SPR experiments in this thesis were performed with a Biacore T200.

The Biacore T200 is a highly automated instrument with an exceptional sensitivity and was designed for high-quality characterization of molecular interactions. The sensor chip together with its protective cassette is introduced into the T200 via the sensor chip port and is placed between the prism and the integrated μ-fluidic cartridge (IFC) within the instrument. Samples and reagents are hold in removable racks that are placed in a temperature-controlled compartment that occupies an autosampler and a sample injection needle. The injection needle is part of the liquid handling system that delivers samples and buffers to the flow cells. Bottles containing the running buffer are placed on the left tray whereas a waste bottle and a bottle containing water are placed on the right tray.

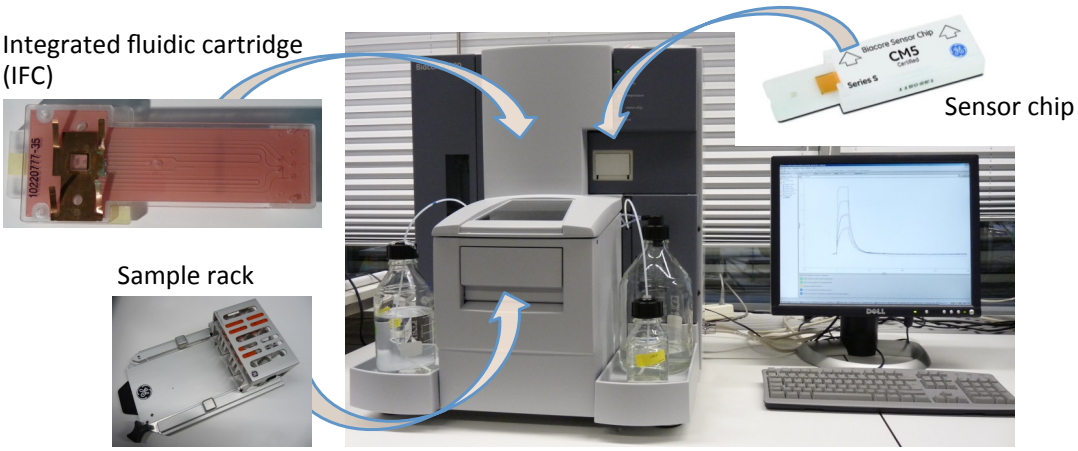


Figure 2: The Biacore T200 device. Prior to a SPR measurement the sample rack is placed in the temperature-controlled sample compartment and the sensor chip is inserted into the instrument over the sensor chip. Within the T200 the sensor chip is pressed against the integrated μ-fluidic cartridge (IFC).

The water bottle is supposed to rinse the injection needle between sample injections. Operation of the instrument, data collection and evaluation is handled from a PC running Biacore T200 Control Software and Biacore T200 Evaluation Software (Figure 2).

The Biacore optical unit is composed of a diode emitting polarized light at 760 nm [25] that passes through a semi-cylindrical prism and hits a sensor chip, a glass surface coated with a thin gold layer (50 nm). The glass side of the sensor chip is in optical contact with the prism, while its gold layer is facing the flow cells, which are formed by the IFC on the chip surface (Figure 3). In case of the T200, the IFC creates 4 parallel flow cells where different biomolecular interactions can take place.

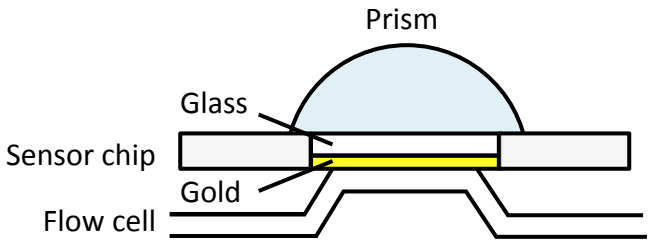


Figure 3: Simplified illustration of a Biacore setup. The prism is a curial component of the optical unit. It directs the incident light onto the sensor chip that consists of a glass layer (detector side) and a gold layer (sample side). The IFC forms the flow cell and is in direct contact with the sensor chip. The figure is modified from Reference [21].

The IFC consist of several micro channels that can be opened and closed through computer-controlled pneumatic microvalves. It is the heart of the microfluidic handling system that controls the delivery of samples and buffers to the sensor chip surface and allows the analyte to pass over the chip surface in a pulse-free, continuous, and automated flow. A regular maintenance of the instrument retards the attrition of liquid handling components, especially of the IFC, and supports the generation of precise and high quality data.[26]

5.4 A typical SPR biosensor experiment and underlying principles

In a biosensor experiment, one binding partner (the ligand) is immobilized on the sensor chip surface whereas the other binding partner (the analyte) is free in solution.[27] A beam of incident light is focused on the ligand-immobilized gold layer and is totally reflected at the glass-gold interface. The generated evanescent wave

penetrates into the flow cell and enables the detection of refractive index properties near the gold layer. During the measurement, a diode array detector [25] continuously records the position of reduced intensity of reflected light and thus is able to calculate the SPR angle (Figure 4-A). Once the analyte is injected into the flow cell, it binds to the immobilized ligand, with the consequence that in turn the refractive index changes close to the sensor chip surface. Since surface plasmon waves are very sensitive to a change in the refractive index, the SPR angle alters and hence the position of the dark band (Figure 4-B).

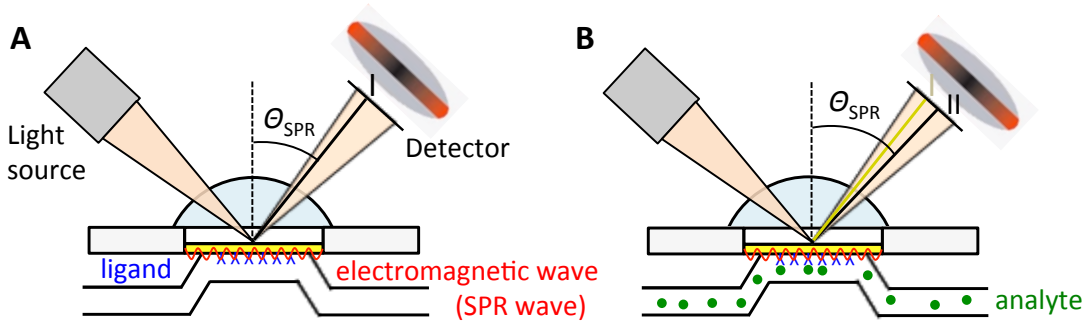


Figure 4: Biacore experiment. The SPR angle shifts from position I (A) to position II (B) when analyte bind to the immobilized ligand. The figure is modified from Reference [21].

The change in SPR angle is proportional to the mass of material bound and is reported as response units (RU), where 1000 RU corresponds to an SPR angle shift of ~ 0,0001° or a change in the refractive index of 10^{-3}.[28] In the case of proteins, 1000 RU represents the binding of 1 ng protein per square nm to the sensor chip surface (Equation 1).[29]

$$1\ RU = 1\ pg/mm^2 \qquad \text{(Equation 1)}$$

A binding curve of a small molecule–protein interaction can be visualized in a sensorgram, a plot of resonance units (RU) against time. A typical sensorgram comprises four phases: an association phase, an equilibrium phase, a dissociation phase, and a regeneration phase. In order to generate high quality data, it is also important to mention that at the beginning of each experiment the system should be primed with running buffer to obtain a stable baseline. The injection of analyte allows the formation of analyte–ligand complexes. In the association phase, analyte binding to the surface-immobilized ligand dominates, resulting in an increase of response units until association balances dissociation. At the end of analyte injection, only

running buffer flows over the surface with the consequence that dissociation of the analyte outweighs, accompanied by a decrease in response units. A regeneration step is indispensable if the analyte is not completely removed from the chip surface (Figure 5).[27,30,31] Regeneration is performed with high ionic strength buffers, detergents, and mild basic (NaOH) or acidic (glycine-HCl) solutions.[32] It is also noteworthy that the ligand solution not only flows over the sample chip surface, but also over a reference cell surface, which is either blank or immobilized with an inactive ligand. Additionally, buffer blanks are included prior to and throughout the measurement in order to get rid of baseline drifts and injection artifacts.

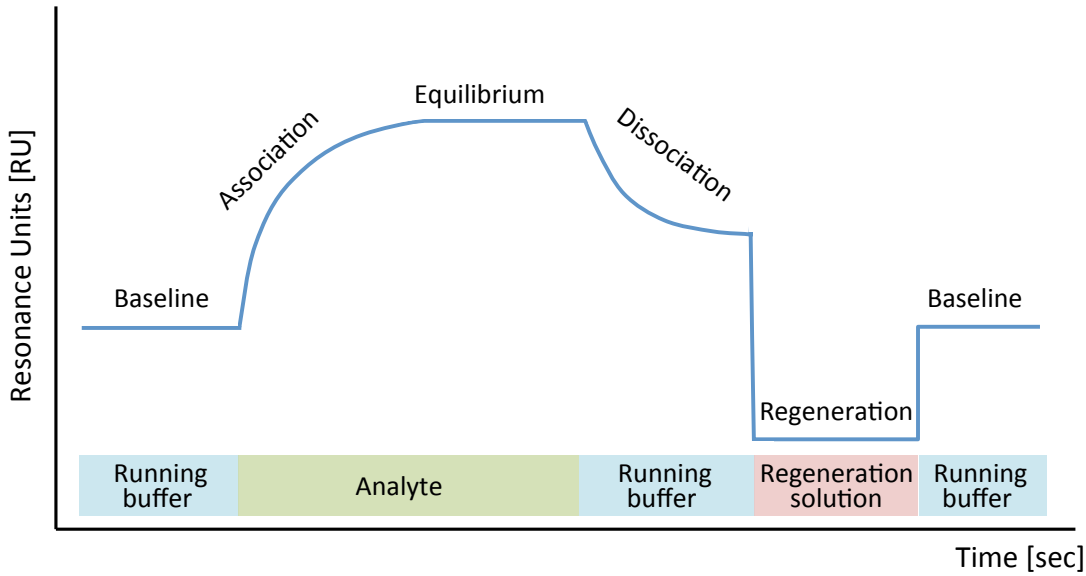


Figure 5: Schematic representation of a typical SPR sensorgram. In the association phase, analyte binds to the immobilized ligand until an equilibrium state is reached. A formed ligand–analyte complex dissociates when analyte injected is completed and only running buffer is in the mobile phase. Regeneration of the chip surface is often applied to restore a stable baseline.

Subsequent to a kinetic run, all sensorgrams are processed using the double referencing procedure that has a favorable impact on the quality of the data.[33] In a first step the reference response is subtracted from the response of the sample cell to compensate for matrix effects, refractive index effects, and non-specific binding of the analyte. In a second data processing step, the response from the average of the blank sensorgrams is subtracted from the already referenced sensorgrams.

The association rate constant (k_{on}, the rate of complex formation) and the dissociation rate constants (k_{off}, the rate of complex decay) can be obtained from a sensorgram by applying suitable curve fitting procedures in the BIAevaluation software.[34]

5.5 Kinetic measurements

In order to obtain a complete kinetic profile of a biomolecular interaction, different analyte concentrations have to be injected in a single experiment due to k_{on}, which is concentration dependent (see also Section 1.6). In order to receive reliable data, analyte concentrations should cover a range between 0.1 to 10 times the K_D, which roughly corresponds to about 10% to 90% of ligand saturation. Prior to a measurement the single concentrations are manually prepared by diluting the analyte in running buffer.[35] A conventional Biacore running buffer consists of 10 mM HEPES at pH 7.4, 150 mM NaCl and 0.005% P20 (polysorbate 20). The purpose of the salt is to suppress electrostatic effects on the carboxylated sensor chip matrix while the surfactant P20 should reduce adsorption of hydrophobic molecules to flow system surfaces. Depending on the assay EDTA is also added to chelate divalent metal ions that might be present in the running buffer.[36]

The kinetics can be determined by multi-cycle kinetics (MCK), the standard kinetic assay comprising several cycles, where each analyte concentration is injected in a separate cycle (Figure 6-A). If the analyte is not completely detached after the dissociation phase, a regeneration step needs to be included to ensure equal starting conditions for the next cycle. To avoid multiple regeneration steps, which sometimes harm or even irreversibly inactivate the immobilized ligand (especially proteins),[37,38] a single-cycle kinetics (SCK) can be performed. During a SCK measurement the analyte is injected with increasing concentrations and short dissociation times in between. A long dissociation period takes place after the last analyte injection.

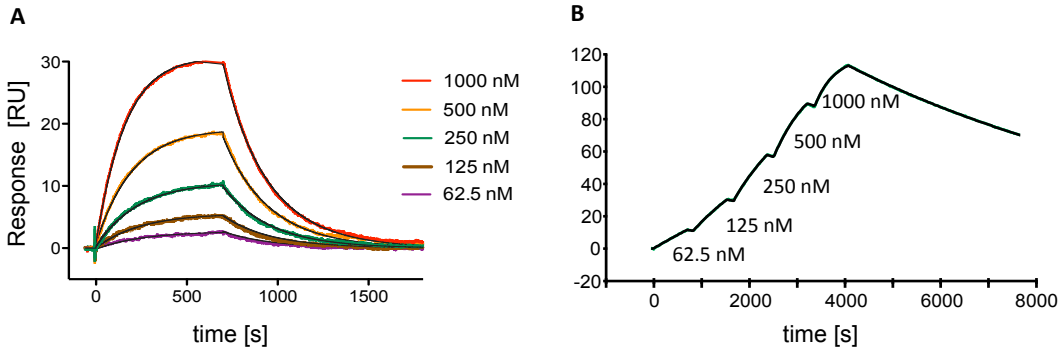


Figure 6. Kinetic SPR measurements. The sensorgrams display a multi-cycle kinetics (MCK) (A) and a single-cycle kinetics (SCK) (B) experiment.

In the end of the measurement, the surface can be regenerated. During the SCK there are no buffer injections or regeneration steps (Figure 6-B).[39]

Compared to MCK, it could be affirmed that SCK provides kinetic results with the same precision.[40] The SCK approach requires less time for a complete analysis, and benefits from the lack of multiple regeneration steps during the measurement. However, SCK is more sensitive to baseline drifts, which are for instance caused by the dissociation of ligand in a capture assay, due to its comparatively long cycle time.

5.6 Interaction mechanisms and kinetic models

There are various kinetic models that operators can apply to calculate k_{on} and k_{off}. The most commonly used model is the Langmuir binding model that describes a 1:1 interaction of an analyte (Y) with a ligand (Z) to form a complex (YZ) (Equation 2). In the Langmuir binding model, all binding sites are considered to be independent and equivalent.[41,42]

$$Y + Z \underset{k_{off}}{\overset{k_{on}}{\rightleftharpoons}} YZ \qquad \text{(Equation 2)}$$

The on-rate k_{on} is a second-order rate constant in units of per molar per second ($M^{-1}s^{-1}$). Generally, the formation of a protein–small molecule complex is rapid. Though, considering that two binding partners need to collide before binding, k_{on} is diffusion-limited and possesses about $5x10^9$ $M^{-1}s^{-1}$ as maximum value for two molecules in solution.[43] Due to favorable electrostatic forces the diffusion limit can also be exceeded.[44] The association of the complex is concentration dependent, while k_{off} does not depend on concentrations of the interaction partners. The off-rate is a first-order rate constant measured in per second (s^{-1}).[45] Based on k_{off}, the dissociative half-life $t_{1/2}$ (time required for 50% of the initially present analyte–ligand complexes to dissociate, Equation 3) or residence time τ (average time an analyte stays bound to its ligand, Equation 4) of a ligand–analyte complex can be calculated.[45,46]

$$t_{1/2} = \ln 2/k_{off} \qquad \text{(Equation 3)}$$

$$\tau = 1/k_{off} \qquad \text{(Equation 4)}$$

In a kinetic Langmuir analysis, k_{off} can be obtained by exponential curve fitting of the dissociation phase. Fast-on and slow-off kinetics are indicative for strong affinities. In a simple bimolecular 1:1 interaction, the equilibrium dissociation constant (K_D) can be directly calculated from the ratio of k_{off} and k_{on} (Equation 5).

$$K_D = k_{off}/k_{on} \qquad \text{(Equation 5)}$$

Additionally, the K_D can be indirectly determined by an equilibrium analysis, providing that the steady state of association and dissociation is reached. If so, the theoretical binding capacity R_{max} is detected at various concentrations.[31,47]

For more complex binding reactions the two-state reaction model is often applied.[48,49,50] This model represents 1:1 binding of analyte to the sensor chip-immobilized ligand with a binding-induced conformational change of one or both reaction partners, where complex YZ undergoes transition into complex YZ^* (Equation 6).

$$Y + Z \underset{k_{off,1}}{\overset{k_{on,1}}{\rightleftarrows}} YZ \underset{k_{off,2}}{\overset{k_{on,2}}{\rightleftarrows}} YZ^* \qquad \text{(Equation 6)}$$

The on and off rate constants of the complex YZ formation and dissociation are $k_{on,1}$ and $k_{off,1}$, whereas $k_{on,2}$ and $k_{off,2}$ describe the conformational change. For reasons of simplification, the model assumes that complex YZ^* solely emerges from and dissociates to complex YZ.[51]

It should be noted that conformational changes usually do not cause a response signal change in Biacore. However, a very pronounced conformational change, e.g. spherical protein adopts an elongated conformation, might imply a refractive index shift pretending a mass concentration alteration on the sensor surface.

In the two-state reaction model, the overall K_D is determined according to equation 7.[51]

$$K_D = k_{off,1}/k_{on,1} \times (k_{off,2}/(k_{off,2}+k_{on,2})) \qquad \text{(Equation 7)}$$

5.7 Immobilization strategies

Depending on its nature, the ligand is attached to the sensor surface by either covalent immobilization, high affinity capture, or hydrophobic absorption (Figure 7).[52]

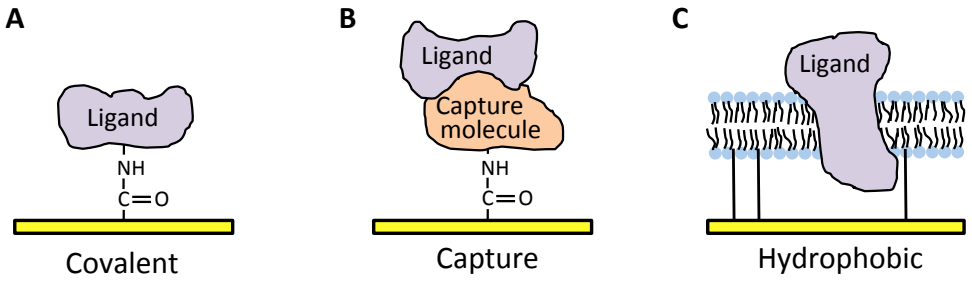


Figure 7: Methods for ligand attachment to the sensor chip surface. The ligand can be covalently immobilized to the sensor chip surface (A) or is captured by a molecule that is immobilized to the surface (B). A ligand and/or a hydrophobic carrier (lipid monolayer or bilayer) are/is attached by hydrophobic adsorption (C). The illustration is modified from Reference [52].

The first two approaches are used in several studies of this thesis and are explained in more details below (Section 1.7.1 & 1.7.2). Especially for covalent immobilization and capturing approaches, sensor chips covered by a 30-100 nm thick and flexible carboxymethyl dextran matrix are used (Figure 8).[53,54] This three-dimensional layer increases the possible immobilization level of ligands and exploits SPR to the greatest feasible extent. Biacore sensor chips possessing only the dextran matrix belong to the CM-series and are listed in Section 1.7.3.[52]

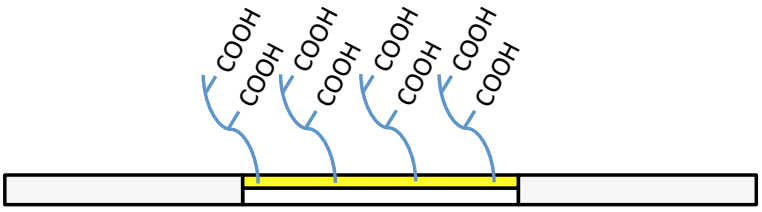


Figure 8: Schematic representation of a carboxymetylated dextran matrix. The matrix covers the surface of a number of Biacore sensor chips (Section 1.7.3).

5.7.1 Covalent immobilization approaches

Covalent immobilization involves irreversible attachment of the ligand to the sensor surface by a covalent chemical link. It is a conventional method used to immobilize ligands and molecules for capturing approaches. The covalent attachment of ligands results in stable surfaces without any detachment of ligands during interaction analysis. This has also the advantage that only bound analyte molecules are removed

during a regeneration cycle. If small molecules are covalently coupled, harsh regeneration solutions are usually unproblematic. The carboxymethylated dextran matrix provides a foundation for numerous covalent immobilization chemistries.

Amine coupling: The most commonly used covalent immobilization method is amine coupling. After the surface is activated with 1-ethyl-3-(3-dimethylaminopropyl)-carbodiimide (EDC) and N-hydroxysuccinimide (NHS) the formed NHS-esters react with primary amine groups of the ligand (Figure 9).

In case of protein ligands, amine groups being part of the protein-binding site often have a negative impact on the activity of the immobilized interaction partner. To avoid this problem, either the binding site can be protected during the immobilization process or another immobilization should be applied. Moreover, amine coupling often gives rise to heterogeneous immobilization results because most proteins have several amine groups.

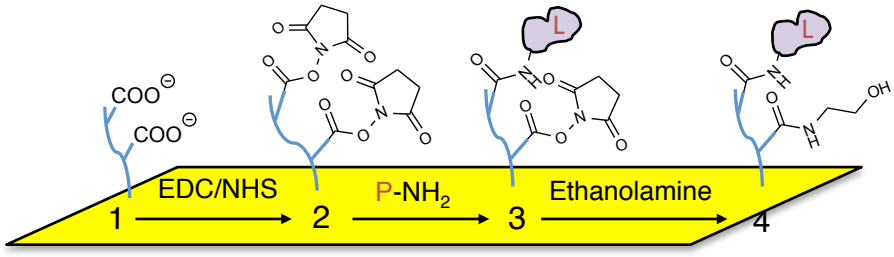


Figure 9. Amine coupling. The dextran surface (1) is activated with EDC/NHS. NHS-esters (2) covalently couple to primary amine groups of a protein (3). Finally, the cell surface is blocked with ethanolamine (4).

Thiol coupling: To obtain a homogenous orientation of the immobilized interaction partner, thiol coupling could be the method of choice. In contrast to amines, thiol groups are usually less common in proteins. Immobilization of thiol-containing ligands can be conducted by the formation of disulfide bridges either with a ligand or a surface thiol coupling approach. In case of ligand thiol coupling, the ligand harbors free thiol groups while the sensor surface has active disulfide groups. In contrast, the chemical conditions for surface thiol coupling are vice versa (Figure 10-A). The formed disulfide bridges are stable under physiological conditions but can be reduced by *ß*-mercaptoethanol or dithiothreitol as well as by a high pH. In this way, a complete regeneration of the chip surface is possible. However, the unstable characteristic under these reducing conditions might negatively influence the

measurements. In such circumstances, operators can make use of maleimide coupling. Maleimides form covalent non-reducible thioester bonds with free thiol groups of the ligand (Figure 10-B).

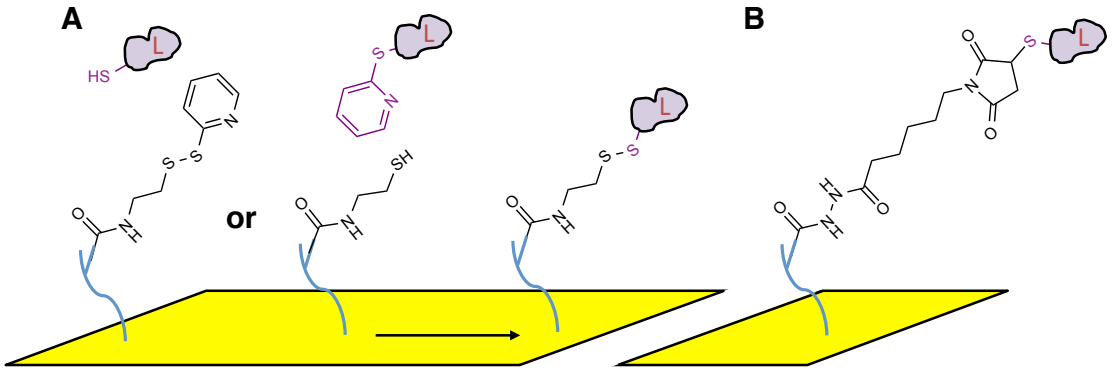


Figure 10. Covalent coupling methods for thiol containing ligands. Ligand and surface thiol coupling (A) and covalent ligand immobilization by maleimide coupling.

Aldehyde coupling: This method is suitable for ligands having aldehyde groups or ligands containing cis-diols and sialic acids residues (e.g. polysaccharides, glycoproteins, and glycoconjugates) that can be easily oxidized to aldehydes.[55] The aldehyde residues are able to couple to a hydrazine-activated surface. To obtain a stable surface, a reduction step with cyanoborohydride is performed at the end of the immobilization procedure (Figure 11). However, ligands might suffer from oxidation and reduction steps with a consequent negative impact on their activity.

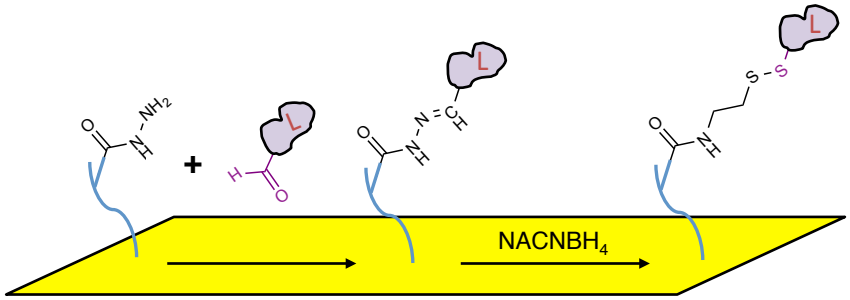


Figure 11. Aldehyde coupling of aldehyde containing ligands.

One should be aware that covalent coupling approaches chemically modify ligands. The modification may influence ligand activity and can result in a decrease or even an affinity loss of the studied interaction. This should be taken into account, especially if small molecules are immobilized. In these unpleasant cases, capturing approaches provide an alternative strategy (Section 1.7.2).

5.7.2 Capturing approaches

Capturing approaches enable that ligands are captured by non-covalent interaction with a capturing molecule. In the majority of cases, the capturing molecule is covalently attached to the sensor surface and mostly ensures that ligands are bound in an orientation-specific manner with the result of a homogeneous chip surface. Additionally, a damage of the ligand is unlikely because the capture can be performed under physiological conditions.

Streptavidin-Biotin Capture: One of the strongest known biological interactions is the non-covalent binding of biotin to streptavidin (K_D about 10^{-15}).[56] The high affinity streptavidin-biotin interaction is rather resistant to extreme pH and temperature conditions, denaturing and detergent agents, and organic solvents.[57] Biacore utilizes this strong interaction to capture biotinylated ligands to a sensor chip surface with covalently attached streptavidin (Figure 12-A). Ready-made streptavidin chips (Sensor Chip SA) can be ordered or streptavidin is covalently immobilized to the sensor chip surface by the operator. The dissociation of streptavidin–biotin caught ligands is usually negligible during a measurement. Biotin can be introduced into a molecule by several methods. One method is the insertion of a unique 15 amino acid peptide taq (AviTaq peptide) into the ligand. Subsequently, the AviTaq peptide can be covalently biotinylated by the biotin ligase (BirA) from *E.coli*.[58]

Antibody-based capture: An antibody based capture uses a specific antibody as capturing molecule. High affinity antibodies against a variety of antigens are commercially available. They can be easily immobilized by amine coupling and their regeneration with low pH containing buffers (e.g. glycine-HCl) is usually unproblematic. After a regeneration cycle ligands and ligand–analyte complexes are not longer available and the capturing molecule needs to bind new ligand for the next cycle. Thus an increased amount of ligand is needed, but one can be sure that the ligand is not damaged by regeneration solutions. Due to the high selectivity for their antigen, antibodies are suitable for the capture of ligands from impure samples or blood plasma.[52]

Instead of covalent immobilization, antibodies from several mammalian species (including human IgG1, IgG2, and IgG4) can be captured by pre-immobilized Protein A. After the antibody (capturing molecule) is bound with its Fc-region to Protein A, the antigen (ligand) can be injected and captured by the antibody (Figure 12-B). Regeneration of the protein A surface is performed using low pH buffers that completely remove the capturing molecule-ligand-analyte complex.[59,60]

Capturing tagged proteins: For purification purposes, proteins are often tagged with poly-histidine that enables binding to Ni^{2+}-nitrilotriacetic acid (NTA) complexes. This principle can also be exploited for capturing His-tagged ligands on a NTA chip surface (Figure 12-C). According to the manufacturer, optimal binding can be obtained by using tags with at least 6 histidine residues. However, the K_D for a single His_6-tagged protein with Ni^{2+}-NTA is only about 10^{-6} M at neutral pH and usually proteins with one His-tag rapidly dissociate from the Ni^{2+}-NTA surface.[61] A more stable surface can be prepared using proteins with a double-His_6 tag (two His_6 separated a short peptide sequence).[61,62] An advantage Ni^{2+}-NTA is the possibility to remove the captured ligands with the chelating agent EDTA. Alternatively, anti-His antibodies can be used to capture His-tagged proteins.

Other prominent protein tags for ligand capturing are FLAG[63] and glutathione-S-transferase (GST) [64] that can be captured by an anti-FLAG mAb and anti-GST antibody surface, respectively.

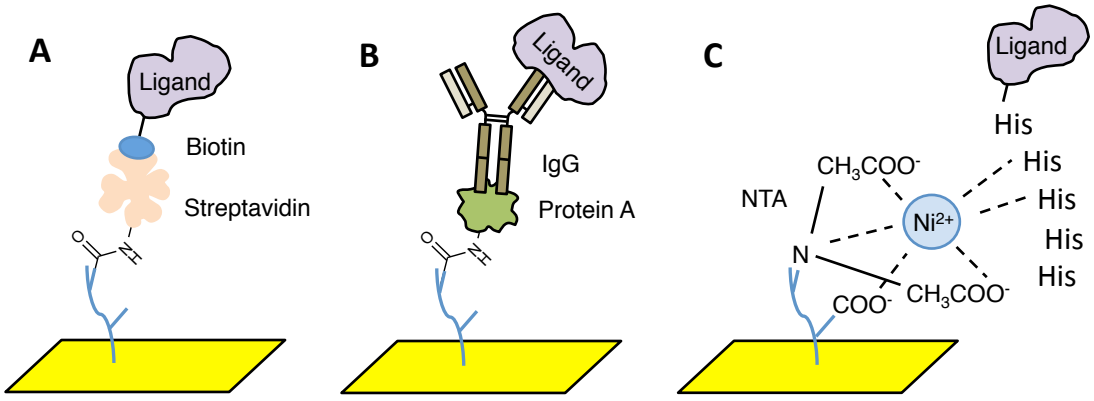


Figure 12. High affinity capturing methods. An SPR method to capture ligands is streptavidin-biotin capturing (A). Protein A binds IgG (capturing molecule) that in turn captures the ligand (B). The right illustration displays binding of a His-tagged protein to a Ni^{2+}-NTA chip surface (C).

5.7.3 Biacore sensor chip surfaces

The following table gives an overview of the different Biacore sensor chips.[52]

Chip	Surface	Application
AU	plain gold surface	custom design
CM3	short dextran matrix	immobilization of large ligands
CM4	standard dextran matrix with a low degree of carboxylation	immobilization of small ligands
CM5	standard dextran matrix	general approach
CM7	standard dextran matrix with a high degree of carboxylation	reaching high immobilization levels
C1	no matrix, low capacity	prevention of dextran interactions
SA	streptavidin immobilized	capture of biotinylated analytes
L1	lipid capturing surface	capture of vesicles and liposomes while maintaining the lipid bilayer
NTA	NTA	capture of His-tagged ligands
HPA	octadecanethiol	hydrophobic adsorption of a polar lipid monolayer
Protein A	Protein A immobilized	capture of IgG1, IgG2, and IgG4

Table 1: Overview of Biacore sensor chips. The table lists chip surfaces and their basic application.[52]

In addition, Xantec bioanalytics also offers a wide range of SPR sensor chips including various polymers, densities, thicknesses, and capture molecules.[65]

5.8 Assay design

The examination of small molecule–protein interactions raises the question whether the small molecule or the protein should be immobilized. According to the scientific issue and properties of the interaction partners, it is an individual decision. In drug discovery, a small molecule has an organic origin and an upper molecular weight of around 900 Da.[66] Proteins are macromolecules that most commonly have molecular weights in the range of 5.5 kDa to 220 kDa.[67] Therefore, higher signal responses are achieved when the small molecule is immobilized.

However, small molecule analytes and high response signals are simultaneously possible by using a sandwich assay. Solution competitive or surface competitive assays can be performed as well. A sandwich assay works with an enhancement molecule (a macromolecule) that specifically binds to the ligand-bound analyte and thus increases the SPR signal (Figure 13-B).[68,69] In a solution competitive assay a macromolecule (detecting molecule) binds to both the analyte and the ligand (the

ligand is equal to the analyte) (Figure 13-C) whereas in a surface competitive assay the macromolecular competitor (detecting molecule) only interacts with the ligand (the ligand is unequal to the analyte) (Figure 13-D). The injection of different analyte concentrations over the ligand, while keeping the concentration of the detecting molecule constant, yields various SPR signals. A higher analyte concentration results in lower changes of response units. In this way, the SPR response can be inversely set in relation to individual analyte concentrations.[36,69,70]

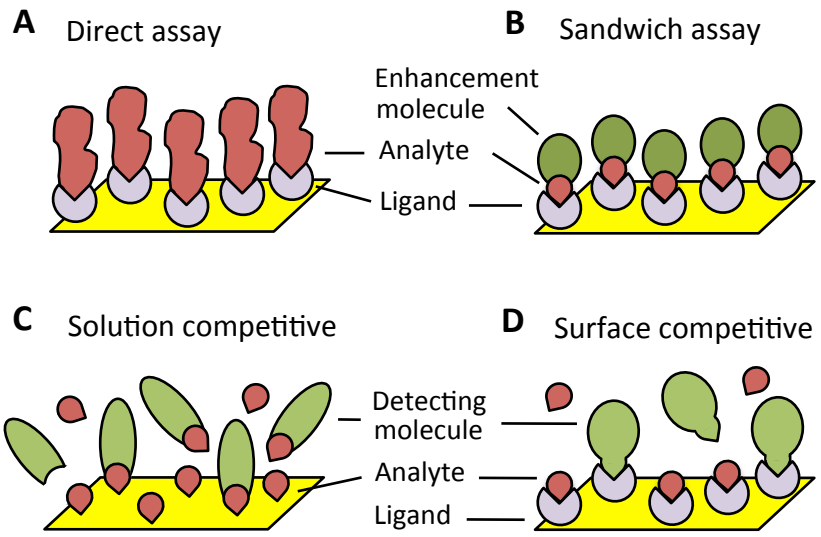


Figure 13: Various SPR assay principles. The schematic illustrations display a direct assay (A), a sandwich assay (B), a solution competitive assay (C), and a surface competitive assay (D). The figure is modified from Reference [69].

In order to estimate eventual experimental signal responses, a pivotal variable is the theoretical analyte binding capacity R_{max}. It can be calculated as follows:

$$R_{max} = (MW_{analyte}/MW_{Ligand}) \times R_L \times S_m \qquad \text{(Equation 8)}$$

where R_{max} is the theoretical analyte binding capacity (RU), $MW_{analyte}$ and MW_{ligand} is the molecular weight of the analyte and the ligand, respectively, R_L is the immobilized level of the ligand (RU), and S_m the stoichiometric ratio.[71]

5.9 Bioanalytical and biophysical applications of SPR

Apart from affinity and kinetic information, the Biacore instrument can also indirectly determine thermodynamic parameters of a molecular interaction. SPR is also a helpful tool to characterize recombinant proteins and monoclonal antibodies in terms of their

structural integrity and activity. Generally, the active concentration of the biotherapeutic protein is tested by calibration-free concentration analysis. In addition, its binding activity to antigens and Fc receptors is verified. When comparing sensorgrams of the latest protein batch with reference sensorgrams, information about drug stability and potency can be obtained.[36]

The application of SPR in protenomics aims for investigations of protein–protein interactions in complex biomolecular networks. In this field, SPR devices are often coupled with MALDI-MS and LC-MS/MS to gain qualitative information about the assayed proteins. In this manner, intramolecular interactions and structure–activity relationships can be ascertained. The newly acquired knowledge may have an important impact on further proteomic research or might facilitate the breakthrough of drug discovery studies.[72]

A newer application area of SPR biosensors is the direct examination of specific ligands in blood plasma. This is a challenging issue due to the high number of biomolecules that might compete with the analyte for binding to the ligand or could unspecifically interact with the sensor chip matrix. Positive experiences have been made with some surface coatings like polymers and peptides that decrease artifactic absorption of unspecific blood plasma proteins and simultaneously function as anchor for the ligand. The usage of metallic nanoparticles as plasmonic materials instead of gold surfaces has enabled a resolution increase and sensitivity gain of SPR measurements. In contrast to the gold layer that induces propagating SPR, nanoparticles are based on localized surface plasmon resonance (LSPR).[73]

In addition, studies exist describing SPR as an effective tool for medical diagnostic analysis in order to verify biological markers. Validated biomarkers for the detection of cancer are for example interleukin-8,[74] the protein vascular endothelial growth factor,[75] and the carcinoembryonic antigen.[76] Moreover, pathogen specific antibodies can be detected and the conduction of both vitamin and hormone analyses are also possible with SPR biosensors.[15] The vision of the future is the usage of miniaturized SPR devices that can be applied for rapid examination of a patient`s blood sample in healthcare facilities.[77]

References

1. Liedberg, B., Nylander, C., and Lunström, I., Surface plasmon resonance for gas detection and biosensing. *Sens. Actuators*, **1983**, *4*, 299-304.
2. Wood, R.W., Xlii. On a remarkable case of uneven distribution of light in a diffraction grating spectrum. *Lond. Edinb. Dubl. Phil. Mag.*, **1902**, *4*(21), 396-402.
3. Otto, A., Excitation of nonradiative surface plasma waves in silver by the method of frustrated total reflection. *Z. Phys.*, **1968**, *216*(4), 398-410.
4. Kretschmann, E. and Raether, H., Notizen: Radiative decay of non radiative surface plasmons excited by light. *Z. Naturforsch. A*, **1968**, *23*(12), 2135.
5. Liedberg, B., Nylander, C., and Lundström, I., Biosensing with surface plasmon resonance – how it all started. *Biosens. Bioelectron.*, **1995**, *10*(8), i-ix.
6. Löfås, S., Malmqvist, M., Rönnberg, I., Stenberg, E., Liedberg, B., and Lundström, I., Bioanalysis with surface plasmon resonance. *Sens. and Actuators B Chem.*, **1991**, *5*(1), 79-84.
7. Malmqvist, M., Biospecific interaction analysis using biosensor technology. *Nature*, **1993**, *361*, 186.
8. Stahelin, R.V., Surface plasmon resonance: A useful technique for cell biologists to characterize biomolecular interactions. *Mol. Biol. Cell*, **2013**, *24*(7), 883-886.
9. Mello, A.J.d., Surface analytical techniques for probing biomaterial processes. 1 ed, ed. Davies, J. **1996**, CRC Press, Boca Raton, New York. 3-5.
10. Tang, Y., Zeng, X., and Liang, J., Surface plasmon resonance: An introduction to a surface spectroscopy technique. *J. Chem. Educ.*, **2010**, *87*(7), 742-746.
11. Markey, F., Principles of surface plasmon resonance, in *Real-time analysis of biomolecular interactions* **2000**, Springer 13-14.
12. Liu, Q., Wu, C., Wang, P., Other cell-based biosensors, in *Cell-based biosensors: Principles and applications*, **2010**, 207-229.
13. Kihm, K., Kim S., Near-fiel thermometry, in *Thermimetry at the nanoscale: Techniques and selected applications*, Royal Society of Chemistry, 315-339.
14. Wijaya, E., Lenaerts, C., Maricota, S., Hastanin, J., Habraken, S., Vilcota J.P., Boukherroub, R., Szunerits, S., Surface plasmon resonance-based biosensors: From the development of different SPR structures to novel surface functionalization strategies. *Curr. Opin. Solid State Mater. Sci.*, **2011**, *15*(5), 208-224.
15. Homola, J., Surface plasmon resonance sensors for detection of chemical and biological species. *Chem. Rev.*, **2008**, *108*(2), 462-493.
16. Nguyen, H.H., Park, J., Kang, S., and Kim, M., Surface plasmon resonance: A versatile technique for biosensor applications. *Sensors (Basel)*, **2015**, *15*(5), 10481-10510.
17. Guo, X., Surface plasmon resonance based biosensor technique: A review. *J. Biophotonics*, **2012**, *5*(7), 483-501.
18. Bakhtiar, R., Surface plasmon resonance spectroscopy: A versatile technique in a biochemist's toolbox. *J. Chem. Educ.*, **2013**, *90*(2), 203-209.
19. Tudos, A.J., Schasfoort, R.B.M, Capter 1: Introduction to surface plasmon resonance, in *Handbook of surface plasmon resonance*, **2008**, Springer, 2.
20. Islam, S., Kouzani, A., J. Dai, X., and Michalski, W., Parameter sensitivity analysis of surface plasmon resonance biosensor through numerical simulation. *Conference Paper: IEEE/ICME*, **2010**, 171-176.
21. Mitchell, J., Small molecule immunosensing using surface plasmon resonance. *Sensors (Basel)*, **2010**, *10*(8), 7323-7346.
22. Michaelis, S., Robelek, R., Wegener J., Studying cell–surface interactions in vitro: A survey of experimental approaches and techniques in *Tissue engineering III: Cell–surface interactions for tissue culture*, **2012**, Springer-Verlag Berlin Heidelberg, 43.
23. Biacore, Surface plasmon resonance. *Technology Note 1*, **2001**.
24. Jonsson, U., Fagerstam, L., Ivarsson, B., Johnsson, B., Karlsson, R., Lundh, K., Lofas, S., Persson, B., Roos, H., Ronnberg, I., and et al., Real-time biospecific interaction analysis using surface plasmon resonance and a sensor chip technology. *Biotechniques*, **1991**, *11*(5), 620-627.
25. Mislovičová J., W.F., and Danielsson B., Chapter 5: Bioanalytical studies based an lectin–carbohydrate interactions measured by ellipsometry and surface plasmon resonance techniques, in *Lectins: Analytical technologies*, **2007**, Elsevier, 105.
26. Biacore T200 instrument handbook, Biacore, GE Healthcare, Uppsala,. **2010**.

27. Hahnefeld, C., Drewianka, S., and Herberg, F.W., Determination of kinetic data using surface plasmon resonance biosensors. *Methods Mol. Med.*, **2004**, *94*, 299-320.
28. Sadana, A., Chapter 12: Surface plasmon resonance biosensors in engineering biosensors: Kinetics and design applications. **2002**, Elevisir. 404.
29. Albrecht, M.T., Wang, W., Shamova, O., Lehrer, R.I., and Schiller, N.L., Binding of protegrin-1 to Pseudomonas aeruginosa and Burkholderia cepacia. *Respir. Res.*, **2002**, *3*, 18.
30. Retra, K., Irth, H., and van Muijlwijk-Koezen, J.E., Surface plasmon resonance biosensor analysis as a useful tool in FBDD. *Drug Discov. Today: Technol.*, **2010**, *7*(3), e181-e187.
31. Ritzefeld, M. and Sewald, N., Real-time analysis of specific protein–DNA interactions with surface plasmon resonance. *J. Amino Acids*, **2012**, *2012*, 816032.
32. Pollack, S.J., Surface plasmon resonance for identifying and characterising small molecule ligands in *Biophysical techniques in drug discovery*, **2018**, The Royal Society of Chemistry 194.
33. Myszka, D.G., Improving biosensor analysis. *J. Mol. Recognit.*, **1999**, *12*(5), 279-284.
34. Karlsson, R. and Fält, A., Experimental design for kinetic analysis of protein–protein interactions with surface plasmon resonance biosensors. *J. Immunol. Methods*, **1997**, *200*(1), 121-133.
35. Marquart, J.A., SPRpages - getting a feeling for the curves, in *Handbook of surface plasmon resonance*, **2017**, 110.
36. Biacore assay handbook, Biacore, GE Healthcare, Uppsala. **2012**.
37. Andersson, K., Areskoug, D., and Hardenborg, E., Exploring buffer space for molecular interactions. *J. Mol. Recognit.*, **1999**, *12*(5), 310-315.
38. Scharenberg, M., Jiang, X., Pang, L., Navarra, G., Rabbani, S., Binder, F., Schwardt, O., and Ernst, B., Kinetic properties of carbohydrate–lectin interactions: FimH antagonists. *ChemMedChem*, **2014**, *9*(1), 78-83.
39. Biacore T200 software handbook. Biacore, GE Healthcare, Uppsala, . **2010**.
40. Karlsson, R., Katsamba, P.S., Nordin, H., Pol, E., and Myszka, D.G., Analyzing a kinetic titration series using affinity biosensors. *Anal. Biochem.*, **2006**, *349*(1), 136-147.
41. Langmuir, I., The constitution and fundamental properties of solids and liquids. Part I. Solids. *JACS*, **1916**, *38*(11), 2221-2295.
42. O'Shannessy, D.J., Brigham-Burke, M., Soneson, K.K., Hensley, P., and Brooks, I., Determination of rate and equilibrium binding constants for macromolecular interactions using surface plasmon resonance: Use of nonlinear least squares analysis methods. *Anal. Biochem.*, **1993**, *212*(2), 457-468.
43. Berg, O.G. and von Hippel, P.H., Diffusion-controlled macromolecular interactions. *Annu. Rev. Biophys. Biophys. Chem.*, **1985**, *14*(1), 131-158.
44. Gabdoulline, R.R. and Wade, R.C., Biomolecular diffusional association. *Curr. Opin. Struct. Biol.*, **2002**, *12*(2), 204-213.
45. Corzo, J., Time, the forgotten dimension of ligand binding teaching. *Biochem. Mol. Biol. Educ.*, **2006**, *34*(6), 413-416.
46. Pan, A.C., Borhani, D.W., Dror, R.O., and Shaw, D.E., Molecular determinants of drug–receptor binding kinetics. *Drug Discov. Today*, **2013**, *18*(13), 667-673.
47. Myszka, D.G., Jonsen, M.D., and Graves, B.J., Equilibrium analysis of high affinity interactions using Biacore. *Anal. Biochem.*, **1998**, *265*(2), 326-330.
48. Muñoz, E., Xu, D., Kemp, M., Zhang, F., Liu, J., and Linhardt, R.J., Affinity, kinetic, and structural study of the interaction of 3-O-sulfotransferase isoform 1 with heparan sulfate. *Biochemistry*, **2006**, *45*(16), 5122-5128.
49. Yakovenko, O., Tchesnokova, V., Sokurenko, E.V., and Thomas, W.E., Inactive conformation enhances binding function in physiological conditions. *PNAS*, **2015**, *112*(32), 9884-9889.
50. Foster, J.S., Williams, A.D., Macy, S., Richey, T., Stuckey, A., Wooliver, D.C., Koul-Tiwari, R., Martin, E.B., Kennel, S.J., and Wall, J.S., A peptide-Fc opsonin with pan-amyloid reactivity. *Front. Immunol.*, **2017**, *8*(1082).
51. Lipschultz, C.A., Li, Y., and Smith-Gill, S., Experimental design for analysis of complex kinetics using surface plasmon resonance. *Methods*, **2000**, *20*(3), 310-318.
52. Biacore sensor surface handbook. Biacore, GE Healthcare, Uppsala. **2016**.
53. Kyprianou, D., Guerreiro, A.R., Chianella, I., Piletska, E.V., Fowler, S.A., Karim, K., Whitcombe, M.J., Turner, A.P., and Piletsky, S.A., New reactive polymer for protein immobilisation on sensor surfaces. *Biosens. Bioelectron.*, **2009**, *24*(5), 1365-1371.
54. Wofsy, C. and Goldstein, B., Effective rate models for receptors distributed in a layer above a surface: Application to cells and Biacore. *Biophys. J.*, **2002**, *82*(4), 1743-1755.

55. O'Shannessy, D.J. and Wilchek, M., Immobilization of glycoconjugates by their oligosaccharides: Use of hydrazido-derivatized matrices. *Anal. Biochem.*, **1990**, *191*(1), 1-8.
56. Green, N.M., Avidin and streptavidin. *Methods Enzymol.*, **1990**, *184*, 51-67.
57. Tytgat, H.L.P., Schoofs, G., Driesen, M., Proost, P., Van Damme, E.J.M., Vanderleyden, J., and Lebeer, S., Endogenous biotin-binding proteins: An overlooked factor causing false positives in streptavidin-based protein detection. *Microb. Biotechnol.*, **2015**, *8*(1), 164-168.
58. Fairhead, M. and Howarth, M., Site-specific biotinylation of purified proteins using BirA. *Methods Mol. Biol.*, **2015**, *1266*, 171-184.
59. Choe, W., Durgannavar, T.A., and Chung, S.J., Fc-binding ligands of immunoglobulin g: An overview of high affinity proteins and peptides. *Materials*, **2016**, *9*(12), 994.
60. Moore, G.L., Bautista, C., Pong, E., Nguyen, D.-H.T., Jacinto, J., Eivazi, A., Muchhal, U.S., Karki, S., Chu, S.Y., and Lazar, G.A., A novel bispecific antibody format enables simultaneous bivalent and monovalent co-engagement of distinct target antigens. *mAbs*, **2011**, *3*(6), 546-557.
61. Nieba, L., Nieba-Axmann, S.E., Persson, A., Hamalainen, M., Edebratt, F., Hansson, A., Lidholm, J., Magnusson, K., Karlsson, A.F., and Pluckthun, A., Biacore analysis of histidine-tagged proteins using a chelating NTA sensor chip. *Anal. Biochem.*, **1997**, *252*(2), 217-228.
62. Khan, F., He, M., and Taussig, M.J., Double-hexahistidine tag with high-affinity binding for protein immobilization, purification, and detection on ni-nitrilotriacetic acid surfaces. *Anal. Chem.*, **2006**, *78*(9), 3072-3079.
63. Einhauer, A. and Jungbauer, A., The flag peptide, a versatile fusion tag for the purification of recombinant proteins. *J. Biochem. Biophys. Methods*, **2001**, *49*(1-3), 455-465.
64. Hutsell, S.Q., Kimple, R.J., Siderovski, D.P., Willard, F.S., and Kimple, A.J., High affinity immobilization of proteins using biotin- and GST-based coupling strategies. *Methods Mol. Biol.*, **2010**, *627*, 75-90.
65. https://www.xantec.com/
66. Veber, D.F., Johnson, S.R., Cheng, H.Y., Smith, B.R., Ward, K.W., and Kopple, K.D., Molecular properties that influence the oral bioavailability of drug candidates. *J. Med. Chem.*, **2002**, *45*(12), 2615-2623.
67. Nagai, R., and Taniguchi, N., Amino achids and proteins, in *Medical biochemistry*, **2014**, Elsevier, 13.
68. Sim, H.R., Wark, A.W., and Lee, H.J., Attomolar detection of protein biomarkers using biofunctionalized gold nanorods with surface plasmon resonance. *Analyst*, **2010**, *135*(10), 2528-2532.
69. Zhang, P., Chen, Y.-P., Wang, W., Shen, Y., and Guo, J.-S., Surface plasmon resonance for water pollutant detection and water process analysis. *Trends Anal. Chem.*, **2016**, *85*, 153-165.
70. Shankaran, D.R., Gobi, K.V., and Miura, N., Recent advancements in surface plasmon resonance immunosensors for detection of small molecules of biomedical, food and environmental interest. *Sens. Actuators B Chem.*, **2007**, *121*(1), 158-177.
71. Brogioni, B. and Berti, F., Surface plasmon resonance for the characterization of bacterial polysaccharide antigens: A review. *MedChemComm*, **2014**, *5*(8), 1058-1066.
72. de Mol, N.J., Surface plasmon resonance for proteomics. *Methods Mol. Biol.*, **2012**, *800*, 33-53.
73. Jatschka, J., Dathe, A., Csáki, A., Fritzsche, W., and Stranik, O., Propagating and localized surface plasmon resonance sensing – a critical comparison based on measurements and theory. *Sens. Biosensing Res.*, **2016**, *7*, 62-70.
74. Yang, C.Y., Brooks, E., Li, Y., Denny, P., Ho, C.M., Qi, F., Shi, W., Wolinsky, L., Wu, B., Wong, D.T., and Montemagno, C.D., Detection of picomolar levels of interleukin-8 in human saliva by SPR. *Lab Chip*, **2005**, *5*(10), 1017-1023.
75. Li, Y., Lee, H.J., and Corn, R.M., Detection of protein biomarkers using RNA aptamer microarrays and enzymatically amplified surface plasmon resonance imaging. *Anal. Chem.*, **2007**, *79*(3), 1082-1088.
76. Chung, J.W., Park, J.M., Bernhardt, R., and Pyun, J.C., Immunosensor with a controlled orientation of antibodies by using neutravidin-protein a complex at immunoaffinity layer. *J. Biotechnol.*, **2006**, *126*(3), 325-333.
77. Puiu, M. and Bala, C., SPR and SPR imaging: Recent trends in developing nanodevices for detection and real-time monitoring of biomolecular events. *Sensors*, **2016**, *16*(6), 870.

6. Isothermal Titration Calorimetry (ITC)

6.1 ITC - A short historical overview

Calorimetry is defined as the measurement of heat at constant pressure.[1] All physical, chemical, and biological processes that either release (exothermic reaction) or absorb (endothermic reaction) heat are amenable to a calorimetric characterization.[2] A calorimeter is used to detect quantities of heat transferred to or from an investigational sample. The first reported calorimeter, an ice calorimeter, dated back to 1783 and was designed by Lavoisier and Laplace. The ice calorimeter measured the heat given off by a guinea pig as the amount of water originating from an ice jacket surrounding the measurement chamber. To prevent an influence from the ambient temperature, an ice-water chamber shielded the ice jacket (Figure 1).[3]

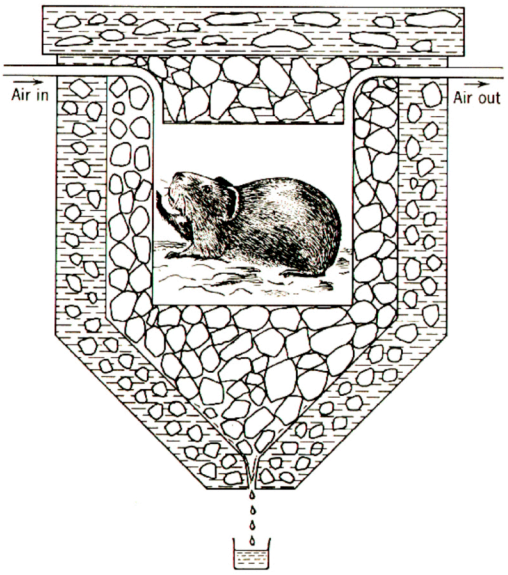


Figure 1: The ice calorimeter. The ice calorimeter is the first reported calorimeter and was developed by Lavoisier and Laplace, who published their invention in 1783. The picture is taken from Reference [4].

Titration calorimetry had its beginnings about 50 years ago when Christensen and Izatt simultaneously determined the equilibrium dissociation constant (K_D) and enthalpy change (ΔH) of metal–cyanide complexation reactions by titrating one component into the other.[5,6] At the same time, titration calorimeters that run under constant temperature were developed.[7,8] Sensitivity improvements of the instruments enabled the study of biological systems in the 1980s.[9,10,11] The first commercially available isothermal titration calorimeter was released by Microcal in 1989. It was called a device for ”determining K_{eq} in minutes”, where K_{eq} is the equilibrium binding

constant. It was specifically designed for the study of biological systems.[12] From this date onward, ITC developed from a specialist technique into a standard tool to directly characterize thermodynamic parameters.[13] Nowadays, high-sensitive isothermal titration calorimeters (VP-ITC and Microcal iTC200 from Malvern as well as Nano ITC Standard Volume and Nano ITC Low Volume from TA Instruments) are routinely used in drug discovery. Recently launched devices (PEAQ ITC from Malvern, and Affinity ITC by TA Instruments) feature even better sensitivity and volume requirements.[14]

6.2 Binding Thermodynamics

ITC is a versatile technique to analyze the physical basis of molecular binding interactions under constant temperature (isothermal conditions). During a measurement, heat generation or absorption is measured, which is equivalent to the change in enthalpy.[15] In addition to ΔH, the equilibrium dissociation constant K_D, and the binding stoichiometry n can be directly determined.[16] K_D ($K_D = 1/K_A$) is related to the change in free energy of binding that is also known as the change in Gibbs free energy (ΔG), which can be calculated by the Gibbs` equation (Equation 1):[17]

$$\Delta G = \Delta G^o - R \cdot T \cdot \ln K_D \qquad \text{(Equation 1)}$$

where R is the gas constant (8.314 $JK^{-1}mol^{-1}$), T the absolute temperature in Kelvin, and $\Delta G°$ the free energy change under standard conditions. Standard conditions are a concentration of 1.0 M for each reactant, a temperature of 25°C, and a pressure of 1 atm.[18] The binding reaction is in equilibrium if $\Delta G = 0$.

$$0 = \Delta G^o - R \cdot T \cdot \ln K_D \qquad \text{(Equation 2)}$$

$$\Delta G^o = R \cdot \mathrm{T} \cdot \ln K_D = - R \cdot T \cdot \ln K_A \qquad \text{(Equation 3)}$$

Under standard conditions the change of entropy $\Delta S°$ can be readily deduced from $\Delta H°$ and $\Delta G°$ (Equation 4) without being dependent on van't Hoff approximations.

$$\Delta S^o = (\Delta H^o - \Delta G^o)/T \qquad \text{(Equation 4)}$$

$$\Delta G^o = \Delta H^o - \mathrm{T}\Delta S^o \qquad \text{(Equation 5)}$$

Changes in both enthalpy and entropy determine whether a non-covalent molecular interaction is spontaneous (thermodynamically favorable). This is the case when $\Delta G°$ is negative ($\Delta G° < 0$). Many different combinations of ΔH and ΔS can result in identical Gibbs free energy and consequently the same binding affinity.[19]

Since enthalpy is commonly dependent on temperature, the change in heat capacity (ΔC_p) of a binding interaction can be determined as well. This is done by repeating a titration at different temperatures under otherwise identical reaction conditions (Equation 8).[20]

$$\Delta C_p = (\partial \Delta H°)/\partial T \qquad \text{(Equation 8)}$$

6.2.1 Enthalpic components

Enthalpically driven forces of a bimolecular interaction involve heat release or consumption. Enthalpic contributions are electrostatic interactions, such as hydrogen bonds, salt bridges, halogen bonds, stacking with π-systems, and van der Waals interactions.[21] The energy of an electrostatic interaction can be approximated using the Coulomb's law (Equation 9):

$$E = \frac{q_1 q_2}{4\pi r \varepsilon_0 \varepsilon_r} \qquad \text{(Equation 9)}$$

where E is the electrostatic energy, q_1 and q_2 are changes of two interacting atoms, r is the distance between q_1 and q_2, ε_0 is the vacuum permittivity, and ε_r is the dielectric constant of the surrounding medium.[22]

Electrostatic interactions involving permanent or induced dipoles are known as van der Waals forces. A single van der Waals interaction typically ranges from −2 kJ mol^{-1} to −4 kJ mol^{-1}. However, van der Waals forces are omnipresent and the impact of several van der Waals bonds can significantly contribute to the stabilization of protein–ligand interactions.[23] Hydrogen bonds are dipole-dipole interactions that are typically stronger than van der Waals forces. They are formed between a hydrogen donor group $X^{\delta-}$–$H^{\delta+}$ and an electronegative hydrogen bond acceptor $A^{\delta-}$.[24] The free energy for hydrogen bonding is in the range between –6.5 kJ mol^{-1} to −19.5 kJ mol^{-1}.[25] Hydrogen bonds are directional and their strength is dependent on

bond length and bond angle. As defined by the IUPAC recommendations of 2011, the strength of the hydrogen bond increases with decreasing distance between the binding partners and an approximation of the donor angle to 180°.[26] Prevalent hydrogen bonds formed in protein–ligand complexes are N-H···O, O−H···O, and N−H···N and feature a median bond length of about 2.8 Å between donor and acceptor heavy atoms.[27] A salt bridge is a hydrogen bond that involves an oppositely and permanently charged ion pair that is within 5 Å. Its strength heavily depends on the environment and particularly buried salt bridges can even exceed an enthalpy contribution of −40 kJ mol^{-1}.[27,28] Halogen bonds can be formed between an electrophilic region of a halogen, the donor (D), and a nucleophile, the acceptor (A), yielding a R−D···A non-covalent contact. The acceptor is commonly a lone pair possessing atom, an anion, or a π-system.[29] In a *π-system*, π-electrons generate a negatively charged periphery that surrounds a positively charged core (*σ*-framework).[30] This electrostatic distribution enables aromatic-aromatic interactions of the π-system-containing amino acids phenylalanine, tyrosine, tryptophan and histidine. These interactions can provide a substantial amount of binding enthalpy. As an example, π-π stacking between neutral histidine and the other aromatic amino acids gives rise to an enthalpy change between -13.0 kJ mol^{-1} to -17.0 kJ mol^{-1}.[31] Furthermore, π-systems of aromatic acids are able to attract cations (cation–π-interaction)[32], halogens, and molecules with permanent dipoles such as water. The (water) OH···π interaction is considered to be a non-conventional hydrogen bond.[33] In addition to their contribution to the binding free energy, interactions involving aromatic rings are also of importance for protein–ligand recognition.[34]

6.2.2 Entropic components

Entropy describes the flexibility of a system and is either favorable or unfavorable.[35,36] The change in entropy related to a small molecule–protein interaction is expressed as the sum of various entropic contributions, as demonstrated in Equation 10.

$$\Delta S = \Delta S_{\text{solvation}} + \Delta S_{\text{conformation}} + \Delta S_{\text{rotational/translational}} \qquad \text{(Equation 10)}$$

Thereby, $\Delta S_{\text{solvation}}$ reflects the solvation entropy that usually favorably contributes to

the binding entropy because it represents the solvent release to bulk upon binding of two interaction partners. The component $\Delta S_{conformation}$ describes the change in conformational entropy and constitutes the alteration of conformational freedom of both binding partners in comparison to their unbound state, which may be a favorable or an unfavorable binding entropy contribution.[19,37] In some cases favorable conformational entropy may decrease the entropic cost of binding. The confinement in translational and rotational freedom is expressed by $\Delta S_{rotational/translational}$. It describes the reduction of free particles in solution and consequently is unfavorable.[19]

6.2.3 Enthalpy-entropy compensation

In the drug development process, a frequent occurrence is enthalpy–entropy (H–S) compensation in aqueous solutions. It describes a linear correlation between enthalpy and entropy changes of interactions between proteins and small molecules.[38,39]

A molecular change in the ligand leading to an additional and/or tighter contact between the interaction partners results in an enthalpic improvement. However, newly formed non-covalent interactions can concurrently lead to a decrease in flexibility in either or both interaction partners. As a consequence, the loss in the overall conformational entropy compensates the gain in enthalpy giving rise to similar values for $\Delta G°$.[40]

Frequently, solvent water molecules are responsible for H–S compensation effects. The hydration shell of both protein and ligand has to be partially desolvated, a enthalpically unfavorable process, to enable direct interactions between the binding partners.[41] In this process, the penalty for desolvation enthalpies is more pronounced for polar groups than for nonpolar groups. However, an associated water-entropy gain ($\Delta S_{solvation}$) due to the release of tightly water molecules into bulk water can compensate the loss of enthalpy.[21]

6.3 Isothermal titration calorimeters

An isothermal calorimeter is equipped with a sample cell containing one binding partner (titrate) and a motor-driven syringe that provides the other binding partner (titrant). The syringe is provided with a rotating paddle that serves as a titrant delivery

and stirring device. In addition to the sample cell, the microcalorimeter possesses a reference cell filled with water. Both cells are identical and made of chemically inert and thermally conducting material.[42] They are surrounded by a thermostated adiabatic jacket and are equilibrated to the same temperature before a measurement. During a run the titrant is injected into the sample cell and heat is released or absorbed in case of complex formation between titrant and titrate. The temperature change in the sample cell is recognized by a heat-sensing component (thermocouple) that determines the temperature differences (ΔT) between the two cells. In order to consistently maintain the sample and reference cell at exactly the same temperature, a proportional-integral-derivative (PID) translates ΔT to the power that should be supplied to a feedback heater located on the sample cell to compensate for this difference [43] while the requirement of cooling occurs passively by heat exchange with the adiabatic jacket (Figure 2-A).[44,45]

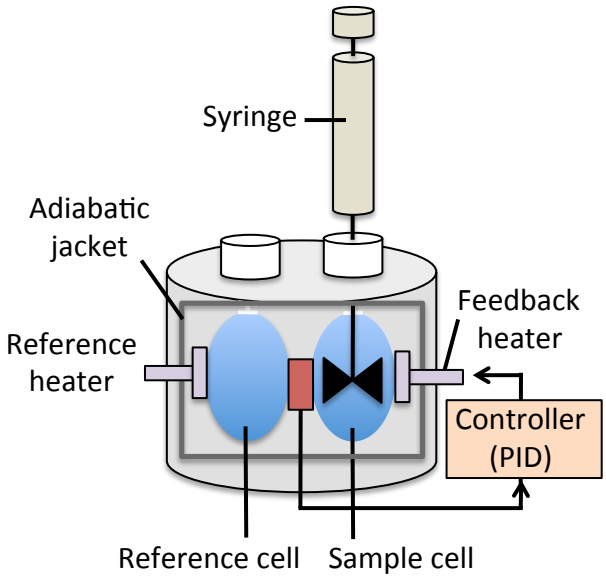


Figure 2: Setup of an isothermal titration calorimeter. The calorimeter is based on the heat-conduction principle, i.e. the power change is directly recorded upon titration of one binding partner into the other. The illustration is modified from Reference [43].

The calorimeters MicroCal iTC200 and MicroCal VP-ITC used in our studies are both from Malvern and execute titrations with the heat-conduction principle as described above. They operate with nanowatt sensitivity.[45] MicroCal iTC200 is a small volume isothermal calorimeter with low sample consumption. The calorimeter is equipped with a semi-automated system that also facilitates the maintenance procedure.[44]

Figure 3: ITC instruments from Malvern. The pictures display a MicroCal iTC200 (left) and a MicroCal VP-ITC (right) isothermal titration calorimeter.

6.4 ITC binding assays

6.4.1 A typical ITC experiment

A careful preparation of samples is a prerequisite to receive reliable thermodynamic parameters that accurately reflect the binding interaction. Prior to every measurement, a soluble titrant aliquot (usually a small molecule if a small molecule–protein interaction is studied) is drawn up into the syringe that is subsequently inserted into the sample cell containing a soluble sample of the other binding partner (commonly a macromolecule). In the course of an experiment, known amounts of the titrant are repetitively injected at isothermal conditions into the sample cell. The formation of a complex gives rise to the release (in case of an exothermic reaction) or the absorption (in the event of an endothermic reaction) of heat.[43,46] Heat alterations are universal and in the most cases they amount to submillions of a degree.[20] The change in the cell feedback power, used to keep a constant temperature, is referred as differential power (DP, with units of μcal sec^{-1}) that is recorded as output signal as a function of time displaying spikes at each titrant injection. (Figure 4-A).[47] To evaluate the measurement, the heat per injection is determined by integrating the DP from the respective spike over time. Afterwards, the single injection heats are normalized by the amount of titrant injected. The integrated and normalized heats per injection are used to fit the binding isotherm as a function of molar ratio [Titrant]/[Titrate] in the sample cell. Based on the binding isotherm, the thermodynamic parameters n, K_D, and ΔH of a binding process can be derived (Figure 4-B).[15]

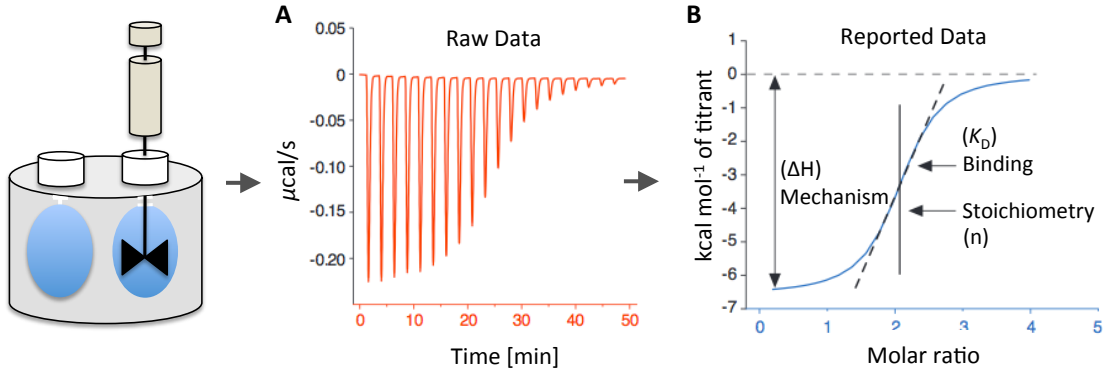


Figure 4: ITC raw data and its analysis. (A) ITC raw data. (B) The blue curve displays the binding isotherm as function of the molar ligand–protein ratio. A simple bimolecular interaction (A + B ⇔ AB) uses a 1:1 binding model to fit the normalized heats per injection. The illustration and plots are modified from References [43,48].

6.4.2 Displacement titration - An assay for tight & very low interactions

ITC can be applied to a wide variety of binding systems were a heat change occurs including protein–protein [49,50] and protein–small molecule [51,52] interactions. However, the dissociation constant of high-affinity interactions ($K_D < 10^{-8}$ M) cannot be determined directly by a single experiment. Due to a very steep slope of the binding isotherm, the direct determination of K_D is unreliable and requires a more complex approach. In this case, a displacement titration experiment is performed, in which a competing weaker ligand artificially reduces the binding affinity of the tight-binding ligand. Thereby, three different titrations have to be conducted (i) a direct titration of the high-affinity ligand (investigational ligand) to the target protein, (ii) a direct titration of the weaker ligand (competing ligand) to the target protein, and (iii) a displacement titration of the high-affinity ligand to the titrate mixture consisting of competing ligand and target protein.[53,54] All three titrations are conducted under the same experimental conditions.[54] The same procedure can be performed with very low-affinity bimolecular interactions, for which the transition of the binding isotherm in a direct titration is ill-defined.[55,56]

6.4.3 Possible emerging undesirable heat effects

The assessed enthalpy in an ITC run not only comprises the binding heat, but also heats arising due to molecular reorganization and conformational changes of one or both interaction partners.[57] Thus, the calculated change in entropy is also incorrect when taking into consideration that $\Delta S°$ is deduced from $\Delta H°$ and $\Delta G°$ (Equation 4).

In some cases, heat effects not involved in the binding interaction itself can also occur. These adverse heats can result from friction and turbulence due to the injection of ligand, the mixing procedure *per se*, or is depended on the dilution of both samples. Unfavorable heats of dilution can either arise when the titrant solution is much higher concentrated than titrate solution or when the pH or ionic strength of titrant and titrate buffer solutions are mismatched. A water into water titration can unveil heat effects due to friction and turbulence and also helps to recognize impurities, for instance due to precipitated protein samples.[54]

6.5 A further ITC application – KinITC

KinITC is a new method developed by Prof. Philippe Dumas to gain not only thermodynamic parameters but also kinetic information, i.e. the association rate constant (k_{on}) and the dissociation constant (k_{off}) of a binding reaction.[58] The implementation of kinITC-ETC (equilibration time curve) into the software Affinimeter made the method commercially accessible. Affinimeter deduces k_{off} from the equilibrium time of each individual peak while classical processing of the corresponding ITC raw data and fitting of the isotherm provides the K_D. Finally, k_{on} can be calculated as the quotient of k_{off}/K_D.[59]

In **publication 1**, "KinITC – One method supports both thermodynamic and kinetic SARs as exemplified on FimH antagonists" the method was independently validated for the first time. The publication also describes the methodical basis of kinITC-ETC and encourages scientists to analyze their ITC raw data with Affinimeter. One interesting aspect of kinITC is the fact that already thermodynamically analyzed and stored ITC data can be reinvestigated to ascertain the kinetic knowledge about the binding process as well.

References

1. Willson, R.J., Calorimetry, in *Principles of thermal analysis and calorimetry*, **2002**, The Royal Society of Chemistry, 120-165.
2. Wadsö, I., Bio-calorimetry. *Trends in Biotechnol.*, **1986**, *4*(2), 45-51.
3. Lavoisier, A.L., de LaPlace, P.S., Mémoire sur la chaleur. *Histoire de l'Académie Royale des Sciences.*, **1780**, 355-408.
4. Kleiber, M., The fire of life. An introduction to animal energetics, ed. Krieger, R.E. **1975**, Publishing Company, Huntington, NY.
5. Christensen, J.J., Izatt, R.M., and Eatough, D., Thermodynamics of metal cyanide coordination. V. Log k, ΔH°, and ΔS° values for the Hg2+-CN-system. *Inorg. Chem.*, **1965**, *4*(9), 1278-1280.

6. Izatt, R.M., Johnston, H.D., Watt, G.D., and Christensen, J.J., Thermodynamics of metal cyanide coordination. VI. Copper(I)- and silver(I)-cyanide systems. *Inorg. Chem.*, **1967**, *6*(1), 132-135.
7. Christensen, J.J., Johnston, H.D., Izatt, R.M., An isothermal titration calorimeter. *Rev. Sci. Instrum.*, **1968**, *39*, 1356–1359.
8. Christensen, J.J., Izatt, R.M., Hansen, L.D., New precision thermometric titration calorimeter. *Rev. Sci. Intr.*, **1965**, *36*, 779-783.
9. Ramsay, G., Prabhu, R., and Freire, E., Direct measurement of the energetics of association between myelin basic protein and phosphatidylserine vesicles. *Biochemistry*, **1986**, *25*(8), 2265-2270.
10. Myers, M., Mayorga, O.L., Emtage, J., and Freire, E., Thermodynamic characterization of interactions between ornithine transcarbamylase leader peptide and phospholipid bilayer membranes. *Biochemistry*, **1987**, *26*(14), 4309-4315.
11. Schon, A. and Freire, E., Thermodynamics of intersubunit interactions in cholera toxin upon binding to the oligosaccharide portion of its cell surface receptor, ganglioside GM1. *Biochemistry*, **1989**, *28*(12), 5019-5024.
12. Wiseman, T., Williston, S., Brandts, J.F., and Lin, L.N., Rapid measurement of binding constants and heats of binding using a new titration calorimeter. *Anal. Biochem.*, **1989**, *179*(1), 131-137.
13. Grolier, J.-P.E. and del Río, J.M., Isothermal titration calorimetry: A thermodynamic interpretation of measurements. *J. Chem. Thermodyn.*, **2012**, *55*, 193-202.
14. Judy, E., Kishore, N., Growing popularity of ultrasensitive microcalorimetry. *Bioenerg.*, **2016**, *5*(2), 1000140.
15. Freire, E., Mayorga, O.L., and Straume, M., Isothermal titration calorimetry. *Anal. Chem.*, **1990**, *62*(18), 950A-959A.
16. Ladbury, J.E., Calorimetry as a tool for understanding biomolecular interactions and an aid to drug design. *Biochem. Soc. Trans.*, **2010**, *38*(4), 888-893.
17. Wangler, A., Schmidt, C., Sadowski, G., and Held, C., Standard Gibbs energy of metabolic reactions: III the 3-phosphoglycerate kinase reaction. *ACS Omega*, **2018**, *3*(2), 1783-1790.
18. Leavitt, S. and Freire, E., Direct measurement of protein binding energetics by isothermal titration calorimetry. *Curr. Opin. Struct. Biol.*, **2001**, *11*(5), 560-566.
19. Du, X., Li, Y., Xia, Y.-L., Ai, S.-M., Liang, J., Sang, P., Ji, X.-L., and Liu, S.-Q., Insights into protein–ligand interactions: Mechanisms, models, and methods. *Int. J. Mol. Sci.*, **2016**, *17*(2), 144.
20. Perozzo, R., Folkers, G., and Scapozza, L., Thermodynamics of protein-ligand interactions: History, presence, and future aspects. *J. Recept. Signal Transduct. Res.*, **2004**, *24*(1-2), 1-52.
21. Bronowska, A.K., Thermodynamics of ligand–protein interactions: Implications for molecular design. **2011**, INTECH Open Access Publisher.
22. Ren, P., Chun, J., Thomas, D.G., Schnieders, M.J., Marucho, M., Zhang, J., and Baker, N.A., Biomolecular electrostatics and solvation: A computational perspective. *Q. Rev. Biophys.*, **2012**, *45*(4), 427-491.
23. Berg, J., Tymoczko JL, Streyer, L, Chemical bonds in biochemistry, in *Biochemistry*, **2002**.
24. John Kuriyan, B.K., David Wemmer, From genes to RNA and proteins, in *The molecules of life: Physical and chemical principles*, **2004**, Garland Science, 5-50.
25. Davis, A.M. and Teague, S.J., Hydrogen bonding, hydrophobic interactions, and failure of the rigid receptor hypothesis. *Angew. Chem. Int. Ed. Engl.*, **1999**, *38*(6), 736-749.
26. Arunan, E., Desiraju, G.R., Klein, R., Sadlej, J., Scheiner, S., Alkorta, I., Clary, D., Crabtree, R., J. Dannenberg, J., Hobza, P., Kjaergaard, H.G., Legon, A., Mennucci, B., and Nesbitt, D., Definition of the hydrogen bond (IUPAC recommendations 2011). Vol. 83. **2011**. 1637-1641.
27. Ferreira de Freitas, R. and Schapira, M., A systematic analysis of atomic protein–ligand interactions in the PDB. *MedChemComm*, **2017**, *8*(10), 1970-1981.
28. Poznanski, J., Poznanska, A., and Shugar, D., A Protein Data Bank survey reveals shortening of intermolecular hydrogen bonds in ligand–protein complexes when a halogenated ligand is an H-bond donor. *PLoS One*, **2014**, *9*(6), e99984.
29. R. Desiraju, G., Ho, P., Kloo, L., Legon, A., Marquardt, R., Metrangolo, P., Politzer, P., Resnati, G., and Rissanen, K., Definition of the halogen bond (IUPAC recommendations 2013). Vol. 85. **2013**. 1711-1713.
30. Anjana, R., Vaishnavi, M.K., Sherlin, D., Kumar, S.P., Naveen, K., Kanth, P.S., and Sekar, K., Aromatic-aromatic interactions in structures of proteins and protein-DNA complexes: A study based on orientation and distance. *Bioinformation*, **2012**, *8*(24), 1220-1224.
31. Liao, S.-M., Du, Q.-S., Meng, J.-Z., Pang, Z.-W., and Huang, R.-B., The multiple roles of histidine in protein interactions. *Chem. Cent. J.*, **2013**, *7*, 44-44.

32. Dougherty, D.A., Cation-π interactions involving aromatic amino acids. *J. Nutr.*, **2007**, *137*(6), 1504S-1508S.
33. Dong, X.-Y., Li, X., Li, B., Zhu, Y.-Y., Zang, S.-Q., and Tang, M.-S., Water sandwiched by a pair of aromatic rings in a proton-conducting metal–organic framework. *Dalton Trans.*, **2016**, *45*(45), 18142-18146.
34. Meyer, E.A., Castellano, R.K., and Diederich, F., Interactions with aromatic rings in chemical and biological recognition. *Angew. Chem. Int. Ed. Engl.*, **2003**, *42*(11), 1210-1250.
35. Diehl, C., Engström, O., Delaine, T., Håkansson, M., Genheden, S., Modig, K., Leffler, H., Ryde, U., Nilsson, U.J., and Akke, M., Protein flexibility and conformational entropy in ligand design targeting the carbohydrate recognition domain of galectin-3. *JACS*, **2010**, *132*(41), 14577-14589.
36. Homans, S.W., Dynamics and thermodynamics of ligand–protein interactions, in *Bioactive conformation I*, **2006**, Springer-Verlag Berlin Heidelberg.
37. MacRaild, C.A., Daranas, A.H., Bronowska, A., and Homans, S.W., Global changes in local protein dynamics reduce the entropic cost of carbohydrate binding in the arabinose-binding protein. *J. Mol. Biol.*, **2007**, *368*(3), 822-832.
38. Lumry, R. and Rajender, S., Enthalpy–entropy compensation phenomena in water solutions of proteins and small molecules: A ubiquitous property of water. *Biopolymers*, **1970**, *9*(10), 1125-1227.
39. Dunitz, J.D., Win some, lose some: Enthalpy–entropy compensation in weak intermolecular interactions. *Chem. Biol.*, **1995**, *2*(11), 709-712.
40. Dragan, A.I., Read, C.M., and Crane-Robinson, C., Enthalpy–entropy compensation: The role of solvation. *Eur. Biophys. J.*, **2017**, *46*(4), 301-308.
41. Cramer, J., Krimmer, S.G., Heine, A., and Klebe, G., Paying the price of desolvation in solvent-exposed protein pockets: Impact of distal solubilizing groups on affinity and binding thermodynamics in a series of thermolysin inhibitors. *J. Med. Chem.*, **2017**, *60*(13), 5791-5799.
42. O'Brien, R., Ladbury, J.E., Chowdry, B.Z., Chapter 10: Isothermal titration calorimetry of biomolecules, in *Protein-ligand interactions: Hydrodynamics and calorimetry* **2000**, Oxford University Press, 263-286.
43. Honarmand Ebrahimi, K., Hagedoorn, P.-L., Jacobs, D., and Hagen, W.R., Accurate label-free reaction kinetics determination using initial rate heat measurements. *Sci. Rep.*, **2015**, *5*, 16380.
44. Microcal iTC200 system, user manual. Malvern instruments ltd., worcestershire **2014**.
45. Demarse, N.A., Quinn, C.F., Eggett, D.L., Russell, D.J., and Hansen, L.D., Calibration of nanowatt isothermal titration calorimeters with overflow reaction vessels. *Anal. Biochem.*, **2011**, *417*(2), 247-255.
46. Saponaro, A., Isothermal titration calorimetry: A biophysical method to characterize the interaction between label-free biomolecules in solution. *Bio-protocol*, **2018**, *8*(15), e2957.
47. Aweda, T.A. and Meares, C.F., Combination of isothermal titration calorimetry and time-resolved luminescence for high affinity antibody–ligand interaction thermodynamics and kinetics. *Methods*, **2012**, *56*(2), 145-153.
48. Song, C., Zhang, S., and Huang, H., Choosing a suitable method for the identification of replication origins in microbial genomes. *Front. Microbiol.*, **2015**, *6*, 1049.
49. Doyle, M.L., Characterization of binding interactions by isothermal titration calorimetry. *Curr. Opin. Biotechnol.*, **1997**, *8*(1), 31-35.
50. Pierce, M.M., Raman, C.S., and Nall, B.T., Isothermal titration calorimetry of protein–protein interactions. *Methods*, **1999**, *19*(2), 213-221.
51. Sager, C.P., Fiege, B., Zihlmann, P., Vannam, R., Rabbani, S., Jakob, R.P., Preston, R.C., Zalewski, A., Maier, T., Peczuh, M.W., and Ernst, B., The price of flexibility - a case study on septanoses as pyranose mimetics. *Chem. Sci.*, **2018**, *9*(3), 646-654.
52. Krell, T., Lacal, J., Garcia-Fontana, C., Silva-Jimenez, H., Rico-Jimenez, M., Lugo, A.C., Darias, J.A., and Ramos, J.L., Characterization of molecular interactions using isothermal titration calorimetry. *Methods Mol. Biol.*, **2014**, *1149*, 193-203.
53. Sigurskjold, B.W., Exact analysis of competition ligand binding by displacement isothermal titration calorimetry. *Anal. Biochem.*, **2000**, *277*(2), 260-266.
54. Velazquez-Campoy, A. and Freire, E., Isothermal titration calorimetry to determine association constants for high-affinity ligands. *Nat. Protoc.*, **2006**, *1*(1), 186-191.
55. Velázquez Campoy, A. and Freire, E., ITC in the post-genomic era…? Priceless. *Biophysical Chemistry*, **2005**, *115*(2), 115-124.
56. Turnbull, W.B. and Daranas, A.H., On the value of c: Can low affinity systems be studied by isothermal titration calorimetry? *JACS*, **2003**, *125*(48), 14859-14866.

57. Salim, N.N. and Feig, A.L., Isothermal titration calorimetry of RNA. *Methods*, **2009**, *47*(3), 198-205.
58. Burnouf, D., Ennifar, E., Guedich, S., Puffer, B., Hoffmann, G., Bec, G., Disdier, F., Baltzinger, M., and Dumas, P., Kinitc: A new method for obtaining joint thermodynamic and kinetic data by isothermal titration calorimetry. *JACS*, **2012**, *134*(1), 559-565.
59. Dumas, P., Ennifar, E., Da Veiga, C., Bec, G., Palau, W., Di Primo, C., Pineiro, A., Sabin, J., Munoz, E., and Rial, J., Extending ITC to kinetics with kinitc. *Methods Enzymol.*, **2016**, *567*, 157-180.

Section II

Publications and Manuscripts

Chemistry, 2018, *24*(49), 13049-13057
doi: 10.1002/chem.201802599

Publication 2

KinITC—One Method Supports both Thermodynamic and Kinetic SARs as Exemplified on FimH Antagonists

Pascal Zihlmann,[1)‡] Marleen Silbermann,[1)‡] Timothy Sharpe,[2)] Xiaohua Jiang,[1)] Tobias Mühlethaler,[1)] Roman P. Jakob,[3)] Said Rabbani,[1)] Christoph P. Sager,[1)] Priska Frei,[1)] Lijuan Pang,[1)] Timm Maier,[3)] Beat Ernst[1)*]

[‡]These authors contributed equally to the project

[1)] Institute of Molecular Pharmacy, University of Basel, Klingelbergstr. 50, 4056 Basel, Switzerland

[2)] Biophysics Facility, University of Basel, Klingelbergstr. 70, 4056 Basel, Switzerland

[3)] Focal Area Structural Biology, University of Basel, Klingelbergstr. 70, 4056 Basel, Switzerland

[*] Corresponding author
Tel.: 0041 61 267 15 51; Fax: 0041 61 207 15 52
E-mail: beat.ernst@unibas.ch

Contributions of Marleen Silbermann

- Manuscript preparation in collaboration with Pascal Zihlmann
- SPR experiments and data evaluation/interpretation
- Cloning, expression, purification, and biotinylation of $FimH_{LD}$-AVITag-6His

Bioanalytical Chemistry

KinITC—One Method Supports both Thermodynamic and Kinetic SARs as Exemplified on FimH Antagonists

Pascal Zihlmann[+],[a] Marleen Silbermann[+],[a] Timothy Sharpe,[b] Xiaohua Jiang,[a] Tobias Mühlethaler,[a] Roman P. Jakob,[c] Said Rabbani,[a] Christoph P. Sager,[a] Priska Frei,[a] Lijuan Pang,[a] Timm Maier,[c] and Beat Ernst*[a]

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 sufficient 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 protein–ligand 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"),[1] little attention has been paid to kinetic aspects until recently. Only after Copeland introduced the drug–target residence time concept in 2006, kinetic considerations gained increasing interest.[2] The key message of this model is that residence time ($\tau = 1/k_{off}$)[3] or half-life ($t_{1/2} = \ln 2/k_{off}$)[2,4] 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 not only with high financial costs but also with a substantial risk of failure, mainly due to insufficient efficacy and unwanted side effects.[5] Several studies have shown that kinetic information is more reliable than affinity data to predict in vivo potency as well as the resistance potential of a drug candidate.[6] One of many examples is the HIV reverse transcriptase inhibitor efavirenz that exhibits only a low affinity for its target (5 μM) but a long dissociation half-life ($t_{1/2} = 2.8$ h).[7] Furthermore, kinetic studies by 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.[8] Nevertheless, early-phase drug discovery still focuses mainly on the optimization of K_Ds, although it has been recognized that kinetic rate constants for association (k_{on}) and dissociation (k_{off}) are a necessary precondition for a comprehensive description of the binding process.[9]

Long dissociation half-lives are a crucial feature of many small molecule drugs on the market, for example, for the neuraminidase inhibitor oseltamivir (47 min),[10] the selective COX-2 inhibitor rofecoxib (9 h)[11] or the HIV-1 protease inhibitor darunavir (>240 h).[12] Generally, a drug–target complex with a long dissociation half-life can compensate for unfavorable pharma-

[a] Dr. P. Zihlmann,[+] M. Silbermann,[+] Dr. X. Jiang, T. Mühlethaler, Dr. S. Rabbani, Dr. C. P. Sager, P. Frei, Dr. L. Pang, Prof. Dr. B. Ernst
Institute of Molecular Pharmacy
Pharmazentrum, University of Basel
Klingelbergstrasse 50, 4056 Basel (Switzerland)
E-mail: beat.ernst@unibas.ch

[b] Dr. T. Sharpe
Biophysics Facility
Biozentrum, University of Basel
Klingelbergstrasse 70, 4056 Basel (Switzerland)

[c] Dr. R. P. Jakob, Prof. Dr. T. Maier
Focal Area Structural Biology
Biozentrum, University of Basel
Klingelbergstrasse 70, 4056 Basel (Switzerland)

[+] These authors contributed equally to this work.

Supporting information and the ORCID identification number(s) for the author(s) of this article can be found under: https://doi.org/10.1002/chem.201802599.

cokinetics, that is, a bimolecular complex can still exist while the unbound drug molecule is already cleared from the body.[13] This is of particular importance for substances with short plasma half-lives.[4] 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).[14] 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.[15] In opposition to off-rates, which are independent of free ligand concentration, on-rates can be increased by the administration of higher doses. However, this requires sufficient oral bioavailability and increases the risk of undesired side effects based on the elevated amount of drug in circulation.[16]

It is surprising that studies on the correlation of molecular structures and their binding kinetics, so-called structure–kinetic relationships (SKRs), are rare.[17] In a rough approximation, more rotatable bonds and higher molecular weights correlate with long complex half-lives.[18] Moreover, water-shielded hydrogen bonds also tend to improve the lifetime of protein–ligand complexes.[19] 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.[9]

Surface plasmon resonance (SPR) spectroscopy has evolved into the method of choice for measuring binding kinetics in drug discovery.[20] It monitors non-covalent interactions in real-time by detecting a mass-dependent change in the refractive index close to the sensor surface.[21] 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 change (ΔC_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.[22]

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 (ΔH°), whereas the change in free energy ($\Delta G^\circ = RT\ln K_D$, with R being the universal gas constant and T the absolute temperature) and the change in entropy (ΔS°, $\Delta G^\circ = \Delta H^\circ - T\Delta S^\circ$) are calculated.[23] Both techniques are the gold-standard in their field of application, SPR for the determination of kinetic and ITC for thermodynamic parameters.[24]

Recent publications from Dumas and co-workers have the potential to transform the role of ITC in drug discovery, since it describes an approach to derive binding kinetics from ITC data (kinITC).[25]

In an ITC experiment, after each injection and mixing of titrant, an association reaction in the calorimeter cell relaxes to the equilibrium position dictated by the new total concentrations of reactants. This relaxation to equilibrium occurs at a rate that is determined by the association and dissociation rate constants for the reaction, together with the concentrations of free and complexed species. This relaxation process gives rise to a peak of heat release or uptake that is measured in the calorimeter. As the concentration of species changes throughout the titration, the relaxation kinetics changes, and so the shape (width) of the peak changes, with narrower peaks at the beginning of the titration and wider peaks nearer the point of stoichiometric equivalence. Since the observed shape of the peak actually depends on both the rate of relaxation to equilibrium and the intrinsic rate constant of the power compensation electronics of the calorimeter, it is necessary to deconvolute these contributions.

KinITC extracts the underlying kinetic information from the ITC thermogram by analysing the shape of each injection peak, and has been described by Dumas.[25] Here we have employed a simplified version of this analysis, kinITC-ETC (Equilibration Time Curve), implemented in the commercially available ITC analysis software AFFINI-meter.[25a] This method calculates the time necessary for the differential power curve to return to the baseline after an injection using the baseline-corrected thermogram, yielding a bell-shaped ETC curve. This curve is fitted as a function of k_{off} and the intrinsic rate constant for the instrument electronics. The enthalpy changes for association and the association constant, derived from fitting of the integrated binding isotherm (or determined in other experiments), are used to constrain this fit, so the kinetic and equilibrium parameters determined by this analysis are necessarily correlated and a reliable fitted value of the equilibrium constant is essential (Figure 1).

Clearly, 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. Although bearing a great potential to expand the positive impact of kinetic investigations in drug discovery, kinITC has only hesitantly found its way into SKR studies yet.[26] To reduce the initial reservation for the application of kinITC, kinetic parameters derived from ITC experiments are independently validated for the first time in the present study. In addition, we demonstrate that kinITC is a powerful tool for "data mining".

Therefore, the binding kinetics of a series of congeneric mannosides binding to the lectin domain of FimH ($FimH_{LD}$) were determined. FimH is a virulence factor located at the tip of type 1 pili of uropathogenic *E. coli* (UPEC) strains.[27] 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 infections (UTI).[28] 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.[29] Re-analyzing data from numerous ITC[30] and SPR[31] 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.

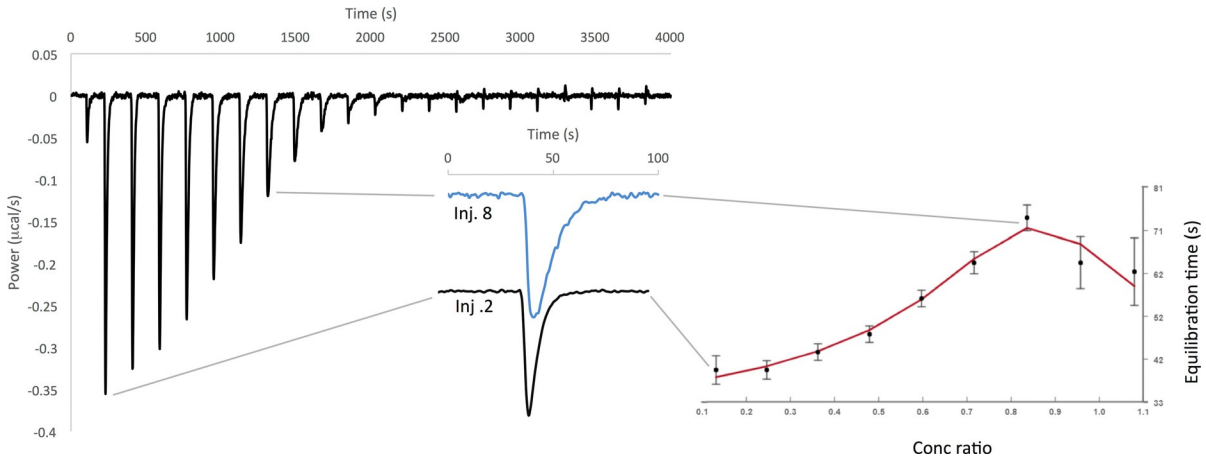


Figure 1. Illustrative representation of the kinITC-ETC obtained from the thermogram of a 1:1 interaction. The peak broadening observed at mid-titration (injection 8) as compared to the initial peaks (injection 2) is an indicative that the thermogram contains kinetic information that can be determined with kinITC. The equilibration time for each peak is plotted against titrant/titrate molar ratio to yield the ETC. Curve fitting of ETC according to the simplified kinITC method yields the values of the off-rate constant (k_{off}) and the response time.

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.[25b] However, with kinITC-ETC integrated into the commercially available ITC analysis software AFFINImeter (Version 1.0415–1.1510) this hurdle could be overcome. It uses fully automated data processing to derive kinetic and thermodynamic parameters 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 Tables 1, 2, and 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 $FimH_{LD}$.

Validation of kinITC-ETC by SPR

Although Dumas et al. proved the potential of kinITC with the evaluation of multistep kinetic RNA folding,[25b] there is, to the best of our knowledge, no independent comparison of data

Table 1. Comparison of kinetic data obtained by kinITC-ETC and SPR for the interaction of $FimH_{LD}$ with the mannose-derived antagonists **1**–**4**.[a]

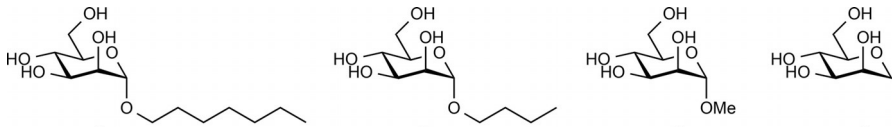


Table 1-A: Association rate constants k_{on} of $FimH_{LD}$ with compound **1**–**4** determined by ITC and SPR.

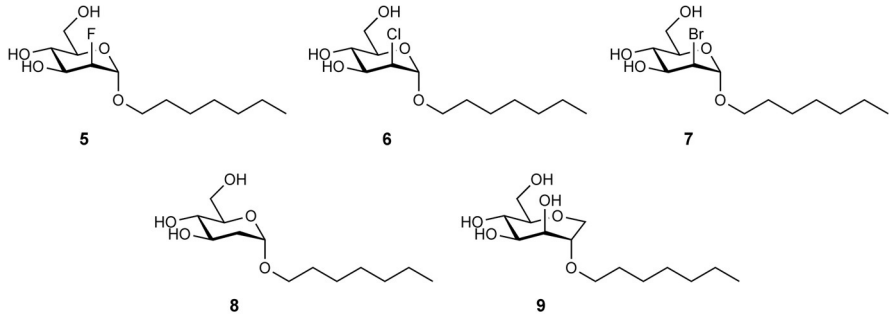
Cmpd	$k_{on}^{kinITC-ETC}$		k_{on}^{SPR}		$k_{on}^{kinITC-ETC}/k_{on}^{SPR}$	
	k_{on} [$M^{-1}s^{-1}$]	rk_{on}	k_{on} [$M^{-1}s^{-1}$]	rk_{on}	Absolute	Relative
1	3.32×10^{4}	1.00	2.45×10^{4}	1.00	1.35	1.00
2	9.71×10^{3}	0.29	5.57×10^{3}	0.23	1.74	1.29
3	2.31×10^{3}	0.07	2.45×10^{3}	0.10	0.94	0.70
4	2.00×10^{3}	0.06	2.67×10^{3}	0.11	0.75	0.55

Table 1-B: Dissociation rate constants k_{off} of FimHLD with compound **1**–**4** determined by ITC and SPR.

Cmpd	$k_{off}^{kinITC-ETC}$			k_{off}^{SPR}			$k_{on}^{kinITC-ETC}/k_{on}^{SPR}$	
	k_{off} [s^{-1}]	rk_{off}	$t_{1/2}$ [min]	k_{off} [s^{-1}]	rk_{off}	$t_{1/2}$ [min]	Absolute	Relative
1	7.27×10^{-4}	1.00	15.9	1.54×10^{-4}	1.00	75.0	4.72	1.00
2	1.15×10^{-3}	1.58	10.1	1.89×10^{-4}	1.23	61.1	6.07	1.29
3	2.86×10^{-3}	3.93	4.1	1.11×10^{-3}	7.21	10.4	2.57	0.54
4	1.82×10^{-3}	2.50	6.3	6.51×10^{-4}	4.23	17.7	2.80	0.59

[a] Absolute (k_{on} and k_{off}) and relative (rk_{on} and rk_{off}) kinetic parameters are reported for both methods (Table 1 A for on-rates and Table 1 B for off-rates). Relative values are normalized to *n*-heptyl α-D-mannopyranoside (**1**), which is set to 1. The discrepancies between kinITC-ETC and SPR for k_{on} (Table 1 A) and k_{off} (Table 1 B) values are displayed in the rightmost columns. K_D values (Table S2) as well as confidence intervals for k_{on} and k_{off} (ITC, Table S4 and SPR, Table S5) are part of the Supporting Information.

Table 2. Kinetic binding parameters for the interaction of $FimH_{LD}$ with mannose derivatives of *n*-heptyl α-D-mannopyranoside (**1**).[a]



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

[a] Confidence intervals of the fitted parameters k_{on}, k_{off}, K_D, and the response time of the calorimeter feedback circuit are part of the Supporting Information (Table S4). Relative changes (rk_{on}, rk_{off}) are compared to *n*-heptyl α-D-mannopyranoside (**1**).

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_{on}, the comparison revealed an excellent correlation with absolute values differing by less than a factor of two (Table 1 A), whereas up to 6-fold deviation was observed for k_{off} values (Table 1 B). These differences might be a consequence of ligand rebinding on the SPR chip surface.[32] However, when we normalized the k_{off} values obtained by SPR and kinITC-ETC to that of *n*-heptyl α-D-mannopyranoside (**1**) (giving relative k_{off}, rk_{off}) differences of less than a factor of 2 arose (Table 1 B). In their comparison of SPR and kinITC, Burnouf et al. reported a similar difference for thiamine pyrophosphate interacting with mRNA.[25b]

Kinetics of $FimH_{LD}$ 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.[14a,33] 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.[17,18,34] 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 ≈1:16, Table 1 A) than of the off-rate constants (for kinITC ≈1:4, Table 1 B).

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)[35] and **2** (PDB-code: 1UWF)[36] (Table S3). For all four ligands, both protein structure and binding mode were found to be identical (Figure 2). 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.

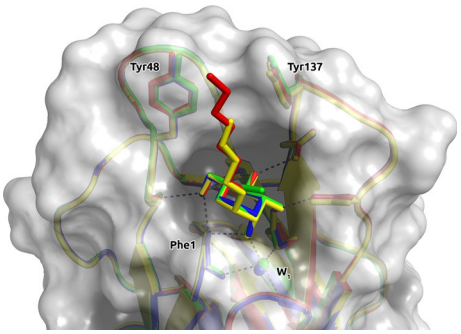


Figure 2. Binding mode of compounds **1**–**4** to $FimH_{LD}$. The co-crystal structures of **1** (red, PDB-code: 4XO8), **2** (yellow, PDB-code: 1UWF), **3** (green, PDB-code: 5JCR), and **4** (blue, PDB-code: 5MUC) show coinciding binding modes. Their mannose moieties form identical hydrogen bond networks with $FimH_{LD}$ and a structural water molecule (W_1), 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 ($TS^{\ddagger}$), and in the bound state (B) characterized by the kinetic rate constants k_{on} and k_{off} is the simplest way to describe a protein–ligand interaction (Figure 3 A). However, more common is a multistep binding mechanism with apparent rate constants composed of multiple elementary rate constants.[9] Rate constants depend upon the free

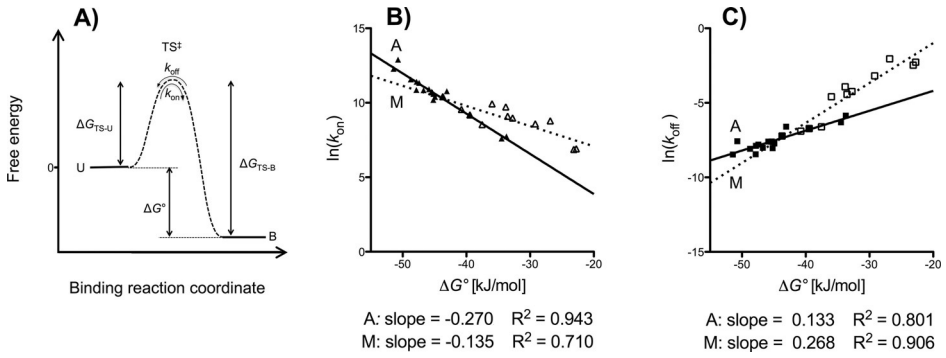


Figure 3. A) The energy diagram illustrates a simple one-step binding process with the transition state ($TS^{\ddagger}$) separating the unbound state (U) from the bound state (B). ΔG_{TS-U} is the activation free energy of association, ΔG_{TS-B} the activation free energy of dissociation, ΔG° the free 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 ΔG° depends on the site of a ligand's modification. Ligands with modified aglycones (**2**–**4** and **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** and **24**–**29**; for structure and binding parameters see Supporting Information, hollow symbols, dotted regression lines): A, aglycone; M, mannose moiety.

energy difference ΔG_{TS-U} (between $TS^{\ddagger}$ and U) for k_{on} and the free energy difference ΔG_{TS-B} (between $TS^{\ddagger}$ and B) for k_{off}, respectively (Figure 3 A). The energy difference between B and U represents the equilibrium free energy of binding (ΔG_{B-U}, or ΔG°). 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 ΔG_{TS-B} (decrease k_{off}) and to a lesser extent decrease ΔG_{TS-U} (increase k_{on}). Concomitantly, the change of ΔG° 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_D value (Table S2). Hence, increased on-rates are related to favorable interactions realized in $TS^{\ddagger}$ or a destabilization of U relative to $TS^{\ddagger}$, 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 $FimH_{LD}$. 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,[37] that is, in a pre-binding state, an interaction with a hydrophobic patch of the enzymatic cavity is formed.[38]

To investigate the kinetic behavior of $FimH_{LD}$ 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 μM) were analyzed by kinITC-ETC. Thermodynamic and kinetic fits (Table S1) and binding parameters (Table S3 and 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 indolinylphenyl aglycone) and a second subset ($n = 11$) representing compounds with modified mannose moieties (deoxy- and deoxy-halogeno derivatives **5**–**8** and **24**–**27**, a 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 ΔG° with k_{on} (Figure 3 B, $p < 0.001$) and k_{off} values (Figure 3 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 $FimH_{LD}$, affect the k_{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,[30c] influence k_{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 $TS^{\ddagger}$, resulting

in the observed stronger effect on the off-rate constant k_{off} (slope = 0.268) than on the on-rate constant k_{on} (slope = −0.135). For structural modifications of the aglycone, however, the impact on k_{on} values (slope = −0.270) is more pronounced than on k_{off} values (slope = 0.133). This suggests that either the $TS^{\ddagger}$ is stabilized relative to the bound state B (e.g. when the aglycone establishes stronger interactions with the receptor in the $TS^{\ddagger}$ than in the bound state) or that the unbound state U is destabilized relative to $TS^{\ddagger}$.

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 glycan or aglycone), neglecting various factors, for example, 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.

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 3B and Figure 3C), interactions of ligands **5**–**9** with $FimH_{LD}$ were analyzed in more detail (Table 2). The 2-hydroxyl of *n*-heptyl α-D-mannopyranoside (**1**) forms two hydrogen bonds with $FimH_{LD}$, once as H-bond donor to a structural water molecule (W_1) and once as H-bond acceptor from the positively charged N-terminal Phe (Figure 2). 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⧧, which enhances on- and off-rates and (ii) the formation of more specific interactions in the bound state, which decreases off-rates. However, mannose modification may also affect the on-rate constant, as observed for the ring extended compound **9**. As shown in the study of Sager et al.,[30a] septulose **9** (PDB-code: 5CGB) and pyranoside **1** (PDB-code: 4XO8) establish an identical H-bond network with $FimH_{LD}$. 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. The present work displays that the increased flexibility in solution of ligand **9** stabilizes U relative to $TS^{\ddagger}$ 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 when compared to **1** (Table 2).

Conclusions

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_{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_{off} in SPR measurements may be due to the rebinding of the analyte 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_{off} may arise from errors associated with the kinITC-ETC method, that is, the observed kinetics of heat evolution in ITC experiments includes a contribution from the intrinsic instrumental response (τITC), which can vary between instruments and for the same instrument depending on cleanliness. Since we re-analyzed existing data for this study, τITC 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 at most in the order of a factor of 6 and therefore in an acceptable range 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_D values not only stand for prolonged complex half-lives, but also for increased on-rates indicating a stabilization of the transition state and destabilization of the unbound state. On a structural level, the kinetic contribution 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 $FimH_{LD}$, mainly influence the off-rate constant k_{off}, while modifications on the aglycone establishing hydrophobic interactions with the two tyrosines at the entrance to the mannose-binding pocket, predominantly affect k_{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. This finding is in excellent agreement with previous observations from Gaspari et al. for the interaction of hCAII with hydrophobic ligands.[38] In general, initial hydrophobic interactions to stabilize protein–ligand interactions established before the final binding mode is reached might facilitate ligand recruitment from solution. Furthermore, electrostatic interactions may beneficially influence the on-rate as exemplified by the replacement 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. In contrast to our in vitro results, rebinding of monovalent antagonists to the CRD of neighboring pili may result in a longer FimH occupancy in vivo.

Our kinetic study with $FimH_{LD}$ is the first systematic analysis of protein–ligand interactions 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 the 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."

Experimental Section

Protein cloning, expression, purification, and biotinylation. For all ITC experiments, $FimH_{LD}$ of *E.coli* K-12 strain was expressed with a C-terminal thrombin cleavage site and a 6His-tag ($FimH_{LD}$-Th-6His, 173 residues) following a previously published protocol.[39] For SPR experiments $FimH_{LD}$-AVITag-6His gene construct was cloned by overlap extension PCR using the $FimH_{LD}$-Th-6His[39] gene construct and an AVITag-6His gene sequence as templates and was subsequently ligated into the plasmid pNT.[40] The histidine-tagged recombinant protein was expressed in the protease-deficient *E. coli* strain HM125 and purified with Ni-NTA affinity chromatography.[39] 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′500 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 $FimH_{LD}$ (100 nM) was immobilized on the surface of a streptavidin chip (sensor chip SA) using the "aim for immobilized level wizard" (1′300 RU, 5 $\mu L\,min^{-1}$). A reference surface with Amino PEG biotin (50 μM, Polypure) was prepared (time and flow rate: 60 s and 10 $\mu L\,min^{-1}$, 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 $\mu L\,min^{-1}$. 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 μl and 15 μl, a reference power of 10 μcal/sec, a stirring speed of 307 rpm, in high feedback mode, and with a filter period of 2 sec. Preceding the measurements, $FimH_{LD}$-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 $24'180\,M^{-1}cm^{-1}$. The active protein concentration was determined by an ITC experiment with $FimH_{LD}$ binding to compound **1**.[41] The thermodynamic parameters K_A (association constant) and ΔH° (change in enthalpy) and the kinetic parameter k_{off} (dissociation rate constant) were measured by ITC. All parameters were evaluated using the fully automated analysis software package from AFFINImeter (Version 1.0415–1.1510, Software for Science Developments, Santiago de Compostela, Spain).[25b] The parameters ΔG° (free energy of binding) and ΔS° (change in entropy) were calculated from [Eq. (1)] and k_{on} (association rate constant) was calculated according to [Eq. (2)]:

$$\Delta G^\circ = \Delta H^\circ - T\Delta S^\circ = -RT\ \ln K_A \quad (1)$$

$$K_A = k_{on}/k_{off} = 1/K_D \quad (2)$$

with T being the absolute temperature and R the universal gas constant (8.314 $Jmol^{-1}K^{-1}$). 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.[42] 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)[43] for the competitive experiments of compounds **19** and **23** are described in Fiege et al. 2015[30c] and for compounds **15**–**18**, **20**, and **21** in Kleeb et al. 2015.[30d] 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.[23] Correlations of the kinetic parameters k_{on} (B) and k_{off} (C) with the change of free energy of binding (ΔG°), 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).

Crystallization and structure. $FimH_{LD}$–**3** and $FimH_{LD}$–**4** (compound **3** and **4** in complex with $FimH_{LD}$) were crystallized by sitting-drop vapor diffusion at 4 °C. $FimH_{LD}$ (residues 1–158) was used at a final concentration of 10 $mg\,mL^{-1}$ (ca. 0.5 mM) with a 5-fold molar ligand excess in HEPES buffer (pH 7.4, 20 mM). After several months, plate like crystals appeared in 1.6 M $(NH_4)_2SO_4$, 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_2SO_4.[44] Data were collected at the PX beamline (X06SA) of the Swiss Light Source (Paul Scherrer Institute, Switzerland), indexed, integrated, and scaled with XDS.[45] The structures were solved by molecular replacement with PHASER[46] with 4X50.pdb[30c] as a search model. The models were built in COOT[47] and refined with the PHENIX software.[48] Geometric ligand restraints were generated with PRODRG.[49] 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.[50] Figure 2 was produced using PyMOL (The PyMOL Molecular Graphics System, Version 1.8 Schrödinger, LLC).

Acknowledgements

We gratefully thank Eva Muñoz from AFFINImeter-S4SD, Edificio Emprendia, Campus Vida, Santiago de Compostela, for providing Figure 1. This work was supported by Swiss National Science Foundation Grants 200020 146202 and R'EQUIP 145023.

Conflict of interest

The authors declare no conflict of interest.

Keywords: analytical methods · calorimetry · kinetics · method validation of kinITC · thermodynamics

[1] a) F. Bosch, L. Rosich, *Pharmacology* **2008**, *82*, 171 – 179; b) P. Ehrlich, *Lancet* **1913**, *182*, 445 – 451.
[2] R. A. Copeland, D. L. Pompliano, T. D. Meek, *Nat. Rev. Drug Discov.* **2006**, *5*, 730 – 739.
[3] R. A. Copeland, *Nat. Rev. Drug Discov.* **2016**, *15*, 87 – 95.
[4] G. Dahl, T. Akerud, *Drug Discovery Today* **2013**, *18*, 697 – 707.
[5] M. Hay, D. W. Thomas, J. L. Craighead, C. Economides, J. Rosenthal, *J. Nat. Biotechnol.* **2014**, *32*, 40 – 51.
[6] a) F. B. E. Garah, J. L. Stigliani, F. Coslédan, B. Meunier, A. Robert, *ChemMedChem* **2009**, *4*, 1469 – 1479; b) J. F. Vanhauwe, M. Ercken, D. van de Wiel, M. Jurzak, J. E. Leysen, *Psychopharmacology (Berlin, Ger.)* **2000**, *150*, 383 – 390; c) S. Nishino, J. Arrigoni, J. Shelton, T. Kanbayashi, W. C. Dement, E. Mignot, *J. Neurosci.* **1997**, *17*, 6401 – 6408.
[7] V. A. Braz, L. A. Holladay, M. D. Barkley, *Biochemistry* **2009**, *49*, 601 – 610.
[8] B. Maschera, G. Darby, G. Palú, L. L. Wright, M. Tisdale, R. Myers, E. D. Blair, E. S. Furfine, *J. Biol. Chem.* **1996**, *271*, 33231 – 33235.
[9] A. C. Pan, D. W. Borhani, R. O. Dror, D. E. Shaw, *Drug Discovery Today* **2013**, *18*, 667 – 673.
[10] W. M. Kati, D. Montgomery, R. Carrick, L. Gubareva, C. Maring, K. McDaniel, K. Steffy, A. Molla, F. Hayden, D. Kempf, *Antimicrob. Agents Chemother.* **2002**, *46*, 1014 – 1021.
[11] C. Chan, S. Boyce, C. Brideau, S. Charleson, W. Cromlish, D. Ethier, J. Evans, A. Ford-Hutchinson, M. Forrest, J. Gauthier, *J. Pharmacol. Exp. Ther.* **1999**, *290*, 551 – 560.
[12] I. Dierynck, M. De Wit, E. Gustin, I. Keuleers, J. Vandersmissen, S. Hallenberger, K. Hertogs, *J. Virol.* **2007**, *81*, 13845 – 13851.
[13] P. J. Tummino, R. A. Copeland, *Biochemistry* **2008**, *47*, 5481 – 5492.
[14] a) H. Lu, P. J. Tonge, *Curr. Opin. Chem. Biol.* **2010**, *14*, 467 – 474; b) S. Kapur, P. Seeman, *J. Psychiatry Neurosci.* **2000**, *25*, 161.
[15] G. Vauquelin, *Expert Opin. Drug Discov.* **2010**, *5*, 927 – 941.
[16] a) R. A. Copeland, *Drug-Target Residence Time in Thermodynamics and Kinetics of Drug Binding*, R. Mannhold, H. Kubinyi, G. Folkers, G. Keserü, D. C. Swinney, Eds. John Wiley & Sons, Weinheim **2015**, Vol. 65, 157 – 167; b) P. Y. Muller, M. N. Milton, *Nat. Rev. Drug Discov.* **2012**, *11*, 751 – 761.
[17] G. Klebe, *ChemMedChem* **2015**, *10*, 229 – 231.
[18] D. C. Miller, G. Lunn, P. Jones, Y. Sabnis, N. L. Davies, P. Driscoll, *MedChemComm* **2012**, *3*, 449 – 452.
[19] P. Schmidtke, F. J. Luque, J. B. Murray, X. Barril, *J. Am. Chem. Soc.* **2011**, *133*, 18903 – 18910.
[20] U. Jönsson, L. Fägerstam, B. Ivarsson, B. Johnsson, R. Karlsson, K. Lundh, S. Löfås, B. Persson, H. Roos, I. Rönnberg, *BioTechniques* **1991**, *11*, 620 – 627.
[21] R. Karlsson, *J. Mol. Recognit.* **2004**, *17*, 151 – 161.
[22] G. A. Holdgate, W. H. Ward, *Drug Discovery Today* **2005**, *10*, 1543 – 1550.
[23] W. B. Turnbull, A. H. Daranas, *J. Am. Chem. Soc.* **2003**, *125*, 14859 – 14866.
[24] G. Holdgate, S. Geschwindner, A. Breeze, G. Davies, N. Colclough, D. Temesi, L. Ward in *Protein-Ligand Interactions, Vol. 1008* (Eds.: M. A. Williams, T. Daviter) Humana Press, **2013**, pp. 327 – 355.
[25] a) P. Dumas, E. Ennifar, C. Da Veiga, G. Bec, W. Palau, C. Di Primo, A. Piñeiro, J. Sabin, E. Muñoz, J. Rial in *Methods in Enzymology, Vol. 567* (Eds.: L. F. Andrew) Academic Press, **2016**, pp. 157 – 180; b) D. Burnouf, E. Ennifar, S. Guedich, B. Puffer, G. Hoffmann, G. Bec, F. Disdier, M. Baltzinger, P. Dumas, *J. Am. Chem. Soc.* **2012**, *134*, 559 – 565.
[26] S. Igde, S. Röblitz, A. Müller, K. Kolbe, S. Boden, C. Fessele, T. K. Lindhorst, M. Weber, L. Hartmann, *Macromol. Biosci.* **2017**, *17*, 1700198.
[27] a) E. V. Sokurenko, V. Chesnokova, D. E. Dykhuizen, I. Ofek, X.-R. Wu, K. A. Krogfelt, C. Struve, M. A. Schembri, D. L. Hasty, *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 8922 – 8926; b) N. Sharon, *FEBS Lett.* **1987**, *217*, 145 – 157.
[28] X. R. Wu, T. T. Sun, J. J. Medina, *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 9630 – 9635.
[29] a) D. Abgottspon, B. Ernst, *Chimia* **2012**, *66*, 166 – 169; b) M, Hartmann, T. K. Lindhorst, *Eur. J. Org. Chem.* **2011**, *2011*, 3583 – 3609.
[30] a) C. P. Sager, B. Fiege, P. Zihlmann, R. Vannam, S. Rabbani, R. P. Jakob, R. C. Preston, A. Zalewski, T. Maier, M. W. Peczuh, B. Ernst, *Chem. Sci.* **2018**, *9*, 646 – 654; b) W. Schönemann, M. Lindegger, S. Rabbani, P. Zihlmann, O. Schwardt, B. Ernst, *Perspecti. Sci.* **2017**, *11*, 53 – 61; c) B. Fiege, S. Rabbani, R. C. Preston, R. P. Jakob, P. Zihlmann, O. Schwardt, X. Jiang, T. Maier, B. Ernst, *ChemBioChem* **2015**, *16*, 1235 – 1246; d) 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, B. Ernst, *J. Med. Chem.* **2015**, *58*, 2221 – 2239; e) L. Pang, S. Kleeb, K. Lemme, S. Rabbani, M. Scharenberg, A. Zalewski, F. Schadler, O. Schwardt, B. Ernst, *ChemMedChem* **2012**, *7*, 1404 – 1422..
[31] M. Scharenberg, X. Jiang, L. Pang, G. Navarra, S. Rabbani, F. Binder, O. Schwardt, B. Ernst, *ChemMedChem* **2014**, *9*, 78 – 83.
[32] M. Gopalakrishnan, K. Forsten-Williams, T. R. Cassino, L. Padro, T. E. Ryan, U. C. Täuber, *Eur. Biophys. J.* **2005**, *34*, 943 – 958.
[33] a) R. A. Copeland, *Expert Opin. Drug Discov.* **2010**, *5*, 305 – 310; b) D. C. Swinney, *Curr. Opin. Drug Discov. Devel.* **2009**, *12*, 31 – 39.
[34] E. V. Schneider, J. Böttcher, R. Huber, K. Maskos, L. Neumann, *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 8081 – 8086.
[35] M. M. Sauer, R. P. Jakob, J. Eras, S. Baday, D. Eriş, G. Navarra, S. Bernèche, B. Ernst, T. Maier, R. Glockshuber, *Nat. Commun.* **2016**, *7*, 10738 – 10738.
[36] 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, H. De Greve, *Mol. Microbiol.* **2005**, *55*, 441 – 455.
[37] a) R. King, A. Burgen, *Proc. R. Soc. London Ser. B* **1976**, *193*, 107 – 125; b) P. W. Taylor, R. W. King, A. S. Burgen, *Biochemistry* **1970**, *9*, 2638 – 2645.
[38] R. Gaspari, C. Rechlin, A. Heine, G. Bottegoni, W. Rocchia, D. Schwarz, J. Bomke, H.-D. Gerber, G. Klebe, A. Cavalli, *J. Med. Chem.* **2016**, *59*, 4245 – 4256.
[39] S. Rabbani, X. Jiang, O. Schwardt, B. Ernst, *Anal. Biochem.* **2010**, *407*, 188 – 195.
[40] 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.
[41] H. Edelhoch, *Biochemistry* **1967**, *6*, 1948 – 1954.
[42] B. W. Sigurskjold, *Anal. Biochem.* **2000**, *277*, 260 – 266.

[43] J. C. Houtman, P. H. Brown, B. Bowden, H. Yamaguchi, E. Appella, L. E. Samelson, P. Schuck, *Protein Sci.* **2007**, *16*, 30 – 42.
[44] K. A. Rubinson, J. E. Ladner, M. Tordova, G. L. Gilliland, *Acta Crystallogr. D* **2000**, *56*, 996 – 1001.
[45] W. Kabsch, *Acta Crystallogr. D* **2010**, *66*, 133 – 144.
[46] A. J. McCoy, R. W. Grosse-Kunstleve, P. D. Adams, M. D. Winn, L. C Storoni, R. J. Read, *J. Appl. Crystallogr.* **2007**, *40*, 658 – 674.
[47] a) P. Emsley, B. Lohkamp, W. G. Scott, K. Cowtan, *Acta Crystallogr. D* **2010**, *66*, 486 – 501; b) P. Emsley, K. Cowtan, *Acta Crystallogr. D* **2004**, *60*, 2126 – 2132.
[48] a) 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. D* **2010**, *66*, 213 – 221; b) P. D. Adams, R. W. Grosse-Kunstleve, L.-W. Hung, T. R. Ioerger, A. J. McCoy, N. W. Moriarty, R. J. Read, J. C. Sacchettini, N. K. Sauter, T. C. Terwilliger, *Acta Crystallogr D.* **2002**, *58*, 1948 – 1954.
[49] A. W. Schuttelkopf, D. M. Van Aalten, *Acta Crystallogr. D* **2004**, *60*, 1355 – 1363.
[50] G. M. Sastry, M. Adzhigirey, T. Day, R. Annabhimoju, W. Sherman, *J. Comput. Aided Mol. Des.* **2013**, *27*, 221 – 234.

Manuscript received: May 23, 2018
Revised manuscript received: June 21, 2018
Accepted manuscript online: June 25, 2018
Version of record online: July 27, 2018

Manuscript 1

Off-rate Screening by Surface Plasmon Resonance – The Search for Promising Lead Structures Targeting Low-affinity FimH

Marleen Silbermann[1], Deniz Eris[1], Oliver Schwardt[1], Wojciech Schönemann[1], Xiaohua Jiang[1], Priska Frei[1], Lijuan Pang[1], Jacqueline Bezençon[1], and Beat Ernst[1*]

[1)] Institute of Molecular Pharmacy, University of Basel,

Klingelbergstr. 50, 4056 Basel, Switzerland

[*] Corresponding author
Tel.: 0041 61 267 15 51; Fax: 0041 61 207 15 52
E-mail: beat.ernst@unibas.ch

Contributions of Marleen Silbermann

- Manuscript preparation
- Methodological development of the off-rate screen
- SPR experiments and data evaluation/interpretation in collaboration with Deniz Eris
- Protein expression and purification of $FimH_{FL-B}$ in collaboration with Deniz Eris

Abbreviations:

CRD - carbohydrate recognition domain

Da - dalton

DMSO - Dimethyl sulfoxide

DsG - donor strand of FimG

DsG^{biotin} - biotinylated donor strand of FimG

EDTA - ethylenediaminetetraacetic acid

ESI - electrospray ionization

$FimH_{FL}$ - full-length FimH

$FimH_{FL\text{-}B}$ - biotinylated full-length FimH

$FimH_{LD}$ - FimH lectin domain

$FimH_{PD}$ - FimH pilin domain

FPA - flourescence polarization assay

FPLC - fast protein liquid chromatography

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

HM - *n*-heptyl α-D-mannopyranoside

IPTG - isopropyl β-D-1-thiogalactopyranoside

K_D - equilibrium dissociation constant

k_{off} - dissociation rate constant

k_{on} - association rate constant

LB - lysogeny broth

MCKs - Multi-cycle kinetics

MOPS - 3-(*N*-morpholino)propanesulfonic acid

NaCl - sodium chloride

NaH_2PO_4 - sodium dihydrogen phosphate

OD_{600} - optical density measured at a wavelength of 600 nm

P20 - Polysorbate 20

PEG - polyethylene glycol

τ - residence time

r = Pearson correlation coefficient

R&D - research and development

RI - refractive index

RU - resonance units

SA - streptavidin sensor chip

SDS-PAGE - sodium dodecyl sulfate-polyacrylamide gel electrophoresis

SPR - surface plasmon resonance

UPECs - uropathogenic *Escherichia coli* strains

UPIa - uroplakin Ia

UTIs - urinary tract infections

$t_{1/2}$ – dissociative half-life

Tris - tris(hydroxymethyl)aminomethane

Ttds - trioxatridecan-succinamic acid

X - time

Y - response units

Y_0 – response units at time zero

Abstract

The development of new medicines is a tedious business with high expenses and an enormous time exposure. Linked to this problem is an extremely high compound failure rate also after completion of the discovery research. Affinity-based methods are routinely employed to determine the binding affinity of a compound–target complex *in vitro*, which decides if the compound stays in the drug discovery pipeline. However, binding affinity frequently does not reflect *in vivo* efficacy of potential drug candidates. It has been shown that the dissociation half-life ($t_{1/2}$) or the residence time (τ) of a bimolecular complex is a reliable measure to predict *in vivo* efficacy. Both rely on the dissociation rate constant (k_{off}), which can be determined by surface plasmon resonance (SPR) experiments. On this account we developed an off-rate screen that can be easily implemented into early drug discovery to identify potential hits and leads. We used the bacterial lectin FimH as therapeutical target and tested a library consisting of various structural compound classes based on mannose with the intention to detect antagonists exhibiting long half-lives. FimH is under investigation as treatment option for urinary tract infections (UTIs) caused by uropathogenic *E. coli* (UPEC) strains. So far, the standard therapy is a short-term administration of antibiotics but due to the growing problem of antibiotic-resistance an alternative medication is required. We concentrated on the low-affinity state of FimH ($FimH_{FL}$), which most probably is the therapeutically relevant conformation that needs to be blocked in order to prevent adhesion of UPECs to the bladder epithelium.

Introduction

The constant need for new drugs is a highly challenging task for the pharmaceutical industry. The overall research and development (R&D) process across all therapeutic areas on average amounts 14 years for new drugs [1] and is accompanied by a extremely high compound attrition rate.[2] A serious issue in this respect is the high probability to fail due to poor *in vivo* efficacy,[3] which is tested only after 4.5 years of discovery research.[4] In order to reduce R&D expenses, companies work on aproaches to avoid the high compound dropout caused by incomparable *in vitro* and *in vivo* results.

Traditional *de novo* drug discovery mostly starts with the identification and the validation of a biological target for therapeutic intervention and is followed by the search for small molecules that bind to the target and elicit a pharmacological effect.[5] The formation of a binary complex of a drug and a protein is of primary importance for *in vivo* drug action because a drug only triggers a pharmacological effect when bound to the target.[6] For this reason, affinity-based techniques are routinely employed to identify active compounds by measuring the binding strength of small molecule–protein complexes *in vitro*.[7,8] Binding affinity is often quantified by measuring the equilibrium dissociation constant (K_D), which is the ratio of k_{off}/k_{on}, where k_{off} is the dissociation rate constant and k_{on} the association rate constant.[9] The K_D of a bimolecular interaction can be determined in vitro by several analytical methods as for example isothermal titration calorimetry (ITC),[10] surface plasmon resonance (SPR),[11] fluorescence polarization (FP),[12] fluorescence energy resonance transfer (FRET),[13] and affinity chromatography.[14]

However, it has been observed that binding affinity, measured *in vitro*, is not comparable with *in vivo* conditions, which are rather described by the lifetime of a binary complex.[15,16] The lifetime of a binary complex *in vivo* principally depends on k_{off} and can be quantified either by the dissociative half-life ($t_{1/2} = (\ln 2)/k_{off}$) or the residence time ($\tau = 1/k_{off}$).[15,17,18] A low k_{off} is associated with an occupation of the binding site over an extended period of time. Kinetic *in vitro* studies exhibited that k_{off} does not always correlate with K_D [16,19] and contradict the general use of binding affinity as surrogate of *in vivo* efficacy.[20] A higher binding affinity does not necessarily yield a better drug. For instance, it could be shown that *in vivo* efficacy of a series of A_{2A} adenosine receptor agonists correlated well with the residence time ($r^2 = 0.90$), while a correlation with the binding affinity ($r^2 = 0.13$) was not observed.[21] A study on epidermal growth factor (EGFR) tyrosine kinase inhibitors (TKIs) showed that lapatinib exhibited a prolonged downregulation of EGFR signaling compared to gefitinib despite a worse apparent inhibitory constant (K_i^{app}) of 3 nM over 0.4 nM. By contrast, the inhibitory efficacy is better reflected by the off-rate. Lapatinib featured a half-life of 300 min while the $t_{1/2}$ of gefitinib was of less than 10 min.[22] Furthermore, the work of Seow and coworkers on antagonists targeting the inflammatory protein complement C5a could even demonstrate that amongst antagonists with comparable *in vitro* potency, the one with the longest residence time exhibited a better oral activity in comparison to more drug-like or orally better bioavailable drug candidates.[23]

The off-rate can be determined by SPR, which is the method of choice to reveal the kinetic fingerprint of small molecule–protein interactions. The performance of an SPR experiment requires the immobilization of one interaction partner (the ligand). During a kinetic measurement, the other interaction partner (the analyte) is injected over the

chip-immobilized ligand. The investigation of both the on-rate and the off-rate is time-consuming and includes several measurement cycles with different concentrations of the analyte due to the concentration dependence of k_{on}. In contrast, the determination of exclusively the off-rate only contains one cycle because k_{off} is a first-order rate constant measured in per second (s^{-1}) and consequently concentrations of the interaction partners do not matter.[24] A study of Murray and coworkers also shows that due to the concentration independence of k_{off} it is also possible to perform an off-rate screen with unpurified compound reaction mixtures containing varying concentrations of the investigational compound. This approach additionally saves time that is usually required for compound separation and purification.[25] Despite the mentioned advantages, off-rate screens, are only hesitantly implemented in discovery research as practically non-existent published studies display. This is surprising, because an off-rate screen is a straightforward approach delivering an appropriate prediction of *in vivo* efficacy and a promising selection of successful lead candidates.

In our study, we originated an SPR off-rate screening assay to spot small molecule antagonists that efficiently target the low-affinity state of the bacterial lectin FimH. The development of antagonists that bind to FimH is a new attempt to treat urinary tract infections (UTIs) that are triggered by a heterogeneous group of uropathogenic *Escherichia coli* (UPEC) strains. For women, the risk to develop at least one symptomatic UTI during life is around 50% to 60%.[26,27] Until now, the only treatment option is a short-term administration of antibiotics. The antibiotic therapy relieves affected patients from symptoms and prevents bacterial expansion to the upper urinary tract, where the infection can progress into a life-threatening pyelonephritis or urosepsis.[28] However, an increasing resistance formation among different bacterial strains against antibiotics requests an effective alternative

medication. FimH is the main virulence factor used by UPECs to provoke an infection by mediating mannose-specific adherence to uroplakin Ia (UPIa) that is expressed on human host cells. The protein is localized at the tip of type 1 pili, filamentous surface organelles of *E. coli*, and consists of a pilin domain ($FimH_{PD}$) and a lectin domain ($FimH_{LD}$). $FimH_{LD}$ acts as a mannose-specific carbohydrate recognition domain (CRD), while $FimH_{PD}$ anchors the protein to FimG, the preceding pilus subunit.

However, the development of FimH antagonists is difficult, mainly due to the conformational nature of FimH. Recent work of Sauer *et al.* demonstrated that FimH mainly adopts three distinct conformational states: a low-affinity, a medium-affinity, and a high-affinity state.[29] Most of the studies performed so far have focused on the development of antagonists binding to high-affinity FimH that can be easily produced by expressing only the isolated lectin domain ($FimH_{LD}$).[30] Cristal structures of $FimH_{LD}$ cocrystallized with diverse monovalent mannosides unveiled a hydrophobic entrance, the tyrosine gate, which is formed by two tyrosines (Tyr48 and Tyr137) and an isoleucine (Ile52). As a consequence several studies concentrated on the identification of high-affinity antagonists with lipophilic aglycones that efficiently interact with the tyrosine gate.[12,31,32,33,34,35,36,37,38,39,40,41,42] Modifications on the mannose moiety resulted in a drop in binding affinity due to an impairment of the sophisticated hydrogen bond network between mannose and amino acids buried in the mannose-binding pocket.[35,43,44] Apart from the α-anomeric position, all hydroxyl groups of mannose form direct hydrogen bonds with the protein (Asp47, Asp54, Asn135, Asp140, and Phe1).[45] However, high-affinity FimH exists in the urinary tract only when the protein is already bound to UPIa and in the presence of tensile forces, which are produced by the bulk flow of urine during micturition. Therapeutically successful FimH antagonists should, however, occupy the low-

affinity state of FimH. When the binding site is blocked, the adhesion of UPECs to the bladder epithelium is prevented leading to bacterial clearance. Under static conditions, the transition between the low- and medium-affinity state is highly dynamic. Unbound full-length FimH ($FimH_{FL}$) is in the low-affinity conformation and adopts a medium-affinity conformation upon binding to UPIa. In contrast to high-affinity FimH, low-affinity FimH has a shallow binding site and displays much weaker affinities against antagonists that were on average 100-fold worse for *n*-heptyl *α*-D-mannopyranoside (HM) and series of biphenyl mannoside-based antagonists.[46,47]

In order to spot promising small molecules for $FimH_{FL}$ with long half-lives, we designed an SPR assay to screen a mannoside-based compound library against $FimH_{FL}$. The study describes both the development and the application of the SPR off-rate screening method. We used biotinylated low-affinity FimH for ligand capture on a streptavidin SPR sensor chip. A production of recombinant full-length FimH ($FimH_{FL}$) is generally sophisticated. It requires a 14-amino acid synthetic peptide donor strand (ADVTITVNGKVVAKR) derived from FimG (DsG) otherwise it is unstable in solution.[29] As in previous studies, we decided to use $t_{1/2}$ as a measure in our experiments.[48]

Methodological development

Protein production and purification

Biotinylated full-length FimH ($FimH_{FL\text{-}B}$) was basically prepared according to Sauer and coworkers.[29] Initially, the plasmid encoding for FimH•FimC was introduced into the *E. coli* K12 strain HM125 that lacks periplasmic proteases.[49] Bacterial cells

harboring the plasmid were grown at 30°C in LB (lysogeny broth) medium supplemented with 100 μg mL^{-1} ampicillin until the density of the bacterial culture reached an OD_{600} of 1.5. At this point, co-expression of FimH and FimC was induced by adding isopropyl β-D-1-thiogalactopyranoside (IPTG) to a final concentration of 1 mM. The bacterial culture was incubated for another 12 to 16 h and subsequently harvested by centrifugation. Bacterial pellets were resuspended in extraction buffer containing 50 mM Tris (tris(hydroxymethyl)aminomethane), pH 7.5, 150 mM NaCl (sodium chloride), 5 mM EDTA (ethylenediaminetetraacetic acid), and 1 mg mL^{-1} polymyxin B sulfate that lysed the bacterial membrane releasing periplasmic proteins. The suspension was prepared with 13 mL extraction buffer per liter of culture medium and stirred for 1.5 h at 4°C. Centrifugation allowed collection of the supernatant containing the released proteins, which are referred to as periplasmic extract. Purification of FimH•FimC and later of $FimH_{FL-B}$ was performed using anion-exchange or cation-exchange fast protein liquid chromatography (FPLC) at 4°C. Before protein mixtures were loaded on either an Uno Q column (anion exchange column from Bio-Rad) or a Mono S column (cation exchange column from GE Healthcare), the columns were equilibrated with the same buffer the protein mixtures were dialyzed against. During the **first** purification step, the supernatant was dialyzed against 20 mM Tris, pH 8.0 and loaded onto the Uno Q column. Subsequently, the flow-through containing the FimH•FimC complex was collected and dialyzed against 10 mM MOPS (3-(*N*-morpholino)propanesulfonic acid), pH 8.0. The **second** purification step comprised loading of the protein mixture onto the Mono S column. FimH•FimC bound to the cation exchange column and was eluted with a linear gradient of increasing NaCl concentrations from 0 to 300 mM. Fractions were collected and those comprising the FimH•FimC complex were pooled and dialyzed

against 20 mM NaH_2PO_4 (sodium dihydrogen phosphate), pH 7.4 and 50 mM NaCl. The FimH•FimC eluate was concentrated to about 40 µM and incubated with a 3-fold molar excess of DsG^{biotin} under constant shaking (70 rpm) for 48 h at 37°C afterwards. The synthetic peptide DsG^{biotin} (purchased from JPT Peptide Technologies) consisted of the donor strand FimG sequence ADVTITVNGKVVAKR (DsG) that was C-terminal amended with a Ttds (trioxatridecan-succinamic acid) linker and a biotinylated lysine. During the incubation time, FimC was substituted by DsG^{biotin} resulting in the formation of FimH•DsG^{biotin} complexes. The mixture was then dialyzed against 20 mM acetic acid, pH 4.5. To separate FimH•DsG^{biotin} ($FimH_{FL\text{-}B}$) from FimC, excess of DsG^{biotin}, and unreacted FimH•FimC, a **third** purification step was performed. On that account, the mixture was loaded onto the Mono S column and $FimH_{FL\text{-}B}$ was eluated with a linear gradient of increasing NaCl concentrations from 0 to 400 mM. Fractions containing $FimH_{FL\text{-}B}$ were pooled and dialyzed against 50 mM NaH_2PO_4. The purity of $FimH_{FL\text{-}B}$ was verified by SDS-PAGE (sodium dodecyl sulfate-polyacrylamide gel electrophoresis) and its mass (calculated mass: 31 291.2 Da) was checked by electrospray ionization (ESI)-mass spectrometry (measured mass: 31 291.0 Da, Functional Genomic Center Zürich; Supporting Information, S1). The yield of purified $FimH_{FL\text{-}B}$ was between 1-2 mg per liter culture medium. The activity of $FimH_{FL\text{-}B}$ was tested using a fluorescence polarization assay (FPA) (Supporting Information, S2). $FimH_{FL\text{-}B}$ was stored at -80°C.

Surface plasmon resonance

Capturing of $FimH_{FL\text{-}B}$:

$FimH_{FL\text{-}B}$ was captured on a streptavidin sensor chip (SA, from GE Healthcare) via its biotinylated peptide. The streptavidin–biotin system is one of the strongest

noncovalent biological interactions, having an exceptionally high affinity (K_D of about 10^{-14}), and is stable under a wide range of chemical conditions.[50,51]

$FimH_{FL-B}$ (100 nM) was pulse-wise delivered to the chip surface by an immobilization wizard using a flow rate of 5 µL min^{-1}. In order to minimize mass transfer during kinetic measurements, the capturing density of the protein was kept rather low (aimed immobilization level: 2.000 RU). Since $FimH_{FL-B}$ is solely biotinylated at the C-terminus of the peptide, the capturing of the protein on the chip is homogenous. The reference surface was blocked by injecting 100 nM biotin-PEG_8-amine (8-units polyethylene glycol spacer) for 60 s at a flow rate of 10 µL min^{-1} since injection of mannosides over an unimmobilized streptavidin surface exhibited unspecific binding events. Unspecific binding of mannose to streptavidin was already observed by Houen and coworkers.[52]

The production and the capturing protocol of $FimH_{FL-B}$ were also applied in another study about target-directed dynamic combinatorial chemistry (tdDCC).[11]

Sample preparation:

Stock compounds were stored at a concentration of 10 mM in 100% DMSO. HBS-EP running buffer (0.01 M HEPES (4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid), pH 7.4, 0.15 M NaCl, 3 mM EDTA, and 0.005% surfactant P20 from GE Healthcare) was used to dilute compounds to intermediate compound dilutions of 100 µM (1% DMSO) at first. These intermediate stocks were further diluted with HBS-EP to 1 µM working solutions (0.01% DMSO).

Off-rate screening:

The off-rate screen was performed at 37°C using HBS-EP as running buffer. According to control measurements the remaining DMSO concentration (0.01%) in the 1 µM working solutions did not affect the refractive index. Thus, an implementation of DMSO solvent correction was not necessary. However, 0.01% DMSO was mixed into the running buffer to keep sample and running buffer conditions as equal as possible. The multi-cycle wizard was used to perform single cycle measurements per mannoside ligand with an association time of 120 s and a dissociation time of 900 s. Positive controls were used to monitor and confirm consistent binding responses. Raw data was processed using Prism (GraphPad Software) by plotting the response units (Y) versus the time of the dissociation phase (X) (Figure 1A). The association phase was cut off and solely the dissociation phase was fitted using the *Dissociation - One phase exponential decay model* (Figure 1B).[53]

The determination of the off-rate with the *Dissociation - One phase exponential decay* model was calculated as follows:

$$Y = (Y_0 - \text{plateau}) \bullet \exp(-k_{off} \bullet X) + \text{plateau} \qquad \text{(Equation 1)}$$

where Y_0 is the binding in response units (RU) at time zero of the dissociation fit and consists of span + plateau. Y decreases to the baseline with the rate constant k_{off} (in inverse units of seconds). The off-rate was used to determine the dissociation half-life ($t_{1/2}$) as follows:

$$t_{1/2} = \ln 2 / k_{off} \qquad \text{(Equation 2)}$$

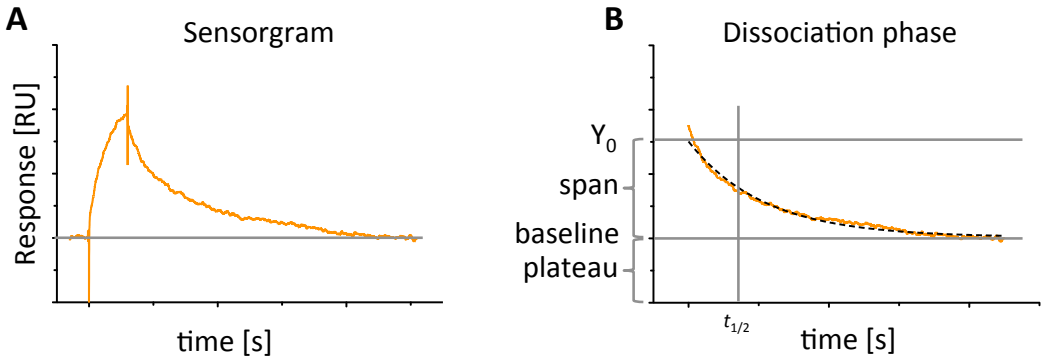


Figure 1. Data processing and evaluation with Prism GraphPad. (A) Single-cycle measurements were performed per compound with an association time of 120 s and a dissociation time of 900 s. (B) The dissociation phase was fitted using the *Dissociation - One phase exponential decay model*.[53]

Multi-cycle kinetics (MCKs) measurements:

Multi-cycle kinetics (MCKs) assays were performed in order to investigate and confirm hits of the off-rate screen. The association and the dissociation phase were adjusted to 180 s (or 240 s) and 1200 s (or 2200 s), respectively. The Biacore T200 Control Software and the Biacore T200 Evaluation Software (both Version 3.0) were applied for data processing and kinetic evaluation. Due to a biphasic behavior of antagonist–$FimH_{FL}$ interactions, which was also investigated in other studies,[54,55] a two-state binding model was applied using double referenced sensorgrams with RI (bulk refraction index) set to 0. Any change in the RI was subtracted out by using the information of the reference cell that lacked $FimH_{FL\text{-}B}$.

Results

1. Off-rate screen

The search for potent monovalent α-D-mannosides over the past decades gave rise to various promising candidates.[12,31,32,33,34,35,36,37,38,39,40,41,42] However, predominantly binding affinity was used as a measure for the identification of lead structures. In addition, until recently, binding affinity studies were performed exclusively with the

isolated FimH lectin domain ($FimH_{LD}$), which is locked in the high-affinity state. To counteract this drawback, a library of 177 *α*-D-mannosides belonging to 10 different structural aglycone classes (Figure 2A-J) was screened against the biotinylated low-affinity state of FimH ($FimH_{FL-B}$) to discover monovalent antagonists that bind to $FimH_{FL-B}$ with long half-lives. Additionally, compounds with an altered mannose moiety (Figure 2K-L) were screened to investigate the impact of an altered hydrogen bond network on the kinetic point of view.

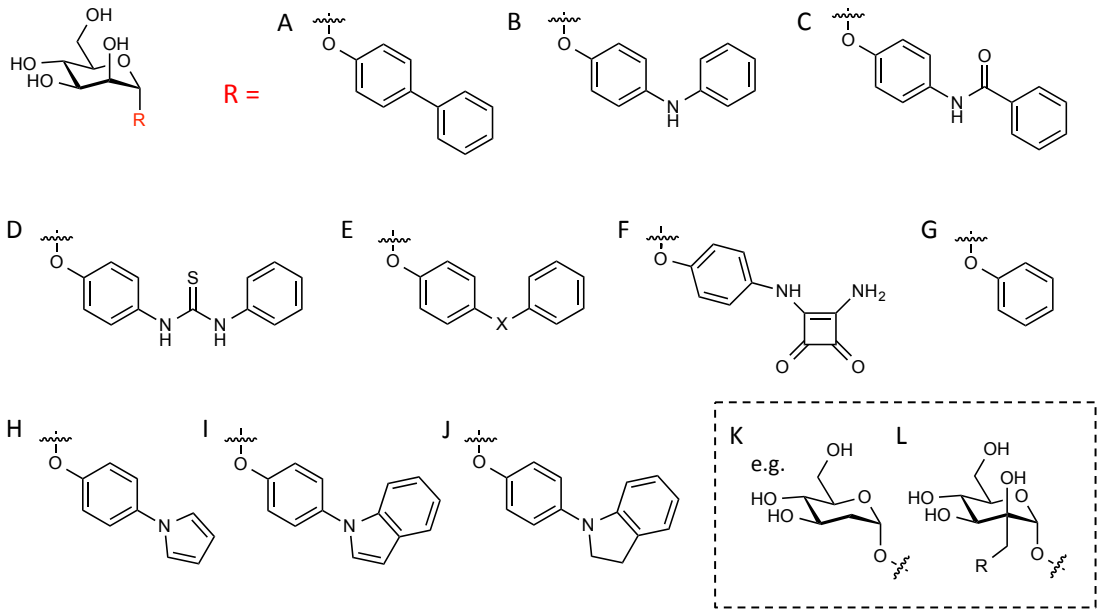


Figure 2. *α*-D-mannoside classes included in the $FimH_{FL-B}$ off-rate screen. Altogether 12 different compound classes were investigated: Biphenyl derivates (A), diphenylamine derivates (B), phenylbenzamide derivates (C), diphenylthiourea derivates (D), spaced bi/triphenyl derivates (E), squaric acid derivates (F), phenyl derivates (G), monocyclic hetroarylphenyl derivates (H), bicyclic hetroarylphenyl derivates (I), indolinphenyl and tetrahydroquinolonephenyl derivates (J), antagonists with modifications on the mannose ring (K), and branched mannosides (L).

Library compounds were assigned to the following structural classes (class size in brackets): biphenyl derivates (42), diphenylamine derivates (16), phenylbenzamide derivates (15), diphenylthiourea derivates (2), spaced bi/triphenyl derivates (15), squaric acid derivates (9), phenyl derivates (5), monocyclic hetroarylphenyl derivates (23), bicyclic hetroarylphenyl derivates (17), indolylphenyl and tetrahydro-quinolinephenyl derivates (20), antagonists with modifications on the mannose ring

(12), and branched mannosides (1). The core structure of the single compound classes is illustrated in Figure 2.

In the following, mannosides that exhibited half-lives longer than 1 min in the screen were defined as hits. In total, 37 mannosides fulfilled this criterion and exhibited $t_{1/2}$s up to 3.3 min. The remaining compounds either displayed dissociation half-lives between 0.1 to 1 min (85 compounds) or did not bind at all (55 compounds). Hits were further assigned to low-response and high-response hits. High-response hit compounds had a response equal or greater than 2.5 RU at the beginning of the dissociation fit whereas low-response compounds displayed lower binding values. In Figure 3, high-response hit compounds are circled.

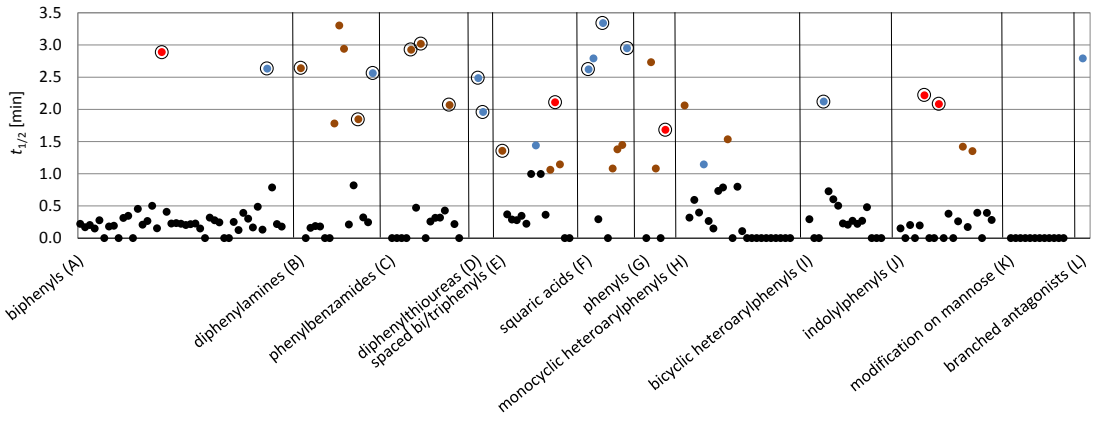


Figure 3. Scatter plot displaying the half-lives of screened mannosides. Compounds were grouped according to their structural classes (Figure 2). Mannosides with a $t_{1/2}$ longer than 1 min are colored according to their sensorgram shapes I-IV, which are described in Figure 4 (red = I, green = II, blue = III, and brown = IV). All hits that exhibited a RU equal or greater than 2.5 at the beginning of the dissociation fit are circled.

Affinity studies of biphenyl derivates binding to $FimH_{LD}$ revealed high-affinity interactions.[34,35,39] This high-affinity antagonism was attributed to π-π staking and hydrophobic interactions between the aromatic phenyl rings and the tyrosine gate.[35] In our study, biphenyl mannosides (Figure 3A) represented the biggest compound class in the off-rate screen existing of 42 distinct molecules. It turned out that all

biphenyl derivates displayed poor dissociation half-lives with the exception of 2 biphenyls (one ester prodrug with a chloride group in *ortho* position of the first phenyl ring and one cyano-substituted in *ortho* and *para* position of the first and second phenyl ring, respectively) that exhibited $t_{1/2}$s of more than 2.5 min. At the first glance, this result may seem surprising because a study of Sauer and coworkers had shown that upon ligand binding the shallow mannose-binding pocket of $FimH_{FL}$ adopts the same conformation (including the tyrosine gate) as $FimH_{LD}$. However, it has also been observed that the pilin domain acts as negative allosteric regulator. Intramolecular conformational dynamics enhance the dissociation of antagonists by more than 100,000-fold and consequently accelerate the release of $FimH_{FL}$.[29]

Bicyclic heteroarylphenyl derivates (Figure 3I), mainly comprising indolylphenyls, and compounds with indolinylphenyl aglycones (Figure 3J) were also under investigation. In search of high-affinity $FimH_{LD}$ antagonists, docking studies predicted an improved fit of the indolyl and the indolinyl moiety with Tyr48 in comparison to the outer phenyl group of biphenyl mannosides.[38] As already observed for biphenyl derivates, the interaction with Tyr48 of $FimH_{FL}$ seems unimportant for long dissociation half-lives. Only 5 hits were detected out of both classes with a total of 37 compounds. Additionally, an *ortho*-chloro substituent on the phenyl ring adjacent to the aromatic oxygen did not improve $t_{1/2}$ as it was the case for the binding affinity of mannoside–$FimH_{LD}$ interactions.[38]

All mannosides with a phenyltriazole aglycone did not bind to $FimH_{FL\text{-}B}$ at all. They belonged to the structural class of monocyclic hetroarylphenyl derivates (Figure 3H) that altogether displayed few hits (3 hits out of 23 compounds). Aglycones of the spaced bi/triphenyl class (Figure 3E) were quite inhomogeneous due to different spacer moieties and the number of phenyl groups. The most promising hit candidate

in this class was a mannoside with a triphenyl aglycone. The class of biphenyls spaced by an amide moiety (Figure 3C) yielded 3 hits out of 15 compounds. All phenylbenzamide hits featured $t_{1/2}$s above 2 min and a response above 2.5 at the beginning of the dissociation fit.

Biphenyls spaced by an aniline moiety (Figure 3B) or a thiourea moiety (Figure 3D) exhibited a relatively high number of hits (6 out of 16 and 2 hits out of 2 compounds, respectively) with dissociation half-lives up to 3.3 min. Another promising monovalent mannoside class binding to $FimH_{FL}$ consisted of compounds with aglycones including a squaric acid moiety (Figure 3F). The screen revealed 7 out of 9 squaric acid derivates that bound to $FimH_{FL\text{-}B}$ with a half-life above 1 min. In a cell-free competitive binding assay, a squaric acid mannoside also showed a high binding affinity towards the mannose-binding site of $FimH_{LD}$.[56] The development of FimH antagonists efficiently targeting the CRD of $FimH_{FL}$ and $FimH_{LD}$ might be beneficial for the treatment of urinary tract infections. A more precise investigation of diphenylamines, squaric acids, and diphenylthiourea derivates seems to be worth the effort in the search of antagonists blocking $FimH_{FL}$ as long as possible.

Binding studies have shown that any modification on the hydroxyl moieties of the pyranose ring considerably diminishes the binding affinity of mannosides to $FimH_{LD}$.[31,44,57] The hydrogen bond network between the mannose moiety and amino acids buried in the CRD seems to be already perfectly aligned. The off-rate screen underlines this assertion. All compounds that were modified on the mannose moiety (Figure 3K) did not bind to $FimH_{FL\text{-}B}$ at all.

The only investigated branched antagonist, a 2-C-branched mannoside (Figure 3L), showed a $t_{1/2}$ of 2.8 min but a low response of only 1.0 RU at the beginning of the

dissociation fit (low-response hit). The same 2-C branched mannoside was already examined in a binding study with $FimH_{LD}$ as target protein. The binding assay unveiled that functional groups in equatorial 2-position of the mannose were not able to occupy a vacant hydrophobic cavity in the binding pocket of $FimH_{LD}$. On the contrary, steric hindrance even deteriorated the affinity compared to the mother compound *n*-heptyl α-D-mannoside (HM).[58] In the present screen, the 2-C-branched mannoside exhibited indeed a considerable half-life, however only a low amount of compound bound to $FimH_{FL\text{-}B}$ during association. The low binding response may probably be connected to the steric constraint.

2. Sensorgram Shapes

Among all antagonists displaying a $t_{1/2}$ above 1 min, diverse sensorgram shapes became apparent indicating different modes of binding. The dissociation phase can be divided into 4 different unbinding mechanism: the off-rate is very slow and the sensorgram does not return to the initial baseline during the investigated dissociation phase (I), the off-rate is slow and the sensorgram returns to the initial baseline during the investigated dissociation phase (II), dissociation is biphasic (very fast dissociation at first and slow unbinding behavior afterwards) and hallmarked by a nonexistent return to the initial baseline during investigated dissociation (III), dissociation is biphasic (very fast dissociation at first and slow unbinding afterwards) and the sensorgram returns to the initial baseline (IV). Figure 4 illustrates the distinct sensorgram shapes, which are also assigned to the compounds in the scatter plot by color coding (Figure 3; I = red, II = green, III = blue, and IV = brown).

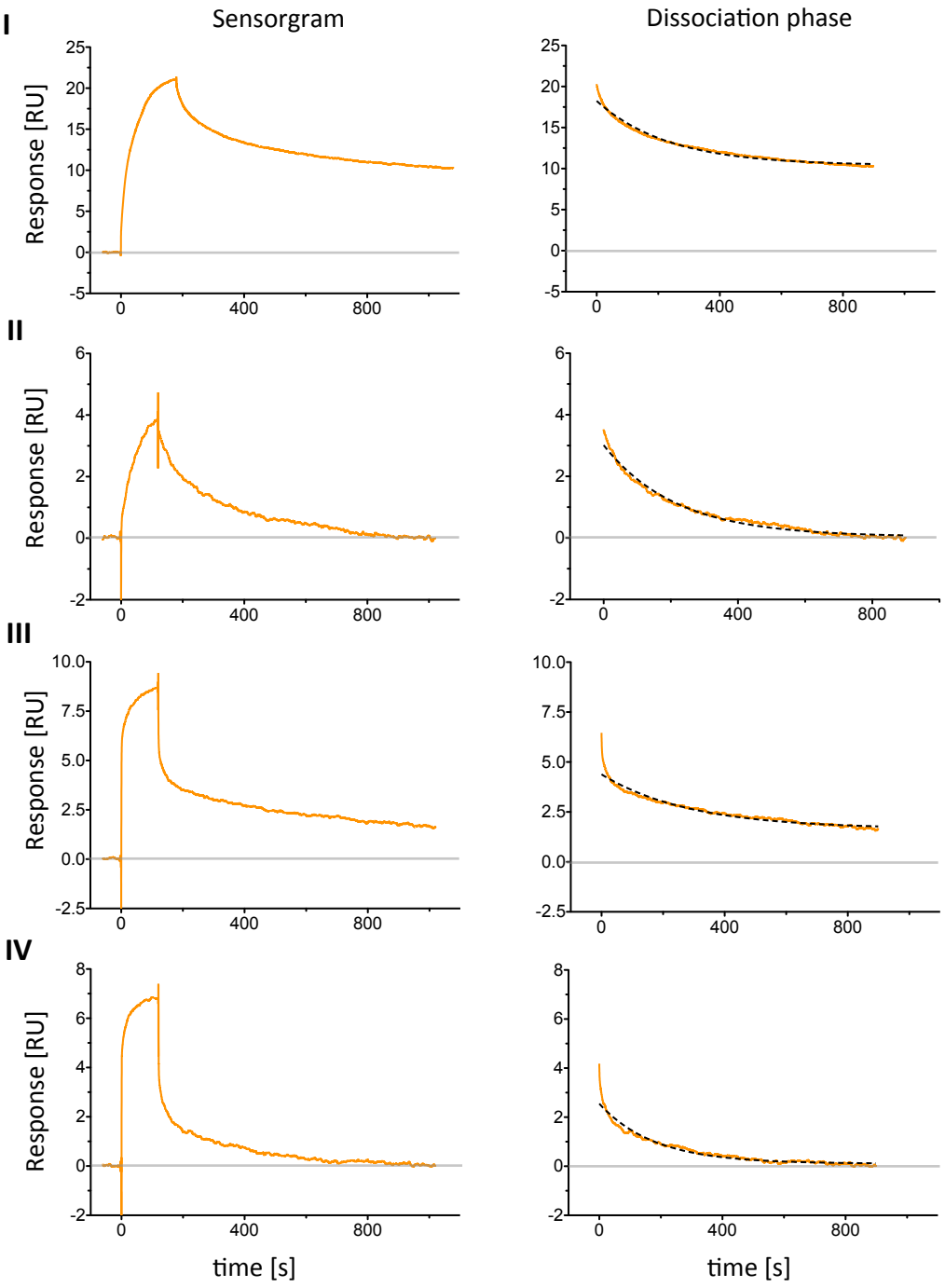


Figure 4. Binding behavior of potential FimH antagonists. Hit compounds displayed 4 various kinetic fingerprints: the off-rate is very slow characterized by a nonexistent return to the initial baseline during the investigated dissociation phase (I), the off-rate is slow and dissociation returns to the initial baseline (II), dissociation is biphasic (very fast dissociation at first and slow unbinding behavior afterwards) and hallmarked by a nonexistent return to the initial baseline during investigated dissociation (III), dissociation is biphasic (very fast dissociation at first and slow unbinding afterwards) and dissociation returns to the initial baseline (IV).

Table 1 lists the hits from the longest $t_{1/2}$ of 3.3 min down to 1.1 min. In addition it shows the sensorgram shape of the respective hit and displays compounds with a response equal or larger than 2.5 at the beginning of the dissociation fit. High-response compounds are indicated in bold.

Table 1. Hit compounds of the $FimH_{FL-B}$ off-rate screen. All compounds with a dissociation half-live above 1 min are ranked by $t_{1/2}$ ($t_{1/2} = (\ln 2/k_{off})$. Additionally, the shape of the dissociation phase is indicated. Hits with a response equal or larger than 2.5 at the beginning of the dissociation fit (high-response hits) are highlighted in bold.

Rank	Class	Shape	$k_{off}\times10^{-3}$	$t_{1/2}$	Rank	Class	Shape	$k_{off}\times10^{-3}$	$t_{1/2}$
			s^{-1}	*min*				s^{-1}	*min*
1	***squaric acid***	III	3.46	3.3	**20**	***phenylbenzamide***	IV	5.59	2.1
2	*diphenylamine*	IV	3.50	3.3	21	*monocycl. HA*	IV	5.61	2.1
3	***phenylbenzamide***	IV	3.83	3.0	**22**	***diphenylthiourea***	II	5.90	2.0
4	***squaric acid***	III	3.92	3.0	23	*diphenylamine*	IV	6.26	1.8
5	*diphenylamine*	IV	3.93	2.9	**24**	***diphenylamine***	IV	6.49	1.8
6	***phenylbenzamide***	IV	3.95	2.9	25	*phenyl*	I	6.87	1.7
7	***biphenyl***	I	4.00	2.9	26	*monocycl. HA*	IV	7.53	1.5
8	*branched mannose*	II	4.14	2.8	27	*squaric acid*	IV	7.99	1.5
8	*squaric acid*	III	4.14	2.8	28	*spaced BiPh*	II	8.03	1.4
10	*phenyl*	IV	4.23	2.7	29	*indolinphenyl*	IV	8.14	1.4
11	***diphenylamine***	I	4.38	2.6	30	*squaric acid*	IV	8.38	1.4
12	***biphenyl***	III	4.39	2.6	**31**	***spaced BiPh***	IV	8.53	1.4
13	***squaric acid***	III	4.41	2.6	32	*indolinphenyl*	IV	8.56	1.4
14	***diphenylamine***	III	4.51	2.6	33	*spaced BiPh*	IV	10.1	1.1
15	***diphenylthiourea***	II	4.65	2.5	33	*monocycl. HA*	III	10.1	1.1
16	***indolinphenyl***	I	5.21	2.2	35	*squaric acid*	IV	10.7	1.1
17	***bicyclic HA***	III	5.45	2.1	35	*phenyl*	IV	10.7	1.1
18	***branched BiPh***	I	5.48	2.1	37	*spaced BiPh*	IV	10.9	1.1
19	***indolinphenyl***	I	5.55	2.1					

The dissociation half-life of antagonists featuring a dissociation phase that does not return to the initial baseline (dissociation type I and III) is probably underestimated. This is due to the fact that the association phase was cut off before the off-rate was determined by the *Dissociation - One phase exponential decay model*.[53] The dissociation phase solely is taken into consideration for the determination of the off-rate. Hence, the information loss can result in a shifted baseline and a misinterpretation of the dissociation rate (Figure 5).

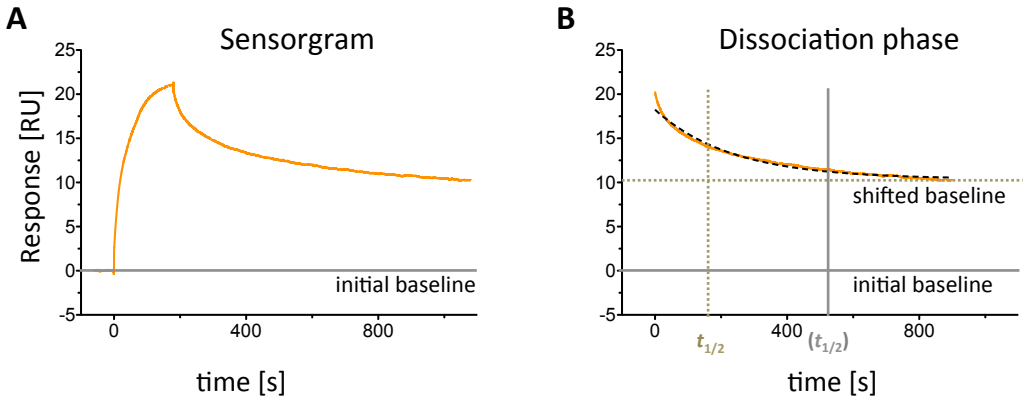


Figure 5. Evaluation of mannosides with type I and II dissociation. A) During the investigated dissociation phase, sensorgrams are not returning to the initial baseline. B) Due to the information loss of the initial baseline, the binding model falsely interprets curve progression of compound dissociation.

3. Multi-cycle kinetics (MCKs)

Multi-cycle kinetics (MCKs) measurements were performed in order to prove the assumption that dissociation half-lives might be underestimated for compounds with type I and III sensorgrams. In addition, MCKs runs allowed a verification of the calculated $t_{1/2}$ in general and an evaluation of the biphasic dissociation behavior in more detail. A detailed kinetic assessment was completed for **1** (type IV dissociation), **2** (type I dissociation), and **3** (type II dissociation) (Figure 6) that displayed half-lives ranging from 2.5 to 3.3 min in the off-rate screen.

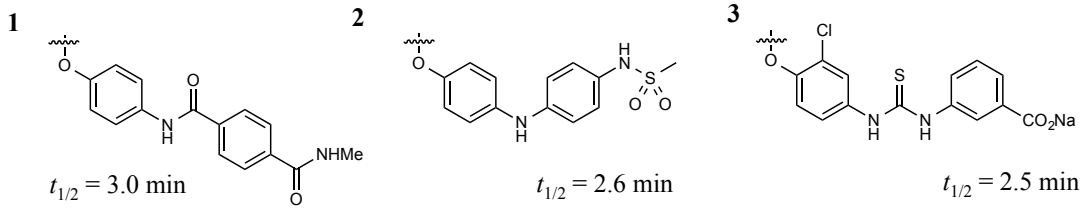


Figure 6. Hit compounds evaluated with multi-cycle kinetics (MCKs). MCKs measurements were performed with phenylbenzamide derivate **1** (group IV, rank 3, OS540), diphenylamine derivate **2** (group I, rank 11, PF28), and diphenylthiourea derivate **3** (group II, rank 15, OS472-2).

As expected, the dissociation phase (type IV) of **1** was characterized by a biphasic behavior (Figure 7A). The same applied to the association phase, with the result that the kinetic data required fitting to a two-state binding model (Equation 3).[55] In this model, the analyte (FimH, A) binds to the ligand (mannoside, B) to form an initial complex (AB) and then undergoes subsequent binding or conformational change to form a more stable complex (A*B).

$$A + B \underset{k_{off,1}}{\overset{k_{on,1}}{\rightleftharpoons}} AB \underset{k_{off,2}}{\overset{k_{on,2}}{\rightleftharpoons}} A^{*}B \qquad \text{(Equation 3)}$$

Against our expectations, it was not possible to fit MCKs measurements of **2** and **3** with a simple Langmuir 1:1 binding model. Although both compounds did not exhibit a biphasic dissociation phase in the off-rate screen (**2** = dissociation type I and **3** = dissociation type II), the kinetic data of **2** and **3** also had to be fitted using the two-state binding model described above (Figure 7B-C). It might be that all FimH antagonists dissociate from the CRD of $FimH_{FL}$ in 2 phases with an initial fast dissociation phase and a slower unbinding phase afterwards resulting in a short $t_{1/2,1}$ and long $t_{1/2,2}$, respectively. The ratio of fast and slow dissociation is variable and depends on the compound binding to $FimH_{FL\text{-}B}$. The studies of Rabbani *et al.* and Yakovenko *et al.* also displayed a biphasic binding behavior of compounds targeting $FimH_{FL}$ and fimbrial tip FimH, respectively.[54,55] The on- and off-rates of **1**, **2**, and **3** determined by MCKs measurements are listed in Table 2.

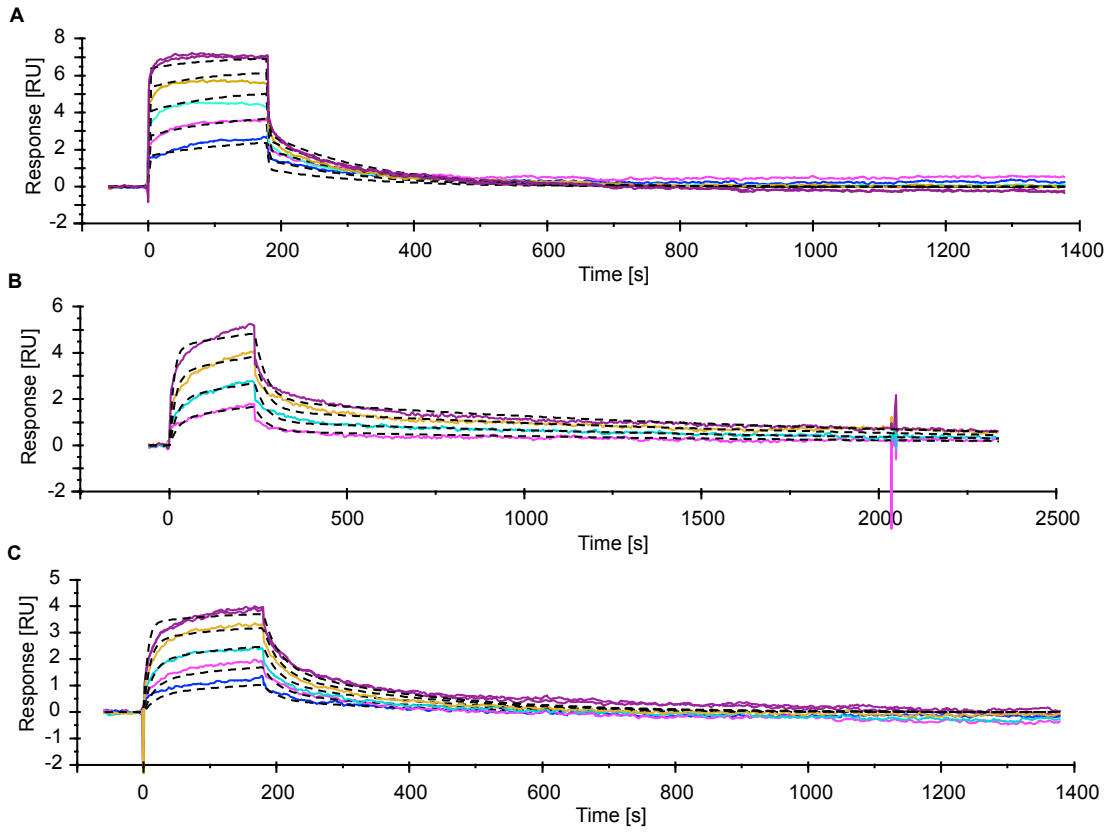


Figure 7. Multi-cycle kinetics (MCKs) measurements of mannosides binding to $FimH_{FL-B}$. MCKs of **1** (A), **2** (B), and **3** (C) binding to $FimH_{FL-B}$ were performed with association times of 180 s or 240 s and a dissociation time of 1200 s or 2200 s using a flow rate of 30 µL min^{-1}. Raw data was fitted using a two-state reaction model with RI (bulk refractive index) set to zero.

Table 2. Kinetic parameters of $FimH_{FL-B}$ with mannosides 1-3 determined by MCKs.
The table shows values of multi-cycle kinetics runs on a SA sensor chip with captured $FimH_{FL-B}$.

Comp.	$k_{on,1}\times10^4$	$k_{off,1}$	$t_{1/2,1}$ [a]	$k_{on,2}\times10^{-4}$	$k_{off,2}\times10^{-4}$	$t_{1/2,2}$	$K_{D,\,kinetic}$ [b]
	$M^{-1}s^{-1}$	s^{-1}	*min*	$M^{-1}s^{-1}$	s^{-1}	*min*	*nM*
1	159	0.932	0.01	57.6	6.7	1.7	314
2	2.4	0.029	0.39	21.9	6.2	18.6	271
3	5.5	0.039	0.29	43.9	50.3	2.3	384

[a] Complex half-lives ($t_{1/2}$) were calculated as follows: $t_{1/2} = \ln 2/k_{off}$.
[b] Kinetic K_D value was determined according to the following equation:
$K_D = k_{off,1}/k_{on,1}\times(k_{off,2}/(k_{off,2} + k_{on,2}))$.

The dissociation half-lives of **1** and **3** determined by the off-rate screen could be confirmed by the MCKs measurements (Table 3). As expected, the $t_{1/2}$ of **2** was even underestimated in the screen due to an incorrect baseline interpretation (Figure 5). The determined dissociation half-life for **2** amounted 2.6 min in the off-rate screen whereas the MCKs measurement calculated a $t_{1/2,2}$ of 18.6 min (Table 3).

Table 3. Comparison of $t_{1/2}$ determined by the off-rate screen (left column) and MCKs measurement (right column).

Compound	$t_{1/2}$ off-rate screen	$t_{1/2,2}$ MCKs
	min	*min*
1	3.0	1.7
2	2.6	18.6
3	2.5	2.3

4. Comparison off-rate versus affinity

Relative affinities (FP_{signal}) of binding to $FimH_{FL}$ are available for all screened mannosides. They were determined by a fluorescence polarization assay (FPA) and were subsequently ranked. The off-rate rank and the relative affinity (FP_{signal}) rank of one compound were plotted to evaluate a potential relative correlation of k_{off} and binding affinity (Figure 8A). The two data sets showed no correlation (Pearson correlation coefficient (r) = -0.43). Solely compounds that did not bind in the off-rate screen at all (not included in the plot, Figure 8A), most notably ligands with modified mannose moiety, had in the majority of cases (70 %) also poor affinities (above 100 nM).

Absolute affinities are available for the 34 best relative affinity-binders. It turned out that all of them had a K_D value below 100 nM. The K_D of these compounds was subsequently plotted versus its k_{off} value (Figure 8B). Only 2 compounds that bound $FimH_{FL}$ with a K_D below 100 nM displayed a half-live above 1 min in the off-rate screen. The most promising of these two compounds concerning its kinetics is a squaric acid derivate (OS420, rank 13, $t_{1/2}$ = 2.6 min, sensorgram shape III, RU at the beginning of the dissociation fit = 4.2).

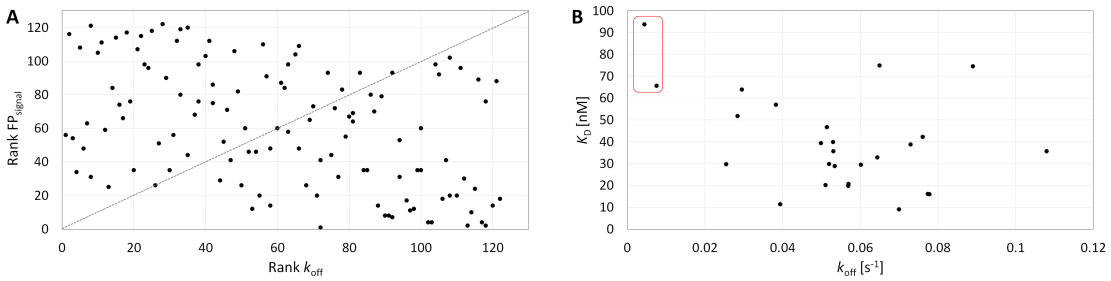


Figure 8. Correlation of affinity and kinetic data of mannosides binding to $FimH_{FL}$. (A) The relative affinity (rank) of mannosides was plotted against its relative off-rate (rank). (B) The K_D of the 34 best relative affinity-binders was determined by fluorescence polarization and plotted against k_{off}. Only 2 of these 34 compounds had a $t_{1/2}$ above 1 min (OS420 (rank 13, type III dissociation, $t_{1/2}$ = 2.6 min) and WS320 (rank 26, type IV dissociation, $t_{1/2}$ = 1.5 min)).

Conclusion

The determination of the dissociation half-life ($t_{1/2}$), which is given as ln2/k_{off}, has been proven as a successful method to predict *in vivo* efficacy at early stages of drug discovery.[21,22,59] For that reason, a surface plasmon resonance (SPR) assay with full-length FimH ($FimH_{FL}$) as target protein was developed to screen the off-rate of a library consisting of 177 α-D-mannosides with various aglycone classes and modified mannose moieties.

The off-rate screen revealed that removal or replacement of individual hydroxyl groups resulted in an inability of all investigated compounds with modified mannose moieties to bind $FimH_{FL}$. The determination of relative binding affinities between $FimH_{FL}$ and ligands with modified mannose moieties resulted in a similar observation. An altered hydrogen bond network implicated significant affinity losses. An inevitable importance of the hydrogen bond network also displayed both affinity [35,43,44] and kinetic [48] investigations of $FimH_{LD}$ interacting with compounds comprising a modified mannose moiety. The hydrogen bond network seems to be the basic condition of complex stability for both mannoside–$FimH_{LD}$ and mannoside–$FimH_{FL}$ complexes. A change in the hydrogen bond network exerts a substantial

influence on both K_D and k_{off}. In the context of $FimH_{LD}$, kinetic data has already been shown that the off-rate is more depended on the maintenance of the hydrogen bond network whereas higher on-rates resulted from a more pronounced interaction of the aglycone with the tyrosine gate.[48]

Biphenyl, phenylindole, and phenylindoline derivates binding to $FimH_{FL}$ showed poor dissociation half-lives indicating a nonexistent or minor support of these aglycones in complex stabilization. In contrast, biphenyl derivates binding to $FimH_{FL}$ performed well in the relative affinity screen. A beneficial impact of aglycones with two aromatic rings was also observed in binding affinity studies with high-affinity FimH as target.[34,35,38,39] An *ortho*-chloro substituent on the first phenyl ring was not identified to prolong $t_{1/2}$ as it contributed to an improvement in binding affinity for both mannoside–$FimH_{FL}$ and mannoside–$FimH_{LD}$ interactions.[38] In general, the performance of mannosides in the off-rate screen compared to relative affinity screen, both having $FimH_{FL}$ as target, revealed no correlation between long half-lives and strong affinities. Among the different aglycone classes, biphenyls spaced by an aniline or a thiourea moiety and mannosides with aglycones including a squaric acid moiety exhibited relatively high numbers of hits with dissociation half-lives up to 3.3 min.

The off-rate screening results also emphasized the importance of a thoroughly considered assay dissociation time. The longer the dissociation time, the longer the measurement time, but the more likely the sensorgram returns to baseline. A return to the initial baseline ensures a correct dissociation fit and thus a reliable estimate of the dissociation half-life. The determined $t_{1/2}$s of compounds, which exhibited no baseline return during the investigated dissociation time, are likely always underestimated.

Especially for these compounds it is worthwhile to reassess $t_{1/2}$ in secondary measurements such as a multi-cycle kinetics (MCKs) assay. For mannoside **2** a $t_{1/2,2}$ of 18.6 min was measured in a MCKs run.

As MCKs measurements have shown, probably all mannosides interact with $FimH_{FL}$ in a biphasic manner. Unbinding of mannoside–$FimH_{FL}$ complexes involved a fast and a slow dissociation component that occurred in varying proportions. As a consequence, an evaluation of the off-rate screen was more complex than expected. Furthermore, the complex dissociation behavior of mannoside–$FimH_{FL}$ complexes presumably made it impossible to find molecular descriptors that correlated with long dissociation half-lives. The descriptor analysis was performed with PaDEL-Descriptor, DRAGON, Schrödinger's QikProp, and CODESSA software and included the off-rates of the whole dataset (177 α-D-mannosides). Nevertheless, a new knowledge basis was generated, including structural elements and insights in the dissociation behavior of mannoside–$FimH_{FL}$ complexes, to develop potent $FimH_{FL}$ antagonists for an efficient treatment of UTIs.

References

1. Paul, S.M., Mytelka, D.S., Dunwiddie, C.T., Persinger, C.C., Munos, B.H., Lindborg, S.R., and Schacht, A.L., How to improve R&D productivity: The pharmaceutical industry's grand challenge. *Nat. Rev. Drug Discov.*, **2010**, *9*(3), 203-214.
2. Carnero, A., High-throughput screening in drug discovery. *Clin. Transl. Oncol.*, **2006**, *8*(7), 482-490.
3. Kola, I. and Landis, J., Can the pharmaceutical industry reduce attrition rates? *Nat. Rev. Drug Discov.*, **2004**, *3*, 711.
4. Schuhmacher, A., Gassmann, O., and Hinder, M., Changing R&D models in research-based pharmaceutical companies. *J. Transl. Med.*, **2016**, *14*(1), 105.
5. Mohs, R.C. and Greig, N.H., Drug discovery and development: Role of basic biological research. *Alzheimers Dement. (N Y)*, **2017**, *3*(4), 651-657.
6. Ehrlich, P., Address in pathology, on chemiotherapy: Delivered before the seventeenth international congress of medicine. *Brit. Med. J.*, **1913**, *2*(2746), 353-359.
7. Geschwinder, S., Affinity-based screening methodologies and their application in the hit-to-lead phase, in *Lead generation*, **2016**, Wiley‐VCH Verlag GmbH.
8. Bergsdorf, C. and Ottl, J., Affinity-based screening techniques: Their impact and benefit to increase the number of high quality leads. *Expert Opin. Drug Discov.*, **2010**, *5*(11), 1095-1107.

9. Fang, Y., Ligand–receptor interaction platforms and their applications for drug discovery. *Expert Opin. Drug Discov.*, **2012**, *7*(10), 969-988.
10. Sager, C.P., Fiege, B., Zihlmann, P., Vannam, R., Rabbani, S., Jakob, R.P., Preston, R.C., Zalewski, A., Maier, T., Peczuh, M.W., and Ernst, B., The price of flexibility - a case study on septanoses as pyranose mimetics. *Chem. Sci.*, **2018**, *9*(3), 646-654.
11. Frei, P., Pang, L., Silbermann, M., Eris, D., Muhlethaler, T., Schwardt, O., and Ernst, B., Target-directed dynamic combinatorial chemistry: A study on potentials and pitfalls as exemplified on a bacterial target. *Chemistry*, **2017**, *23*(48), 11570-11577.
12. Kleeb, S., Pang, L., Mayer, K., Eris, D., Sigl, A., Preston, R.C., Zihlmann, P., Sharpe, T., Jakob, R.P., Abgottspon, D., Hutter, A.S., Scharenberg, M., Jiang, X., Navarra, G., Rabbani, S., Smiesko, M., Lüdin, N., Bezençon, J., Schwardt, O., Maier, T., and Ernst, B., FimH antagonists: Bioisosteres to improve the in vitro and in vivo PK/PD profile. *J. Med. Chem.*, **2015**, *58*(5), 2221-2239.
13. Lee, M.M. and Peterson, B.R., Quantification of small molecule–protein interactions using FRET between tryptophan and the pacific blue fluorophore. *ACS Omega*, **2016**, *1*(6), 1266-1276.
14. Zheng, X., Li, Z., Beeram, S., Podariu, M., Matsuda, R., Pfaunmiller, E.L., White, C.J., Carter, N., and Hage, D.S., Analysis of biomolecular interactions using affinity microcolumns: A review. *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci.*, **2014**, *0*, 49-63.
15. Copeland, R.A., Pompliano, D.L., and Meek, T.D., Drug–target residence time and its implications for lead optimization. *Nat. Rev. Drug Discov.*, **2006**, *5*(9), 730-739.
16. Lu, H. and Tonge, P.J., Drug–target residence time: Critical information for lead optimization. *Curr. Opin. Chem. Biol.*, **2010**, *14*(4), 467-474.
17. Tummino, P.J. and Copeland, R.A., Residence time of receptor–ligand complexes and its effect on biological function. *Biochemistry*, **2008**, *47*(20), 5481-5492.
18. Copeland, R.A., Conformational adaptation in drug–target interactions and residence time. *Future Med. Chem.*, **2011**, *3*(12), 1491-1501.
19. Markgren, P.O., Schaal, W., Hamalainen, M., Karlen, A., Hallberg, A., Samuelsson, B., and Danielson, U.H., Relationships between structure and interaction kinetics for HIV-1 protease inhibitors. *J. Med. Chem.*, **2002**, *45*(25), 5430-5439.
20. Nunez, S., Venhorst, J., and Kruse, C.G., Target–drug interactions: First principles and their application to drug discovery. *Drug Discov. Today*, **2012**, *17*(1-2), 10-22.
21. Guo, D., Mulder-Krieger, T., Ijzerman, A.P., and Heitman, L.H., Functional efficacy of adenosine A(2A) receptor agonists is positively correlated to their receptor residence time. *Br. J. of Pharmacol.*, **2012**, *166*(6), 1846-1859.
22. Wood, E.R., Truesdale, A.T., McDonald, O.B., Yuan, D., Hassell, A., Dickerson, S.H., Ellis, B., Pennisi, C., Horne, E., Lackey, K., Alligood, K.J., Rusnak, D.W., Gilmer, T.M., and Shewchuk, L., A unique structure for epidermal growth factor receptor bound to GW572016 (Lapatinib): Relationships among protein conformation, inhibitor off-rate, and receptor activity in tumor cells. *Cancer Res.*, **2004**, *64*(18), 6652-6659.
23. Seow, V., Lim, J., Cotterell, A.J., Yau, M.-K., Xu, W., Lohman, R.-J., Kok, W.M., Stoermer, M.J., Sweet, M.J., Reid, R.C., Suen, J.Y., and Fairlie, D.P., Receptor residence time trumps drug-likeness and oral bioavailability in determining efficacy of complement C5a antagonists. *Sci. Rep.*, **2016**, *6*, 24575.
24. Corzo, J., Time, the forgotten dimension of ligand binding teaching. *Biochem. Mol. Biol. Educ.*, **2006**, *34*(6), 413-416.
25. Murray, J.B., Roughley, S.D., Matassova, N., and Brough, P.A., Off-rate screening (ORS) by surface plasmon resonance. An efficient method to kinetically sample hit to lead chemical space from unpurified reaction products. *J. Med. Chem.*, **2014**, *57*(7), 2845-2850.
26. Foxman, B., Epidemiology of urinary tract infections: Incidence, morbidity, and economic costs. *Am. J. Med.*, **2002**, *113 Suppl 1A*, S5-13.
27. Leigh, D., Urinary tract infections, in *Topley and Wilson's principles of bacteriology, virology and immunity*, **1990**, Butler and Tanler Ltd., Frome and London.
28. Hooton, T.M. and Stamm, W.E., Diagnosis and treatment of uncomplicated urinary tract infection. *Infect. Dis. Clin. North Am.*, **1997**, *11*(3), 551-581.
29. Sauer, M.M., Jakob, R.P., Eras, J., Baday, S., Eriş, D., Navarra, G., Bernèche, S., Ernst, B., Maier, T., and Glockshuber, R., Catch-bond mechanism of the bacterial adhesin FimH. *Nat. Commun.*, **2016**, *7*, 10738.

30. Rabbani, S., Jiang, X., Schwardt, O., and Ernst, B., Expression of the carbohydrate recognition domain of FimH and development of a competitive binding assay. *Anal. Biochem.*, **2010**, *407*(2), 188-195.
31. Bouckaert, J., Berglund, J., Schembri, M., De Genst, E., Cools, L., Wuhrer, M., Hung, C.S., Pinkner, J., Slattegard, R., Zavialov, A., Choudhury, D., Langermann, S., Hultgren, S.J., Wyns, L., Klemm, P., Oscarson, S., Knight, S.D., and De Greve, H., Receptor binding studies disclose a novel class of high-affinity inhibitors of the Escherichia coli FimH adhesin. *Mol. Microbiol.*, **2005**, *55*(2), 441-455.
32. 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., and Gouin, S.G., Thiazolylaminomannosides as potent antiadhesives of type 1 piliated Escherichia coli isolated from Crohn's disease patients. *J. Med. Chem.*, **2013**, *56*(13), 5395-5406.
33. Chalopin, T., Alvarez Dorta, D., Sivignon, A., Caudan, M., Dumych, T.I., Bilyy, R.O., Deniaud, D., Barnich, N., Bouckaert, J., and Gouin, S.G., Second generation of thiazolylmannosides, FimH antagonists for E. coli-induced Crohn's disease. *Org. Biomol. Chem.*, **2016**, *14*(16), 3913-3925.
34. Cusumano, C.K., Pinkner, J.S., Han, Z., Greene, S.E., Ford, B.A., Crowley, J.R., Henderson, J.P., Janetka, J.W., and Hultgren, S.J., Treatment and prevention of urinary tract infection with orally active FimH inhibitors. *Sci. Transl. Med.*, **2011**, *3*(109), 109ra115.
35. Han, Z., Pinkner, J.S., Ford, B., Obermann, R., Nolan, W., Wildman, S.A., Hobbs, D., Ellenberger, T., Cusumano, C.K., Hultgren, S.J., and Janetka, J.W., Structure-based drug design and optimization of mannoside bacterial FimH antagonists. *J. Med. Chem.*, **2010**, *53*(12), 4779-4792.
36. Han, Z., Pinkner, J.S., Ford, B., Chorell, E., Crowley, J.M., Cusumano, C.K., Campbell, S., Henderson, J.P., Hultgren, S.J., and Janetka, J.W., Lead optimization studies on FimH antagonists: Discovery of potent and orally bioavailable ortho-substituted biphenyl mannosides. *J. Med. Chem.*, **2012**, *55*(8), 3945-3959.
37. 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., and Janetka, J.W., Antivirulence isoquinolone mannosides: Optimization of the biaryl aglycone for FimH lectin binding affinity and efficacy in the treatment of chronic UTI. *ChemMedChem*, **2016**, *11*(4), 367-373.
38. Jiang, X., Abgottspon, D., Kleeb, S., Rabbani, S., Scharenberg, M., Wittwer, M., Haug, M., Schwardt, O., and Ernst, B., Antiadhesion therapy for urinary tract infections: A balanced PK/PD profile proved to be key for success. *J. Med. Chem.*, **2012**, *55*(10), 4700-4713.
39. 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., and Ernst, B., FimH antagonists for the oral treatment of urinary tract infections: From design and synthesis to in vitro and in vivo evaluation. *J. Med. Chem.*, **2010**, *53*(24), 8627-8641.
40. Pang, L., Kleeb, S., Lemme, K., Rabbani, S., Scharenberg, M., Zalewski, A., Schadler, F., Schwardt, O., and Ernst, B., FimH antagonists: Structure-activity and structure-property relationships for biphenyl alpha-D-mannopyranosides. *ChemMedChem*, **2012**, *7*(8), 1404-1422.
41. Schwardt, O., Rabbani, S., Hartmann, M., Abgottspon, D., Wittwer, M., Kleeb, S., Zalewski, A., Smiesko, M., Cutting, B., and Ernst, B., Design, synthesis and biological evaluation of mannosyl triazoles as FimH antagonists. *Bioorg. Med. Chem.*, **2011**, *19*(21), 6454-6473.
42. Sperling, O., Fuchs, A., and 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*(21), 3913-3922.
43. Fiege, B., Rabbani, S., Preston, R.C., Jakob, R.P., Zihlmann, P., Schwardt, O., Jiang, X., Maier, T., and Ernst, B., The tyrosine gate of the bacterial lectin FimH: A conformational analysis by NMR spectroscopy and X-ray crystallography. *Chembiochem*, **2015**, *16*(8), 1235-1246.
44. Old, D.C., Inhibition of the interaction between fimbrial haemagglutinins and erythrocytes by D-mannose and other carbohydrates. *J. Gen. Microbiol.*, **1972**, *71*(1), 149-157.
45. Hung, C.S., Bouckaert, J., Hung, D., Pinkner, J., Widberg, C., DeFusco, A., Auguste, C.G., Strouse, R., Langermann, S., Waksman, G., and Hultgren, S.J., Structural basis of tropism of Escherichia coli to the bladder during urinary tract infection. *Mol. Microbiol.*, **2002**, *44*(4), 903-915.

46. Mayer, K., Eris, D., Schwardt, O., Sager, C.P., Rabbani, S., Kleeb, S., and Ernst, B., Urinary tract infection: Which conformation of the bacterial lectin FimH is therapeutically relevant? *J. Med. Chem.*, **2017**, *60*(13), 5646-5662.
47. Schönemann, W., Cramer, J., Mühlethaler, T., Fiege, B., Silbermann, M., Rabbani, S., Dätwyler, P., Zihlmann, P., Jakob, R.P. Sager, C.P, Smieško, M., Schwardt, O., Maier, T., Ernst B., The different affinity stages of FimH and their thermodynamics of binding. *Manuscript*.
48. Zihlmann, P., Silbermann, M., Sharpe, T., Jiang, X., Muhlethaler, T., Jakob, R.P., Rabbani, S., Sager, C.P., Frei, P., Pang, L., Maier, T., and Ernst, B., KinITC - One method supports both thermodynamic and kinetic sars as exemplified on FimH antagonists. *Chemistry*, **2018**, *24*(49), 13049-13057.
49. Meerman, H.J. and Georgiou, G., Construction and characterization of a set of E. coli strains deficient in all known loci affecting the proteolytic stability of secreted recombinant proteins. *Bio/Technology*, **1994**, *12*, 1107.
50. Green, N.M., Avidin and streptavidin. *Methods Enzymol.*, **1990**, *184*, 51-67.
51. Sano, T. and Cantor, C.R., Cooperative biotin binding by streptavidin. Electrophoretic behavior and subunit association of streptavidin in the presence of 6 M urea. *J. Biol. Chem.*, **1990**, *265*(6), 3369-3373.
52. Houen, G. and Hansen, K., Interference of sugars with the binding of biotin to streptavidin and avidin. *J. Immunol. Methods*, **1997**, *210*(2), 115-123.
53. Graphpad software. Equation: One phase decay. *GraphPad Curve Fitting Guide*.
54. Yakovenko, O., Tchesnokova, V., Sokurenko, E.V., and Thomas, W.E., Inactive conformation enhances binding function in physiological conditions. *PNAS*, **2015**, *112*(32), 9884-9889.
55. Rabbani, S., Fiege, B., Eris, D., Silbermann, M., Jakob, R.P., Navarra, G., Maier, T., and Ernst, B., Conformational switch of the bacterial adhesin FimH in the absence of the regulatory domain: Engineering a minimalistic allosteric system. *J. Biol. Chem.*, **2018**, *293*(5), 1835-1849.
56. Scharenberg, M., Schwardt, O., Rabbani, S., and Ernst, B., Target selectivity of FimH antagonists. *J. Med. Chem.*, **2012**, *55*(22), 9810-9816.
57. Zihlmann, P.J., X.; Sager, C. P.; Fiege, B.; Jakob, R. P.; Siegrist, S.; Zalewski, A.; Rabbani, S.; Eriş, D.; Silbermann, M.; Pang, L.; Mühlethaler, T.; Sharpe, T.; Maier, T.; Ernst, B., High-affinity carbohydrate-lectin interaction: How nature makes it possible. *unpublished results*.
58. Schönemann, W., Lindegger, M., Rabbani, S., Zihlmann, P., Schwardt, O., and Ernst, B., 2-C-branched mannosides as a novel family of FimH antagonists - synthesis and biological evaluation. *Perspect. Sci.*, **2017**, *11*, 53-61.
59. Lu, H., England, K., am Ende, C., Truglio, J.J., Luckner, S., Reddy, B.G., Marlenee, N.L., Knudson, S.E., Knudson, D.L., Bowen, R.A., Kisker, C., Slayden, R.A., and Tonge, P.J., Slow-onset inhibition of the FabI enoyl reductase from francisella tularensis: Residence time and in vivo activity. *ACS Chem Biol*, **2009**, *4*(3), 221-231.

Supporting Information

S.1 Molecular mass of $FimH_{FL-B}$

In order to check the molecular mass of $FimH_{FL-B,}$ the molecular weight of the purified protein was determined by electrospray ionisation mass spectrometry (ESI-MS) analysis (Functional Genomics Center Zurich). The measured mass of $FimH_{FL-B}$ amounted to 31 291.0 Da and is in accordance with the calculated mass of 31 291.2 Da.

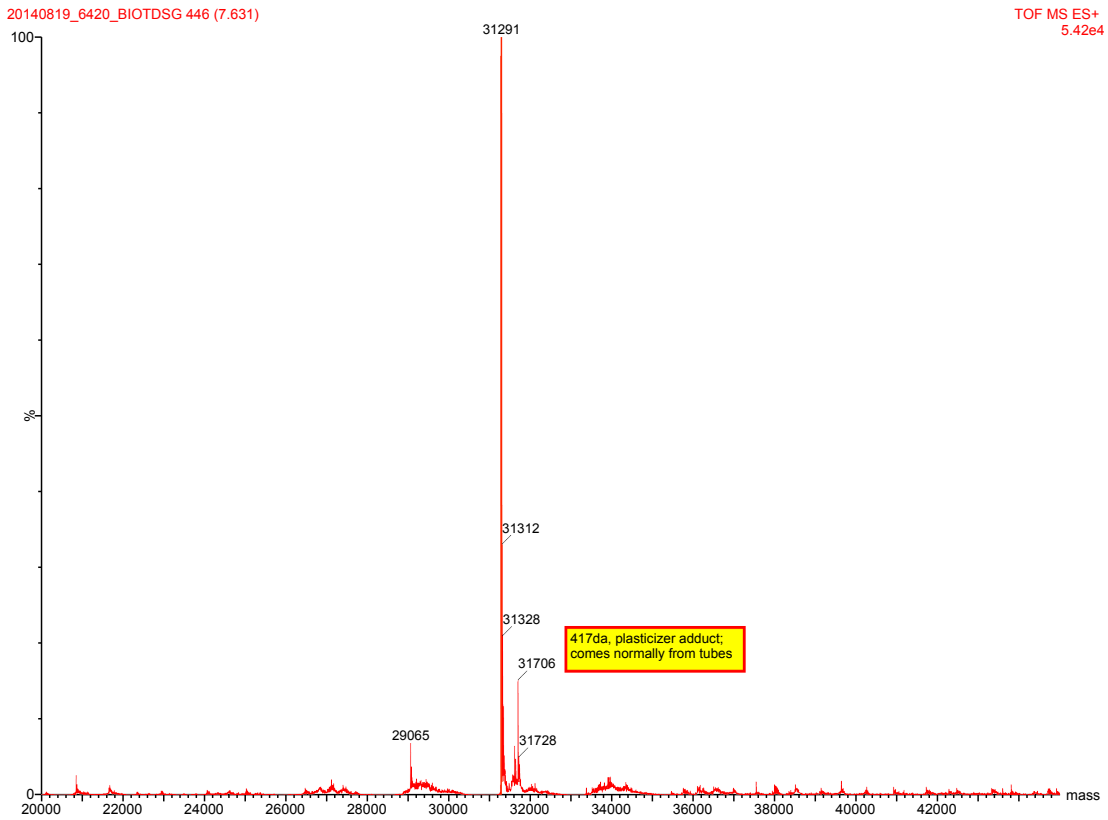


Figure S1. Mass spectrum of the purified $FimH_{FL-B}$ revealed a peak (100%) at 31291 kDa.

S2 Activity of $FimH_{FL-B}$

S2.1 $FimH_{FL-B}$ activity compared to $FimH_{FL}$ using a FPA assay

The activity of the protein was determined by a fluorescence polarization assay (FPA) first described for $FimH_{LD}$ (isolated lectin domain) [1] and later adapted for $FimH_{FL}$ (full-length FimH using a peptide without biotin).[2] The binding affinity of $FimH_{FL-B}$ was compared with $FimH_{FL}$ using a FITC-labeled tracer (constant concentration) as interaction partner. Therefore, dilution series of the proteins (final concentration: beginning with 10 µM and then decreasing with a dilution factor of 2) mixed with tracer (final concentration: 50 nM) were prepared. For this purpose, the protein batches and a tracer stock (10 mM in 100% DMSO) were diluted with assay buffer (20 mM HEPES, 150 mM NaCl, 50 µg/mL BSA, pH 7.4). Samples were incubated for 1 hour before fluorescence polarization was measured with the Synergy™ H1 Multi-Mode microplate reader (BioTek Intruments). K_D values of $FimH_{FL-B}$ and $FimH_{FL-B}$ were determined using Prism (GraphPad Software) and were comparable (K_D of $FimH_{FL-B}$: 129.4 nM versus K_D of $FimH_{FL}$: 111.0 nM).

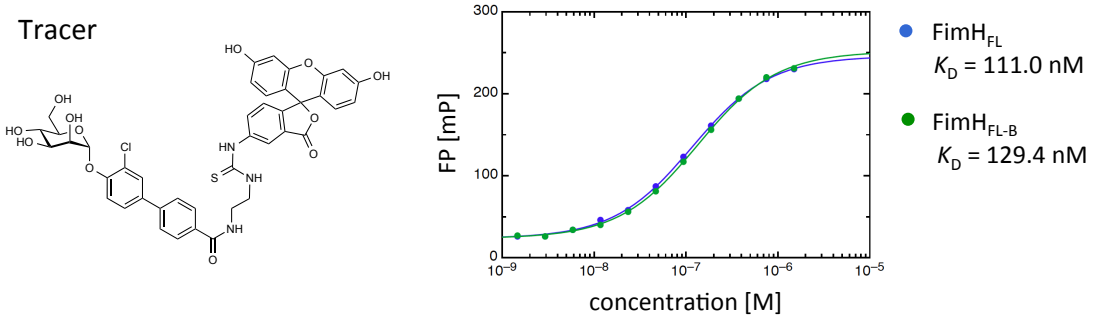
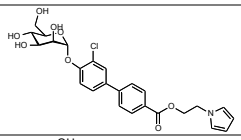
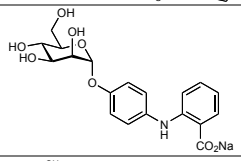
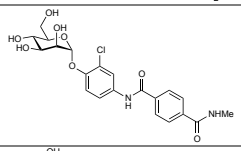
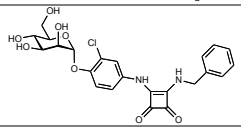
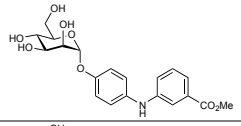
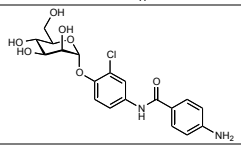
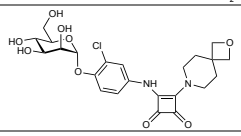
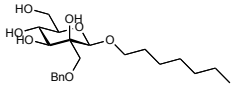
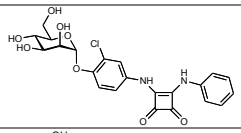
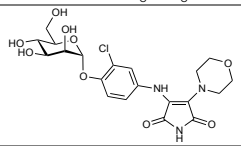
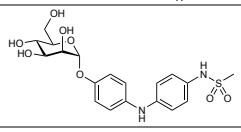
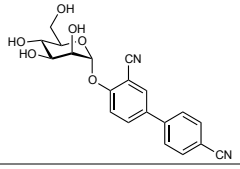
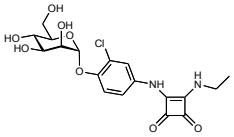
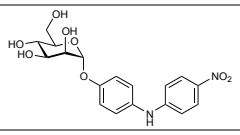
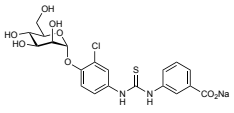
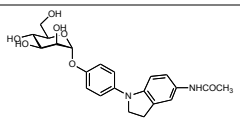
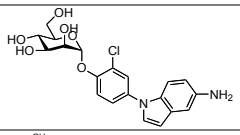
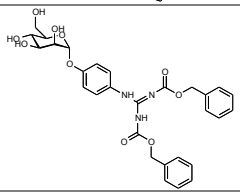
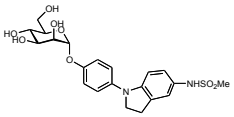
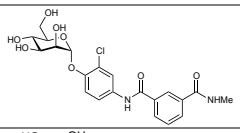
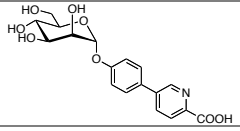
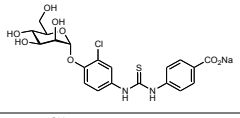
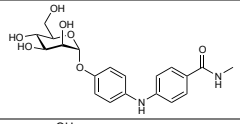
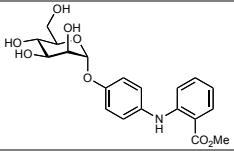


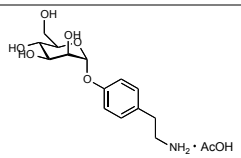
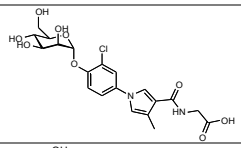
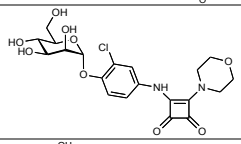
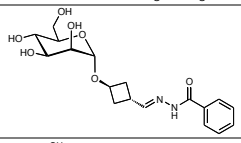
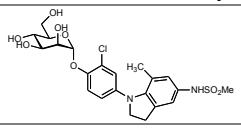
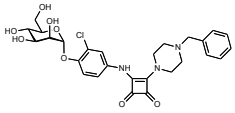
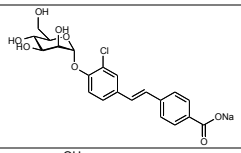
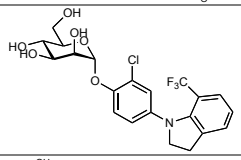
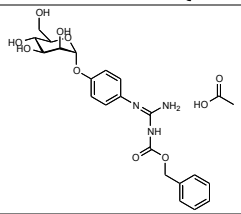
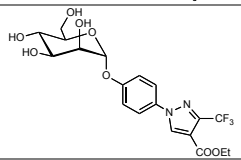
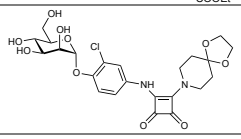
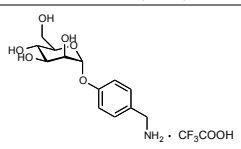
Figure S2. Structure of the FITC-labeled tracer and binding curves of $FimH_{FL}$ and $FimH_{FL-B}$

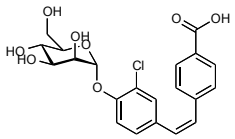
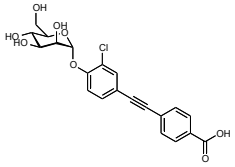
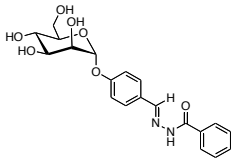
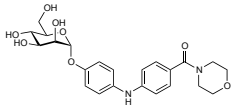
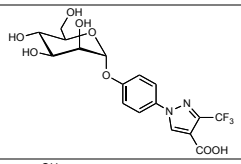
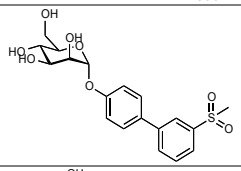
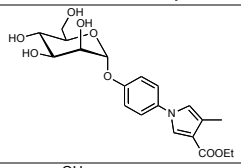
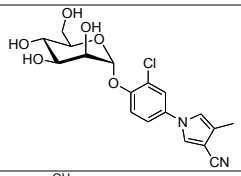
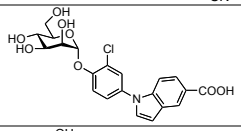
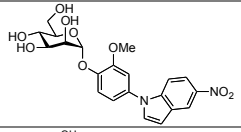
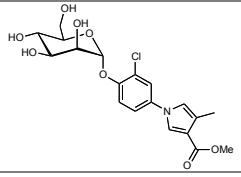
S3 Chemical structures and general information

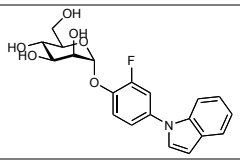
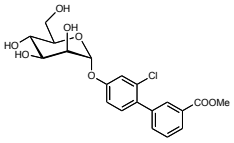
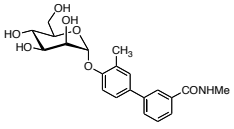
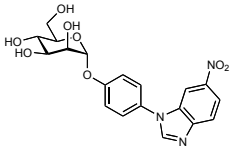
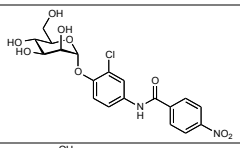
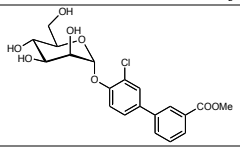
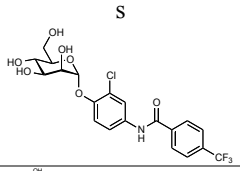
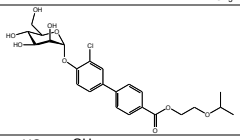
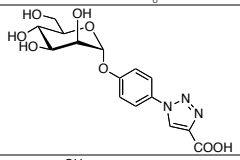
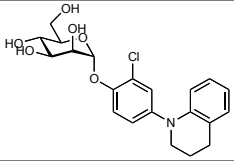
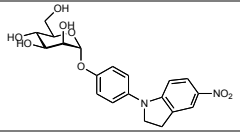
Table S1. All compounds included in the off-rate screen are ranked according to their k_{off}. In the table the chemical structure, the dissociation half-life ($t_{1/2}$), and the binding response at the beginning (RU_B) and the end (RU_E) of the dissociation phase are indicated. In addition, the $logD_{7.4}$ if available is listed and the fluorescence polarization (FP) value, which was determined in a FP assay that measured the relative affinity of mannosides.

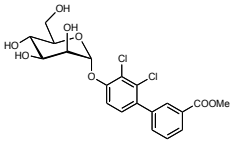
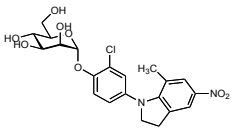
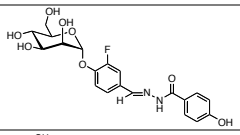
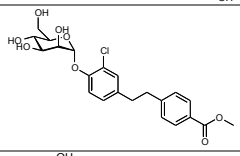
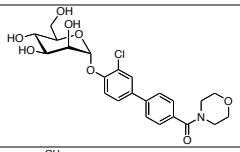
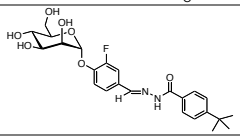
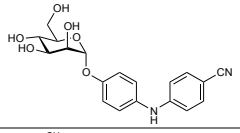
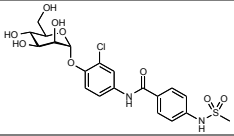
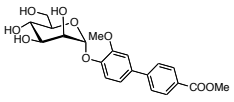
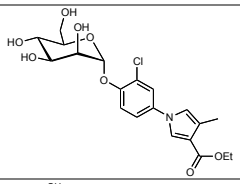
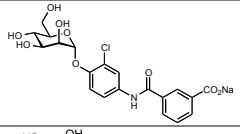
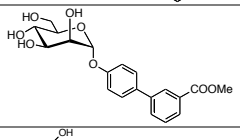
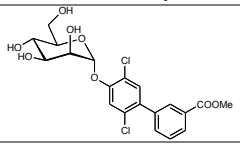
Compound		RU_B	RU_E	k_{off}	$t_{1/2}$ [min]	R^2	$logD_{7.4}$	FP_{Signal}
Rank: 1 OS395		4.4	1.7	3.46E-03	3.3	1.0	1.1±0.0	82
Rank: 2 OS465.2		2.3	0.2	3.50E-03	3.3	1.0	< -1.5	163
Rank: 3 OS540		4.7	0.0	3.83E-03	3.0	1.0	0.4±0.1	80
Rank: 4 OS556		9.7	2.9	3.92E-03	3.0	1.0	-1.1±0.2	69
Rank: 5 OS550		2.2	0.1	3.93E-03	2.9	1.0	1.7±0.0	146
Rank: 6 OS494		2.7	0.4	3.95E-03	2.9	1.0	0.2±0.1	76
Rank: 7 WS176		6.5	2.5	4.00E-03	2.9	1.0	n.d.	86
Rank: 8 OS349		1.0	0.2	4.14E-03	2.8	1.0	3.0±0.0	189
Rank: 8 OS397		2.1	0.6	4.14E-03	2.8	0.9	1.1±0.0	68
Rank: 10 HS030		1.5	0.0	4.23E-03	2.7	1.0	0.3±0.2	140
Rank: 11 PF28		21.1	12.5	4.38E-03	2.6	1.0	-0.4±0.0	153

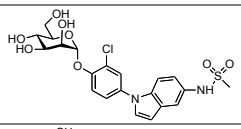
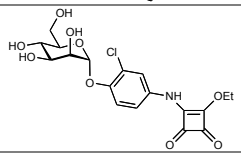
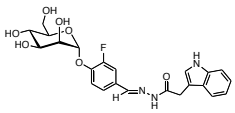
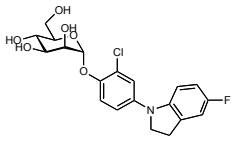
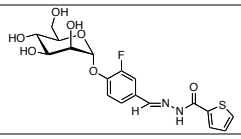
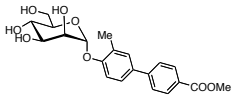
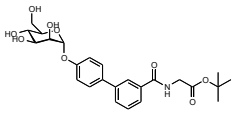
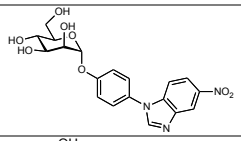
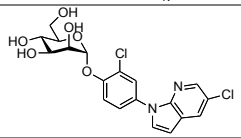
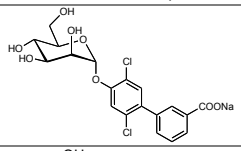
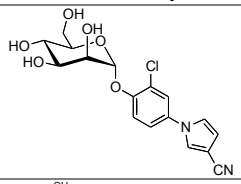
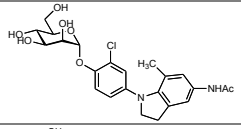
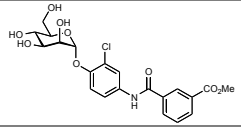
Compound		RU_B	RU_E	k_{off}	$t_{1/2}$ [min]	R^2	$logD_{7.4}$	FP_{Signal}
Rank: 12 PF57		2.5	-0.2	4.39E-03	2.6	0.9	n.d.	84
Rank: 13 OS420		4.2	2.3	4.41E-03	2.6	0.9	-0.5±0.1	62
Rank: 14 PF18		13.2	4.6	4.51E-03	2.6	1.0	1.5±0.1	117
Rank: 15 OS472.2		3.0	0.0	4.65E-03	2.5	1.0	n.d.	158
Rank: 16 JXH-III-057		9.1	4.7	5.21E-03	2.2	1.0	1.2±0.0	102
Rank: 17 JXH-@090		3.8	1.7	5.45E-03	2.1	1.0	n.d.	90
Rank: 18 WS268		2.7	1.2	5.48E-03	2.1	1.0	n.d.	164
Rank: 19 JXH-III-0712		7.0	3.3	5.55E-03	2.1	1.0	1.1±0.1	106
Rank: 20 OS542		2.5	0.1	5.59E-03	2.1	1.0	0.6±0.1	71
Rank: 21 PLJ01068A		2.4	0.3	5.61E-03	2.1	1.0	n.d.	143
Rank: 22 OS470.2		3.1	0.4	5.90E-03	2.0	1.0	n.d.	159
Rank: 23 OS562		2.7	0.7	6.26E-03	1.8	0.9	0.1±0.1	133
Rank: 24 OS465.1		2.0	0.6	6.49E-03	1.8	0.9	2.4±0.1	130

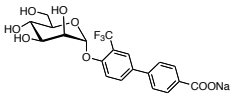
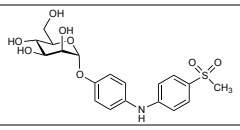
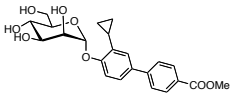
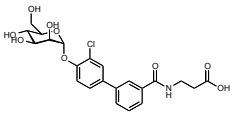
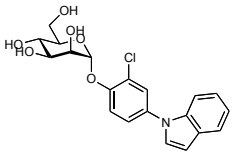
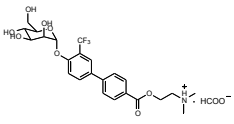
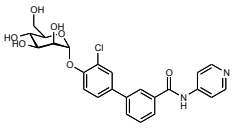
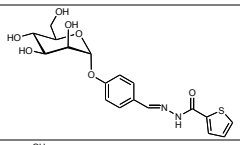
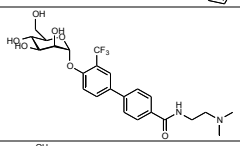
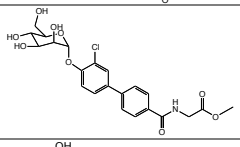
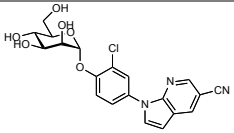
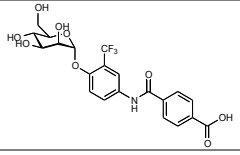
Compound		RU_B	RU_E	k_{off}	$t_{1/2}$ [min]	R^2	$logD_{7.4}$	FP_{Signal}
Rank: 25 WS302A		4.3	1.4	6.87E-03	1.7	1.0	1.3±0.2	169
Rank: 26 WS320		1.4	0.4	7.53E-03	1.5	0.9	n.d.	64
Rank: 27 OS464		1.5	-0.1	7.99E-03	1.5	0.9	-1.1±0.1	77
Rank: 28 PLJ02174B		2.0	-0.7	8.03E-03	1.4	1.0	n.d	190
Rank: 29 JXH-IV-073		1.1	0.3	8.14E-03	1.4	0.9	1.4±0.1	123
Rank: 30 OS396		1.5	0.0	8.38E-03	1.4	0.9	0.8±0.1	71
Rank: 31 WS384		5.2	-0.2	8.53E-03	1.4	0.9	n.d.	8
Rank: 32 JXH-IV-090		1.1	0.1	8.56E-03	1.4	1.0	n.d.	157
Rank: 33 WS265A		1.3	0.6	1.01E-02	1.1	0.9	n.d.	171
Rank: 33 PL2238A		1.4	0.6	1.01E-02	1.1	0.9	2.1±0.0	110
Rank: 35 OS555		2.3	0.9	1.07E-02	1.1	0.9	-0.7±0.1	73
Rank: 35 WS296		0.7	0.3	1.07E-02	1.1	0.8	n.d.	177

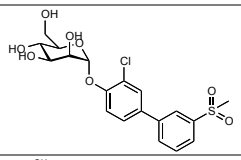
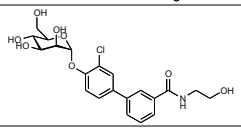
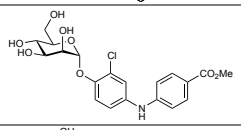
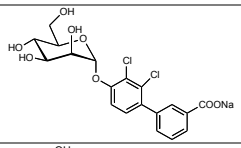
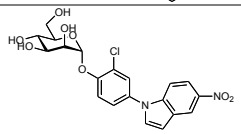
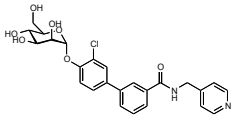
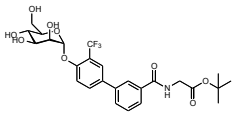
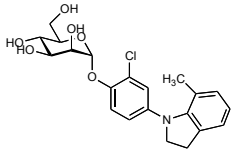
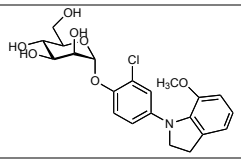
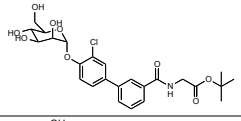
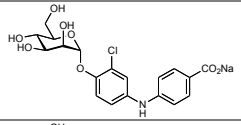
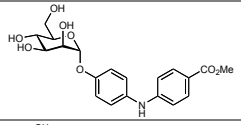
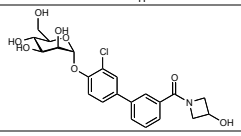
Compound		RU_B	RU_E	k_{off}	$t_{1/2}$ [min]	R^2	$logD_{7.4}$	FP_{Signal}
Rank: 37 WS400		2.2	0.1	1.09E-02	1.1	1.0	n.d	93
Rank: 38 WS370		1.0	0.2	1.16E-02	1.0	0.9	n.d.	106
Rank: 38 PLJ02052B		0.6	0.0	1.16E-02	1.0	0.8	n.d.	133
Rank: 40 OS559		1.2	0.3	1.41E-02	0.8	0.9	0.1±0.1	135
Rank: 41 PLJ02238B		0.6	0.1	1.45E-02	0.8	0.8	< -1.5	157
Rank: 42 WS283		0.5	0.0	1.47E-02	0.8	0.8	0.5±0.0	105
Rank: 42 PLJ02173A		0.5	-0.2	1.47E-02	0.8	0.9	2.3±0.0	118
Rank: 44 PLJ03135A		0.8	-0.4	1.58E-02	0.7	0.9	2.0±0.1	65
Rank: 45 JX4-@0682		1.1	-0.2	1.59E-02	0.7	0.9	1.1*	78
Rank: 46 JXH-@065		2.0	-0.1	1.92E-02	0.6	0.9	n.d.	98
Rank: 47 PLJ02115C		1.0	-0.3	1.95E-02	0.6	0.8	n.d.	72

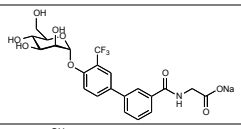
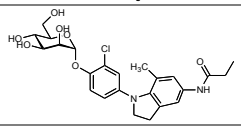
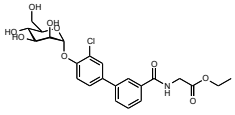
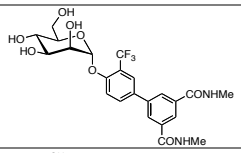
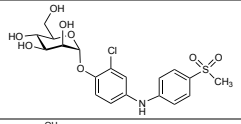
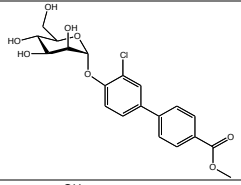
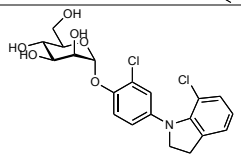
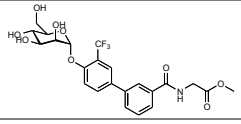
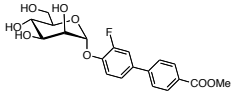
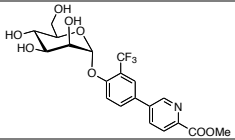
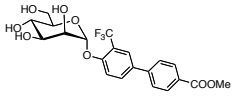
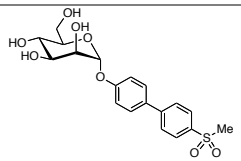
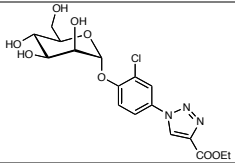
Compound		RU_B	RU_E	k_{off}	$t_{1/2}$ [min]	R^2	$logD_{7.4}$	FP_{Signal}
Rank: 48 JXH-2432		0.9	0.2	2.30E-02	0.5	0.7	2.4*	141
Rank: 49 PLJ01303A		1.2	1.1	2.31E-02	0.5	0.9	2.7±0.1	111
Rank: 50 PLJ02123A		1.1	-0.7	2.38E-02	0.5	0.9	1.2±0.0	64
Rank: 51 JXH-V0321		1.1	0.3	2.41E-02	0.5	0.9	n.d.	85
Rank: 52 OS492		1.8	-0.1	2.45E-02	0.5	0.9	1.6±0.1	75
Rank: 53 PLJ01285A		0.8	-0.5	2.55E-02	0.5	0.9	2.7±0.1	54
Rank: 54 OS525		1.6	0.0	2.71E-02	0.4	0.9	2.6±0.1	75
Rank: 55 WS180		1.1	-0.9	2.84E-02	0.4	0.8	n.d.	59
Rank: 56 PLJ01159A		1.2	-0.4	2.92E-02	0.4	0.9	n.d.	152
Rank: 57 JXH-IV 003		0.8	0.0	2.94E-02	0.4	0.9	3.1±0.1	125
Rank: 58 JXH@200		1.7	0.1	2.95E-02	0.4	0.9	1.9*	76

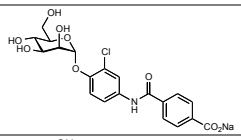
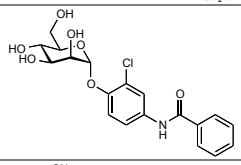
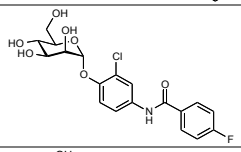
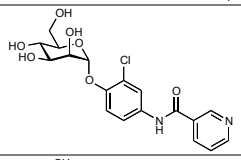
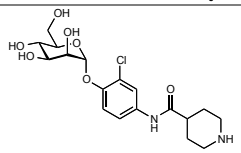
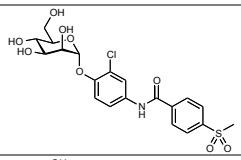
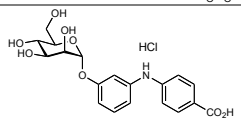
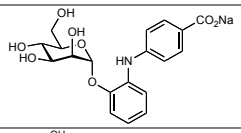
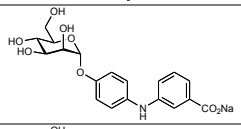
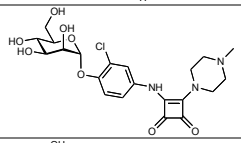
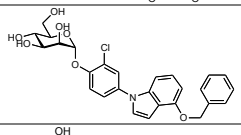
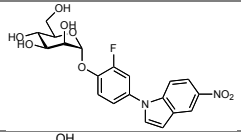
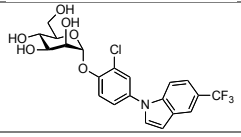
Compound		RU_B	RU_E	k_{off}	$t_{1/2}$ [min]	R^2	$logD_{7.4}$	FP_{Signal}
Rank: 58 PLJ10301A		3.8	-0.6	2.95E-02	0.4	0.9	3.2±0.1	56
Rank: 60 JXH-IV-028		1.3	0.0	3.06E-02	0.4	0.8	2.8±0.0	85
Rank: 61 PLJ03061B		1.0	0.1	3.14E-02	0.4	0.9	n.d.	119
Rank: 62 WS395		1.0	0.0	3.20E-02	0.4	0.9	2.7±0.1	117
Rank: 63 PLJ01128A		1.4	0.5	3.34E-02	0.3	0.9	0.9±0.1	83
Rank: 63 PLJ03061E		1.2	0.0	3.34E-02	0.3	0.9	n.d.	133
Rank: 65 PF11		1.7	0.4	3.60E-02	0.3	0.9	1.6±0.1	139
Rank: 66 OS529		2.0	0.0	3.64E-02	0.3	0.9	0.4±0.1	76
Rank: 66 PLJ01178A		1.3	0.0	3.64E-02	0.3	0.9	1.8±0.1	147
Rank: 68 PLJ02115A		1.8	-0.2	3.65E-02	0.3	0.9	2.8±0.1	64
Rank: 69 OS531_2		1.1	0.0	3.67E-02	0.3	0.8	< -1.5	88
Rank: 70 PLJ01126A		1.2	0.5	3.70E-02	0.3	0.8	2.0±0.1	101
Rank: 71 PLJ01302A		3.7	-0.3	3.83E-02	0.3	1.0	3.0±0.0	59

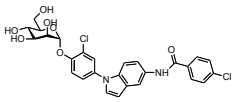
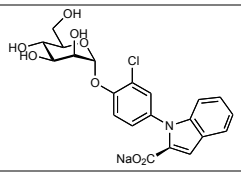
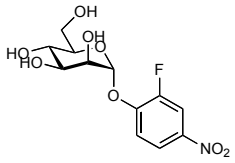
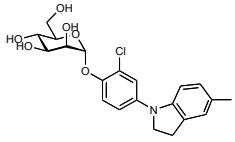
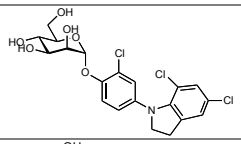
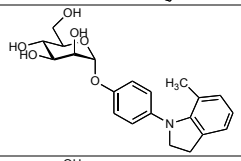
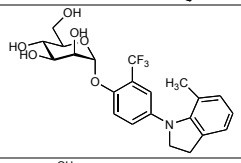
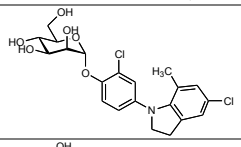
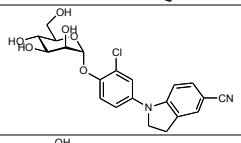
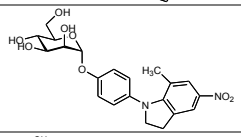
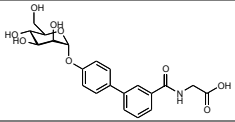
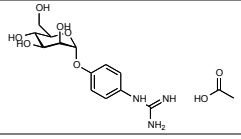
Compound		RU_B	RU_E	k_{off}	$t_{1/2}$ [min]	R^2	$logD_{7.4}$	FP_{Signal}
Rank: 72 JXH-@101		2.1	-0.2	3.94E-02	0.3	0.9	1.8*	72
Rank: 72 OS394		3.2	-0.7	3.94E-02	0.3	0.9	0.1±0.0	34
Rank: 74 PLJ03061C		1.2	0.1	4.02E-02	0.3	0.9	n.d.	128
Rank: 75 JXH-III-034		2.2	0.3	4.10E-02	0.3	0.9	3.2±0.3	73
Rank: 76 PLJ03061D		1.2	0.0	4.15E-02	0.3	0.8	n.d.	100
Rank: 77 PLJ01179A		1.4	-0.1	4.17E-02	0.3	0.9	2.4±0.2	68
Rank: 78 WS280A		1.1	-0.1	4.18E-02	0.3	0.9	2.0±0.1	113
Rank: 79 JXH-V0322		1.5	0.3	4.31E-02	0.3	0.9	n.d.	8
Rank: 80 JXH-@0751		1.6	0.2	4.32E-02	0.3	0.9	3.4*	91
Rank: 81 PLJ01302B		1.2	0.4	4.37E-02	0.3	0.8	n.d.	94
Rank: 81 PLJ03136A		1.0	-0.2	4.37E-02	0.3	0.9	1.5±0.1	87
Rank: 83 JXH-IV-069		0.7	0.1	4.45E-02	0.3	0.7	1.7±0.1	128
Rank: 84 OS531_1		2.0	-0.4	4.50E-02	0.3	0.9	1.8±0.1	71

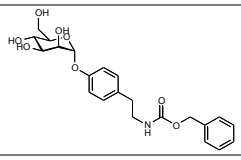
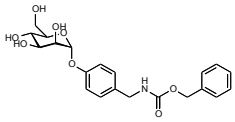
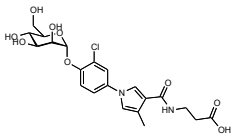
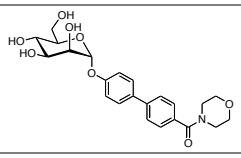
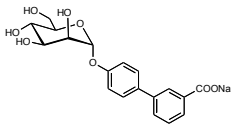
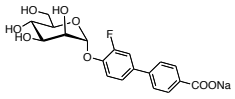
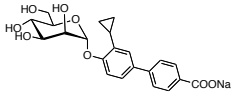
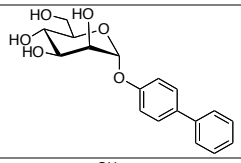
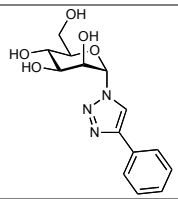
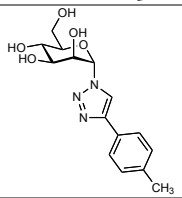
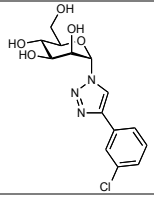
Compound		RU_B	RU_E	k_{off}	$t_{1/2}$ [min]	R^2	$logD_{7.4}$	FP_{Signal}
Rank: 85 PLJ01194B		0.7	-0.2	4.62E-02	0.2	0.8	-0.8±0.1	71
Rank: 86 PF14.1		1.9	0.5	4.71E-02	0.2	0.8	-0.1±0.0	110
Rank: 87 PLJ01181A		1.4	0.0	4.74E-02	0.2	0.9	2.9±0.1	97
Rank: 88 WS291		1.9	-0.1	4.99E-02	0.2	0.9	n.d.	56
Rank: 89 JXH-232		1.2	0.1	5.08E-02	0.2	0.9	1.8*	108
Rank: 90 WS248		3.1	-0.6	5.11E-02	0.2	0.9	1.7±0.0	51
Rank: 91 WS312		6.6	-0.8	5.14E-02	0.2	1.0	2.4±0.1	51
Rank: 92 PLJ03127A		0.9	0.0	5.20E-02	0.2	0.8	n.d.	128
Rank: 92 WS294		3.5	0.3	5.20E-02	0.2	0.9	0.7±0.0	46
Rank: 94 WS063		2.4	0.0	5.24E-02	0.2	0.8	1.7±0.1	68
Rank: 94 JXH-@073		0.9	0.2	5.24E-02	0.2	0.8	1.6*	79
Rank: 96 WS260		0.9	0.0	5.30E-02	0.2	0.8	n.d.	57

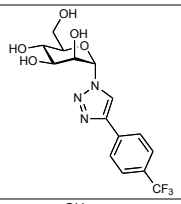
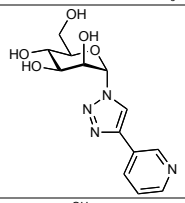
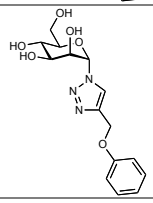
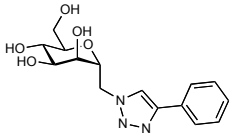
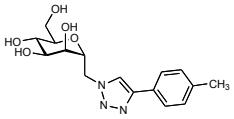
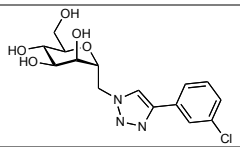
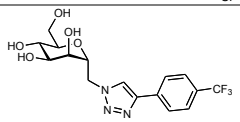
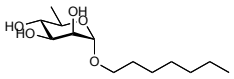
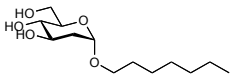
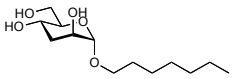
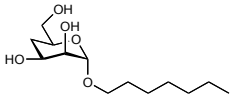
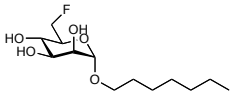
Compound		RU_B	RU_E	k_{off}	$t_{1/2}$ [min]	R^2	$logD_{7.4}$	FP_{Signal}
Rank: 97 WS284		1.3	-0.1	5.31E-02	0.2	0.9	1.0±0.0	53
Rank: 98 WS311		3.1	0.4	5.34E-02	0.2	0.9	0.9±0.1	54
Rank: 99 OS553		5.1	-0.3	5.46E-02	0.2	1.0	2.4±0.0	71
Rank: 100 PLJ02019C		1.1	0.4	5.56E-02	0.2	0.8	n.d.	85
Rank: 100 JHX-224		1.8	0.1	5.56E-02	0.2	0.9	1.8*	71
Rank: 102 WS310		4.2	-0.5	5.69E-02	0.2	0.9	1.8±0.1	45
Rank: 103 WS253		3.5	-0.4	5.70E-02	0.2	0.9	2.8±0.0	45
Rank: 104 JXH-III-044		0.9	0.2	5.72E-02	0.2	0.7	2.8±0.0	133
Rank: 105 JXH-III-047		1.1	0.1	5.89E-02	0.2	0.8	2.2±0.0	127
Rank: 106 WS322		3.4	-0.4	6.02E-02	0.2	0.9	n.d.	58
Rank: 107 OS471		3.1	0.5	6.23E-02	0.2	0.9	-1.4±.0.3	72
Rank: 108 OS461		1.8	0.9	6.44E-02	0.2	0.5	1.6±0.0	134
Rank: 108 WS317		3.1	0.0	6.44E-02	0.2	0.9	0.8±0.0	59

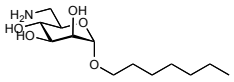
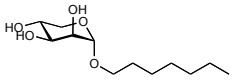
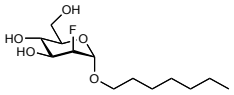
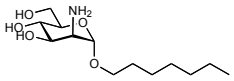
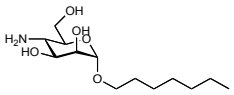
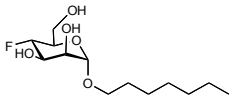
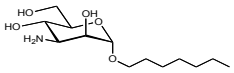
Compound		RU_B	RU_E	k_{off}	$t_{1/2}$ [min]	R^2	$logD_{7.4}$	FP_{Signal}
Rank: 110 WS253B		1.9	0.0	6.50E-02	0.2	1.0	n.d.	59
Rank: 111 JXH-IV-075		0.9	0.1	6.71E-02	0.2	0.6	2.00±0.0	130
Rank: 112 WS215B		3.3	-0.1	6.87E-02	0.2	1.0	1.8±0.1	66
Rank: 113 PLJ02126A		4.7	-0.2	7.00E-02	0.2	1.0	1.1±0.0	41
Rank: 114 PF14.2		3.1	-2.1	7.30E-02	0.2	0.9	0.7±0.1	52
Rank: 115 WS011		1.7	0.0	7.61E-02	0.2	0.9	2.3**	60
Rank: 116 JXH-III-032		1.4	0.4	7.69E-02	0.2	0.7	3.00±0.0	122
Rank: 117 WS257		3.8	-0.5	7.74E-02	0.1	0.9	1.7±0.1	45
Rank: 118 PLJ01175A		0.8	0.0	7.78E-02	0.1	0.6	2.7±0.1	106
Rank: 118 PLJ02045A		2.6	-0.4	7.78E-02	0.1	0.9	1.3±0.1	41
Rank: 120 PLJ01194A		1.3	0.2	8.90E-02	0.1	0.6	2.8±0.1	56
Rank: 121 PLJ01129A		0.9	0.3	9.32E-02	0.1	0.3	0.4±0.0	120
Rank: 122 PLJ02135A		1.4	0.1	1.08E-01	0.1	0.7	0.7±0.0	58

Compound		RU_B	RU_E	k_{off}	$t_{1/2}$ [min]	R^2	$logD_{7.4}$	FP_{Signal}
Rank: 123 OS487		0.8	-0.5	3.79E-02	0.0	0.8	-1.1±0.2	84
Rank: 129 OS488		0.5	-0.1	5.00E-03	0.0	0.8	1.5±0.1	85
Rank: 123 OS493		0.3	-0.6	7.72E-03	0.0	0.8	1.7±0.1	72
Rank: 123 OS498		0.8	-0.4	2.50E-02	0.0	0.8	0.3±0.1	86
Rank: 123 OS500		0.2	-0.5	5.45E-03	0.0	0.8	<-1.5	89
Rank: 123 OS534		0.9	-0.9	2.66E-02	0.0	0.8	0.1±0.1	65
Rank: 123 OS456		0.7	-0.2	2.15E-03	0.0	1.0	<-1.5	145
Rank: 123 OS457		1.1	0.1	2.48E-03	0.0	1.0	<-1.5	180
Rank: 123 OS459		0.3	0.1	5.12E-03	0.0	0.6	<-1.5	169
Rank: 123 OS421		0.3	-0.5	4.63E-03	0.0	0.8	-1.3±0.1	83
Rank: 123 JXH-@041		0.2	-0.1	9.69E-03	0.0	0.8	2.7*	172
Rank: 123 JXH-2431		0.8	0.1	1.00E-02	0.0	0.9	1.9*	90
Rank: 123 JXH-@099		0.2	-0.3	1.17E-02	0.0	0.9	n.d.	124

Compound		RU_B	RU_E	k_{off}	$t_{1/2}$ [min]	R^2	$logD_{7.4}$	FP_{Signal}
Rank: 123 JXH-@097		0.1	-0.3	1.42E-02	0.0	0.8	n.d.	171
Rank: 123 JXH-III-020		0.4	0.0	2.98E-02	0.0	0.7	n.d.	144
Rank: 123 JXH-281		0.1	-2010.0	3.02E-08	0.0		n.d.	169
Rank: 123 JXH-III-033		0.2	-3861.0	2.19E-08	0.0		3.4±0.1	87
Rank: 123 JHX-III-045		0.5	0.1	2.28E-02	0.0	0.6	n.d.	135
Rank: 123 JHX-III-061		0.1	-0.9	1.01E-03	0.0	0.9	2.1±0.0	172
Rank: 123 JHX-III-066		0.3	-0.7	5.13E-04	0.0	0.8	3.4±0.1	109
Rank: 123 JHX-III-078		0.2	-0.4	2.39E-03	0.0	0.9	3.5±0.1	152
Rank: 123 JHX-III-084		0.2	~ -257.7	~ 1.15E-06	0.0	0.8	n.d.	n.d.
Rank: 123 JXH-IV-051		0.4	0.0	9.45E-03	0.0	0.8	2.3±0.0	123
Rank: 123 WS280B		1.3	0.1	5.04E+00	0.0	0.9	n.d.	124
Rank: 123 WS269A		0.2	-0.1	3.97E-03	0.0	0.9	n.d.	182

Compound		RU_B	RU_E	k_{off}	$t_{1/2}$ [min]	R^2	$logD_{7.4}$	FP_{Signal}
Rank: 123 WS301		0.1	-0.1	6.60E-03	0.0	0.4	1.6±0.1	161
Rank: 123 WS239		0.1	-0.1	7.87E-03	0.0	0.3	n.d.	172
Rank: 123 WS319		0.4	-0.2	9.04E-03	0.0	0.8	n.d.	n.d.
Rank: 123 PLJ01076A		0.5	0.1	5.42E-02	0.0	0.5	0.2±0.1	143
Rank: 123 PLJ01278B		0.2	-0.4	1.61E-04	0.0	0.3	<-1.5	149
Rank: 123 PLJ01175B		0.1	-0.1	7.36E-03	0.0	0.5	<-1.5	135
Rank: 123 PLJ01181B		0.6	-0.2	4.62E-02	0.0	0.8	0.8±0.1	114
Rank: 123 PLJ01192A		0.0	-0.2	3.69E-02	0.0	0.2	2.1±0.1	160
Rank: 123 OS373		-0.4	-8.1	2.23E-04	0.0	1.0	0.2	178
Rank: 123 O376		-0.4	~ -577.8	1.33E-06	0.0	0.9	0.9	175
Rank: 123 OS377		-0.2	-1.0	1.76E-03	0.0	0.9	1.2	182

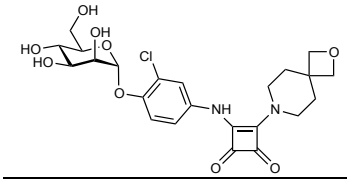
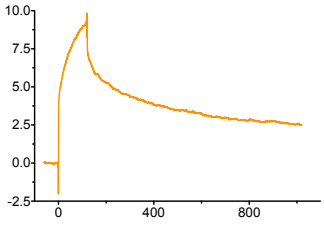
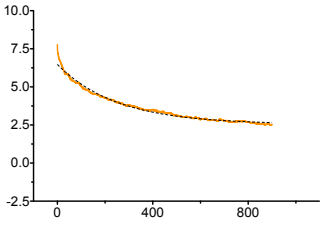
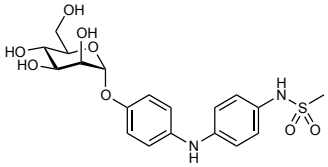
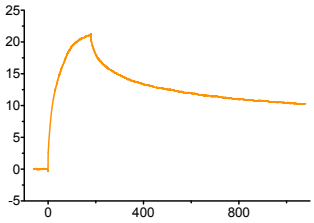
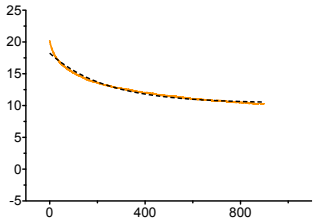
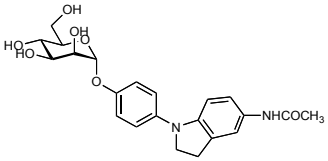
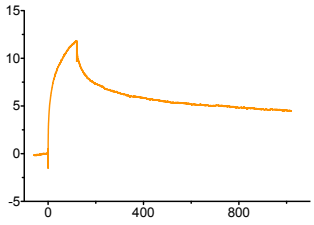
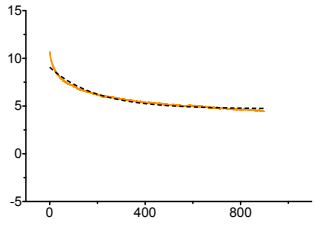
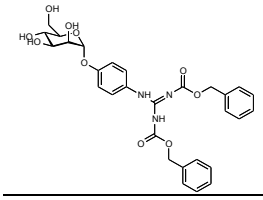
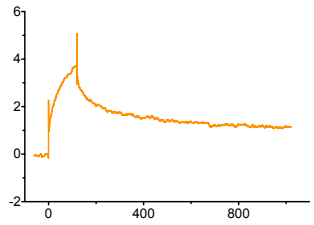
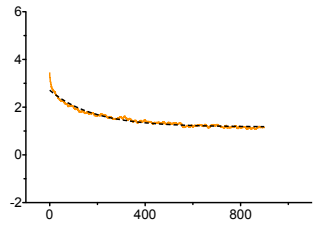
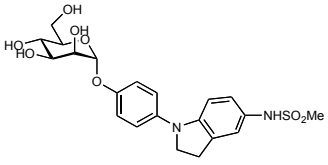
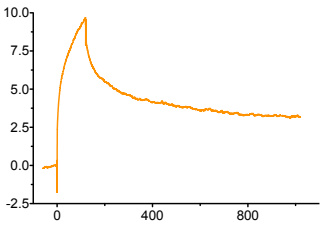
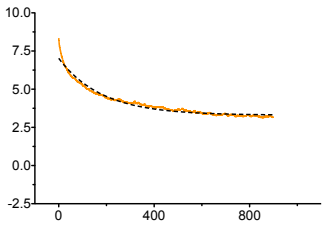
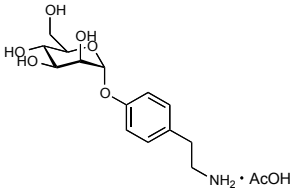
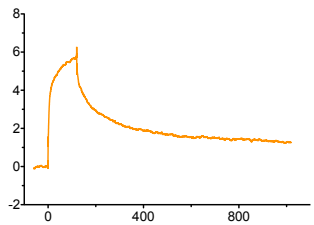
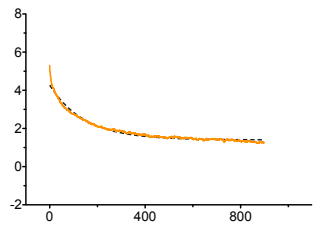
Compound		RU_B	RU_E	k_{off}	$t_{1/2}$ [min]	R^2	$logD_{7.4}$	FP_{Signal}
Rank: 123 OS380		-0.2	-0.7	1.72E-03	0.0	0.8	1.5	175
Rank: 123 OS381		-0.1	~ -288.7	9.86E-07	0.0	0.5	-1.0	183
Rank: 123 OS383		0.0	~ -17.86	1.65E-05	0.0	0.7	-0.2	
Rank: 123 OS405		-0.4	0.0	1.41E-01	0.0	0.2	-0.4	185
Rank: 123 OS412		-0.5	0.0	4.04E-01	0.0	0.2	0.2	186
Rank: 123 OS413		-0.6	-0.1	2.00E-01	0.0	0.3	0.7	188
Rank: 123 OS416		-0.4	0.2	3.91E-01	0.0	0.2	1.1	185
Rank: 123 JXH@0192		0.0	-0.1	1.12E-03	0.0	0.1	n.d.	186
Rank: 123 JXH@031		0.0	0.1	4.65E-03	0.0	0.2	n.d.	194
Rank: 123 JXH@039		0.0	-5051.0	9.41E-09	0.0		n.d.	186
Rank: 123 JXH@13		-1.0	-1.2	3.45E-03	0.0	0.6	n.d.	192
Rank: 123 JXH@137		0.0	-0.1	1.28E-02	0.0	0.1	n.d.	195

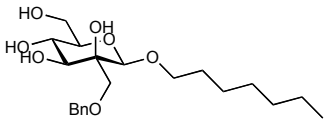
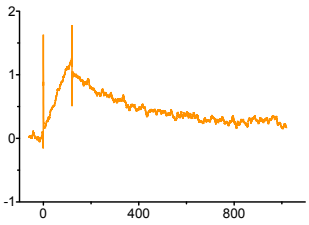
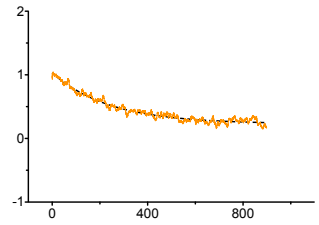
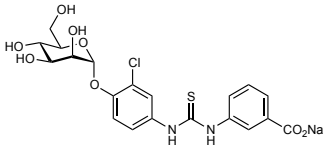
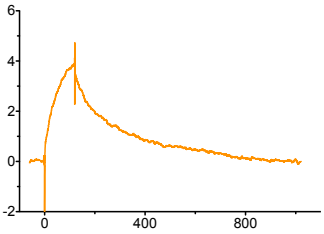
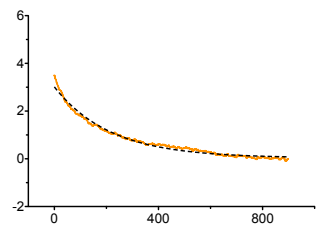
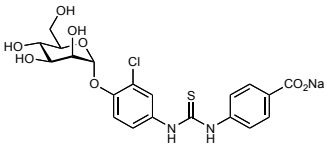
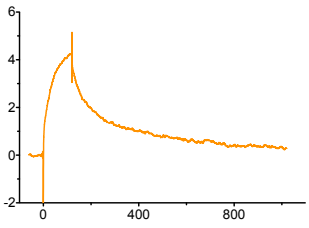
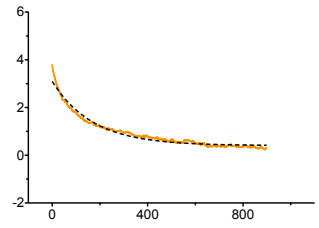
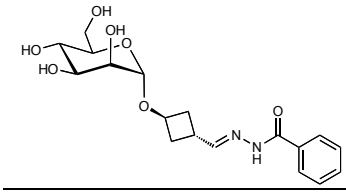
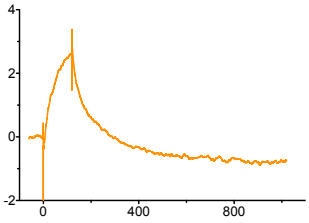
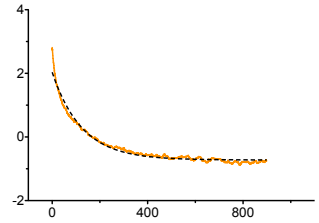
Compound		RU_B	RU_E	k_{off}	$t_{1/2}$ [min]	R^2	$logD_{7.4}$	FP_{Signal}
Rank: 123 JXH@138		0.1	0.1	5.89E-03	0.0	0.1	n.d.	196
Rank: 123 JXH@149		0.2	0.3	8.11E-04	0.0	0.1	n.d.	190
Rank: 123 JXH@158		0.1	~ -5.782	2.42E-05	0.0	0.4	n.d.	189
Rank: 123 JXH@163		0.1	1123.0	2.94E-08	0.0		n.d.	192
Rank: 123 JXH@175		0.1	0.0	2.67E-03	0.0	0.2	n.d.	192
Rank: 123 JXH@235		0.3	-2599.0	8.27E-09	0.0		n.d.	195
Rank: 123 JXH@209					0.0		n.d.	191

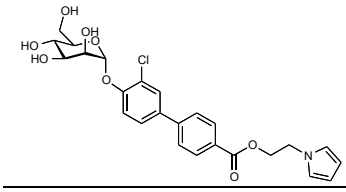
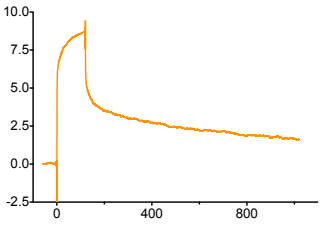
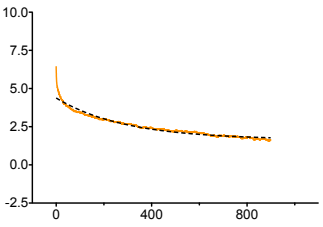
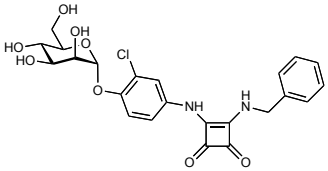
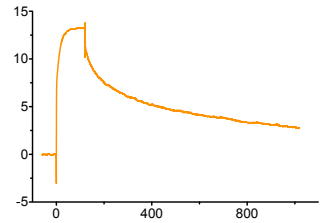
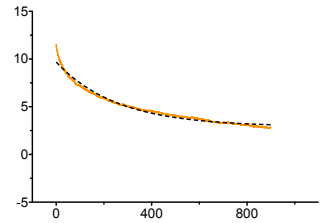
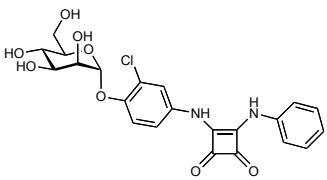
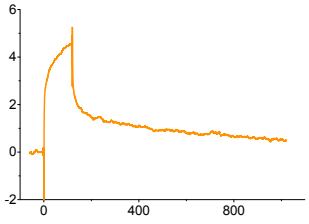
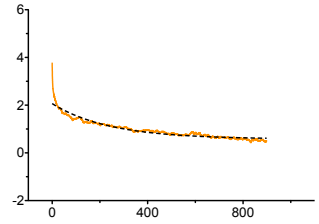
* published value (Jang *et al.*, 2012, *J. Med. Chem.*)
** published value (Klein et al., 2010, *J. Med. Chem*)

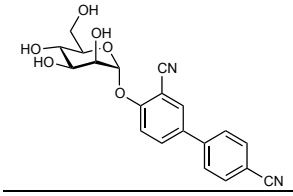
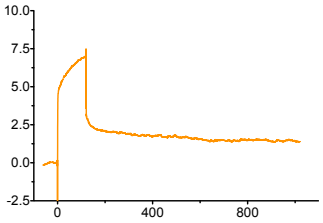
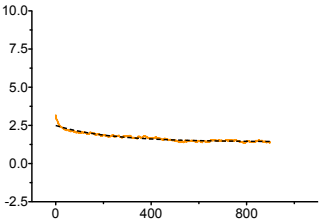
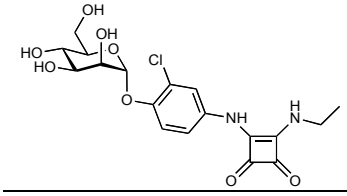
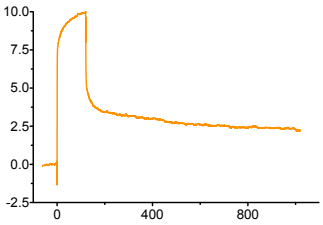
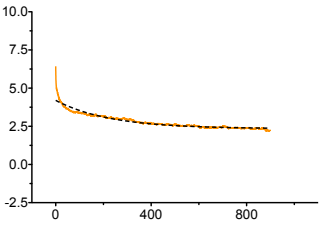
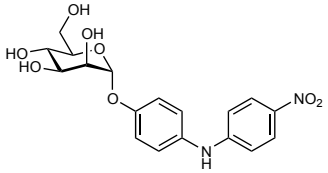
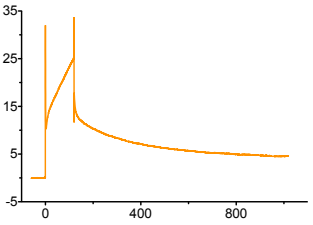
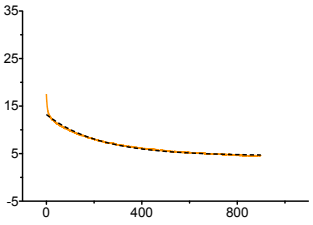
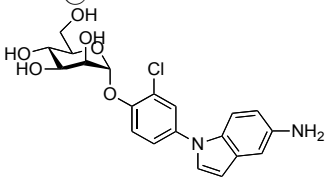
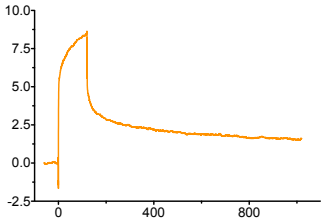
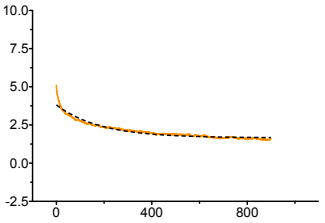
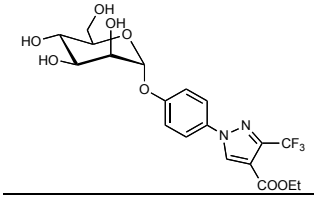
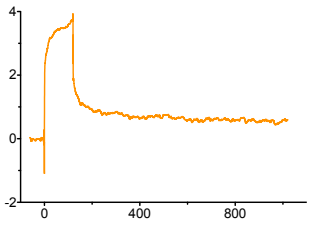
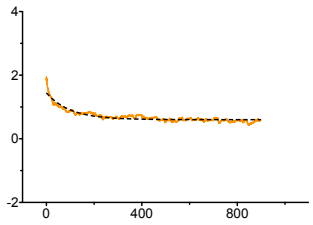
S4 Sensorgrams different shape types of dissociation

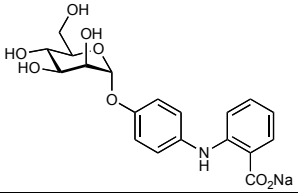
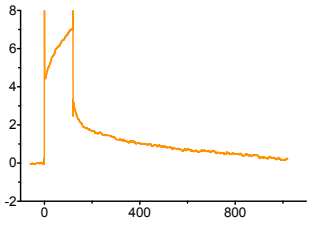
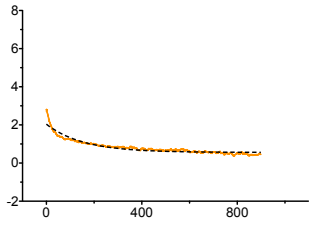
The following figure displays the sensorgrams and dissociation phases (fitted with the 1:1 Langmuir binding model) of hit compounds determined by the off-rate screen. Hit compounds are grouped according to their dissociation type (I-V).

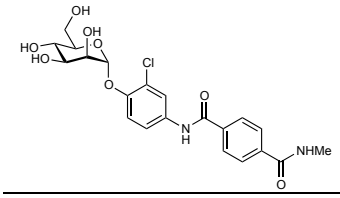
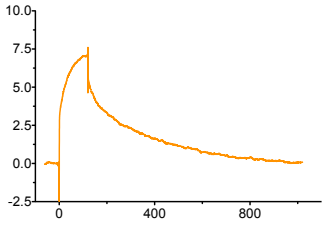
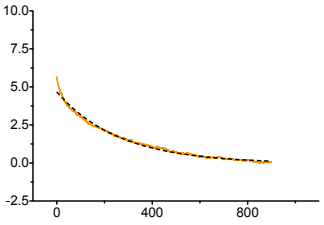
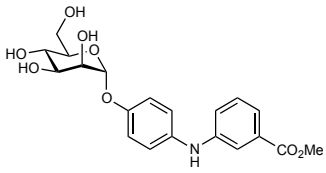
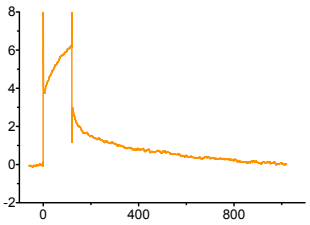
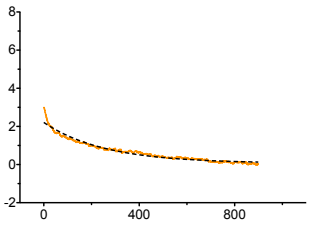
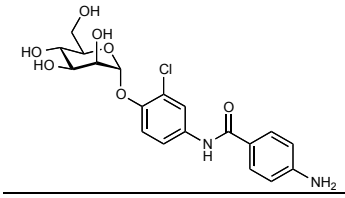
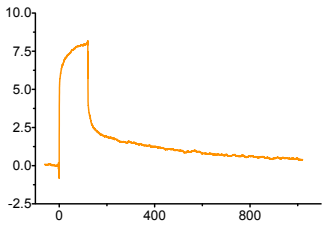
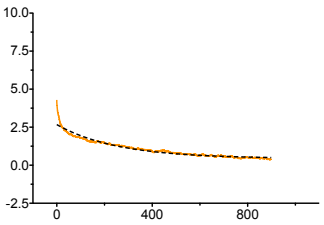
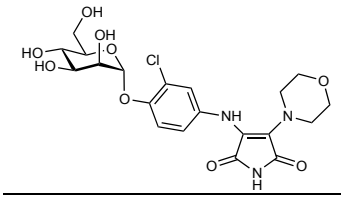
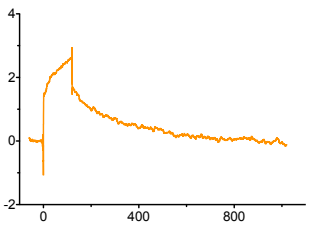
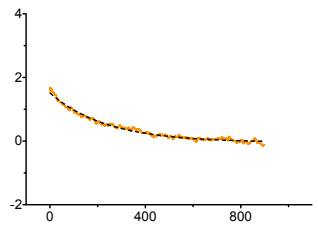
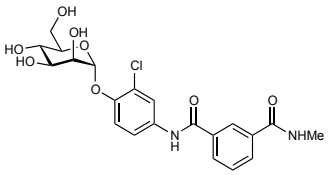
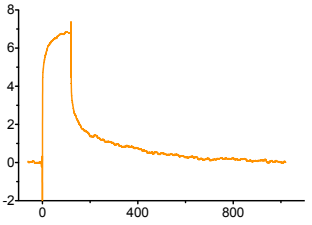
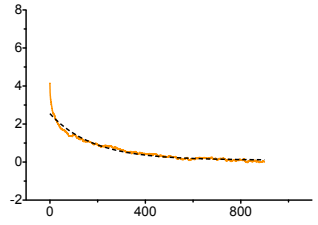
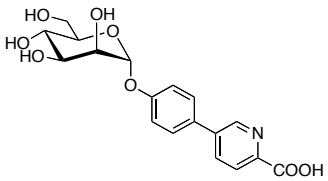
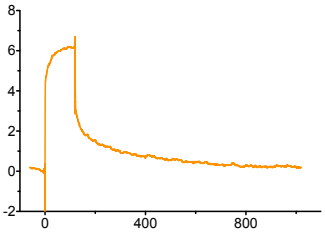
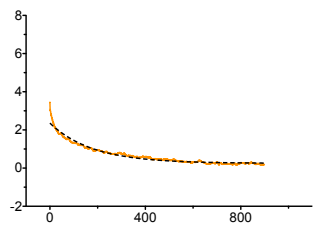
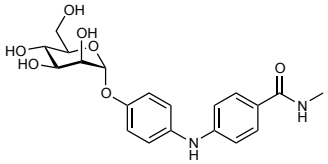
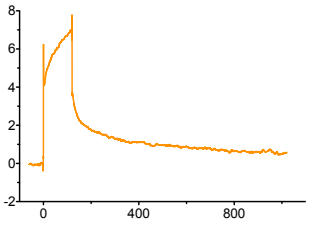
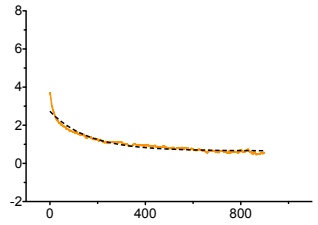
GROUP I	Sensorgram	Disossiation phase
Rank: 7 **WS176** 		
Rank: 11 **PF28** 		
Rank: 16 **JXH-III-057** 		
Rank: 18 **WS268** 		
Rank: 19 **JXH-III-0712** 		
Rank: 25 **WS302** 		

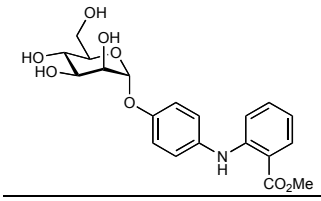
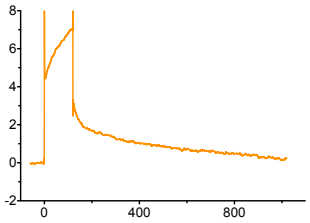
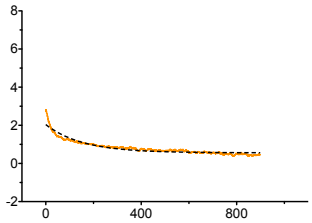
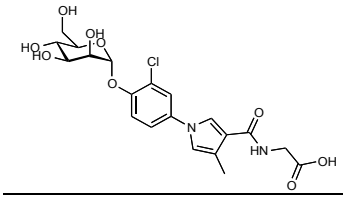
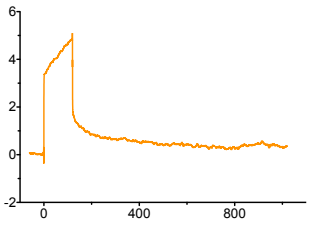
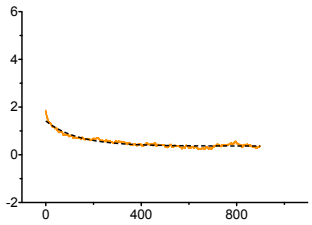
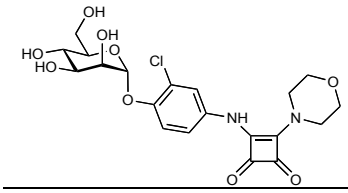
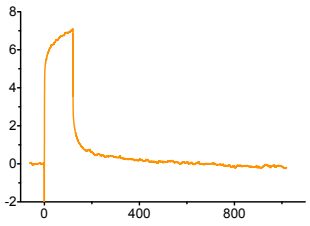
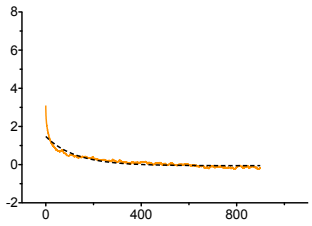
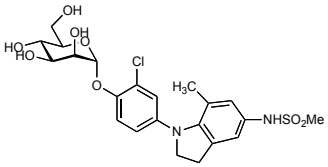
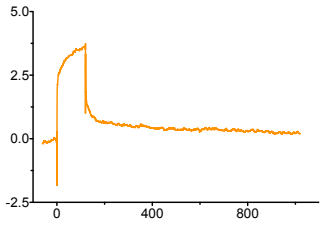
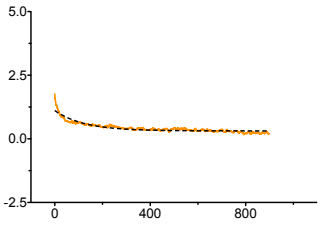
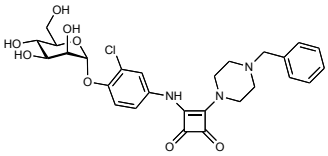
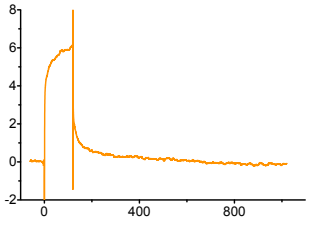
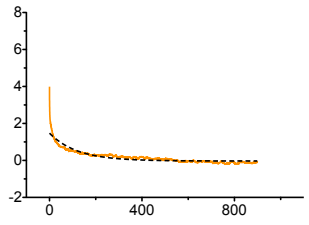
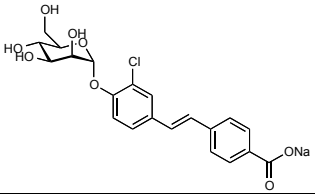
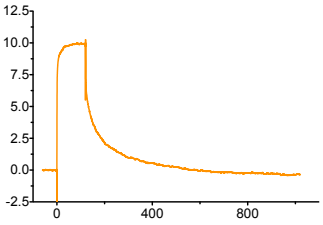
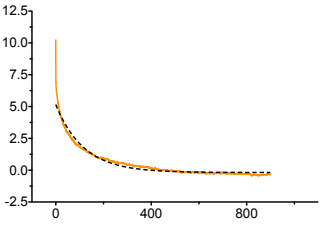
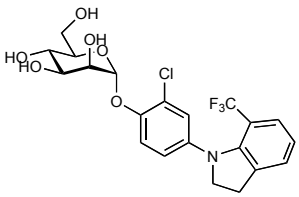
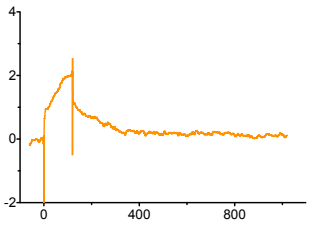
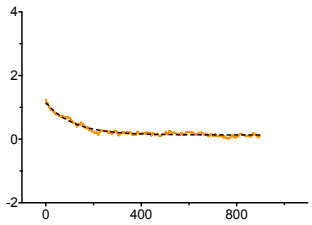
GROUP II	Sensorgram	Disossiation phase
Rank: 8 **OS349** 		
Rank: 15 **OS472.2** 		
Rank: 22 **OS470.2** 		
Rank: 28 **PLJ02174B** 		

GROUP III	Sensorgram	Disossiation phase
Rank: 1 **OS395** 		
Rank: 4 **OS556** 		
Rank: 8 **OS397** 		

GROUP III	Sensorgram	Disossiation phase
Rank: 12 **PF57** 		
Rank: 13 **OS420** 		
Rank: 14 **PF18** 		
Rank: 17 **JXH-@090** 		
Rank: 33 **PLJ2238A** 		

GROUP IV	Sensorgram	Disossiation phase
Rank: 2 **OS465.2** 		

GROUP IV	Sensorgram	Disossiation phase
Rank: 3 **OS540** 		
Rank: 5 **OS550** 		
Rank: 6 **OS494** 		
Rank: 10 **HS030** 		
Rank: 20 **OS542** 		
Rank: 21 **PLJ01068A** 		
Rank: 23 **OS562** 		

GROUP IV	Sensorgram	Disossiation phase
Rank: 24 **OS465.1** 		
Rank: 26 **WS320** 		
Rank: 27 **OS464** 		
Rank: 29 **JXH-IV-073** 		
Rank: 30 **OS396** 		
Rank: 31 **WS384** 		
Rank: 32 **JXH-IV-090** 		

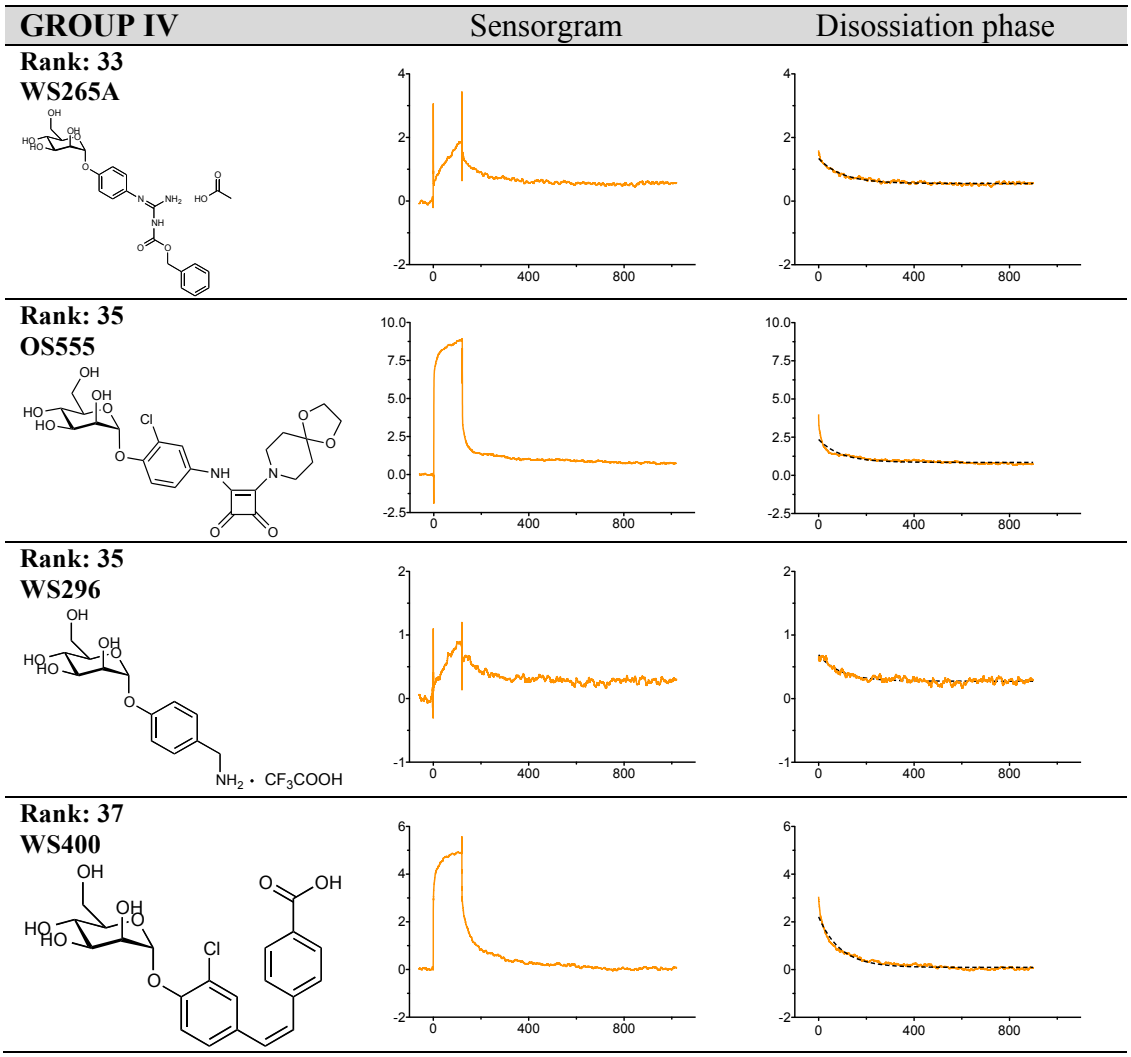


Figure S3. Sensorgrams and dissociation phased of hit compounds grouped by the type of dissociation (I-V).

References:

1. Kleeb, S., Pang, L., Mayer, K., Eris, D., Sigl, A., Preston, R.C., Zihlmann, P., Sharpe, T., Jakob, R.P., Abgottspon, D., Hutter, A.S., Scharenberg, M., Jiang, X., Navarra, G., Rabbani, S., Smiesko, M., Lüdin, N., Bezençon, J., Schwardt, O., Maier, T., and Ernst, B., FimH antagonists: Bioisosteres to improve the in vitro and in vivo PK/PD profile. *J. Med. Chem.*, **2015**, *58*(5), 2221-2239.
2. Mayer, K., Eris, D., Schwardt, O., Sager, C.P., Rabbani, S., Kleeb, S., and Ernst, B., Urinary tract infection: Which conformation of the bacterial lectin FimH is therapeutically relevant? *J. Med. Chem.*, **2017**, *60*(13), 5646-5662.

Journal of Biological Chemistry, 2018, *293*(5), 1835-1849
doi: 10.1074/jbc.M117.802942

Publication 3

Conformational Switch of the Bacterial Adhesin FimH in the Absence of the Regulatory Domain - Engineering a Minimalistic Allosteric System

Said Rabbani,[1)‡] Brigitte Fiege,[1)‡] Deniz Eris,[1)] Marleen Silbermann,[1)]
Roman Peter Jakob,[2)] Giulio Navarra,[1)] Timm Maier,[2)] and Beat Ernst[1)*]

[‡]These authors contributed equally to the project

[1)] Institute of Molecular Pharmacy, University of Basel,
Klingelbergstr. 50, 4056 Basel, Switzerland

[2)] Structural Biology, Biozentrum, University of Basel,
Klingelbergstr. 70, 4056 Basel, Switzerland

[*]Corresponding author
Tel.: 0041 61 267 15 51; Fax: 0041 61 207 15 52
E-mail: beat.ernst@unibas.ch

Contributions of Marleen Silbermann

- SPR experiments and data evaluation/interpretation
- Protein expression and purification of $FimH_{FL}$
- Manuscript preparation of the kinetic part
- Aiding in manuscript revision

Conformational switch of the bacterial adhesin FimH in the absence of the regulatory domain: Engineering a minimalistic allosteric system

Received for publication, June 19, 2017, and in revised form, November 23, 2017 Published, Papers in Press, November 27, 2017, DOI 10.1074/jbc.M117.802942

Said Rabbani[‡1], Brigitte Fiege[‡1,2], Deniz Eris[‡], Marleen Silbermann[‡], Roman Peter Jakob[§], Giulio Navarra[‡], Timm Maier[§], and Beat Ernst[‡3]

From the [‡]Department of Pharmaceutical Sciences, Pharmacenter of the University of Basel, Klingelbergstrasse 50 and the [§]Department Biozentrum, Focal Area Structural Biology, University of Basel, Klingelbergstrasse 70, 4056 Basel, Switzerland

Edited by Wolfgang Peti

For many biological processes such as ligand binding, enzymatic catalysis, or protein folding, allosteric regulation of protein conformation and dynamics is fundamentally important. One example is the bacterial adhesin FimH, where the C-terminal pilin domain exerts negative allosteric control over binding of the N-terminal lectin domain to mannosylated ligands on host cells. When the lectin and pilin domains are separated under shear stress, the FimH–ligand interaction switches in a so-called catch-bond mechanism from the low- to high-affinity state. So far, it has been assumed that the pilin domain is essential for the allosteric propagation within the lectin domain that would otherwise be conformationally rigid. To test this hypothesis, we generated mutants of the isolated FimH lectin domain and characterized their thermodynamic, kinetic, and structural properties using isothermal titration calorimetry, surface plasmon resonance, nuclear magnetic resonance, and X-ray techniques. Intriguingly, some of the mutants mimicked the conformational and kinetic behaviors of the full-length protein and, even in absence of the pilin domain, conducted the cross-talk between allosteric sites and the mannoside-binding pocket. Thus, these mutants represent a minimalistic allosteric system of FimH, useful for further mechanistic studies and antagonist design.

The key step in urinary tract infections (UTI)[4] is the adhesion of uropathogenic *Escherichia coli* (UPEC) to urothelial cells of the host (1, 2). This initial step enables the bacteria to invade and colonize host cells but also to withstand clearance by the bulk flow of urine. Adherence to the urothelial surface is mediated by the adhesin FimH located at the tip of bacterial type 1 pili (3, 4). Full-length FimH ($FimH_{FL}$) is composed of two domains: the N-terminal lectin domain ($FimH_{LD}$) which is connected to the C-terminal pilin domain ($FimH_{PD}$) by a short linker (5). $FimH_{LD}$ contains the carbohydrate recognition domain, which is responsible for binding to the highly mannosylated uroplakin 1a (UP_{1a}) on the urothelial cell surface (6). $FimH_{PD}$ is anchored to the core of the pilus via a donor strand of the subsequent FimG subunit, a process termed donor strand complementation (7, 8).

In the bladder, under static conditions, FimH exhibits only weak interactions with the urothelial surface, which is highly beneficial for bacterial motility and host cell invasion (9). However, when shear forces originating from urine flow occur, a significant enhancement of the strength of the interaction (5, 10) due to a catch bond mechanism was observed (11, 12). Catch bonds play pivotal roles in the regulation and fine-tuning of cellular adhesion events, *e.g.* upon leukocyte recruitment by selectins (13) and integrins (14), by cadherins controlling tissue integrity (15, 16), in the epithelial adhesion of cancer cells (17), or in T cell receptor interaction with peptide-bound major histocompatibility complexes on antigen-presenting cells (18, 19).

Conformation and ligand-binding properties of $FimH_{LD}$ are under the allosteric control of $FimH_{PD}$ (5, 20, 21). Recent X-ray data of a recombinant $FimH_{FL}$ has provided structural evidence for the different conformational states (22, 23). Under static conditions, the interaction of the two domains of $FimH_{FL}$ stabilizes the lectin domain in the low-affinity state, which is characterized by a shallow binding pocket (Fig. 1). Binding to a mannoside ligand induces a conformational change leading to the medium-affinity state, in particular by the displacement of the so-called clamp loop toward the binding pocket. This medium-affinity state, where lectin and pilin domain are still in close contact, is characterized by micromolar affinities and fast off-rates (5, 23). However, upon shear stress, $FimH_{LD}$ and $FimH_{PD}$ separate, inducing the high-affinity state with an up to 300-fold enhanced affinity (20). Using donor–strand complemented $FimH_{FL}$ and combining X-ray and small angle x-ray scattering analyses, it was suggested that FimH is present in a two-state conformational ensemble of low- and high-affinity states in solution (22). This equilibrium can be influenced by sequence

The authors declare that they have no conflicts of interest with the contents of this article.

This article contains Figs. S1–S11, Tables S1–S2, supporting Extended Experimental procedures, and supporting Refs. 1–5.

The atomic coordinates and structure factors (code 5MCA) have been deposited in the Protein Data Bank (http://wwpdb.org/).

[1] Both authors contributed equally to this work.

[2] Recipient of a scholarship from the German Academic Exchange Service (DAAD).

[3] To whom correspondence should be addressed. E-mail: beat.ernst@unibas.ch.

[4] The abbreviations used are: UTI, urinary tract infection; CSP, chemical shift perturbation; LD, lectin domain; PD, pilin domain; FL, full length; FP, fluorescence polarization; ITC, isothermal titration calorimetry; SPR, surface plasmon resonance; UPEC, uropathogenic *Escherichia coli*; BisTris, 2-[bis(2-hydroxyethyl)amino]-2-(hydroxymethyl)propane-1,3-diol; PDB, Protein Data Bank; r.m.s.d., root mean square deviation; MCK, multicycle kinetic; MD, molecular dynamics; Th, thrombin.

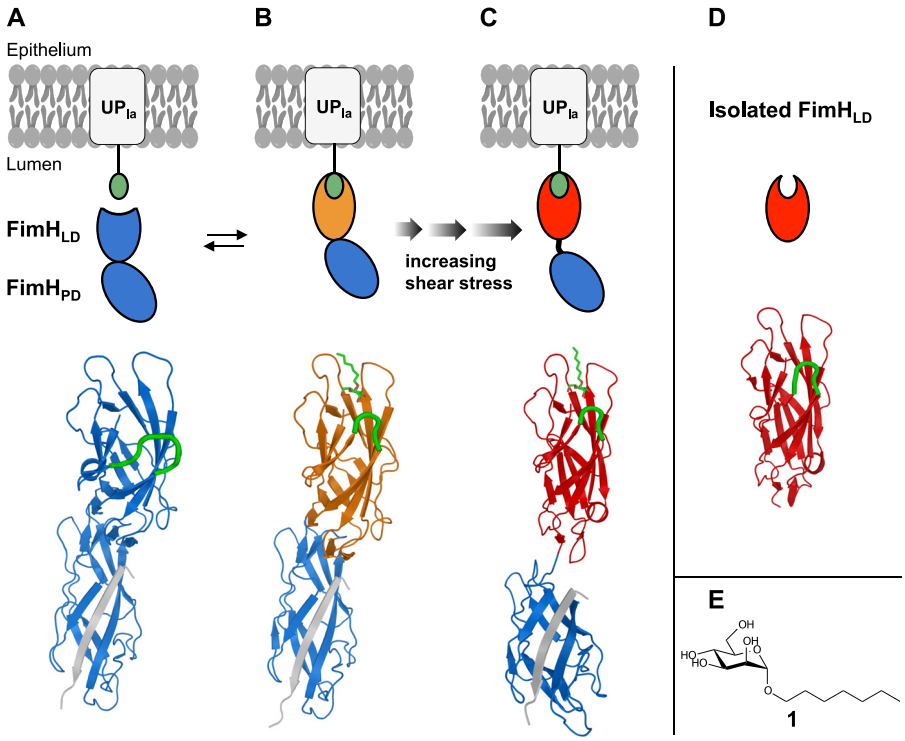


Figure 1. Schematic representation of the conformational states of $FimH_{FL}$ upon binding to mannosylated uroplakin 1a (UP1a) on the urothelium (*top*), and the corresponding X-ray structures in the absence or presence of the model ligand *n*-heptyl-α-D-mannopyranoside (*bottom*). *A,* in the unbound form, $FimH_{LD}$ and $FimH_{PL}$ interact and form the low-affinity state featuring a shallow binding pocket (PDB code 4XOD, donor strand of FimG in *gray*). *B,* upon binding to UP1a in the absence of shear stress, $FimH_{FL}$ adopts the medium-affinity state, where $FimH_{LD}$ and $FimH_{PD}$ are still interacting but the clamp loop (highlighted in *green*) is moved toward the mannoside ligand (in *green*) creating a deep, well-defined binding pocket (PDB code 4XOE). *C,* flow-induced shear stress leads to a separation of $FimH_{LD}$ and $FimH_{PD}$, causing the switch to the high-affinity state (PDB code 4XOB). *D,* isolated lectin domain is constitutively present in the high-affinity state, both in the absence or presence of mannoside ligands (PDB code 4AUU, with ethanediol in the binding pocket, nearly identical to the NMR solution structure, PDB code 3ZPD). *E,* structure of *n*-heptyl-α-D-mannopyranoside (1).

variations, thereby modulating the apparent binding affinity as well as the binding mechanism (induced fit or conformational selection), and thus influencing the infective potential of UPEC.

Because of the absence of negative allosteric regulation exerted by the pilin domain, isolated $FimH_{LD}$ is locked in the high-affinity state (24, 25). Mutations in the interdomain region altering the interactions to the pilin domain clearly demonstrated its contribution to the conformational change within the lectin domain (20). Interestingly, the medium-affinity state (domains interacting) and the high-affinity state (domains separated) exhibit nearly identical mannose-binding pockets, yet possess a strikingly 10^5-fold difference in the off-rate for mannosides (23). This structural information in combination with MD simulations and kinetics experiments (23, 26) led to the concept of dynamic allosteric regulation, *i.e.* the allosteric signal of the pilin domain is transmitted to the proximal regions of the lectin domain, whereas the binding pocket is regulated purely by protein dynamics (27).

The isolated, recombinant $FimH_{LD}$, which is locked in the high-affinity state, exhibits high temporal and thermal stability (24, 28, 29), and consequently, it was mostly used to generate affinity data from antagonist screening and structural data from X-ray crystallography (30–35). However, recent data on the conformational and dynamic regulation of $FimH_{FL}$ upon mannoside binding raised doubt on the applicability of the conformationally locked $FimH_{LD}$ for mimicking urinary tract infections *in vivo* (36, 37).

In a comparison of the low- and high-affinity states of $FimH_{LD}$, Le Trong *et al.* (5) identified four regions undergoing conformational changes upon ligand binding, namely the pocket zipper/clamp loop (Phe-1–Ile-11/Gly-8–Gly-16), the β-bulge (Gln-59–Ser-63), the α-switch (Tyr-64–Phe-71), and the interdomain swing, insertion, and linker loops (Ala-25–Leu-34, Pro-111–Ala-119, Asn-152–Thr-158) (Fig. 2). Later, Rodriguez *et al.* (38) used RosettaDesign and MODIPDesign (computer-based tools for protein structural analysis) to predict specific point mutations within the aforementioned regions, causing the largest difference in energy between the low- and high-affinity conformations and hence should stabilize one of the affinity states.

The mutation A10P was predicted to favor the low-affinity conformation through perturbation of the backbone hydrogen bond of residue Ile-11 in the pocket zipper. In the β-bulge region, the point mutation R60P was predicted to prevent the transition of the portion of Gln-59–Ser-63 from a small loop to a β-strand and thus destabilize the high-affinity conformation. In the α-switch (Tyr-64–Phe-71), the mutation V67K was expected to prevent the conversion of this region to a α-helix and thus lock FimH in low-affinity conformation. The point mutation A119L and the double mutation V27C/L34C at the proximal end of the lectin domain were expected to preserve a conformation of the interdomain loops that allows interaction with the pilin domain and hence stabilizes the low-affinity conformation (5).

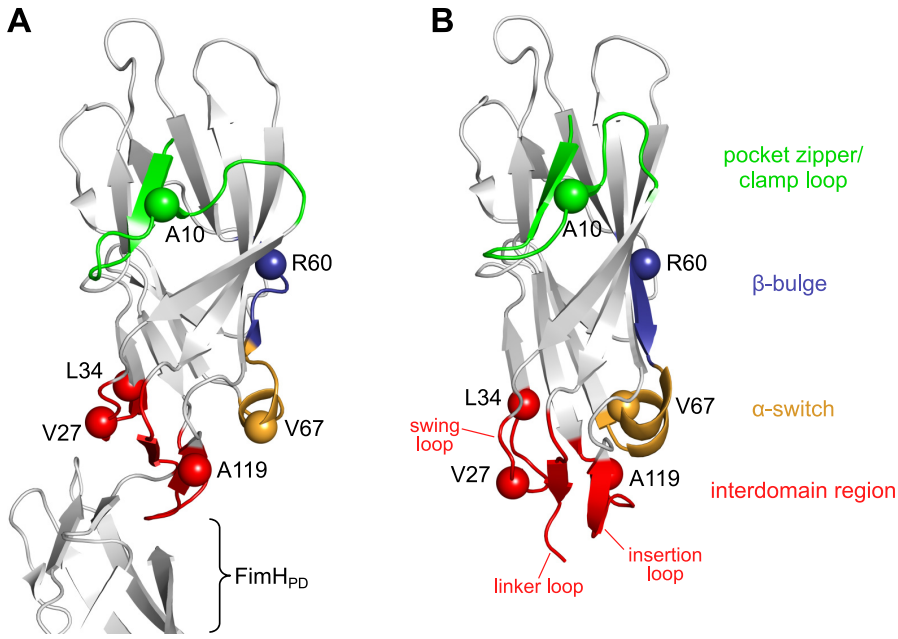


Figure 2. *A,* proposed allosteric regions of $FimH_{LD}$ mapped onto the X-ray structure of $FimH_{FL}$ in the low-affinity conformation within the assembled fimbrial tip (PDB code 3JWN) (5). *B,* same regions in the X-ray structure of the isolated $FimH_{LD}$ (PDB code 4AUU) (41). Point mutations are shown as *spheres*. Data are modified from Ref. 38.

The five FimH variants selected upon this analysis (A10P, R60P, V67K, A119L, V27C/L34C) were expressed in *E. coli* in the context of the entire pilus, and their affinity state was evaluated based on binding to mono- and trimannoside ligands as well as their recognition by the conformation-specific monoclonal antibody mAb21 (38). This antibody recognizes an epitope located in the interdomain region of the isolated WT $FimH_{LD}$ (21, 39). Therefore, binding of mAb21 indicates that FimH is in the high-affinity state, because only then is the interdomain accessible. Furthermore, bacterial adhesion is enhanced in the presence of mAb21, which stabilizes the high-affinity state (21). In contrast, a monoclonal blocking antibody directed against the $FimH_{LD}$ mutant V27C/L34C, presumably locked in the low-affinity state, protected mice from bacterial infection, thus demonstrating the relevance of considering all conformational states for the design of therapeutic FimH antagonists (40). Despite the intensive studies of the FimH variants mentioned above, structural, thermodynamic, and kinetic data regarding the low-affinity conformation were not reported so far.

To further elucidate the degree of conformational heterogeneity and the allosteric effects of mutants in the context of the whole pilus (38), we recombinantly expressed five selected variants (A10P, R60P, V67K, V27C/L43C, and A119L) of isolated $FimH_{LD}$ predicted to be stabilized in the low-affinity state, and we analyzed their thermodynamic and kinetic binding properties, as well as their structural characteristics by NMR and X-ray crystallography with the aim to reveal the allosteric crosstalk between the regulatory regions within $FimH_{LD}$. Furthermore, $FimH_{LD}$ mutants could also serve as easy-accessible mimics of $FimH_{FL}$ for mechanistic studies as well as for the development of FimH antagonists applicable for the anti-adhesive therapy of UTI.

Results

Five $FimH_{LD}$ variants were generated with single- or double-point mutations within the four allosteric regions assumed to stabilize or even lock FimH in the low-affinity state (5, 38, 40). The A10P mutation is located in the pocket zipper/clamp loop; R60P and V67K are part of the β-bulge and α-switch regions, respectively; and A119L and V27C/L34C are located within the interdomain loops (Fig. 2). The purity and the monomeric state of the WT and all variants were confirmed by native PAGE analysis (data not shown).

Folding and stability of $FimH_{LD}$ mutants

The impact of these mutations on protein folding and stability was evaluated by thermal shift-based differential-scanning fluorimetry (42). In this assay, the fluorescence intensity of SYPRO orange dye was recorded during heat denaturation of the protein (42). The melting temperatures (T_m), which indicate the thermal unfolding event, were in the same region for WT $FimH_{LD}$ and the variants A10P, A119L, and R60P (Fig. S1 and Table S1) and were reduced by less than 8 °C for the mutants V67K and V27C/L34C. Similar as for WT $FimH_{LD}$, the addition of *n*-heptyl α-D-mannopyranoside (**1**) considerably increased the thermal stability of all mutants, indicating binding of antagonist **1**. However, for two mutants, R60P and V27C/L34C, the extent of the thermal shift was significantly reduced, suggesting a lower degree of stabilization, *i.e.* less efficient binding than for WT $FimH_{LD}$ (Table S1).

Correlation of structural and affinity changes of $FimH_{LD}$ mutants

To examine the secondary and tertiary structures of WT $FimH_{LD}$ and the five variants in the absence and presence of mannoside **1**, circular dichroism (CD) spectra were recorded

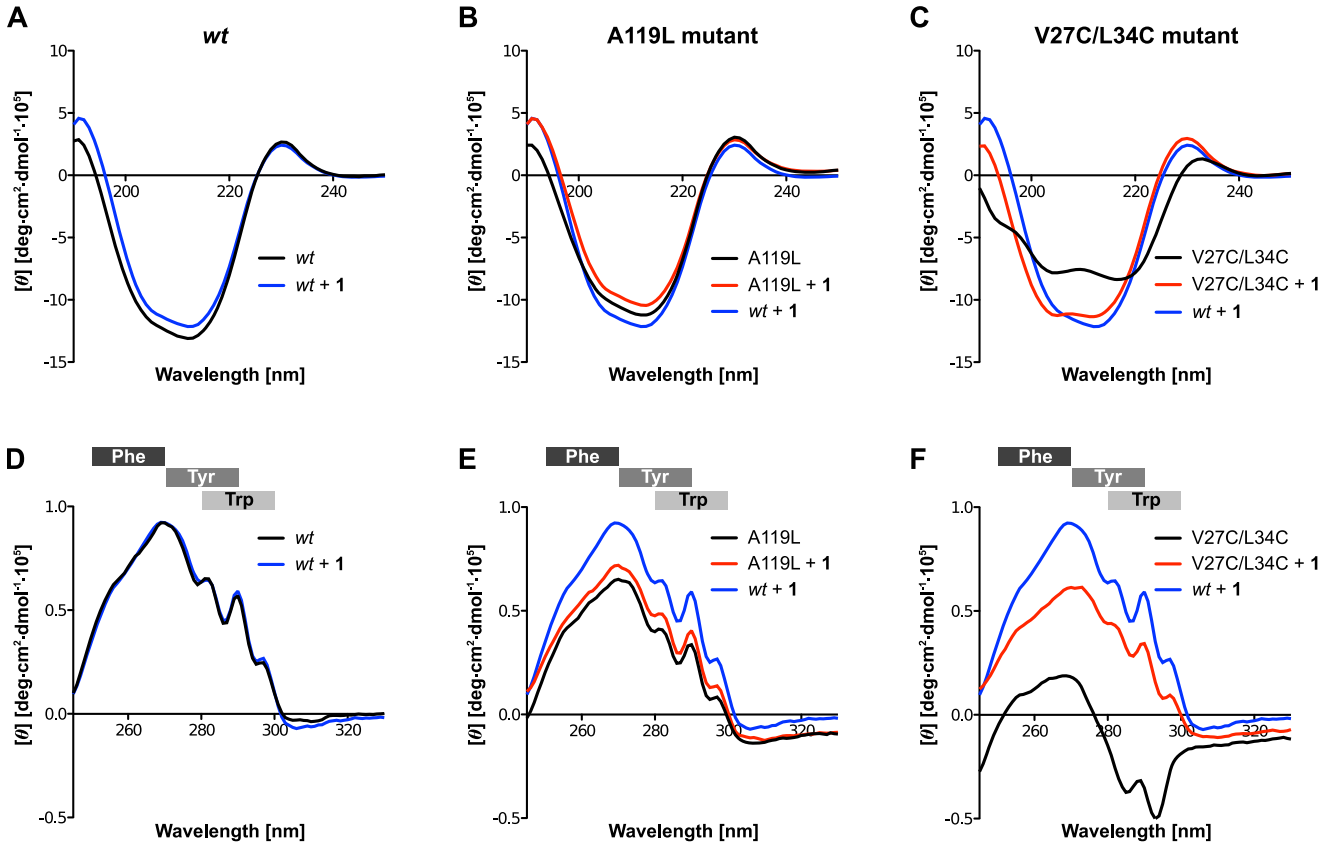


Figure 3. Informative regions of CD spectra in the far-UV (*A–C*) and near-UV region (*D–F*) of WT $FimH_{LD}$ (*A* and *D*) and of the mutants A119L (*B* and *E*) and V27C/L34C (*C* and *F*) in the absence and presence of mannoside 1. For comparison, the CD spectrum of WT $FimH_{LD}$ with **1** (*blue*) is shown in all graphs. *Bars* in *gray* indicate the typical near-UV absorbance range for aromatic residues in $FimH_{LD}$: Phe (250–270 nm, 6 residues), Tyr (270–290 nm, 10 residues), and Trp (280–300 nm, 2 residues).

(Fig. 3 and Figs. S2 and S3). The far-UV (190–250 nm) CD spectrum of WT $FimH_{LD}$ is characteristic for a protein with a high β-sheet content (Fig. 3*A*) (43). Furthermore, upon addition of antagonist **1**, the CD spectrum remained unchanged, indicating a stable fold. Similar conformations are suggested for the mutants A119L and A10P because their CD spectra in the absence and presence of antagonist **1** were almost identical to the WT (Fig. 3*B* and Fig. S2*A*). In contrast, the CD spectra of the mutants V27C/L34C, R60P, and V67K in the absence of ligand were clearly altered with an overall lower molar ellipticity, indicating significant changes in the global protein conformation and a reduction of the β-sheet proportion (Fig. 3*C* and Fig. S2). Surprisingly, the addition of mannoside **1** to these three mutants re-established CD spectra nearly identical to WT $FimH_{LD}$.

Next, the geometry of aromatic side chains was analyzed with near-UV CD (260–320 nm) (Fig. 3, *D–F*, and Fig. S3). Considering the abundance of phenylalanine, tyrosine, and tryptophan residues, many of them in close vicinity of the mannose-binding pocket (Fig. S3, *D–F*), even small changes in the local protein conformation should be detectable (44). Because near-UV CD spectra of WT $FimH_{LD}$ in the absence and presence of mannoside **1** were again superimposable, it can be assumed that the protein conformation in the neighborhood of the aromatic residues is highly conserved (Fig. 3*D*). For the mutants A10P and A119L, the general shape and the curve maxima of the spectra were similar to WT $FimH_{LD}$, and the spectra of the mutants V27C/L34C (Fig. 3*F*), R60P (Fig. S3*B*), and V67K (Fig. S3*C*) were significantly altered with shifted maxima and partly inverted signs of the molar ellipticity. Again, addition of antagonist **1** to these three mutants restored the near-UV CD spectra compared with WT $FimH_{LD}$ to a large extent. These results clearly indicate that mannose binding induces a significant conformational change toward the high-affinity conformation of the so far uncharacterized apo-forms of R60P, V67K, and V27C/L34C.

The dissociation constant K_D of mannoside **1** binding to WT $FimH_{LD}$ and the five mutants were determined with a fluorescence polarization (FP) assay (32). While WT $FimH_{LD}$ and the mutants A119L and A10P exhibited affinities in the low nanomolar range, the mutants R60P and V27C/L34C show reduced affinity, namely by factors of 17 and 35, respectively (Table 1). Finally, despite its altered CD spectra, the mutant V67K showed only a 2-fold reduction in binding affinity. Thus, at this point it cannot be speculated that the conformational changes observed in the CD spectra result in an entropy penalty.

Distinct thermodynamic profiles of conformationally different mutants

To assess the impact of the observed conformational changes upon ligand binding on the thermodynamic fingerprint of the mutants, isothermal titration calorimetry (ITC) experiments were performed with mannoside **1** (Table 2) leading to K_D values in the same range (within a factor of 2) as those obtained

with the FP assay (Table 1). All mutants showed an enthalpy-driven binding profile. Compared with WT $FimH_{LD}$, they exhibit a more unfavorable entropy term. In respect to enthalpy, the mutants A119L and V67K benefit from an improvement of up to 2.8 kJ/mol, whereas R60P and V27C/L34C display penalties in a comparable magnitude (Fig. S4).

We hypothesized that the conformational changes of the mutants R60P and V27C/L34C induced by ligand binding are responsible for unfavorable conformational entropy terms. Therefore, the entropy term obtained for R60P was dissected into solvation, conformational and mixing entropy (Equation 1) (Table 3).

$$\Delta S^0_{obs} = \Delta S^0_{solv} + \Delta S^0_{conf} + \Delta S^0_{mix} \quad \text{(Eq. 1)}$$

The change in mixing entropy (ΔS_{mix}) reflecting the loss in translational and rotational degrees of freedom of both interaction partners upon binding is equal for WT $FimH_{LD}$ and the R60P mutant (Equation 2, where R is the universal gas constant and 55.6 is the molarity of water in molar).

$$\Delta S^0_{mix} = R \ln\left(\frac{1}{55.6}\right) \quad \text{(Eq. 2)}$$

The change in solvation entropy (ΔS_{solv}) was calculated based on the change in heat capacity (ΔC_p, Equation 3),

$$\Delta C_p = \left(\frac{\partial \Delta H^0_{obs}}{\partial T}\right) \quad \text{(Eq. 3)}$$

which was determined by measuring the change of enthalpy (ΔH) as a function of temperature (Fig. 4).

$$\Delta S^0_{solv,\ 298.15\ K} = \Delta C_p \ln\left(\frac{298.15\ K}{385\ K}\right) \quad \text{(Eq. 4)}$$

By inserting the experimentally determined value for ΔC_p into Equation 4 at 385 K as the temperature of zero hydration (T_R), the change in solvation entropy (ΔS_{solv}) at room temperature was obtained (47–49). Interestingly, the change of solvation entropy for the interaction of R60P with mannoside **1** was less favorable compared with WT $FimH_{LD}$ ($-T\Delta\Delta S_{solv}$ = 15.9 kJ/mol) (Table 3). Considering that their ligand-bound structures are highly similar, a possible explanation may be a lower degree of ordering of interfacial water molecules in the apo-R60P, which presumably exhibits a broad and flexible binding pocket. In contrast, the well-defined and narrow binding pocket of WT $FimH_{LD}$ can establish a highly structured hydrogen bond network of water molecules supporting a more beneficial solvation entropy upon ligand binding. Surprisingly, the

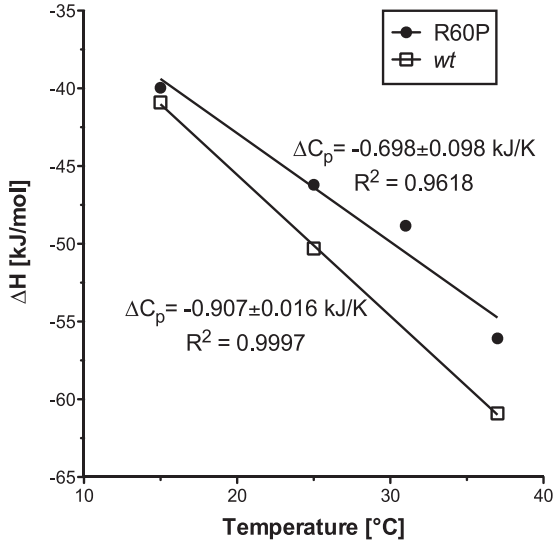


Figure 4. Determination of the heat capacity (ΔC_p) of WT $FimH_{LD}$ and mutant R60P binding to *n*-heptyl α-D-mannopyranoside (1) by ITC measurements at three and four different temperatures, respectively.

Table 1
Affinity of mannoside 1 for WT $FimH_{LD}$ and five mutants, as well as for $FimH_{FL}$ obtained from the FP assay

	K_D (nM)[a]	Relative K_D
Variants of $FimH_{LD}$		
WT	19 ± 2	1
A119L	5.2 ± 0.5	0.27
A10P	16.5 ± 1.5	0.87
V67K	37 ± 3	1.9
R60P	322 ± 52	17
V27C/L34C	661 ± 56	35
Full-length FimH		
WT $FimH_{FL}$[b]	2300 ± 100	121

[a] The assay was performed twice in duplicate measurements.
[b] Full-length WT FimH stabilized by FimG donor strand (23).

Table 2
Thermodynamic profiles of WT $FimH_{LD}$ and mutants thereof from ITC measurements with mannoside 1 at 298 K
The ITC assay was performed twice in duplicate measurements.

Variant	K_D	ΔG^0	ΔH^0	$-T\Delta S^0$	*c*-value	*N*
	nM	*kJ/mol*	*kJ/mol*	*kJ/mol*		
WT	28.9 ± 4.6	−43.0 ± 0.4	−50.3 ± 0.35	7.3 ± 0.5	320	1.00
A119L	24.2 ± 10.53	−43.5 ± 1.0	−52.4 ± 1.5	8.9 ± 1.8	414	1.01
V67K	76.8 ± 16.82	−40.6 ± 0.5	−53.1 ± 0.98	12.5 ± 1.1	131	1.03
R60P	583 ± 122.8	−35.6 ± 0.5	−46.2 ± 2.05	10.6 ± 2.1	17.1	1.05
V27C/L34C	1117 ± 271.03	−34.0 ± 0.6	−45.3 ± 2.54	11.3 ± 2.6	13.4	0.93

Table 3
Entropy dissection of the binding of mannoside 1 to WT $FimH_{LD}$ and the mutant R60P

Variant	ΔC_p	$-T\Delta S_{solv}$	$-T\Delta S_{conf}$	$-T\Delta S_{mix}$[a]	$-T\Delta S_{obs}$
	kJ/K	*kJ/mol*	*kJ/mol*	*kJ/mol*	*kJ/mol*
WT $FimH_{LD}$	−0.907 ± 0.016	−69.1 ± 1.22	66.4 ± 1.24	10.0	7.27 ± 0.245
R60P	−0.698 ± 0.098	−53.2 ± 7.47	53.8 ± 7.51	10.0	10.6 ± 0.724

[a] ΔS_{mix} quantifies the loss in translational and rotational degrees of freedom upon complex formation: calculated as $\Delta S_{mix} = R \cdot \ln(1/55.6)$, with R being the universal gas constant and 55.6 being the molarity of water in molar (47, 49). The assay was performed twice in duplicate measurements.

Table 4

Kinetic parameters of WT $FimH_{LD}$ and $FimH_{LD}$ mutants and $FimH_{FL}$ with mannoside 5 determined by SPR

The table shows the mean value of 2–4 independent measurements on a CM4 sensor chip with immobilized compound **5**.

FimH variant	$K_{D,\,kinetic}$	$k_{on,\,1} \times 10^3$	$k_{off,\,1} \times 10^{-4}$	$t_{½,\,1}$ [a]	$k_{on,\,2} \times 10^{-3}$	$k_{off,\,2} \times 10^{-4}$	$t_{½,\,2}$ [a]
	nM	$M^{-1}s^{-1}$	s^{-1}	*min*	$M^{-1}s^{-1}$	s^{-1}	*min*
WT $FimH_{LD}$	28.8 ± 3.8[b]	4.8 ± 0.4	1.4 ± 0.1	83.6 ± 8.6	–	–	–
$FimH_{LD}$ A10P	41.6 ± 1.7[b]	4.6 ± 0.1	1.9 ± 0.05	60.4 ± 1.4	–	–	–
$FimH_{LD}$ A119L	20.4 ± 0.5[b]	10 ± 0.2	2.1 ± 0.03	54.8 ± 0.7	–	–	–
$FimH_{LD}$ V67K	179 ± 7[b]	5.2 ± 0.2	9.3 ± 0.2	12.5 ± 0.2	–	–	–
$FimH_{LD}$ R60P	1656 ± 203[b]	2.4 ± 0.3	41 ± 1.7	2.9 ± 0.1	–	–	–
$FimH_{LD}$ V27C/L34C	4361 ± 471[b]	92 ± 9	3998 ± 186	1.7 ± 0.1	–	–	–
WT $FimH_{FL}$[c]	6373 ± 198[d]	33 ± 0.3	5116 ± 134	1.4 ± 0.04	1.2 ± 0.01	8.7 ± 0.1	13 ± 0.1

[a] Complex half-lives ($t_{½}$) were calculated as follows: $t_{½} = \ln 2/k_{off}$.
[b] Kinetic K_D value was determined by the following equation: $K_D = k_{off,\,1}/k_{on,\,1}$.
[c] Full-length WT FimH was stabilized by the FimG donor strand (23).
[d] Kinetic K_D value was determined according to the following equation: $K_D = k_{off,\,1}/k_{on,\,1} \times (k_{off,\,2}/(k_{off,\,2} + k_{on,\,2}))$.

conformational entropy $T\Delta S_{conf}$ of R60P was less unfavorable compared with WT $FimH_{LD}$ ($-T\Delta\Delta S_{conf} = -12.6$ kJ/mol). This counter-intuitive observation remains to be investigated and may require NMR dynamics studies that report on the different time scales on which ligand binding and protein conformational changes can occur.

Similar binding kinetics for V27C/L34C mutant and WT $FimH_{FL}$

To gain a detailed insight into the binding kinetics of the interaction of FimH with mannoside ligands, surface plasmon resonance (SPR) measurements were performed (Table 4). The carboxyl-functionalized mannoside **5** (see supporting information for structure and synthesis) was immobilized via an ethylenediamine linker on a CM4 chip. WT $FimH_{LD}$, mutants, and $FimH_{FL}$ were injected in the mobile phase, and their association and dissociation were monitored (Fig. 5 and Fig. S5). WT $FimH_{LD}$ and mutants could be fitted with a one-state binding model (Equation 5).

$$A + B \underset{k_{off,1}}{\overset{k_{on,1}}{\rightleftharpoons}} AB \qquad \text{(Eq. 5)}$$

Compared with other carbohydrate–lectin interactions (50), the on-rate constant k_{on} for WT $FimH_{LD}$ is in the typical range of 10^3–10^4 M^{-1} s^{-1}, whereas the off-rate constant k_{off} is rather low, resulting in a complex half-life $t_{½}$ of 83.6 min. The binding kinetic of the mutants A10P and A119L was similar to the WT $FimH_{LD}$ (Table 4). For V67K, the k_{on} was comparable with the WT $FimH_{LD}$, whereas the k_{off} was 7-fold increased. In case of $FimH_{FL}$, k_{on} was slightly increased, whereas the k_{off} was drastically enhanced by a factor of 3700 resulting in a $t_{½}$ of only 1.4 s. The association and dissociation of $FimH_{FL}$ were characterized by a multiphasic behavior (Fig. 5*B*) requiring fitting to a two-state binding model (Equation 6).

$$A + B \underset{k_{off,1}}{\overset{k_{on,1}}{\rightleftharpoons}} AB \underset{k_{off,2}}{\overset{k_{on,2}}{\rightleftharpoons}} A \cdot B \qquad \text{(Eq. 6)}$$

In this model, the analyte (FimH, A) binds to the ligand (mannoside, B) to form an initial complex (AB) and then undergoes subsequent binding or conformational change to form a more stable complex (A·B).

When the multicycle kinetics measurement was repeated on a chip surface with less densely immobilized ligand, the multiphasic behavior could be confirmed and steric hindrance as a possible artifact excluded (data not shown). Similar findings have recently been reported for FimH within the fully assembled type I pili (51). A two-state binding process of $FimH_{FL}$ featuring a second binding kinetic with a slow off-rate in the range of WT $FimH_{LD}$ (Table 4) likely represents the slow conversion of full-length FimH to the high-affinity state by separation of the lectin and pilin domains.

For the R60P mutant, k_{on} was comparable with WT $FimH_{LD}$, whereas k_{off} was increased 29-fold leading to a $t_{½}$ of 2.9 min and placing its kinetic profile between WT $FimH_{LD}$ and $FimH_{FL}$. Notably, the kinetic profile of the double-cysteine mutant V27C/L34C was very similar to full-length $FimH_{FL}$, except that only a monophasic binding behavior was observed (Fig. 5*A*). In comparison with WT $FimH_{LD}$, the k_{on} of V27C/L34C was increased by a factor of 19 and k_{off} by a factor of almost 2900, resulting in a $t_{½}$ of 1.7 s, *i.e.* very close to that of WT $FimH_{FL}$. In summary, the loss in affinity for R60P and especially for V27C/L34C mainly originates from enhanced k_{off} values, which is in agreement with the concept of a dynamic allostery, *i.e.* the binding pockets in the medium- and high-affinity states of FimH have nearly identical geometries, although the off-rates are modulated by the flexibility of the binding pocket and the clamp loop (23, 26). Consequently, we focused in the following sections on the two most interesting variants (R60P and V27C/L34C) and the mutant A119L for comparison with an unaffected variant.

Conformational fingerprints of selected $FimH_{LD}$ variants by NMR

^1H,^{15}N HSQC experiments of uniformly ^{15}N-labeled WT $FimH_{LD}$ and the mutants R60P, V27C/L34C, and A119L were measured in the absence and presence of mannoside **1** to assess structural differences on a residue-based level. The backbone assignment of WT $FimH_{LD}$ was available from previous studies (25, 45). The ^1H,^{15}N HSQC spectrum of WT $FimH_{LD}$ showed an excellent signal dispersion with numerous clearly isolated peaks, especially those of the binding pocket residues (Fig. 6*A*). Addition of mannoside **1** led to selective signal shifts of residues around the binding pocket (Fig. 6*B* and Fig. S6) as expected from direct ligand interactions (hydrogen bond formation) and from indirect effects (small changes in the binding pocket geometry, *e.g.* a conformational switch of the Tyr-48 side chain when interacting with the *n*-heptyl aglycone) (45). Pronounced

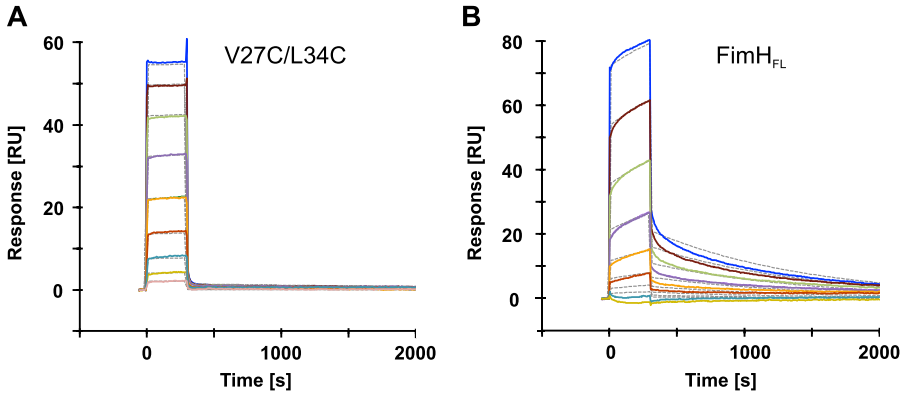


Figure 5. SPR sensorgrams (*colored lines*) of the analytes FimH$_{LD}$ mutant V27C/L34C (*A*) and FimH$_{FL}$ (*B*) WT, with immobilized mannoside 5 (see supporting information). *Gray dashed lines* show kinetic fitting to a one-state binding model (*A*) or a two-state binding model (*B*). The measurements were performed as detailed under "Experimental procedures."

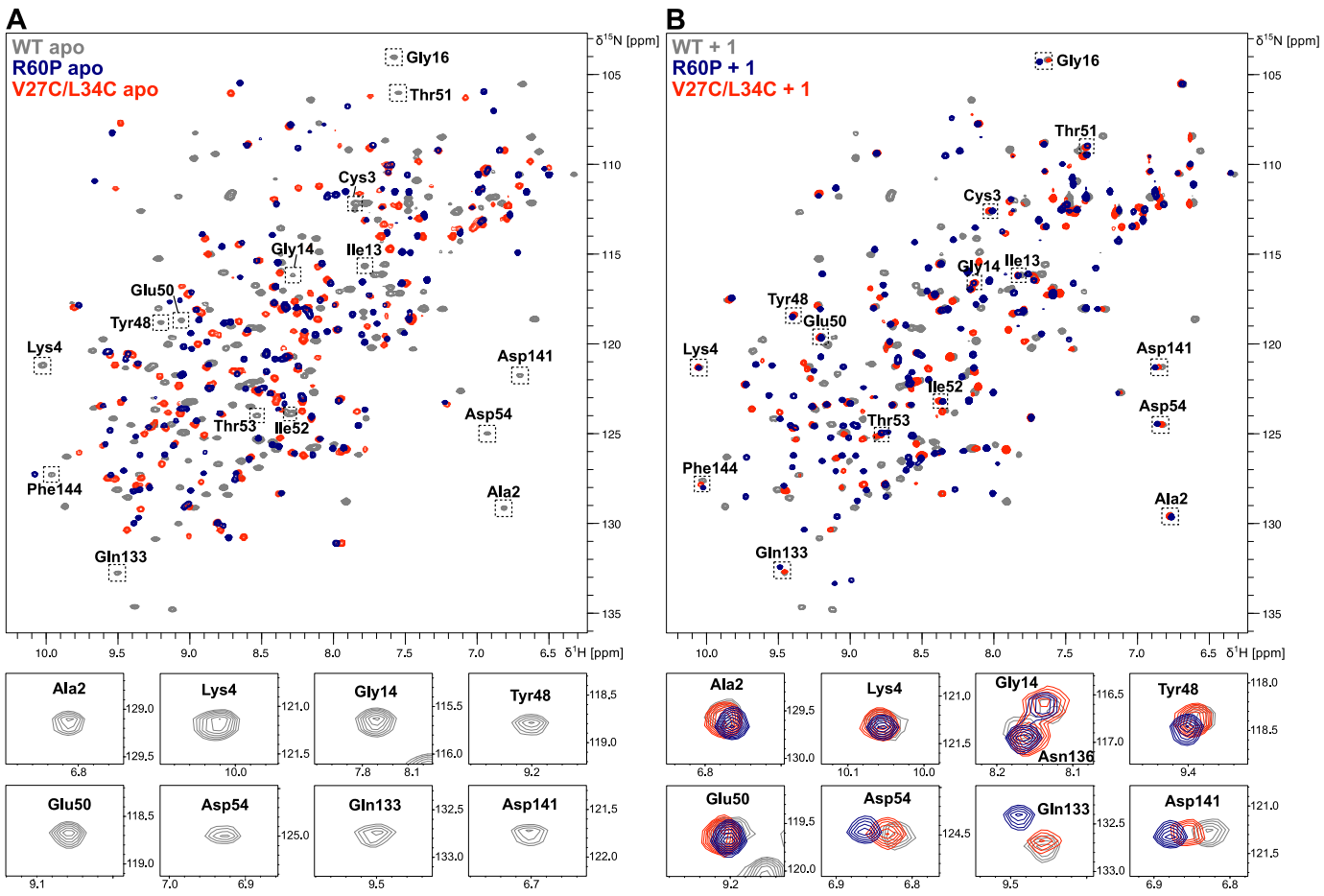


Figure 6. ^{1}H,^{15}N HSQC spectra of WT FimH$_{LD}$ (*gray*) and of the mutants R60P (*blue*) and V27C/L34C (*red*) in the absence (*A*) and presence of mannoside 1 (*B*). Spectral regions with residues of the binding pocket (indicated with *dashed-line boxes*) are expanded in the *lower panel*.

global conformational changes or allosteric effects upon mannoside binding could be excluded in agreement with the X-ray structural data (24, 25, 41, 45) and CD experiments (Fig. 3).

The ^{1}H,^{15}N HSQC spectra of the FimH$_{LD}$ mutant A119L in the absence and presence of mannoside **1** were almost superimposable to that of the WT with only some deviations around the point of mutation (Fig. S7) and a certain degree of peak heterogeneity in the binding pocket region in the absence of ligand (Fig. S8, see "Discussion").

In contrast, significant changes in the ^{1}H,^{15}N HSQC fingerprint spectra were observed for the R60P and V27C/L34C variants in comparison with the WT, suggesting drastically altered apo conformations (Fig. 6*A*). In particular, both mutants showed no peaks in the near vicinity of isolated signals characteristic for the binding pocket in the high-affinity state (Fig. 6*A*, *lower panel*). As we observed, roughly the expected number of peaks in the NMR spectra of the two mutants, this is unlikely to be the result of extensive line broadening due to chemical

exchange processes. Overlapping peaks in isolated spectral regions suggest that the two mutants have similar apo conformations (Fig. S9).

Strikingly, for both mutants, most signals of the binding pocket were recovered upon addition of mannoside **1** (Fig. 6*B*). Chemical shift maps for R60P and V27C/L34C, showing the shifts of the peaks relative to WT $FimH_{LD}$ in the mannoside **1**-bound form, indicate nearly identical chemical shifts in the binding pocket (Fig. 7). This suggests similar conformations for residues in direct contact with the mannosyl moiety (Fig. 7*E*: backbone signals of Gly-14, Asp-47, Asp-54, and Asp-140 and side chain signals of Gln-133 and Asn-135) as well as for the two loops encompassing Tyr-48 and Tyr-137 involved in hydrophobic interactions with the *n*-heptyl aglycone of mannoside **1** (Fig. 7*F*). In chemical shift perturbation (CSP) experiments with R60P and three different mannoside ligands (**1**-**3**, Fig. S10), characteristic peak positions observed in the WT $FimH_{LD}$ with each ligand could be restored, indicating identical ligand recognition. As a note, all mannoside complexes of $FimH_{LD}$ WT and mutants were in the slow or slow-intermediate regime as would also be expected from the off-rate constants measured with SPR (Table 4). Furthermore, for R60P and V27C/L34C, peak positions of the N-terminal pocket zipper and clamp-loop region (residues 2–16) and parts of the β-sheets were also recovered upon addition of mannoside **1**. However, both mutants exhibit significant chemical shift differences compared with WT $FimH_{LD}$ encompassing large parts of the β-sheets distal from the binding pocket, the α-switch, and β-bulge regions as well as the interdomain loops (Fig. 7, *C* and *D*).

The $^{1}H,^{15}N$ HSQC spectrum of the V27C/L34C mutant upon treatment with the reducing agent DTT closely resembled that of WT $FimH_{LD}$, in the absence or presence of mannoside **1** (Fig. S11). Thus, in agreement with previous reports (5, 46), we confirmed that the additional disulfide bridge Cys-27–Cys-34 in the interdomain region of the double mutant V27C/L34C promotes the low-affinity conformation of $FimH_{LD}$.

Low-affinity conformation of mutant R60P results from a slightly altered clamp loop

X-ray crystallography was attempted with the $FimH_{LD}$ mutants R60P and V27C/L34C to characterize their conformation on an atomic level. Although we did not succeed in growing crystals of the double mutant V27C/L34C, we obtained the X-ray crystal structure of R60P at a resolution of 1.6 Å (PDB code 5MCA) (Fig. 8 and Table S2). Whereas the structure was only poorly superimposable (r.m.s.d. of 2.2 Å) (52) with WT $FimH_{LD}$ (PDB code 4AUU) (41), it resembles the low-affinity state of $FimH_{FL}$ (r.m.s.d. of 0.6 Å) as found in the crystal structure of the assembled fimbrial tip (PDB code 3JWN) (5). In both structures, a twist in the β-sandwich fold of the lectin domain leads to a more compressed overall shape and, by a move of the clamp loop by ~3 Å, to a wider binding pocket in comparison with the high-affinity conformation present in WT $FimH_{LD}$ (Fig. 8*A*). In more detail, the conformation of the clamp loop of the mutant R60P is located between the closed conformation present in WT $FimH_{LD}$ and the completely open conformation of the full-length FimH in the assembled fimbrial tip (Fig. 8*B*). According to MD simulations of $FimH_{FL}$, the clamp loop is flexible in the absence of ligand (23), suggesting that its deviation in R60P from the low-affinity structure of FimH of the assembled fimbrial tip (PDB code 3JWN) might be incidental or results from crystallization conditions.

Discussion

The catch bond behavior of the bacterial adhesin FimH represents a fascinating example of the allosteric regulation of protein's conformational and ligand-binding properties (53, 54). Although the structural basis for this regulation has been elucidated (5, 23, 55), there is still a lack of experimental data on the transmission of the allosteric signal from the interdomain region to the distal binding pocket within $FimH_{LD}$. Four regions of $FimH_{LD}$ exhibit significant structural differences in their low- and high-affinity conformation (Fig. 2) and are considered to play an active role in this signal transmission (5). Based on this information, the impact of site-specific mutations on the conformation of the FimH lectin domain has been previously investigated (38). Our own interest was initially focused on mutations expected to stabilize the lectin domain in the low-affinity state, because such mutants would allow a further evaluation of already available FimH antagonists until now mainly tested against the high-affinity state (30–33). To obtain such mutants and to further unravel the allosteric regulation mechanism, we expressed and analyzed the $FimH_{LD}$ variants A10P, R60P, V67K, A119L, and V27C/L34C (Fig. 2), and compared their properties to WT $FimH_{LD}$ and WT $FimH_{FL}$.

Our investigations show that not all point mutations were effecting protein conformation and binding properties. For instance, the mutations A10P (pocket zipper region) and A119L (insertion loop of the interdomain region) had no influence on affinity and kinetics when compared with WT $FimH_{LD}$ (Tables 1 and 4). In a previous study using full-length FimH within the assembled pili of isogenic *E. coli* (38), these two mutants stabilized the low-affinity state, indicated by a low mono- *versus* trimannoside binding ratio. However, they retained the ability to adopt the high-affinity state similar to the WT variant as demonstrated by binding to the monoclonal antibody mAb21 in the presence of mannose. Therefore, for both mutants, these regions were assumed to show only weak allosteric coupling (38). In our study with A10P and A119L, CD analysis (Fig. 3 and Figs. S2 and S3) did not show any evidence of a conformational switch, but instead they strongly supported a high-affinity state similar to WT $FimH_{LD}$, both in the absence and presence of mannoside **1**. In the $^{1}H,^{15}N$ HSQC spectrum of apo-A119L (Fig. S7), double or triple peaks of binding pocket residues very close to the position of the WT signals were observed, indicating conformational flexibility. Upon addition of mannoside **1**, single peaks at positions identical to WT signals were observed, suggesting a rigidification of the binding pocket (Fig. S8). In line with these results, ITC experiments revealed an entropic penalty for A119L in comparison with WT $FimH_{LD}$, which, however, was compensated by a more favorable enthalpy term (Table 2). Despite this altered thermodynamic profile, the binding kinetics were not affected (Fig. 9 and Table 4). To sum up, similar to previous reports (38), we observed only a weak coupling between residue Ala-119 of the

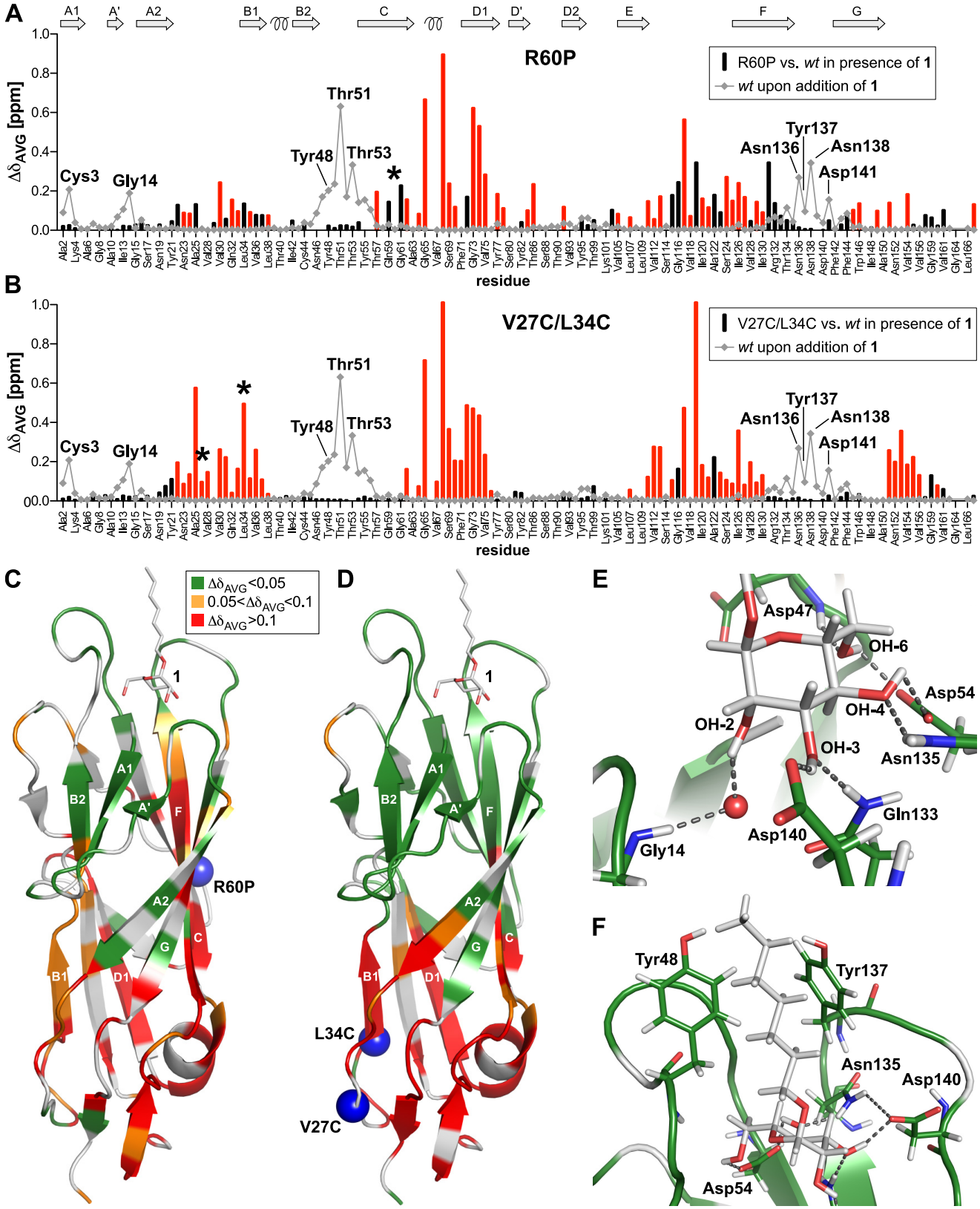


Figure 7. Chemical shift differences between WT FimH$_{LD}$ and the mutants R60P (*A* and *C*) and V27C/L34C (*B*, *D*, *E*, and *F*) in the mannoside 1-bound forms. *A* and *B*, CSP maps of the mutants relative to WT FimH$_{LD}$ in the presence of **1**; *red bars* indicate lower limits according to tentative assignments to the closest unassigned peak of the mutants; CSP effects of WT FimH$_{LD}$ upon addition of mannoside **1** are shown for reference to indicate the location of the binding pocket (*gray diamonds*); secondary structure elements are indicated above; the positions of the mutations are marked with *asterisks*. *C–F*, X-ray structure of WT FimH$_{LD}$ with **1** (PDB code 4BUQ) (35) colored according to CSP effects between the **1**-bound forms of WT FimH$_{LD}$ and R60P (*C*) or WT FimH$_{LD}$ and V27C/L34C (*D–F*).

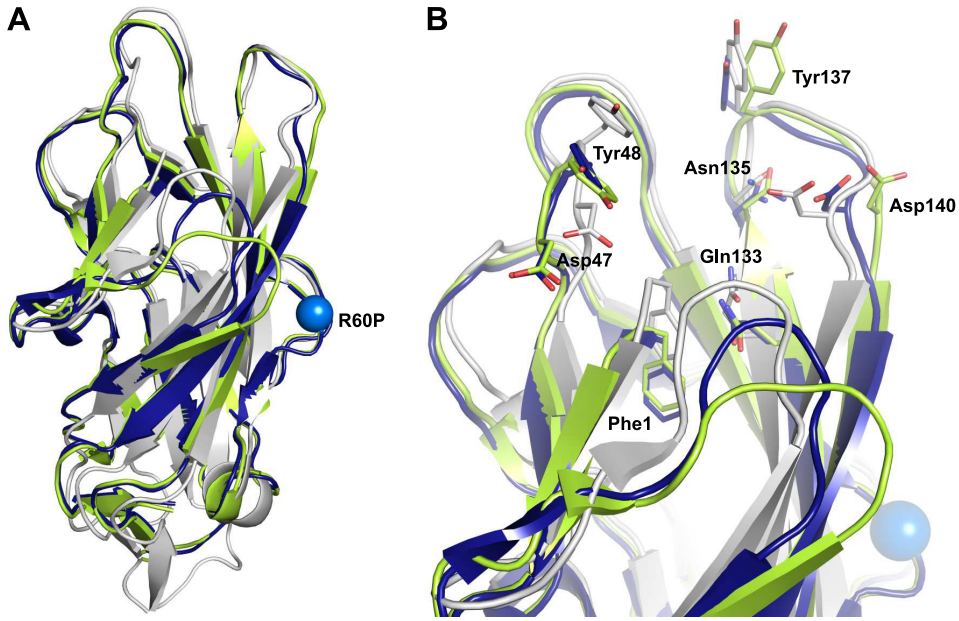


Figure 8. *A,* X-ray structures of $FimH_{LD}$ mutant R60P (*blue,* mutation indicated by *blue sphere*), WT $FimH_{LD}$ (*gray,* PDB code 4AUU), and full-length FimH from the assembled fimbrial tip (*lime,* PDB code 3JWN). *B,* expansion of the mannose-binding pocket and the clamp loop with key residues shown as *sticks*.

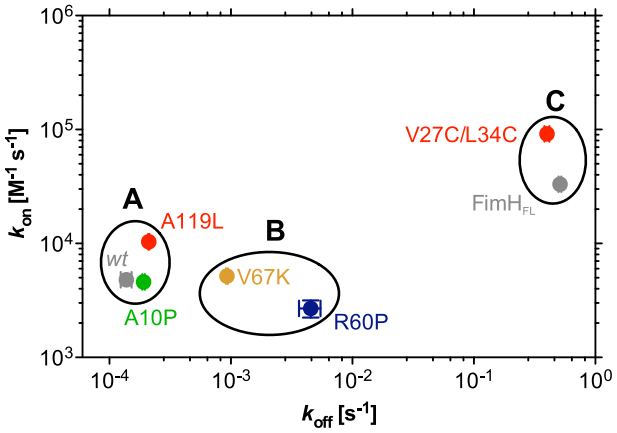


Figure 9. Kinetic rate constants obtained from SPR measurements of FimH variants with immobilized mannoside 5 as ligand: WT $FimH_{LD}$ (*gray*) and mutants (color-coded as in Fig. 2) and $FimH_{FL}$ (*gray*). *A,* variants with rate constants and K_D values similar to WT $FimH_{LD}$. *B,* variants with k_{on} values similar to WT $FimH_{LD}$ but with increased k_{off} and decreased K_D values. *C,* mutant V27C/L34C and $FimH_{FL}$ have similar rate constants, which are, however, drastically different from WT $FimH_{LD}$.

interdomain region with the binding pocket that, however, does not affect the affinity or kinetics of the interaction.

CD analysis for V67K, R60P, and V27C/L34C (Fig. 3 and Figs. S2 and S3) and NMR spectra for R60P and V27C/L34C mutants (Figs. 6 and 7) revealed an altered protein conformation in the ligand-free state that was converted to a conformation similar to WT $FimH_{LD}$ upon mannoside addition. In contrast to previously suggested lock in the low-affinity conformation of these mutants (5, 38, 40), this demonstrates for the first time a considerable conformational change of the isolated $FimH_{LD}$ in the absence of the regulatory pilin domain. For the mutant V67K, designed to destabilize the helical α-switch region of the high-affinity $FimH_{LD}$, binding affinity and kinetics were only mildly affected, suggesting solely a weak coupling between the α-switch region and the binding pocket. However, in comparison with WT $FimH_{LD}$, R60P (β-bulge) and V27C/L34C (interdomain region) showed drastically reduced binding affinities (Table 1) as well as altered conformational (Figs. 6 and 8) and kinetic properties (Table 4). The high resolution X-ray structure of R60P in the absence of ligand (PDB code 5MCA) clearly showed that the mutant adopts a low-affinity conformation similar to full-length FimH as present in the assembled pilus (5). In particular, the clamp loop was not folded over the mannose-binding pocket as observed in WT $FimH_{LD}$ (*e.g.* PDB code 4AUU) but was instead located several angstroms away similar to the full-length structure (PDB code 3JWN, and Fig. 8). $^{1}H,^{15}N$ HSQC NMR spectra of R60P and V27C/L34C confirmed the drastically altered apo conformations, because nearly all peaks were shifted relative to WT $FimH_{LD}$ (Fig. 6*A*). However, addition of mannoside **1** led to a complete recovery of NMR signals of the binding pocket, the clamp loop/pocket zipper, and large parts of the β-sheets, whereas the signals of the α-switch, β-bulge, and interdomain regions remained altered (Fig. 7). Interestingly, such a conformation accurately resembles the medium-affinity state observed in the crystal structure of the mannoside-bound $FimH_{FL}$ (23). Indirectly, this is evidenced by the observation that an inhibitory antibody raised against the $FimH_{LD}$ mutant V27C/L34C, with an epitope in the binding pocket, strongly binds to FimH in the high-affinity state (40), which displays a nearly identical geometry of the binding pocket compared with the medium-affinity state (23). However, the exact conformation of V27C/L34C and R60P in the bound state remains to be confirmed, *e.g.* by X-ray structures of the mutants in complex with mannoside or comparative NMR studies with ^{15}N-labeled $FimH_{FL}$.

The similarity of the $FimH_{LD}$ mutant V27C/L34C with $FimH_{FL}$ was further substantiated by their almost identical kinetic profiles (Fig. 9). Finally, the reduction of the disulfide bond in the interdomain region of the V27C/L34C mutant restores a high-affinity state similar to WT $FimH_{LD}$ (Fig. S11),

thus confirming the involvement of the interdomain region in the allosteric regulation of mannose binding (5).

Our data for the three mutants V67K, R60P, and V27C/L34C are in good agreement with MD simulations of $FimH_{LD}$, indicating a weak coupling of the binding pocket to the α-switch region (V67K) but a strong coupling to the β-bulge (R60P) and to the interdomain regions (V27C/L34C) (26). These couplings are supposed to regulate the dynamics in the clamp loop near the binding pocket and to cause the renowned catch-bond behavior of FimH. A higher flexibility of the clamp loop in the medium-affinity state could be responsible for the dramatically enhanced off-rates for mannosides relative to the high-affinity state, despite nearly identical conformation of the binding pocket (23, 26).

In future work, NMR dynamics experiments (56, 57) could be used to address the higher flexibility of the clamp loop in the medium-affinity state. However, in the case of labeled $FimH_{FL}$, the required expression in minimal media is hampered by the cumbersome expression and purification procedure (23) as well as the low yield.[5] Therefore, the characterized mutants represent useful surrogates of full-length FimH as they can be readily expressed with uniform ^{15}N and ^{13}C labeling in preparative amounts (4–8 mg/liter) and benefit from their smaller size in terms of spectral dispersion and relaxation properties.

Based on the global conformational changes in the mutants of $FimH_{LD}$, V67K, R60P, and V27C/L34C, we expected drastic entropic consequences. Although all three mutants displayed entropic penalties relative to WT $FimH_{LD}$, large enthalpic penalties were the major reason for the loss of affinity of R60P and V27C/L34C. Obviously, an interpretation of the thermodynamics of WT $FimH_{LD}$ and mutants thereof is highly speculative in view of the many contributing factors. Thus, the introduction of mutations may lead to local thermodynamic strain within the protein and to unexpected entropic effects of coordinated water molecules as a result of altered hydrophobic regions. Consequently, the entropic penalty caused by the conformational changes may be partially compensated.

In summary, our data clearly demonstrate that by a single mutation in the β-bulge (R60P) or by a double-point mutation introducing an additional disulfide bridge in the interdomain region (V27C/L34C), the isolated $FimH_{LD}$ can be converted from the high- into the low-affinity state and moreover can retain the same intricate allosteric regulation as observed in the full-length FimH. In particular, the kinetic profile of full-length FimH as the hallmark of the catch bond behavior was almost perfectly reproduced by the $FimH_{LD}$ mutant V27C/L34C (Fig. 9), thus clearly demonstrating the cross-talk between allosteric sites within the lectin domain in the absence of the pilin domain. Neither mutant was conformationally locked in the low-affinity state, but instead a switch to the medium-affinity state upon mannoside binding similar to the full-length protein could be demonstrated. It remains to be seen whether a combination of different point mutations would lead to a complete lock of $FimH_{LD}$ in the low-affinity state, even in the presence of a mannoside ligand.

[5] S. Rabbani, B. Fiege, D. Eris, M. Silbermann, R. P. Jakob, G. Navarra, T. Maier, and B. Ernst, unpublished data.

The mutants characterized herein therefore represent an easily accessible mimic of full-length FimH and may serve as valuable tools for further mechanistic investigations of the allosteric regulation. Ultimately, this knowledge will be required for the design of effective anti-adhesive drugs and antibody-based therapeutics for the treatment of UTI.

Experimental procedures

Cloning of $FimH_{LD}$ mutants

The $FimH_{LD}$ construct linked to a thrombin cleavage site (Th) and a His_6 tag on the C terminus was generated as described (28). The mutants A10P, R60P, V67K, A119L, and V27C/L34C were generated by the overlap extension PCR method (58) using the WT-encoding plasmid as template. The inserts were digested with HindIII and XbaI restriction enzymes, gel-purified, and subsequently ligated into the corresponding cloning site of pDsbA3 expression vector (59). The vectors were then transformed separately into *E. coli* DH5α chemo-competent cells (Novagen, Lucerne, Switzerland). After plasmid isolation and restriction control, the correctness of the constructs was confirmed by DNA sequencing (Microsynth, Balgach, Switzerland). Finally, the vectors were transformed into the protease-deficient *E. coli* strain HM125 (60) for protein expression and export into the periplasmic space.

Protein expression and purification

Bacterial clones were grown at 30 °C with vigorous shaking (300 rpm) in M9 minimal medium (61) supplemented with 2 mM $MgSO_4$, 0.1 mM $CaCl_2$, 20% glucose, 10 μM of each amino acid, basal medium Eagle vitamin mix (Sigma, Buchs, Switzerland), and 100 μg/ml ampicillin. When an A_{600} of 0.8 was reached, the cells were induced with 1 mM isopropyl 1-thio-β-D-galactopyranoside (Applichem, Germany) and further cultivated for 16 h at 30 °C and 300 rpm. Then the cells were cooled on ice for 5 min and harvested by centrifugation at 5000 rpm for 20 min at 4 °C. The pellet was suspended in a cold solution of 50 mM Tris-HCl, pH 7.5, 150 mM NaCl, 5 mM EDTA, 1 mg/ml polymyxin B sulfate (lysis buffer) and stirred for 2 h at 4 °C. After centrifugation at 11,000 rpm for 20 min at 4 °C, the supernatant (periplasmic extract) was dialyzed overnight against 50 mM NaH_2PO_4, 300 mM NaCl, and 10 mM imidazole, pH 8 (binding buffer), and applied to a nickel-nitrilotriacetic acid column (Sigma, Buchs, Switzerland) attached to a BioLogic fast protein liquid chromatography system (Bio-Rad, Reinach BL, Switzerland). The column was washed with binding buffer and afterward eluted with 50 mM NaH_2PO_4, 300 mM NaCl, and 250 mM imidazole, pH 8 (elution buffer). The fractions containing $FimH_{LD}$-Th-His_6 were pooled and dialyzed against 20 mM HEPES buffer, pH 7.4, with 150 mM NaCl and 1 mM $CaCl_2$ (assay buffer). The purity of the protein was verified by SDS-PAGE analysis, and the monomeric state was confirmed by native PAGE. The amount (4–10 mg/liter of each of the five mutants) was determined by direct absorption at 280 nm on an ND-1000 nanodrop spectrophotometer (Thermo Fisher Scientific). Aliquots of the proteins could be stored for up to 3 months at 4 °C. For long-term storage, the proteins were frozen at −80 °C without additives. For protein crystallization, $FimH_{LD}$ constructs (amino acids 1–158), lacking the His_6 tag, were

generated following the same procedure described above and purified by ion-exchange chromatography (29). $FimH_{FL}$ was expressed and purified as described (23).

Differential scanning fluorimetry

Differential scanning fluorimetry was performed using a Rotor-Gene Q 6plex quantitative PCR instrument according to the procedure described previously (42). SYPRO Orange dye was diluted 1:50 in the final samples and was used to monitor the protein-denaturing profiles. The assay was performed in assay buffer with 5 μM protein and *n*-heptyl-α-D-mannopyranoside (**1**) (45, 62) at a final concentration of 5 mM.

Circular dichroism analysis

CD spectra of WT $FimH_{LD}$ and mutants were recorded in the absence and presence of 1 mM mannoside **1** (63, 64). Far-UV (180–260 nm) and near-UV CD spectra (250–350 nm) were recorded with a Chirascan CD spectrometer (Applied Photophysics, Leatherhead, UK). For far-UV analysis, the proteins were diluted to a concentration of 5 μM in 10 mM sodium phosphate buffer at pH 7.4 and measured at 20 °C using a thermostat-controlled cell with 0.1-cm path length. The CD spectra were recorded between 190 and 250 nm with 2 nm bandwidth and 10 scans per sample. For near-UV analysis, all samples were measured at a protein concentration of 1 mg/ml using a thermostat-controlled cell with a path length of 1 cm and between 260 and 320 nm. Mannoside **1** was added to a final concentration of 1 mM for both near- and far-UV CD analysis. CD spectra were corrected by buffer subtraction, and the CD signal was converted to molar ellipticity ([θ]/degree cm^2 $dmol^{-1}$).

FP assay

The fluorescence polarization assay was performed as described previously (32). Briefly, a serial dilution of each $FimH_{LD}$ mutant and WT (0.1 nM to 2.6 μM) was mixed with a constant concentration of **1** (5 nM with WT and mutants V67K and A119L; 10 nM with mutants V27C/L34C and R60P) in a final reaction volume of 100 μl in 20 mM HEPES, pH 7.4, with 150 mM NaCl, and 50 μg/ml BSA. The mixture was incubated at room temperature in black, flat bottom, non-binding surface 96-well microtiter plates (Corning) for at least 24 h. Binding of the tracer ligand **4** (3′-chloro-*N*-(2-(3-(3′,6′-dihydroxy-3-oxo-3*H*-spiro-[isobenzofuran-1,9′-xanthen]-5-yl)-thioureido)-ethyl)-4′-(α-D-mannopyranosyloxy)-biphenyl-4-carboxamide), synthesized as described previously (32) to each $FimH_{LD}$ mutant was monitored via an increase in FP using a Synergy H1 hybrid microplate reader (BioTek). Samples were excited with polarized light at 485 nm and emission was measured at 528 nm. FP data were fitted to a single-site binding model accounting for ligand depletion (65) to determine the K_D value of **4** for each $FimH_{LD}$ variant. The K_D value for **1** was determined competitively via the displacement of **4** by **1** and the associated decrease in FP. For these competition assays, a serial dilution of competitor **1** (7.6 nM to 500.0 μM) was titrated into constant concentrations of $FimH_{LD}$ and **4** (10 nM). The concentration of each mutant was fixed such that it would be above K_D and the competitor concentration at 50% signal change would be in the same approximate range for all mutants. Specifically, it was fixed at 200 nM for V27C/L34C, 100 nM for R60P and V67K, and 50 nM for A119L and WT. FP was measured as described above, and data were fitted to an equilibrium competition binding function (66). Data fitting and analysis were performed using Prism (GraphPad Software).

Isothermal titration calorimetry (ITC)

Prior to analysis, the proteins were dialyzed against 20 mM HEPES buffer, pH 7.4, with 150 mM NaCl. ITC experiments were performed at 15, 25, 31, and 37 °C using a VP-ITC microcalorimeter (Malvern Instruments, UK). Volumes ranging from 6 to 8 μl of 150 μM compound **1** were injected at 600-s intervals into a reaction cell containing 1.45 ml of 10–15 μM $FimH_{LD}$ WT or mutant. Baseline correction and peak integration were completed using Origin (OriginLab). The resulting binding isotherm was imported into SEDPHAT (National Institutes of Health) (69) for baseline subtraction and data fitting. The initial 2-μl injection was not included for data analysis. Data were fit to a single-site binding model to derive K_D and ΔH (change in binding enthalpy) values from the binding isotherm. ΔS^0_{obs} (change in entropy) was calculated from the known thermodynamic quantities and further dissected into conformational, mixing, and solvation entropy according to the methodology as described in Ref. 70. A concentration correction factor was included as fitting parameter to account for the uncertainty in determining $FimH_{LD}$ concentration. Statistical analysis to determine best-fit values and 95% confidence intervals was performed with SEDPHAT (69).

SPR

Kinetic data were obtained by SPR-based measurements using a Biacore T200 system (GE Healthcare, Uppsala, Sweden). Compound **5** (see supporting information for synthesis and structure) was immobilized on a carboxy-dextran-coated CM4 sensor chip (GE Healthcare) in a two-step process. In the first step, 1,2-diamino-ethane (0.1 M in borate buffer, pH 8.5) was covalently coupled to the chip matrix through NHS-ester chemistry using the amine coupling kit (GE Healthcare). The carboxy-functionalized compound **5** was activated by NHS/EDC. After raising the pH to 8.5 with 0.1 M borate buffer, the derivatized ligand was injected into the system (0.5 mM for 30 s) and reacted with the free amine groups of 1,2-diamino-ethane. Ethanolamine at a 1 M concentration was used to block any remaining activity on the surface. A reference cell was prepared by immobilizing 8-hydroxyoctanoic acid using the amine coupling kit (GE Healthcare) followed by blocking with 1 M ethanolamine.

All SPR measurements were performed 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 μl min^{-1}. Because of baseline drifts and incomplete chemical regeneration (50 mM NaOH, 180 s), complete dissociation of the protein (analyte) was achieved by sufficiently long dissociation times. Because of the resulting long measurement times for multicycle kinetic (MCK) assays, MCK was only performed if the k_{off} was above 10^{-3} s^{-1}. Otherwise, the single cycle kinetics mode was used that includes a single

dissociation phase to obtain kinetic information. Data processing and kinetic evaluations were performed with the Biacore T200 control software and the Biacore T200 evaluation software (both version 1.0), respectively. The Langmuir 1:1 model of interaction was used to determine k_{on}, k_{off}, and K_D values. For $FimH_{FL}$, data were fitted using a two-state reaction model. The bulk effect on the refractive index was set to zero because of the reference cell subtraction.

NMR spectroscopy

All experiments were performed at 298 K on a Bruker Avance III 600 MHz NMR spectrometer equipped with a 5-mm TXI room temperature probe head. Samples contained 120–240 μM ^{15}N-labeled WT $FimH_{LD}$ or mutant in 20 mM phosphate buffer, pH 7.4, with 7% D_2O. Mannosides **1**, *O*-methyl-α-D-mannopyranoside (**2**), and *n*-heptyl 4-deoxy-4-fluoro-α-D-mannopyranoside (**3**) (synthesized as described before (45, 62), see Fig. S10 for structures) were solved in D_2O and added stepwise to the protein up to a 10-fold molar excess. The mutant V27C/L34C (150 μM) was additionally measured in the presence of DTT (5 mM) in the absence or presence of compound **1** (2 mM). ^{1}H,^{15}N HSQC experiments were acquired and processed with Topspin 3.2 (Bruker BioSpin, Switzerland) and analyzed with CcpNmr Analysis (version 2.2) (67). The backbone assignment of $FimH_{LD}$ was available from previous studies (25, 45). Partial assignments of mutant proteins were done on the basis of chemical shift proximity. Where applicable, combined chemical shift differences, $\Delta\delta_{AV}$, between the signals of WT and mutant were calculated (68) as shown in Equation 7,

$$\Delta\delta_{AV} = \sqrt{(\Delta\delta^{1}H^{N})^2 + (0.2\Delta\delta^{15}N)^2} \quad \text{(Eq. 7)}$$

As all tested compounds bound to all $FimH_{LD}$ variants in the slow to intermediate–slow exchange regime on the NMR time scale, partial assignments of the bound states could not be easily transferred to the free states. For residues that cannot be assigned unambiguously in the spectra of the mutants, a lower limit of the $\Delta\delta_{AVG}$ value was determined that corresponds to the chemical shift difference between the WT peak and the nearest unassigned peak in the mutant spectrum (chemical shift maps in Fig. 7).

Protein crystallization and structure determination

$FimH_{LD}$ R60P crystals were grown (10 mg/ml in 20 mM HEPES, pH 7.4) in 1.5 M $(NH_4)_2SO_4$, 0.1 M BisTris, pH 6.5, at 20 °C. Crystals were cryopreserved by a quick soak in 2.5 M Li_2SO_4 (71) and flash-cooled in liquid nitrogen. Data were collected with synchrotron radiation at the X06DA beamline at the Swiss Light Source (Paul Scherrer Institut, Switzerland). Data were indexed, integrated, and scaled with the XDS package (72). The structure was solved by molecular replacement with PHASER (73) using the $FimH_{LD}$–biphenyl α-D-mannopyranoside complex (PDB code 4X50) as a search model (45). The structure was iteratively built using the COOT software (74), refined with the PHENIX software (75), and validated using MolProbity (76). The structure has been deposited in the Protein Data Bank under the accession code 5MCA.

Author contributions—S. R. performed cloning, protein production, stability testing, and CD measurements. B. F. conducted and analyzed NMR experiments. D. E. and M. S. performed ITC and SPR measurements. R. P. J. and T. M. obtained the X-ray structure of mutant R60P. G. N. synthesized compounds. S. R., B. F., and B. E. designed the work and wrote the manuscript, and all other authors provided helpful discussions.

Acknowledgments—We thank Prof. Stephan Grzesiek (Biozentrum, University of Basel) for the opportunity to measure on a 600 MHz NMR spectrometer. We acknowledge Dr. Timothy Sharpe (Biozentrum, Biophysics Facility, University of Basel) for general support.

References

1. Mulvey, M. A., Schilling, J. D., Martinez, J. J., and 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 CrossRef Medline
2. Schilling, J. D., Mulvey, M. A., and 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,** S36–S40 CrossRef
3. Jones, C. H., Pinkner, J. S., Roth, R., Heuser, J., Nicholes, A. V., Abraham, S. N., and Hultgren, S. J. (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,** 2081–2085 CrossRef Medline
4. Sokurenko, E. V., Chesnokova, V., Doyle, R. J., and Hasty, D. L. (1997) Diversity of the *Escherichia coli* type 1 fimbrial lectin. Differential binding to mannosides and uroepithelial cells. *J. Biol. Chem.* **272,** 17880–17886 CrossRef Medline
5. 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., and Thomas, W. E. (2010) Structural basis for mechanical force regulation of the adhesin FimH via finger trap-like β-sheet twisting. *Cell* **141,** 645–655 CrossRef Medline
6. Sharon, N. (2006) Carbohydrates as future anti-adhesion drugs for infectious diseases. *Biochim. Biophys. Acta* **1760,** 527–537 CrossRef Medline
7. Choudhury, D., Thompson, A., Stojanoff, V., Langermann, S., Pinkner, J., Hultgren, S. J., and Knight, S. D. (1999) X-ray structure of the FimC-FimH chaperone-adhesin complex from uropathogenic *Escherichia coli*. *Science* **285,** 1061–1066 CrossRef Medline
8. Sauer, F. G., Fütterer, K., Pinkner, J. S., Dodson, K. W., Hultgren, S. J., and Waksman, G. (1999) Structural basis of chaperone function and pilus biogenesis. *Science* **285,** 1058–1061 CrossRef Medline
9. Nilsson, L. M., Thomas, W. E., Trintchina, E., Vogel, V., and Sokurenko, E. V. (2006) Catch bond-mediated adhesion without a shear threshold: trimannose versus monomannose interactions with the FimH adhesin of *Escherichia coli*. *J. Biol. Chem.* **281,** 16656–16663 CrossRef Medline
10. Thomas, W. E., Trintchina, E., Forero, M., Vogel, V., and Sokurenko, E. V. (2002) Bacterial adhesion to target cells enhanced by shear force. *Cell* **109,** 913–923 CrossRef Medline
11. Isberg, R. R., and Barnes, P. (2002) Dancing with the host; flow-dependent bacterial adhesion. *Cell* **110,** 1–4 CrossRef Medline
12. Marshall, B. T., Long, M., Piper, J. W., Yago, T., McEver, R. P., and Zhu, C. (2003) Direct observation of catch bonds involving cell-adhesion molecules. *Nature* **423,** 190–193 CrossRef Medline
13. Phan, U. T., Waldron, T. T., and Springer, T. A. (2006) Remodeling of the lectin-EGF-like domain interface in P- and L-selectin increases adhesiveness and shear resistance under hydrodynamic force. *Nat. Immunol.* **7,** 883–889 CrossRef Medline
14. Chen, W., Lou, J., Evans, E. A., and Zhu, C. (2012) Observing force-regulated conformational changes and ligand dissociation from a single integrin on cells. *J. Cell Biol.* **199,** 497–512 CrossRef Medline
15. Akiyoshi, B., Sarangapani, K. K., Powers, A. F., Nelson, C. R., Reichow, S. L., Arellano-Santoyo, H., Gonen, T., Ranish, J. A., Asbury, C. L., and

Biggins, S. (2010) Tension directly stabilizes reconstituted kinetochore-microtubule attachments. *Nature* **468,** 576–579 CrossRef Medline

16. Buckley, C. D., Tan, J., Anderson, K. L., Hanein, D., Volkmann, N., Weis, W. I., Nelson, W. J., and Dunn, A. R. (2014) The minimal cadherin-catenin complex binds to actin filaments under force. *Science* **346,** 1254211 CrossRef Medline
17. Fiore, V. F., Ju, L., Chen, Y., Zhu, C., and Barker, T. H. (2014) Dynamic catch of a Thy-1-α5β1+syndecan-4 trimolecular complex. *Nat. Commun.* **5,** 4886 CrossRef Medline
18. Das, D. K., Feng, Y., Mallis, R. J., Li, X., Keskin, D. B., Hussey, R. E., Brady, S. K., Wang, J. H., Wagner, G., Reinherz, E. L., and Lang, M. J. (2015) Force-dependent transition in the T-cell receptor β-subunit allosterically regulates peptide discrimination and pMHC bond lifetime. *Proc. Natl. Acad. Sci. U.S.A.* **112,** 1517–1522 CrossRef Medline
19. Liu, B., Chen, W., Evavold, B. D., and Zhu, C. (2014) Accumulation of dynamic catch bonds between TCR and agonist peptide-MHC triggers T cell signaling. *Cell* **157,** 357–368 CrossRef Medline
20. Aprikian, P., Tchesnokova, V., Kidd, B., Yakovenko, O., Yarov-Yarovoy, V., Trinchina, E., Vogel, V., Thomas, W., and Sokurenko, E. (2007) Interdomain interaction in the FimH adhesin of *Escherichia coli* regulates the affinity to mannose. *J. Biol. Chem.* **282,** 23437–23446 CrossRef Medline
21. Tchesnokova, V., Aprikian, P., Kisiela, D., Gowey, S., Korotkova, N., Thomas, W., and Sokurenko, E. (2011) Type 1 fimbrial adhesin FimH elicits an immune response that enhances cell adhesion of *Escherichia coli*. *Infect. Immun.* **79,** 3895–3904 CrossRef Medline
22. Kalas, V., Pinkner, J. S., Hannan, T. J., Hibbing, M. E., Dodson, K. W., Holehouse, A. S., Zhang, H., Tolia, N. H., Gross, M. L., Pappu, R. V., Janetka, J., and Hultgren, S. J. (2017) Evolutionary fine-tuning of conformational ensembles in FimH during host-pathogen interactions. *Sci. Adv.* **3,** e1601944 Medline
23. Sauer, M. M., Jakob, R. P., Eras, J., Baday, S., Eriş, D., Navarra, G., Bernèche, S., Ernst, B., Maier, T., and Glockshuber, R. (2016) Catch-bond mechanism of the bacterial adhesin FimH. *Nat. Commun.* **7,** 10738 CrossRef Medline
24. Bouckaert, J., Berglund, J., Schembri, M., De Genst, E., Cools, L., Wuhrer, M., Hung, C. S., Pinkner, J., Slättegård, R., Zavialov, A., Choudhury, D., Langermann, S., Hultgren, S. J., Wyns, L., Klemm, P., *et al.* (2005) Receptor binding studies disclose a novel class of high-affinity inhibitors of the *Escherichia coli* FimH adhesin. *Mol. Microbiol.* **55,** 441–455 Medline
25. Vanwetswinkel, S., Volkov, A. N., Sterckx, Y. G., Garcia-Pino, A., Buts, L., Vranken, W. F., Bouckaert, J., Roy, R., Wyns, L., and van Nuland, N. A. (2014) Study of the structural and dynamic effects in the FimH adhesin upon α-D-heptyl mannose binding. *J. Med. Chem.* **57,** 1416–1427 CrossRef Medline
26. Interlandi, G., and Thomas, W. E. (2016) Mechanism of allosteric propagation across a β-sheet structure investigated by molecular dynamics simulations. *Proteins* **84,** 990–1008 CrossRef Medline
27. Tsai, C. J., del Sol, A., and Nussinov, R. (2008) Allostery: absence of a change in shape does not imply that allostery is not at play. *J. Mol. Biol.* **378,** 1–11 CrossRef Medline
28. Rabbani, S., Jiang, X., Schwardt, O., and Ernst, B. (2010) Expression of the carbohydrate recognition domain of FimH and development of a competitive binding assay. *Anal. Biochem.* **407,** 188–195 CrossRef Medline
29. Vetsch, M., Sebbel, P., and Glockshuber, R. (2002) Chaperone-independent folding of type 1 pilus domains. *J. Mol. Biol.* **322,** 827–840 CrossRef Medline
30. Han, Z., Pinkner, J. S., Ford, B., Chorell, E., Crowley, J. M., Cusumano, C. K., Campbell, S., Henderson, J. P., Hultgren, S. J., and 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 CrossRef Medline
31. Hartmann, M., and 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.* 3583–3609
32. Kleeb, S., Pang, L., Mayer, K., Eris, D., Sigl, A., Preston, R. C., Zihlmann, P., Sharpe, T., Jakob, R. P., Abgottspon, D., Hutter, A. S., Scharenberg, M., Jiang, X., Navarra, G., Rabbani, S., *et al.* (2015) FimH antagonists: bioisosteres to improve the *in vitro* and *in vivo* PK/PD profile. *J. Med. Chem.* **58,** 2221–2239 CrossRef Medline
33. 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., and 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 CrossRef Medline
34. Pang, L., Kleeb, S., Lemme, K., Rabbani, S., Scharenberg, M., Zalewski, A., Schädler, F., Schwardt, O., and Ernst, B. (2012) FimH antagonists: structure-activity and structure-property relationships for biphenyl α-D-mannopyranosides. *ChemMedChem* **7,** 1404–1422 CrossRef Medline
35. Roos, G., Wellens, A., Touaibia, M., Yamakawa, N., Geerlings, P., Roy, R., Wyns, L., and Bouckaert, J. (2013) Validation of reactivity descriptors to assess the aromatic stacking within the tyrosine gate of FimH. *ACS Med. Chem. Lett.* **4,** 1085–1090 CrossRef Medline
36. Eris, D., Preston, R. C., Scharenberg, M., Hulliger, F., Abgottspon, D., Pang, L., Jiang, X., Schwardt, O., and Ernst, B. (2016) The conformational variability of FimH: which conformation represents the therapeutic target? *ChemBioChem* **17,** 1012–1020 CrossRef Medline
37. Mayer, K., Eris, D., Schwardt, O., Sager, C. P., Rabbani, S., Kleeb, S., and Ernst, B. (2017) Urinary tract infection–which conformation of the bacterial lectin FimH is therapeutically relevant? *J. Med. Chem.* **60,** 5646–5662 CrossRef Medline
38. Rodriguez, V. B., Kidd, B. A., Interlandi, G., Tchesnokova, V., Sokurenko, E. V., and Thomas, W. E. (2013) Allosteric coupling in the bacterial adhesive protein FimH. *J. Biol. Chem.* **288,** 24128–24139 CrossRef Medline
39. Tchesnokova, V., Aprikian, P., Yakovenko, O., Larock, C., Kidd, B., Vogel, V., Thomas, W., and Sokurenko, E. (2008) Integrin-like allosteric properties of the catch bond-forming FimH adhesin of *Escherichia coli*. *J. Biol. Chem.* **283,** 7823–7833 CrossRef Medline
40. Kisiela, D. I., Rodriguez, V. B., Tchesnokova, V., Avagyan, H., Aprikian, P., Liu, Y., Wu, X. R., Thomas, W. E., and Sokurenko, E. V. (2013) Conformational inactivation induces immunogenicity of the receptor-binding pocket of a bacterial adhesin. *Proc. Natl. Acad. Sci. U.S.A.* **110,** 19089–19094 CrossRef Medline
41. Wellens, A., Lahmann, M., Touaibia, M., Vaucher, J., Oscarson, S., Roy, R., Remaut, H., and Bouckaert, J. (2012) The tyrosine gate as a potential entropic lever in the receptor-binding site of the bacterial adhesin FimH. *Biochemistry* **51,** 4790–4799 CrossRef Medline
42. Niesen, F. H., Berglund, H., and Vedadi, M. (2007) The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability. *Nat. Protoc.* **2,** 2212–2221 CrossRef Medline
43. Grzymajlo, K., Kuzminska-Bajor, M., Jaworski, J., Dobryszycki, P., and Ugorski, M. (2010) The high-adhesive properties of the FimH adhesin of *Salmonella enterica* serovar *Enteritidis* are determined by a single F118S substitution. *Microbiology* **156,** 1738–1748 CrossRef Medline
44. Kelly, S. M., and Price, N. C. (2000) The use of circular dichroism in the investigation of protein structure and function. *Curr. Protein Pept. Sci.* **1,** 349–384 CrossRef Medline
45. Fiege, B., Rabbani, S., Preston, R. C., Jakob, R. P., Zihlmann, P., Schwardt, O., Jiang, X., Maier, T., and 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 CrossRef Medline
46. Nilsson, L. M., Yakovenko, O., Tchesnokova, V., Thomas, W. E., Schembri, M. A., Vogel, V., Klemm, P., and Sokurenko, E. V. (2007) The cysteine bond in the *Escherichia coli* FimH adhesin is critical for adhesion under flow conditions. *Mol. Microbiol.* **65,** 1158–1169 CrossRef Medline
47. Baker, B. M., and Murphy, K. P. (1997) Dissecting the energetics of a protein-protein interaction: the binding of ovomucoid third domain to elastase. *J. Mol. Biol.* **268,** 557–569 CrossRef Medline
48. Baldwin, R. L. (1986) Temperature dependence of the hydrophobic interaction in protein folding. *Proc. Natl. Acad. Sci. U.S.A.* **83,** 8069–8072 CrossRef Medline
49. Murphy, K. P., Xie, D., Thompson, K. S., Amzel, L. M., and Freire, E. (1994) Entropy in biological binding processes: estimation of translational entropy loss. *Proteins* **18,** 63–67 CrossRef Medline

50. Scharenberg, M., Jiang, X., Pang, L., Navarra, G., Rabbani, S., Binder, F., Schwardt, O., and Ernst, B. (2014) Kinetic properties of carbohydrate-lectin interactions: FimH antagonists. *ChemMedChem* **9,** 78–83 CrossRef Medline
51. Yakovenko, O., Tchesnokova, V., Sokurenko, E. V., and Thomas, W. E. (2015) Inactive conformation enhances binding function in physiological conditions. *Proc. Natl. Acad. Sci. U.S.A.* **112,** 9884–9889 CrossRef Medline
52. Krissinel, E., and Henrick, K. (2004) Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta Crystallogr. D Biol. Crystallogr.* **60,** 2256–2268 CrossRef Medline
53. Kern, D., and Zuiderweg, E. R. (2003) The role of dynamics in allosteric regulation. *Curr. Opin. Struct. Biol.* **13,** 748–757 CrossRef Medline
54. Motlagh, H. N., Wrabl, J. O., Li, J., and Hilser, V. J. (2014) The ensemble nature of allostery. *Nature* **508,** 331–339 CrossRef Medline
55. Sokurenko, E. V., Vogel, V., and Thomas, W. E. (2008) Catch-bond mechanism of force-enhanced adhesion: counterintuitive, elusive, but. widespread? *Cell Host Microbe* **4,** 314–323 CrossRef Medline
56. Salmon, L., Bouvignies, G., Markwick, P., and Blackledge, M. (2011) Nuclear magnetic resonance provides a quantitative description of protein conformational flexibility on physiologically important time scales. *Biochemistry* **50,** 2735–2747 CrossRef Medline
57. Torchia, D. A. (2011) Dynamics of biomolecules from picoseconds to seconds at atomic resolution. *J. Magn. Reson.* **212,** 1–10 CrossRef Medline
58. Heckman, K. L., and Pease, L. R. (2007) Gene splicing and mutagenesis by PCR-driven overlap extension. *Nat. Protoc.* **2,** 924–932 CrossRef Medline
59. Hennecke, J., Sebbel, P., and Glockshuber, R. (1999) Random circular permutation of DsbA reveals segments that are essential for protein folding and stability. *J. Mol. Biol.* **286,** 1197–1215 CrossRef Medline
60. Meerman, H. J., and Georgiou, G. (1994) Construction and characterization of a set of *E. coli* strains deficient in all known loci affecting the proteolytic stability of secreted recombinant proteins. *Biotechnology* **12,** 1107–1110 CrossRef Medline
61. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd Ed., Appendix A.3, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
62. Oscarson, S., and Tidén, A. K. (1993) Syntheses of the octyl and tetradecyl glycosides of 3,6-di-*O*-α-D-mannopyranosyl-α-D-mannopyranose and of 3,4-di-*O*-α-D-mannopyranosyl-α-D-mannopyranose. A new way for 2,4-di-*O*-protection of mannopyranosides. *Carbohydr. Res.* **247,** 323–328 CrossRef Medline
63. Beychok, S. (1966) Circular dichroism of biological macromolecules. *Science* **154,** 1288–1299 CrossRef Medline
64. Greenfield, N. J. (2006) Using circular dichroism spectra to estimate protein secondary structure. *Nat. Protoc.* **1,** 2876–2890 Medline
65. Cooper, A. (2011) *Biophysical Chemistry*, RSC Publishing, Cambridge, UK
66. Wang, Z. X. (1995) An exact mathematical expression for describing competitive binding of two different ligands to a protein molecule. *FEBS Lett.* **360,** 111–114 CrossRef Medline
67. Vranken, W. F., Boucher, W., Stevens, T. J., Fogh, R. H., Pajon, A., Llinas, M., Ulrich, E. L., Markley, J. L., Ionides, J., and Laue, E. D. (2005) The CCPN data model for NMR spectroscopy: development of a software pipeline. *Proteins* **59,** 687–696 CrossRef Medline
68. Pellecchia, M., Sebbel, P., Hermanns, U., Wüthrich, K., and Glockshuber, R. (1999) Pilus chaperone FimC-adhesin FimH interactions mapped by TROSY-NMR. *Nat. Struct. Biol.* **6,** 336–339 CrossRef Medline
69. Houtman, J. C., Brown, P. H., Bowden, B., Yamaguchi, H., Appella, E., Samelson, L. E., and Schuck, P. (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,** 30–42 CrossRef Medline
70. Sager, C. P., Fiege, B., Zihlmann, P. P., Vannam, R., Rabbani, S., Jakob, R. P., Preston, R. C., Zalewski, A., Maier, T., Peczuh, M. W., and Ernst, B. (2018) The price of flexibility–a case study on septanoses as pyranose mimetics. *Chem. Sci.* CrossRef
71. Rubinson, K. A., Ladner, J. E., Tordova, M., and Gilliland, G. L. (2000) Cryosalts: suppression of ice formation in macromolecular crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **56,** 996–1001 CrossRef Medline
72. Kabsch, W. (2010) Integration, scaling, space-group assignment and post-refinement. *Acta Crystallogr. D Biol. Crystallogr.* **66,** 133–144 CrossRef Medline
73. McCoy, A. J. (2007) Solving structures of protein complexes by molecular replacement with Phaser. *Acta Crystallogr. D Biol. Crystallogr.* **63,** 32–41 CrossRef Medline
74. Emsley, P., and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60,** 2126–2132 CrossRef Medline
75. Adams, P. D., Afonine, P. V., Bunkóczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L. W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., *et al.* (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* **66,** 213–221 CrossRef Medline
76. Chen, V. B., Arendall, W. B., 3rd, Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S., and Richardson, D. C. (2010) MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **66,** 12–21 CrossRef Medline

Supporting Information

Conformational Switch of the Bacterial Adhesin FimH in the Absence of the Regulatory Domain - Engineering a Minimalistic Allosteric System

Said Rabbani,[1] Brigitte Fiege,[1] Deniz Eris,[1] Marleen Silbermann,[1] Roman Peter Jakob,[2] Giulio Navarra,[1] Timm Maier,[2] and Beat Ernst[1*]

Supporting information for this article can be found under:

http://www.jbc.org/content/suppl/2017/11/27/M117.802942.DC1/jbc.M117.802942-1.pdf

Chemistry, 2017, *23*(48), 11570-11577
doi: 10.1002/chem.201701601

Publication 4

Target-directed Dynamic Combinatorial Chemistry: A Study on Potentials and Pitfalls as Exemplified on a Bacterial Target

Priska Frei,[1)] Lijuan Pang,[1)] Marleen Silbermann, Deniz Eris,[1)] Xiaohua Jiang,[1)] Tobias Mühlethaler,[1)] Tobias Mühlethaler,[1)] Oliver Schwardt,[1)] and Beat Ernst[1)*]

[1)] Institute of Molecular Pharmacy, University of Basel, Klingelbergstr. 50, 4056 Basel, Switzerland

[*]Corresponding author
Tel.: 0041 61 267 15 51; Fax: 0041 61 207 15 52
E-mail: beat.ernst@unibas.ch

Contributions of Marleen Silbermann

- SPR experiments and data evaluation/interpretation in collaboration with Priska Frei
- Protein expression and purification of $FimH_{FL-B}$
- Proofreading of the manuscript

Medicinal Chemistry |Hot Paper|

Target-directed Dynamic Combinatorial Chemistry: A Study on Potentials and Pitfalls as Exemplified on a Bacterial Target

Priska Frei, Lijuan Pang, Marleen Silbermann, Deniz Eriş, Tobias Mühlethaler, Oliver Schwardt, and Beat Ernst*[a]

Abstract: Target-directed dynamic combinatorial chemistry (DCC) is an emerging technique for the efficient identification of inhibitors of pharmacologically relevant targets. In this contribution, we present an application for a bacterial target, the lectin FimH, a crucial virulence factor of uropathogenic *E. coli* being the main cause of urinary tract infections. A small dynamic library of acylhydrazones was formed from aldehydes and hydrazides and equilibrated at neutral pH in presence of aniline as nucleophilic catalyst. The major success factors turned out to be an accordingly adjusted ratio of scaffolds and fragments, an adequate sample preparation prior to HPLC analysis, and the data processing. Only then did the ranking of the dynamic library constituents correlate well with affinity data. Furthermore, as a support of DCC applications especially to larger libraries, a new protocol for improved hit identification was established.

Introduction

Dynamic combinatorial chemistry (DCC) is a method for generating compound libraries by reversible reactions of scaffolds and fragments. These libraries are under thermodynamic control and their composition is a result of the thermodynamic stability of each member. Because of the continuous interconversion of their constituents, dynamic libraries are adaptive and external stimuli or templates can influence their composition.[1] Target-directed DCC was first performed two decades ago[2] and has since emerged as a valuable method for in situ synthesis and affinity ranking of ligands for a wide range of protein targets.[3] When the target protein binds and thus stabilizes selected members of a dynamic library, a shift of its composition occurs, that is, the target acts as a template that biases the library composition, leading to the amplification of ligands with high affinity at the expense of weaker ones (Figure 1).

Major drawbacks of target-directed DCC are the limited number of suitable reactions necessary for the assembly of dynamic libraries and the large consumption of the target protein. To establish the equilibrium in the presence and absence of the target protein, the reaction of the scaffold with the fragments should be reversible. On the other hand, for analytical purposes, this equilibrium must be frozen. Furthermore, to guarantee the stability of the target protein, near-physiological reaction conditions should be applied. Because high protein concentrations are generally necessary for target-directed DCC, protein supply might be a significant issue.[3c]

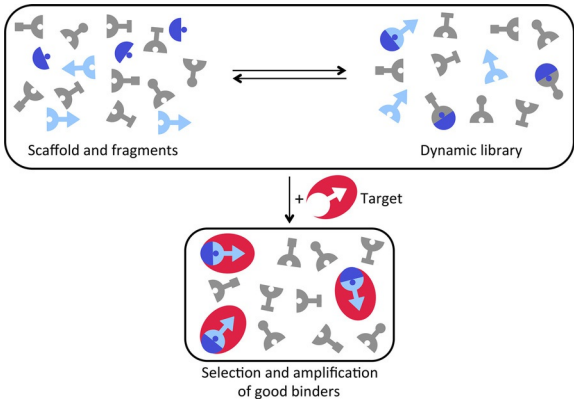


Figure 1. Schematic representation of target-directed DCC approach, indicating the different composition of a library at equilibrium in the absence and presence of a target protein.

Finally, an additional, but barely addressed challenge is the analysis of the library composition. So far, besides NMR,[4] MS,[5] and X-ray crystallography,[2a] analysis by high-performance liquid chromatography (HPLC) has been predominantly applied.[2b,c,5e,h,6] In an ideal setup, HPLC delivers a separate signal for each member of a library and therefore also information regarding its concentration. Generally, a comparison of HPLC traces of the same library equilibrated in the absence and presence of the target protein allows the identification of binders. Before HPLC analysis, bound ligands have to be released from the target protein, which then has to be removed to avoid affecting the result of the analysis.[1a] Standard techniques for protein removal include microfiltration, microdialysis, denaturation and aggregation with organic solvents, acids, bases, heat,

[a] P. Frei, L. Pang, M. Silbermann, D. Eriş, T. Mühlethaler, O. Schwardt, Prof. em. Dr. B. Ernst
Institute of Molecular Pharmacy
Pharmacenter, University of Basel
Klingelbergstrasse 50, 4056 Basel (Switzerland)
E-mail: beat.ernst@unibas.ch

Supporting information for this article can be found under: https://doi.org/10.1002/chem.201701601.

or chaotropic agents.[7] However, these manipulations can potentially alter the library composition, an effect that should be clarified in advance. Whereas some authors thoroughly describe sample preparation prior to HPLC analysis,[6c,k] in most reports detailed information is missing.

As target protein for our investigation, we used the bacterial adhesin FimH, which plays an important role in urinary tract infections, predominantly caused by uropathogenic *Escherichia coli* (UPEC).[8] The interaction of FimH, located at the distal tip of UPEC's type 1 pili, with the highly mannosylated glycoprotein uroplakin 1a on urothelial cells of the host is the first step of the infection.[9] Blocking FimH with aryl mannosides is a promising strategy to prevent the initial bacterial adhesion and therefore colonization of the bladder.[10] Since such FimH antagonists are not bactericidal and therefore do not exert any selection pressure, the probability of developing antimicrobial resistance is expected to be reduced.[11]

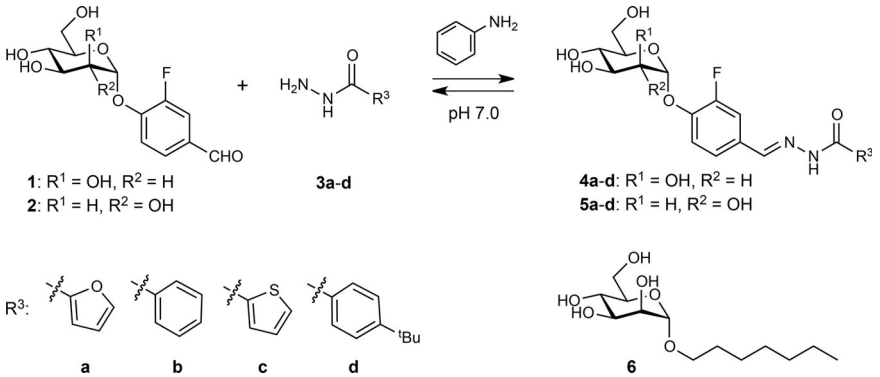
To explore possible solutions to the above mentioned challenges, we decided to use only small dynamic libraries for our proof-of-concept study, consisting of acylhydrazones formed from aldehyde scaffolds **1** and **2** and the commercial hydrazides **3a–d** (Scheme 1).[6h,i,12] Under acidic conditions, the equilibrium is rapidly established. However, when physiological conditions are applied, that is, neutral pH, the equilibrium is only slowly established. Therefore, nucleophilic catalysis with aniline was necessary, facilitating acylhydrazone formation with aldehydes via an intermediate Schiff base.[6h,12b] Finally, a pH increase to basic conditions leads to the required irreversibility and renders the library pseudostatic. To verify whether changes in library composition provoked by FimH lead to a selective amplification of those acylhydrazones exhibiting higher affinities, the K_D values of all library members were determined by surface plasmon resonance (SPR). Furthermore, since the equilibration could depend on deviating binding kinetics of the acylhydrazones, half-lives $t_{1/2}$ of the ligand–protein complexes were derived from SPR data. Finally, the rates of the chemical formation of acylhydrazones were examined to exclude that differences in reaction kinetics might influence equilibrium formation as well. Only when all members show comparable kinetic behavior, possible alterations by non-equilibrium states can be excluded.

Results and Discussion

Due to the mannose specificity of FimH[13] and previous studies with *O*-aryl α-D-mannosides,[10g] the mannose scaffold **1** is expected to yield high-affinity FimH antagonists upon reaction with the hydrazides **3a–d** (→**4a–d**). In contrast, the acylhydrazones **5a–d** derived from the glucose scaffold **2** should exhibit strongly diminished affinities and were therefore used as controls. As target protein, the full-length FimH lectin ($FimH_{FL}$)[14] was used. In each DCC experiment, the library was equilibrated in the absence (blank library) and presence of the target protein (template library) and analyzed by UV-HPLC at 310 nm.

For sample preparation, that is, for the release of bound ligands and the removal of the target protein prior to HPLC analysis, an adequate method is essential for a successful DCC experiment. One possibility is microfiltration, a technique, which was previously applied.[2b,6c,g,h,k] However, in our case, HPLC analysis after filtration with various commercially available filters revealed massive alterations in composition. Since these could originate from the adherence of constituents of the library to the filter material, addition of surfactant (Tween 20) as well as rinsing with EtOH were explored, but turned out to have been in vain. Further efforts to remove the protein, involving denaturation and precipitation by organic solvents (MeOH, EtOH, and MeCN at 33.3% v/v), urea (to a final concentration of 8 M) and NaOH (pH 12) did not lead to the desired separation of $FimH_{FL}$. Further, addition of MeOH even resulted in precipitation of non-protein constituents, most likely buffer salts. Attempts using heat were discarded, as the acylhydrazones were not stable at elevated temperatures.

Finally, the problem could be solved by using biotinylated full-length FimH ($FimH_{FL\text{-}B}$), which consists of two domains, the FimH lectin domain and the FimH pilin domain. In the assembled pilus, the N-terminal donor strand of the adjacent FimG subunit is essential for the stabilization of $FimH_{FL}$. For the generation of stable $FimH_{FL}$, this donor strand consisting of 15 amino acids has been used.[14] Incorporation of a donor strand equipped with a C-terminally linked biotin yielded $FimH_{FL\text{–}B}$. After the dynamic library experiment, the biotinylated protein was coupled to commercially available streptavidin–agarose beads and removed by centrifugation.



Scheme 1. Aldehydes **1** and **2** were used as scaffolds and the hydrazides **3a–d** as fragments for the generation of acylhydrazone libraries. To support dissociation of library members from the target protein FimH, *n*-heptyl α-D-mannopyranoside (**6**)[13] was applied in the work-up procedure (see Figures 2 and 5).

As depicted in Figure 2, in the first step of the optimized work-up protocol, the pH was raised to 8.5 to freeze the acylhydrazone composition (see Supporting Information). Then, most of the bound ligand was displaced from the target protein by adding a large excess of *n*-heptyl α-D-mannoside (**6**; Scheme 1), a known ligand of FimH.[13] Next, the protein was captured by addition of streptavidin–agarose. A further pH-increase to 12 guaranteed the complete displacement of bound library members from the target protein. Both acylhydrazones as well as the streptavidin–biotin binding were stable under these harsh basic conditions. Finally, $FimH_{FL-B}$ coupled to streptavidin–agarose beads could be removed quantitatively by centrifugation and the supernatant was subjected to UV-HPLC analysis. Absorption was measured at 310 nm, because at this wavelength acylhydrazones strongly absorb whereas unreacted scaffolds and fragments, aniline, and **6** do not. Blank and template libraries were treated equally to avoid any influence of the sample preparation falsifying the library composition.

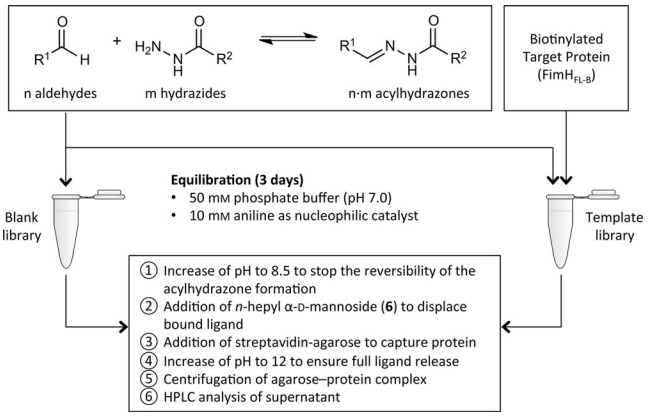


Figure 2. Experimental protocol for target-directed DCC using $FimH_{FL-B}$ and streptavidin-agarose. For increasing the pH, 1 M aqueous NaOH was added.

Mannose library

The efficiency of our workup protocol was tested with a library of four acylhydrazones generated from mannoside **1** and the hydrazides **3a–d** (→**4a–d**). Equimolar amounts (200 μM) of mannoside **1** and each fragment were equilibrated in the absence and presence of 100 μM $FimH_{FL-B}$ for three days. With longer equilibration times, no further change of the composition could be detected. After sample preparation as depicted in Figure 2, the analysis was performed by UV-HPLC. For peak assignment, scaffold **1** was reacted with each fragment individually in the presence of aniline, and the HPLC retention times of the formed acylhydrazones were determined as reference values (see Supporting Information). Because all acylhydrazones exhibit similar extinction coefficients (see Supporting Information), the library composition could be assessed with the relative peak area (RPA), that is, the sum of all peak areas was set to 100% and each peak was assigned its percentage (Figure 3A). For quantifying the influence of $FimH_{FL-B}$, the normalized change of RPA was used [Eq. (1); Figure 3B].

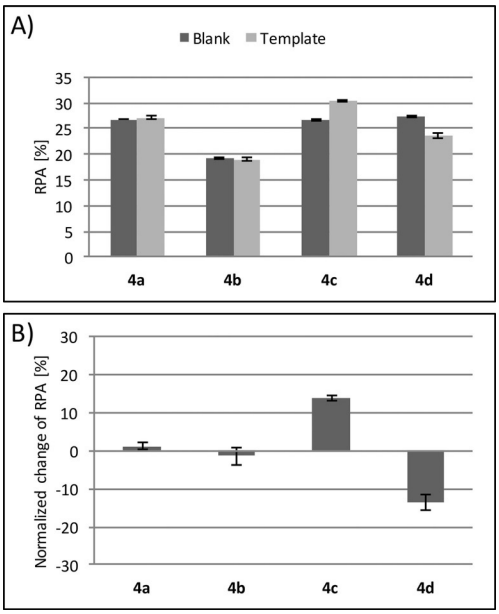


Figure 3. Result of the DCC-library with mannoside **1** (200 μM), hydrazides **3a–d** (each at 200 μM), and aniline (10 mM). A) Mean relative peak area (RPA) of each acylhydrazone of the library in the absence (blank) and presence (template) of $FimH_{FL-B}$ (100 μM). Error bars indicate standard deviation (SD) over three experiments. B) Normalized change of RPA, see Equation (1). Error bars indicate Gaussian error propagation (see Supporting Information).

$$\text{Normalized change of RPA} = \frac{RPA_{template} - RPA_{blank}}{RPA_{blank}} \quad (1)$$

A positive bar indicates amplification of an acylhydrazone in the template sample, whereas a negative bar indicates depletion. Next, the normalized changes of RPA of **4a–d** were compared to their K_D values determined with SPR (Table 1). Additionally, the affinity of scaffold **1** was determined to ensure that reaction with **3a–d** yields antagonists with higher affinity, so that potentially unreacted **1** does not interfere with target-binding of **4a–d**. For these SPR measurements, $FimH_{FL-B}$ was immobilized on a streptavidin coated sensor chip (for details see Supporting Information).

Table 1. Dissociation constants K_D and complex half-lives $t_{1/2}$ by SPR.

Compound	K_D [μM]	$t_{1/2}$ [s]
1	3.16	46.3
4a	0.52	2.57
4b	0.55	2.04
4c	0.33	3.51
4d	0.76	1.98
2	194	5.26
5a	491[a]	0.01
5b	735[a]	0.01
5c	405[a]	0.16
5d	152[a]	0.40
6	7.20	0.22

[a] K_D values are approximations, as the dissociations of **5a–d** were too fast to be accurately detected by SPR.

In the template library, the acylhydrazone **4c** with the highest affinity for the target (K_D of 0.33 μM) experienced the highest amplification, whereas the worst binder **4d** (K_D of 0.76 μM) was depleted. The concentrations of the acylhydrazones **4a** and **4b** with intermediate K_D values (0.52 and 0.55 μM, respectively) were neither amplified nor reduced in the presence of the target protein. To exclude that different kinetic behavior of the library members has an impact on the equilibration rate (for association (k_{on}) and dissociation rate constants (k_{off}) see Supporting Information), the half-lives ($t_{1/2} = \ln 2/k_{off}$) of the complexes formed by the antagonists and $FimH_{FL-B}$ were calculated (Table 1). Since **4a**–**d** show comparable half-lives, that is, between 1.98 and 3.51 s, an influence on the equilibrium formation can be excluded. Furthermore, **1** showed a 3- to 10-fold lower affinity than **4a**–**d** (Table 1). Thus, a considerable competition for the target's binding site by unreacted scaffold **1** can be ruled out. Theoretically, the Schiff base formed from **1** with aniline might also possess affinity for the target. However, since this intermediate is short-lived,[12b] an influence can be neglected. Additionally, when the aniline-catalyzed formation of the various acylhydrazones **4a**–**d** was monitored by UV-HPLC, similar chemical reaction rates were obtained (see Supporting Information).

Glucose library

A key question to be addressed is how the target protein affects the equilibrium concentrations of acylhydrazones with much lower FimH-affinities. Thus, a library formed from glucoside **2** and the hydrazides **3a**–**d** (→**5a**–**d**) was examined. Due to the mannose-specific binding of FimH, these acylhydrazones show a much lower affinity for the target (Table 1). Their K_D values are approximations, since the dissociations of **5a**–**d** are too fast to be accurately detected by SPR, as displayed by the short $t_{1/2}$. Overall, the observed much smaller differences between the blank and template library reflect the smaller competition among the acylhydrazones (Figure 4A).

Mixed libraries

When a combined library from equimolar amounts of the scaffolds **1** and **2** and the fragments **3a-d** (each at 200 μM) was generated, the analysis of the template library exhibited changes of the mannose containing components **4a**–**d** similar to the previous experiment. However, the glucose containing members of the library **5a**–**d** were, contradictory to their low affinities, not decreased but hardly affected (Figure 4B). This can be explained by the excess of the hydrazide fragments **3a**–**d** that renders a real competition between aldehyde **1** and **2** impossible. Although aldehyde **1** yields better binders (→ **4a**–**d**), hydrazides **3a**–**d** are still available to react with aldehyde **2**, leaving the concentrations of the glucosides **5a**–**d** unaffected. Consequently, depletion of a glucoside can only occur when one of its fragments is incorporated into a better binder. Following the same argumentation, **4d** was reduced because scaffold **1** was consumed by hydrazide **3c** to form the better inhibitor **4c**. It is important to note that depletion of a compound might not coincide with weakness of affinity, but could also be the result of a competition for the scaffolds.

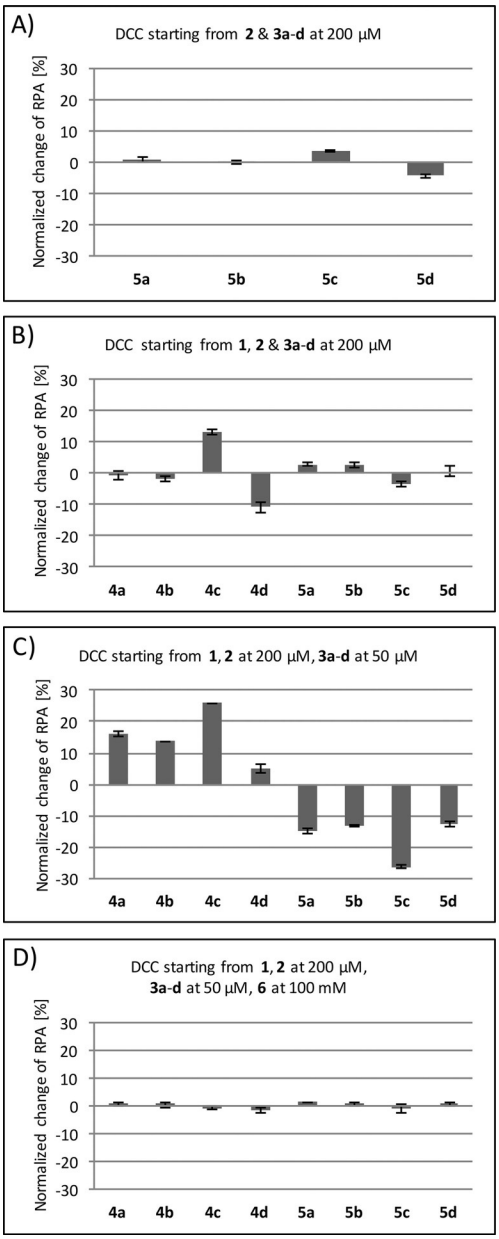


Figure 4. Results from libraries with different scaffold and fragment ratios; A) **2** and **3a**–**d** at 200 μM; B) **1**, **2**, and **3a**–**d** at 200 μM; C) **1** and **2** at 200 μM, **3a**–**d** at 50 μM; D) **1** and **2** at 200 μM, **3a**–**d** at 50 μM, **6** at 100 mM. All samples contained aniline (10 mM) and were incubated in the presence and absence of 100 μM $FimH_{FL-B}$. Bars indicate normalized change of relative peak area (RPA), see Equation (1), over three measurements. Error bars indicate Gaussian error propagation (see Supporting Information).

To further evaluate this hypothesis, we adjusted the composition of the library components to allow a real competition by keeping the concentrations of the aldehydes **1** and **2** at 200 μM and reducing the hydrazide concentrations to 50 μM. After equilibration, the analysis revealed an amplification of better binders **4a**–**d**, but a decrease of the low affinity glucosides **5a**–**d** (Figure 4C). The extent, to which the RPA of the mannosides **4a**–**d** increased, correlates well with their K_D

values (Table 1). This result clearly underlines that the ratio of scaffolds and fragments largely influences the outcome of a target-directed DCC experiment.

Finally, to clarify whether the composition of template libraries is resulting from specific interactions of the acylhydrazones with the binding site of $FimH_{FL-B}$ and is not biased by concurrent non-specific interactions, a control experiment was conducted. For this purpose, a library composed of the scaffolds **1** and **2**, each at 200 μM, the hydrazides **3a–d** at 50 μM, and a large excess, that is, 100 mM, of the known FimH antagonist **6**[13] (Table 1) was explored. When the heptyl mannoside (**6**) blocked the binding site of $FimH_{FL-B}$, its interaction with members of the acylhydrazone library is prevented and therefore the library composition remained unchanged (Figure 4D), a clear indication that the shift in composition of the template library is solely caused by the target's binding site and not distorted by non-specific interactions.

Capturing protocol

A target-directed DCC approach is only broadly applicable when there is no need for detailed information regarding the target's binding site. However, for the protocols discussed above, prior knowledge is required for adjusting scaffold and fragment concentrations.

Therefore, a broader applicable protocol for larger and more complex libraries was developed. Based on DCC approaches employing immobilized targets[15] or building blocks[16] and biotin-streptavidin supported separation techniques[17] a so-called capturing protocol (Figure 5) was explored. The basis for this approach is the fact that at pH 8.5 the acylhydrazone equilibrium is effectively frozen (Figure 5A), whereas the binding to $FimH_{FL-B}$ is only slightly influenced (see Supporting Information). Streptavidin–agarose beads added to the target library interact with the $FimH_{FL-B}$–ligand complex, allowing its separation from other library constituents by centrifugation (Figure 5B). After removal of the supernatant containing unbound ligand, the precipitate is suspended anew and the bound ligands are released by addition of *n*-heptyl mannoside (**6**). To complete the ligand release, the pH is increased to 12. After a second centrifugation step (Figure 5C), the supernatant containing previously bound ligand is analyzed. As before, blank and template libraries were treated equally.

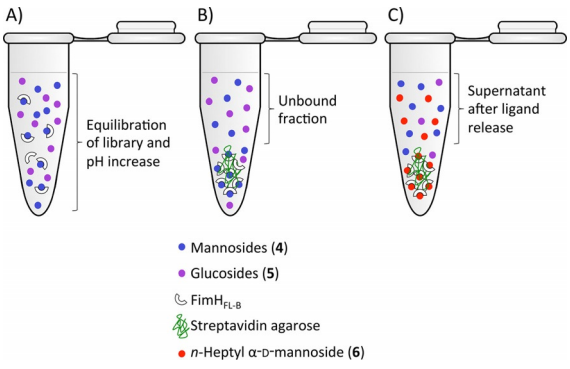


Figure 5. Capturing protocol for a template library consisting of acylhydrazones **4** and **5** and $FimH_{FL-B}$. A) Library after three days of equilibration and pH increase to 8.5; B) addition of streptavidin–agarose beads and centrifugation; and C) after removal of the unbound fraction, addition of a large excess of *n*-heptyl α-D-mannopyranoside (**6**), pH increase to 12 and centrifugation, the acylhydrazones released to the supernatant are analyzed by UV-HPLC.

To validate the new protocol, a library consisting of the scaffolds **1** and **2** and the hydrazides **3a–d** (each at 200 μM) was equilibrated. The analysis of the supernatant of the template library after ligand release, as compared to the blank library, showed a clear depletion of glucosides **5a–d** whereas mannosides **4a–d** were altered according to their K_D values (Figure 6A). Compared to the protocol missing the capturing step (Figure 4B), the distinction between mannosides and glucosides was vastly improved.

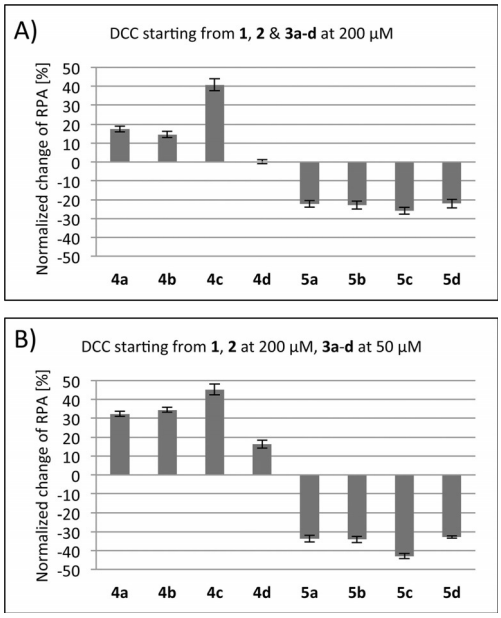


Figure 6. Results of the capturing experiments; Libraries were composed of A) **1** and **2** at 200 μM and **3a–d** at 200 μM and B) **1** and **2** at 200 μM and **3a–d** at 50 μM, respectively. Equilibration was catalyzed by aniline (10 mM) and carried out in presence and absence of $FimH_{FL-B}$ (100 μM). Bars indicate normalized change of relative peak area (RPA) over three measurements, see Equation (1). Error bars indicate Gaussian error propagation (see Supporting Information).

Additionally, a second library with reduced concentrations of hydrazides (now each at 50 μM) was evaluated according to the capturing protocol. Again, when comparing supernatants after ligand release, the template sample showed an increase of the mannosides **4a–d**, whereas the concentrations of the glucosides **5a–d** were decreased (Figure 6B).

A comparison of these two experiments reveals that the increments in amplification of the mannosides **4a–d** were less pronounced when restricting concentrations instead of an excess of hydrazides were used (Figure 6B and 6A, respectively). In the former case, the restricted amount of hydrazides limits the augmentation of the acylhydrazones' concentration. The best binder **4c**'s amplification is confined by **3c**, leaving more aldehyde **1** to form **4a**, **4b**, and **4d**. In the latter case, in

which **3c** is available in excess, a substantial increase of **4c** and, consequently, decrease of the other mannosides is observed, allowing a more distinct ranking. Given these points, the rank order of mannosides is not merely the result of the interaction with the target during the capturing step, but effectuated by their amplification during library equilibration. Therefore, the combination of a capturing step, where target-bound ligands are separated from the unbound fraction, and target-directed DCC synergistically leads to an improved hit identification.

Conclusion

Employing target-directed DCC, antagonists of the bacterial adhesin $FimH_{FL}$ could be identified and ranked according to their affinities. In the presence and absence of the target protein, dynamic libraries of acylhydrazones from reversibly reacting aldehydes as scaffolds and hydrazides as fragments were equilibrated and then analyzed by HPLC. Using the relative peak area (RPA), the composition of blank and template libraries could easily be compared and by calculating the normalized change of RPA, the influence of the target assessed.

It was particularly interesting that sample preparation before HPLC analysis turned out to be a crucial prerequisite for success, especially since the majority of DCC applications has not explicitly focused on this issue so far. The removal of the target protein by microfiltration and various denaturation methods failed because the composition of the library was altered or $FimH_{FL}$ could not be separated from the sample. Therefore, $FimH_{FL}$ equipped with a biotin tag ($FimH_{FL-B}$) was used, enabling the removal of the target protein with streptavidin-agarose beads without affecting the library composition. This protocol turned out to be highly sensitive. When applied to a mannose library, the ranking of members with K_D values differing by less than a factor of 2 could be achieved.

Because the glucosides **5a–d** exhibit much lower affinities than their mannose counterparts **4a–d**, $FimH_{FL-B}$ had a much smaller impact on their library composition. Furthermore, with a mixed library containing both the high-affinity mannosides and the low-affinity glucosides, amplification and decrease of individual members was heavily depending on the concentration ratio of scaffolds and fragments. Amplification of the better binders, that is, the mannosides **4a–d** at the expense of the worse binders, that is, the glucosides **5a–d** was only observed, when the carbohydrate scaffolds vied for limited amount of fragments. Lacking this competition, the mannosides were still ranked according to their affinity, but glucosides stayed unaffected by the target.

However, since this approach requires knowledge of the preferred binding motif, a broader applicable capturing protocol was developed. There, non-bound ligands are removed in a first step, followed by release and analysis of the library members bound to the target. This protocol allows a considerably improved distinction between tight and weak binders.

As previously noted,[1a] the establishment of suitable analytical protocols should be considered a major challenge of target-directed DCC, especially with increasing library sizes and complexities. The herein presented approach is a contribution to this challenge and a support for the design of future DCC experiments. Obviously, hits identified by a DCC approach need to be further optimized according to their physicochemical, pharmacokinetic as well as pharmacodynamics properties.

Experimental Section

Synthesis

A detailed description can be found in the Supporting Information.

Materials

UV-HPLC measurements were performed on an Agilent 1100/1200 system (Agilent). The synthetic non-biotinylated FimG donor strand (DsG, sequence ADVTITVNGKVVAKR) and biotinylated version (DsG^{biotin}) were purchased from JPT and had a >95% purity. To obtain DsG^{biotin}, the original peptide sequence of DsG[14,18] was amended on the C-term by a Ttds (trioxatridecan-succinamic acid) linker and a biotinylated lysine. Affinity values and binding constants were determined via SPR using a Biacore T200 system (GE Healthcare).

Protein production and purification

$FimH_{FL}$ and $FimH_{FL-B}$ were purified essentially as described previously.[14] Briefly, FimH and FimC were co-expressed in *E. coli* HM125[19] harboring the corresponding co-expression plasmid. Cells were grown at 30 °C in LB medium containing ampicillin (100 μg mL^{-1}). When OD_{600} reached 1.5, co-expression of FimH and FimC was induced by addition of isopropyl-ß-D-thiogalactoside (IPTG) to a final concentration of 1 mM. Cells were grown for 12 to 16 h after induction and harvested by centrifugation before resuspension in 50 mM Tris (tris(hydroxymethyl)aminomethane; pH 7.5), 150 mM NaCl, 5 mM EDTA (ethylenediaminetetraacetic acid) and 1 mg mL^{-1} polymyxin B sulfate (13 mL per liter of culture medium). The suspension was stirred for 1.5 h at 4 °C and centrifuged to collect the supernatant (periplasmic extract). All subsequent purification steps were performed at 4 °C. The periplasmic extract was dialyzed against 20 mM Tris (pH 8.0) and loaded onto an Uno Q (Bio-Rad) column equilibrated with the same buffer. The flow-through containing FimH·FimC complex was collected, dialyzed against 10 mM MOPS (3-(*N*-morpholino)propanesulfonic acid; pH 7.0) and loaded onto a Mono S column (GE Healthcare), which was equilibrated with the same buffer. FimH·FimC was eluted with a linear gradient from 0 to 300 mM NaCl. Fractions containing the complex were pooled and dialyzed against a buffer containing 20 mM NaH_2PO_4 (pH 7.4) and 50 mM NaCl. FimH·FimC complex was concentrated to 40 μM and incubated (48 h at 37 °C) with a 3-fold molar excess of the DsG for $FimH_{FL}$ and DsG^{biotin} for $FimH_{FL-B}$, respectively. The resulting reaction mixture contained $FimH_{FL}$ (FimH·DsG) or $FimH_{FL-B}$ (FimH·DsG^{biotin}), isolated FimC displaced from the FimH·FimC complex by DsG or DsG^{biotin}, potentially unreacted FimH·FimC complex, and excess DsG or DsG^{biotin}. This mixture was dialyzed against 20 mM acetic acid (pH 4.5) and loaded onto a Mono S column equilibrated with the same buffer. Full-length FimH was eluted with a linear gradient from 0 to 400 mM NaCl. Fractions containing either $FimH_{FL}$ or $FimH_{FL-B}$ were pooled, dialyzed against 50 mM NaH_2PO_4 (pH 7.0), analyzed by SDS–PAGE (sodium dodecyl sulfate–polyacrylamide gel electrophoresis) and ESI-mass spectrometry (calculated mass for $FimH_{FL}$: 30635.3 Da, measured mass: 30635.0 Da; calculated mass for $FimH_{FL-B}$:

31 291.2 Da, measured mass: 31 291.0 Da, Functional Genomic Center Zürich) and stored at −80 °C. The yields of purified $FimH_{FL}$ and $FimH_{FL-B}$ were between 1–2 mg per liter culture medium.

DCC experiments

In a typical experiment, stock solutions of aldehydes and hydrazides in DMSO (dimethylsulfoxide) were combined appropriately. From this combined stock, aliquots where taken and mixed with buffer (50 mM NaH_2PO_4, pH 7.0) for the blank libraries or $FimH_{FL-B}$ stock and buffer for the template libraries. Acylhydrazone formation was catalyzed by the addition of aniline (1 M stock in DMSO) to a final concentration of 10 mM. The final volume of each sample was set to 30 μL and the maximal DMSO concentration below 10%. The libraries were equilibrated on a shaker (400 rpm, 25 °C) for three days. Then, 0.6 μL of aq. NaOH (→pH 8.5) and 3 μL of 2 M aq. **6** were added and the mixture incubated for 30 min at RT. Afterwards, 50 μL of High Capacity streptavidin–agarose slurry (Thermo Scientific) were added. The mixture was incubated for 25 min at RT. Then, the pH was increased to 12 by addition of 2.4 μL aq. NaOH and the mixture incubated for 5 min before centrifugation (10 min, 4000 g, 25 °C). The supernatant was subjected to UV-HPLC analysis: Column: Waters Altlantis T3, 3 μm, 2.1×150 mm (Waters Corporation), A: H_2O + 0.01% TFA (trifluoroacetic acid); B: MeCN + 0.01% TFA. Detection: UV absorption at 310 nm. Gradient: 5% B→20% B (30 min)→60% B (54 min), flow rate: 0.3 mL min^{-1}, injection volume: 20 μL.

In the capturing protocol, the pH was increased to 8.5, 50 μL streptavidin–agarose slurry were added, and the mixture incubated for 25 min. After centrifugation, 50 μL of supernatant containing unbound ligand were removed. The sedimented streptavidin beads were suspended in 50 μL buffer (pH 8.5), and 3 μL of 2 M aq. **6** were added. After 30 min incubation, the pH was increased to 12, the mixture centrifuged, and the second supernatant analyzed analogously.

Surface plasmon resonance

$FimH_{FL-B}$ (100 nM) was pulse-wise delivered to the surface of a streptavidin sensor chip (SA) by an immobilization wizard (aimed immobilization level: 2.000 RU; flow rate: 5 μL min^{-1}). A reference surface with biotin-poly(ethylenglycol)amine (100 nM) was prepared (time: 60 s; flow rate: 10 μL min^{-1}) in order to remove unspecific binding events of the compounds with the streptavidin surface. Multi-cycle kinetics with two-fold increasing compound concentrations (see Supporting Information) were performed at 25 °C using HBS-EP (0.01 M HEPES (4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid) pH 7.4, 0.15 M NaCl, 3 mM EDTA (ethylenediaminetetraacetic acid), 0.005% surfactant P20, from GE Healthcare) as running buffer. Compounds were dissolved in running buffer and injected over the flow cells for 180 s at a flow rate of 30 μL min^{-1}. Thereafter, a dissociation phase followed, in which buffer flowed over the chip surface for 800 s with the same flow rate. Due to fast dissociation rates, chemical regeneration of the senor chip surface was not required. The Biacore T200 Control Software and the Biacore T200 Evaluation Software (both Version 3.0) were applied for data processing and kinetic evaluation. A 1:1 binding model was applied using double referenced sensorgrams with RI (bulk refraction index) set to 0.

Acknowledgements

Financial Support for Deniz Eriş by the Swiss National Foundation (SNF 31003A 144183) is kindly acknowledged.

Conflict of interest

The authors declare no conflict of interest.

Keywords: acylhydrazone · drug discovery · dynamic combinatorial chemistry · FimH antagonists · supramolecular chemistry

[1] a) S. Ladame, *Org. Biomol. Chem.* **2008**, *6*, 219–226; b) F. B. L. Cougnon, J. K. M. Sanders, *Acc. Chem. Res.* **2012**, *45*, 2211–2221; c) J.-M. Lehn, *Angew. Chem. Int. Ed.* **2015**, *54*, 3276–3289; *Angew. Chem.* **2015**, *127*, 3326–3340.

[2] a) B. A. Katz, J. Finer-Moore, R. Mortezaei, D. H. Rich, R. M. Stroud, *Biochemistry* **1995**, *34*, 8264–8280; b) I. Huc, J.-M. Lehn, *Proc. Natl. Acad. Sci. USA* **1997**, *94*, 2106–2110; c) S. Sakai, Y. Shigemasa, T. Sasaki, *Tetrahedron Lett.* **1997**, *38*, 8145–8148.

[3] a) O. Ramström, J.-M. Lehn, *Nat. Rev. Drug Discovery* **2002**, *1*, 26–36; b) M. Mondal, A. K. H. Hirsch, *Chem. Soc. Rev.* **2015**, *44*, 2455–2488; c) R. Huang, I. Leung, *Molecules* **2016**, *21*, 910.

[4] a) P. Vongvilai, M. Angelin, R. Larsson, O. Ramström, *Angew. Chem.* **2007**, *119*, 946–966; b) Y. Zhang, M. Angelin, R. Larsson, A. Albers, A. Simons, O. Ramström, *Chem. Commun.* **2010**, *46*, 8457–8459; c) R. Larsson, Z. Pei, O. Ramström, *Angew. Chem. Int. Ed.* **2004**, *43*, 3716–3718; *Angew. Chem.* **2004**, *116*, 3802–3804; d) M. Mondal, N. Radeva, H. Köster, A. Park, C. Potamitis, M. Zervou, G. Klebe, A. K. H. Hirsch, *Angew. Chem. Int. Ed.* **2014**, *53*, 3259–3263; *Angew. Chem.* **2014**, *126*, 3324–3328.

[5] a) Z. Yang, Z. Fang, W. He, Z. Wang, H. Gan, Q. Tian, K. Guo, *Bioorg. Med. Chem. Lett.* **2016**, *26*, 1671–1674; b) M. Sindelar, T. A. Lutz, M. Petrera, K. T. Wanner, *J. Med. Chem.* **2013**, *56*, 1323–1340; c) M. Demetriades, I. K. H. Leung, R. Chowdhury, M. C. Chan, M. A. McDonough, K. K. Yeoh, Y.-M. Tian, T. D. W. Claridge, P. J. Ratcliffe, E. C. Y. Woon, C. J. Schofield, *Angew. Chem. Int. Ed.* **2012**, *51*, 6672–6675; *Angew. Chem.* **2012**, *124*, 6776–6779; d) M. T. Cancilla, M. M. He, N. Viswanathan, R. L. Simmons, M. Taylor, A. D. Fung, K. Cao, D. A. Erlanson, *Bioorg. Med. Chem. Lett.* **2008**, *18*, 3978–3981; e) J. Leclaire, L. Vial, S. Otto, J. K. M. Sanders, *Chem. Commun.* **2005**, 1959–1961; f) T. Hotchkiss, H. B. Kramer, K. J. Doores, D. P. Gamblin, N. J. Oldham, B. G. Davis, *Chem. Commun.* **2005**, 4264–4266; g) S.-A. Poulsen, *J. Am. Soc. Mass Spectrom.* **2006**, *17*, 1074–1080; h) B. Danieli, A. Giardini, G. Lesma, D. Passarella, B. Peretto, A. Sacchetti, A. Silvani, G. Pratesi, F. Zunino, *J. Org. Chem.* **2006**, *71*, 2848–2853.

[6] a) O. Ramström, J.-M. Lehn, *ChemBioChem* **2000**, *1*, 41–48; b) M. Hochgürtel, R. Biesinger, H. Kroth, D. Piecha, M. W. Hofmann, S. Krause, O. Schaaf, C. Nicolau, A. V. Eliseev, *J. Med. Chem.* **2003**, *46*, 356–358; c) L. Milanesi, C. A. Hunter, S. E. Sedelnikova, J. P. Waltho, *Chem. Eur. J.* **2006**, *12*, 1081–1087; d) B. Shi, R. Stevenson, D. J. Campopiano, M. F. Greaney, *J. Am. Chem. Soc.* **2006**, *128*, 8459–8467; e) S. Zameo, B. Vauzeilles, J.-M. Beau, *Eur. J. Org. Chem.* **2006**, *24*, 5441–5444; f) A. Valade, D. Urban, J.-M. Beau, *J. Comb. Chem.* **2007**, *9*, 1–4; g) D. E. Scott, G. J. Dawes, M. Ando, C. Abell, A. Ciulli, *ChemBioChem* **2009**, *10*, 2772–2779; h) V. T. Bhat, A. M. Caniard, T. Luksch, R. Brenk, D. J. Campopiano, M. F. Greaney, *Nat. Chem.* **2010**, *2*, 490–497; i) A. J. Clipson, V. T. Bhat, I. McNae, A. M. Caniard, D. J. Campopiano, M. F. Greaney, *Chem. Eur. J.* **2012**, *18*, 10562–10570; j) C. Saiz, V. Castillo, P. Fontán, M. Bonilla, G. Salinas, A. Rodríguez-Haralambides, S. G. Mahler, *Mol. Diversity* **2014**, *18*, 1–12; k) R. J. Lins, S. L. Flitsch, N. J. Turner, E. Irving, S. A. Brown, *Tetrahedron* **2004**, *60*, 771–780; l) M. Mondal, N. Radeva, H. Fanlo-Virgós, S. Otto, G. Klebe, A. K. H. Hirsch, *Angew. Chem.* **2016**, *55*, 9422–9426.

[7] a) M. C. McMaster, HPLC: A Practical User's Guide, 2nd ed., John Wiley & Sons Inc., Hoboken, New Jersey, **2007**, p. 144; b) J. Pawliszyn in Wilson & Wilson's Comprehensive Anal. Chem. Vol. 37 Sampling and Sample

Preparation for Field and Laboratory, 1st edition (Ed. D. Barceló), Eilseev Science B. V., Amsterdam, **2002**, pp. 800 – 801.
[8] B. Foxman, *Am. J. Med.* **2002**, *113*, 5 – 13.
[9] T. J. Wiles, R. R. Kulesus, M. A. Mulvey, *Exp. Mol. Pathol.* **2008**, *85*, 11 – 19.
[10] 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. Lüthi, M. Scharenberg, J. Bezençon, E. Gubler, L. Pang, M. Smieško, B. Cutting, O. Schwardt, B. Ernst, *J. Med. Chem.* **2010**, *53*, 8627 – 8641; c) C. K. Cusumano, J. S. Pinkner, Z. Han, S. E. Greene, B. A. Ford, J. R. Crowley, J. P. Henderson, J. W. Janetka, S. J. Hultgren, *Sci. Transl. Med.* **2011**, *3*, 109ra115; d) O. Schwardt, S. Rabbani, M. Hartmann, D. Abgottspon, M. Wittwer, S. Kleeb, A. Zalewski, M. Smieško, B. Cutting, B. Ernst, *Bioorg. Med. Chem.* **2011**, *19*, 6454 – 6473; e) 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; f) L. Pang, S. Kleeb, K. Lemme, S. Rabbani, M. Scharenberg, A. Zalewski, F. Schädler, O. Schwardt, B. Ernst, *ChemMedChem* **2012**, *7*, 1404 – 1422; g) 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, B. Ernst, *J. Med. Chem.* **2015**, *58*, 2221 – 2239.
[11] N. Sharon, *Biochim. Biophys. Acta Gen. Subj.* **2006**, *1760*, 527 – 537.
[12] a) P. Crisalli, E. T. Kool, *J. Org. Chem.* **2013**, *78*, 1184 – 1189; b) A. Dirksen, S. Dirksen, T. M. Hackeng, P. E. Dawson, *J. Am. Chem. Soc.* **2006**, *128*, 15602 – 15603.
[13] 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, H. De Greve, *Mol. Microbiol.* **2004**, *55*, 441 – 455.
[14] M. M. Sauer, R. P. Jakob, J. Eras, S. Baday, D. Eriş, G. Navarra, S. Berneche, B. Ernst, T. Maier, R. Glockshuber, *Nat. Commun.* **2016**, *7*, 10738.
[15] A. V. Eliseev, M. I. Nelen, *Chem. Eur. J.* **1998**, *4*, 825 – 834.
[16] B. R. McNaughton, B. L. Miller, *Org. Lett.* **2006**, *8*, 1803 – 1806.
[17] A. Bugaut, J.-J. Toulme, B. Rayner, *Org. Biomol. Chem.* **2006**, *4*, 4082 – 4088.
[18] M. Vetsch, C. Puorger, T. Spirig, U. Grauschopf, E. U. Weber-Ban, R. Glockshuber, *Nature* **2004**, *431*, 329 – 333.
[19] H. J. Meerman, G. Georgiou, *Nat. Biotechnol.* **1994**, *12*, 1107 – 1110.

Manuscript received: April 10, 2017
Accepted manuscript online: June 27, 2017
Version of record online: August 4, 2017

Arkivoc, 2019*,* part iv, 143-167
doi: 10.24820/ark.5550190.p010.913

Publication 5

Comparison of Affinity Ranking by Target-directed Dynamic Combinatorial Chemistry and Surface Plasmon Resonance

Priska Frei, Marleen Silbermann, Tobias Mühlethaler, Xiaohua Jiang,
Oliver Schwardt, Rachel Hevey, and Beat Ernst*

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

*Corresponding author
Tel.: 0041 61 267 15 51; Fax: 0041 61 207 15 52
E-mail: beat.ernst@unibas.ch

Contributions of Marleen Silbermann

- SPR experiments and data evaluation/interpretation in collaboration with Priska Frei
- Protein expression and purification of $FimH_{FL-B}$
- Proofreading of the manuscript

The Free Internet Journal
for Organic Chemistry

Paper

Archive for
Organic Chemistry

Arkivoc **2019**, part iv, 0-0
to be inserted by editorial office

Comparison of affinity ranking by target-directed dynamic combinatorial chemistry and surface plasmon resonance

Priska Frei, Marleen Silbermann, Tobias Mühlethaler, Xiaohua Jiang, Oliver Schwardt, Rachel Hevey, and Beat Ernst*

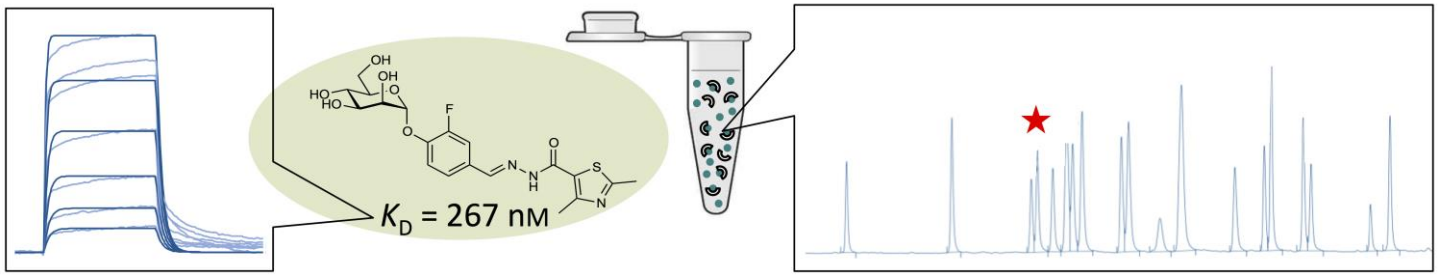
Department of Pharmaceutical Sciences, University of Basel, Klingelbergstr. 50, CH-4056 Basel, Switzerland
E-mail: Beat.Ernst@unibas.ch

Dedicated to Steven Hanessian, for his friendship and support over many years

Received 02-20-2019 **Accepted** 04-21-2019 **Published on line** 05-02-2019

Abstract

Target-directed dynamic combinatorial chemistry (tdDCC) is a powerful method to screen ligands for pharmacologically relevant targets. Generating a dynamic library from reversibly reacting building blocks in the presence of a target protein leads to the amplification of the most potent library constituents. In previous studies on tdDCC, these compounds were identified in a qualitative “hit/no-hit”-manner. However, the precise relationship between the degree of amplification and the affinity of the library constituent has not yet been evaluated. To study the amplification–affinity relationship, we compared tdDCC experiments, employing reversible acylhydrazone formation and the bacterial adhesin FimH as a target, with affinities of the library constituents as determined by surface plasmon resonance.



Keywords: Acylhydrazone, FimH antagonists, dynamic combinatorial chemistry, supramolecular chemistry, drug discovery

Introduction

Dynamic combinatorial chemistry (DCC) describes the generation of dynamic compound libraries from reversibly reacting building blocks. These libraries, which are under thermodynamic control, remain adaptive by continuous interconversion of building blocks and products. Therefore, addition of a protein target alters their equilibrium composition by binding, thereby stabilizing, and ultimately amplifying specific library constituents. Target-directed DCC (tdDCC) exhibits a self-screening ability, leading to the amplification of those members of the library with the highest affinity for the protein target, as depicted for an acylhydrazone library in Figure 1. This makes tdDCC a valuable tool for drug discovery.[1-6]

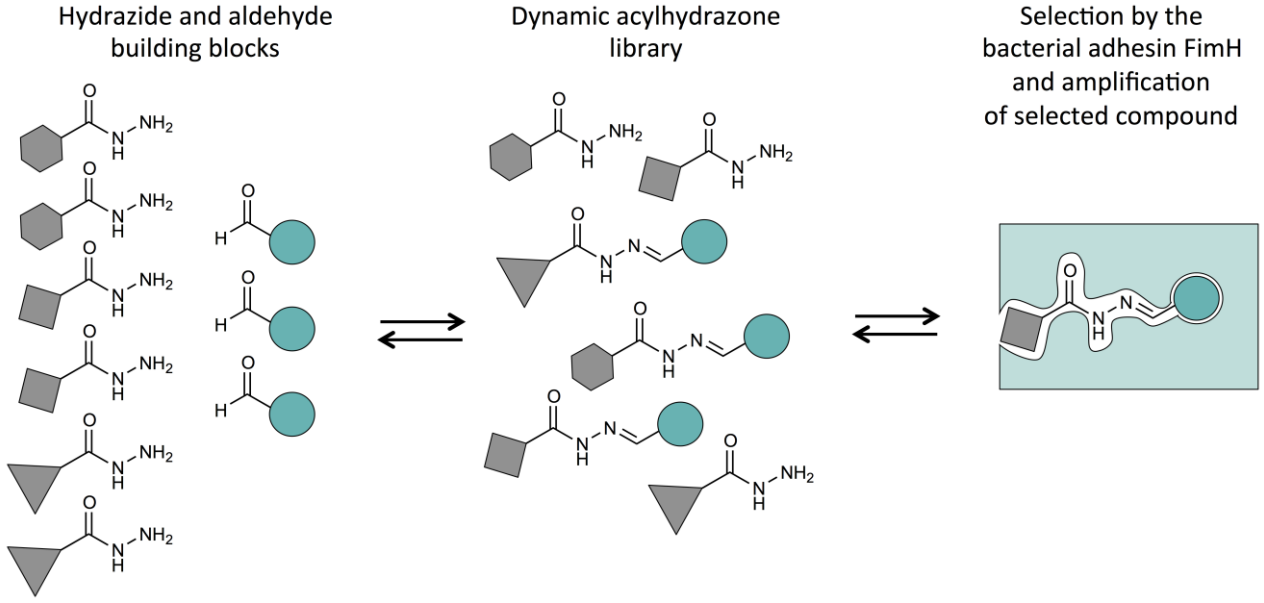


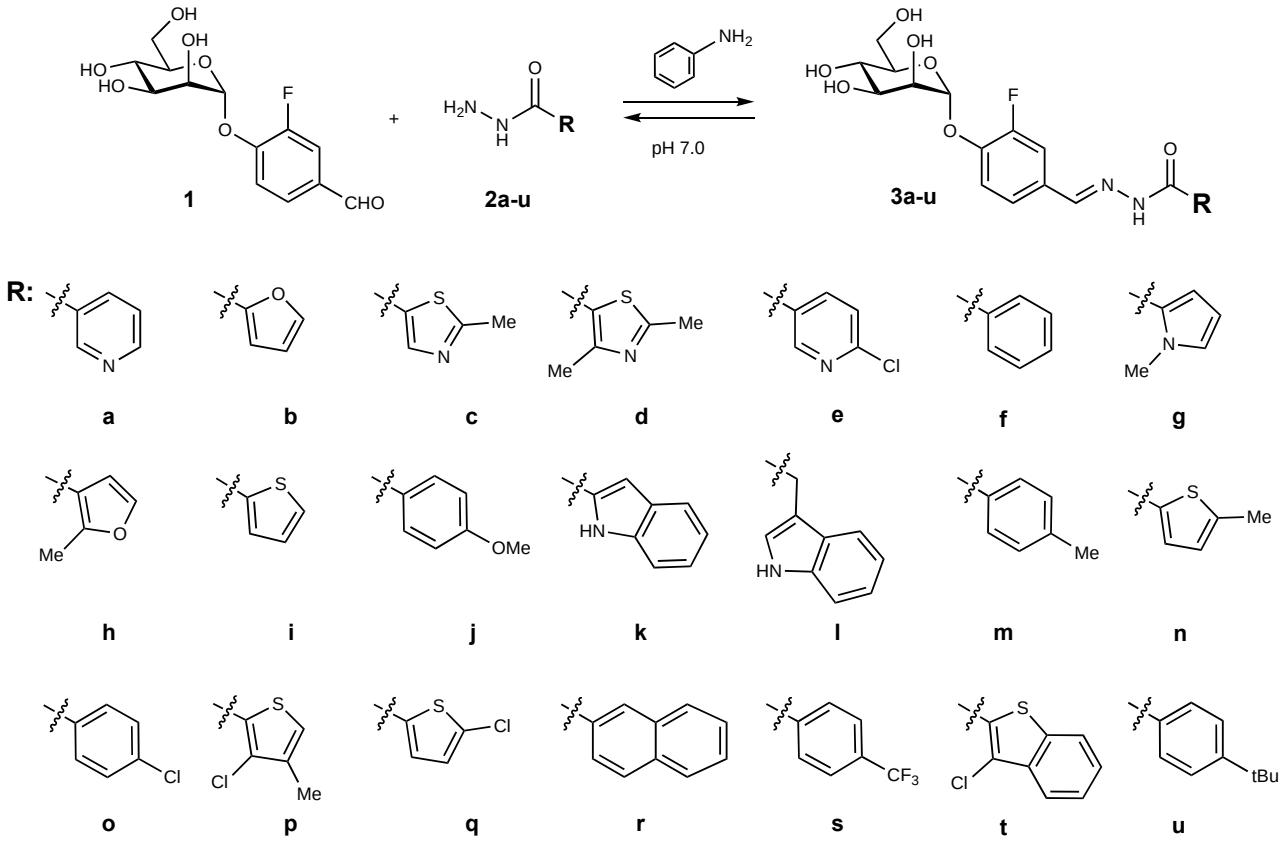
Figure 1. Schematic illustration of tdDCC. Reversibly reacting hydrazide and aldehyde building blocks generate a dynamic acylhydrazone library, which, when challenged with the target FimH, responds with a shift in its equilibrium composition including amplification of selected high-affinity ligands.

Whereas earlier reports focused on small libraries in a qualitative "hit/no-hit"-manner,[3-6] a more precise affinity ranking is required to increase the value of tdDCC. So far, the relationship between the extent of amplification and affinity was addressed in detail only by Nasr *et al*.[7] Furthermore, it has been noted that if several library constituents exhibit high affinity, the identification of the best binders might be difficult.[8] Finally, theoretical considerations regarding the relationship between host-ligand interactions and extent of amplification have been reported.[9]

In this communication, we have examined in greater detail the relationship between the amplification of library members in tdDCC experiments and their dissociation constants (K_D). As target protein we selected the bacterial adhesin FimH, located at the distal tip of type 1 pili of uropathogenic *E. coli* (UPEC) strains, which are the cause of the majority of urinary tract infections (UTI).[10-12] In the initial step of infection, FimH binds to the highly mannosylated surface protein uroplakin 1a on urothelial host cells,[13] a process which can be prevented with FimH antagonists such as aryl mannosides.[14-20]

We have reported previously that in addition to ligand affinity, also the ratio of scaffold to fragment building blocks, sample preparation, analysis, and method of data processing can be crucial factors in dictating experiment success.[21] Here, we extend the reported protocol for tdDCC using the bacterial adhesin FimH as

target protein to a larger scale. When aldehyde scaffold **1**[21] reacts reversibly with hydrazide fragments **2a**-**u** (→ **3a**-**u**; Scheme 1), the acylhydrazone library **3a**-**u** is obtained. At neutral pH in aqueous media, equilibrium formation is facilitated by aniline as a nucleophilic catalyst.[22]



Scheme 1. Aldehyde scaffold **1** and hydrazide fragments **2a**-**u** used for the reversible generation of acylhydrazones **3a**-**u** at neutral pH with aniline promotion.

Equilibration was carried out in the absence (blank library) or presence (template library) of biotinylated full-length FimH ($FimH_{FL-B}$)[23] for three days.[21] The reversibility of the acylhydrazone exchange reaction was then blocked by an adjustment in pH, which effectively locks in the library composition and renders the library suitable for analysis.[24] Prior to UV-HPLC analysis, the protein–ligand complex was captured using commercially available streptavidin agarose, and any unbound ligand removed from the sample (Figure 2A and 2B). The protein-bound ligand was then released by a further increase in pH together with addition of the FimH antagonist *n*-heptyl α-D-mannopyranoside[25] (**4**, Figure 2C). The supernatant was then analyzed with HPLC and the chromatograms of template libraries compared to those of equally treated blank samples.

On account of toxicological and stability concerns over the acylhydrazone moiety, we subsequently explored its potential for bioisosteric replacement.

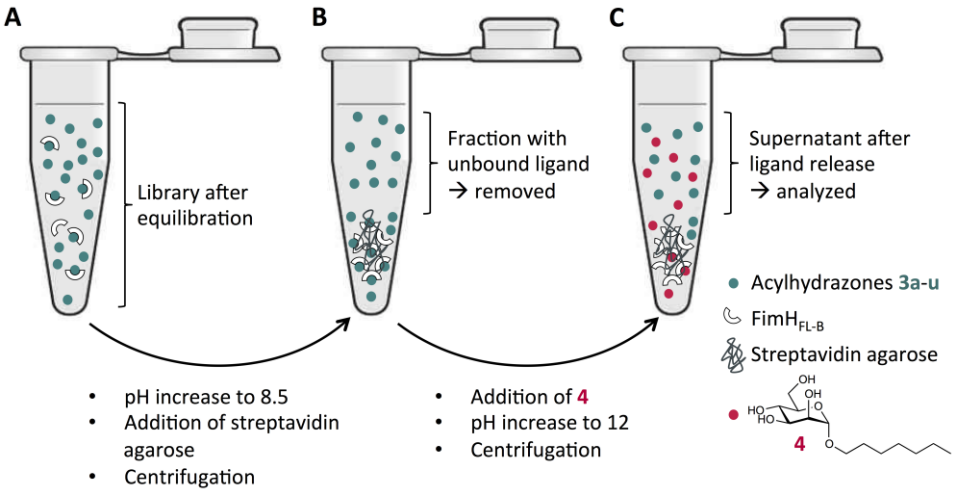


Figure 2. Schematic representation of the experimental tdDCC set-up.[21]

A. After sufficient equilibration, the library composition is made static through a pH increase (pH 8.5), and then the biotinylated target protein – including bound ligands – is captured with commercially available streptavidin agarose.
B. After centrifugation, the supernatant (containing unbound ligands) is discarded and the agarose-protein-ligand complex resuspended.
C. Bound ligands are released from the captured FimH into bulk solvent by the addition of excess of *n*-heptyl α-D-mannopyranoside (**4**) [25] and a further pH increase (pH 12). After a final centrifugation step, the supernatant is removed and analyzed by HPLC.

Results and Discussion

Target-directed dynamic combinatorial chemistry (tdDCC)

Because the outcomes of tdDCC experiments are influenced not only by the ratio of scaffolds to fragments,[21] but also by the ratio of library constituents to target protein,[26] we studied this issue in more detail by employing different amounts of building blocks, while keeping the protein concentration constant. Dynamic libraries were generated with varying concentrations of aldehyde scaffold **1** (600 μM, 400 μM, and 200 μM) together with hydrazide fragments **2a**-**u** each at 100 μM. All libraries were equilibrated in the presence and absence of 100 μM FimH$_{FL-B}$ (measured in triplicates). In this experimental design, the amount of scaffold **1** is determining the maximum attainable acylhydrazone concentration. Assuming full conversion of the aldehyde, facilitated by the addition of excess hydrazides, concentration ratios of 6:1, 4:1, and 2:1 between acylhydrazones and target protein should be reached. With a 1:1 ratio, accurate detection was not possible due to an insufficient amount of acylhydrazones, while unsatisfactory solubility became an issue at higher protein concentrations. The libraries were analyzed by UV-HPLC, detecting optical density at 310 nm. Conveniently, only acylhydrazones absorb UV light at this particular wavelength, but not other library constituents such as aniline, unreacted hydrazides, or *n*-heptyl α-D-mannopyranoside (**4**). Peaks in the resulting chromatograms were assigned using reference samples. Because both **3f** and **3g**, as well as **3m** and **3n**, gave overlapping signals which could not be resolved, we treated these signals as containing equal amounts of both constituents.

Composition of each library was determined based on the relative peak area (RPA), where the summated peak areas in each chromatogram were set to 100% and each peak assigned its fraction. Based on this information, the *normalized change of RPA*[21] between template and blank samples could be calculated and the influence of $FimH_{FL-B}$ on the library composition assessed (Figure 3). A positive value indicates amplification of a compound in the presence of $FimH_{FL-B}$, whereas a negative value indicates depletion.

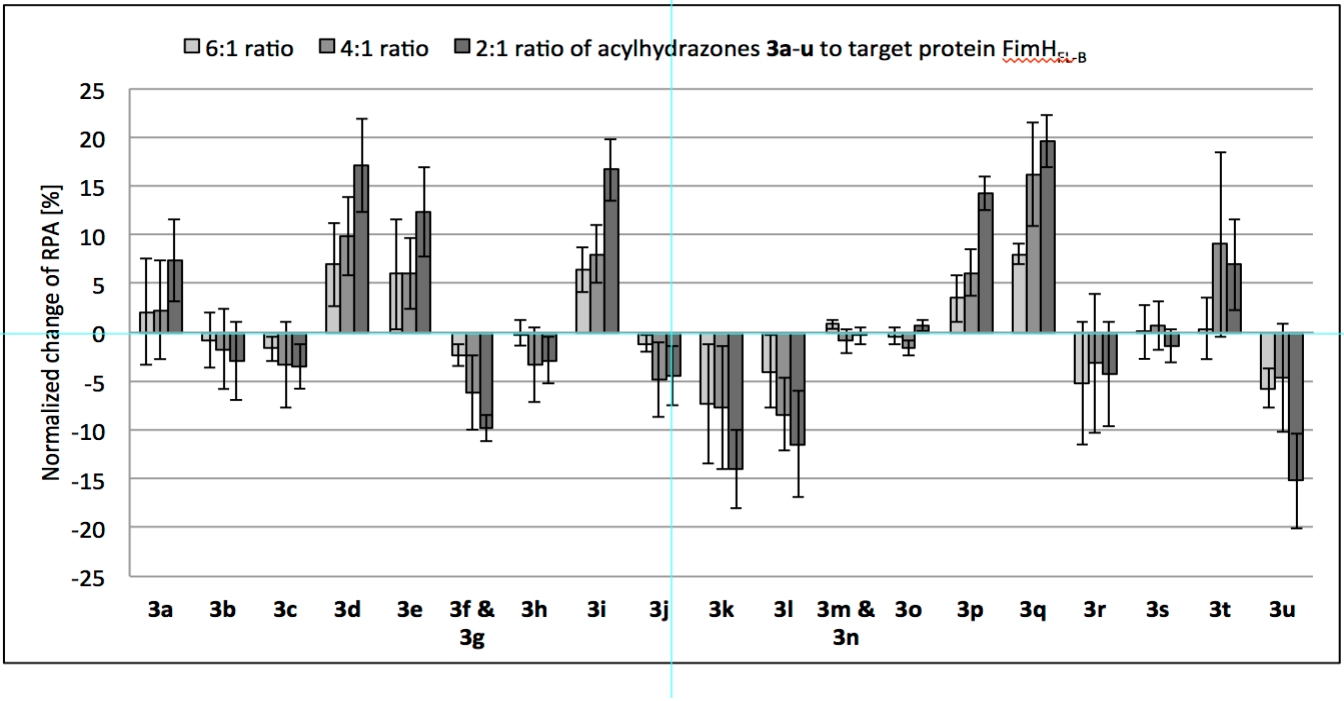


Figure 3. *Normalized change of RPA*[21] between template and blank libraries. Bars from light to dark grey indicate 6:1, 4:1, and 2:1 concentration ratios of total acylhydrazones to $FimH_{FL-B}$. Triplicates of all libraries were generated in the presence and absence of 100 μM $FimH_{FL-B}$. Error bars indicate error propagation of standard deviations of RPAs over three measurements.[21]

The most pronounced influence of $FimH_{FL-B}$ on the library composition was observed when core aldehyde **1** was present at 200 μM, resulting in a 2:1 ratio of acylhydrazones to target protein. When the scaffold was employed at higher concentrations (400 μM and 600 μM), changes in composition of the libraries were less pronounced. To evaluate if the *normalized changes of RPA* correlated with the affinities of the corresponding acylhydrazones, K_D values for **3a**-**u** were determined by surface plasmon resonance (SPR) using a previously established procedure.[21] Owing to the structural similarity of the compounds, differences in observed K_D values were not drastic (267 nM to 760 nM; Table 1). Obviously, this narrow distribution of affinities places highest demands on the applied analytical tools.

Table 1. Affinities measured by surface plasmon resonance.

Compound	K_D [nM]	Compound	K_D [nM]
1	3160[21]	**3k**	427
3a	359	**3l**	508
3b	520[21]	**3m**	484
3c	358	**3n**	390
3d	267	**3o**	440
3e	492	**3p**	377
3f	550[21]	**3q**	286
3g	461	**3r**	337
3h	642	**3s**	536
3i	330[21]	**3t**	376
3j	462	**3u**	760[21]

Results of tdDCC experiments are commonly reported in a "hit/no-hit"-manner. Amplification ("hit") and depletion ("no-hit") were only considered significant when the propagated error did not cross the baseline (Figure 3). Otherwise, compound concentrations were regarded as unchanged. Ligands **3d**, **3i**, **3p** and **3q** were amplified by the target in each of the three described experimental designs, whereas **3a** and **3t** were only amplified in the replicates with a 2:1 acylhydrazones to $FimH_{FL-B}$ ratio. The investigation of compounds with a negative *normalized change of RPA* revealed that signals **3l**, **3u**, and overlapping **3f** and **3g** were decreased, in accordance with lower affinities. For acylhydrazones with intermediate affinities, the combined signal of **3m** and **3n** stayed unchanged with 200 μM and 400 μM, but was slightly increased with 600 μM of core aldehyde **1**. Finally, **3j** (K_D of 462 nM) and **3o** (K_D of 440 nM) remained unchanged in all libraries.

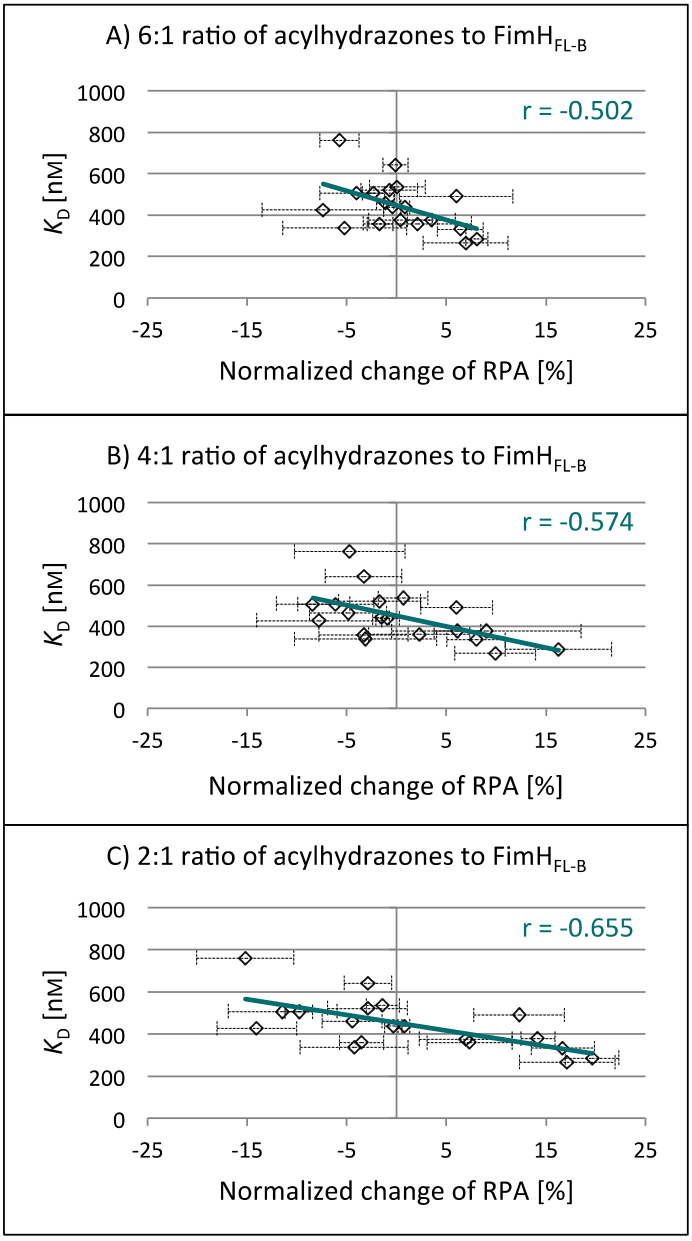


Figure 4. Quantitative relationship between the normalized change of relative peak area (RPA) and the dissociation constant (K_D) for different acylhydrazones. Pearson's correlation coefficient r is given as a measure of alignment.

Even though the majority of *normalized changes of RPA* correlated with the affinity data obtained by SPR, deviations were found for some acylhydrazones: albeit exhibiting rather high affinities of 358 nM and 337 nM, the *normalized changes of RPA* for **3c** and **3r** remained unchanged, while **3e** with a K_D of only 492 nM was amplified. Furthermore, **3b**, **3h**, and **3s** with rather low affinities of 523 nM, 642 nM, and 536 nM, respectively, remained unchanged, whereas **3k** with an affinity of 427 nM was decreased. Some of these aberrations could potentially result from the fast binding kinetics (both on- and off-rates) between acylhydrazones and FimH$_{FL-B}$, as was observed in the SPR experiments (see Supporting Information). Since the kinetic constants could not be uniquely determined in all cases, some of the reported K_D values may be erroneous. Apart from the fast binding kinetics[21] which may have impeded the SPR measurements, we currently have no explanation for these outliers. However, it is important to keep in mind that when different techniques for affinity

measurements are utilized, deviations resulting from some inherent errors of measurements are often unavoidable.

To assess the quantitative relationship between tdDCC and SPR results, the *normalized change of RPA* for each acylhydrazone was plotted against its K_D value. In the case of overlapping signals in the tdDCC experiment, the K_D value used was the average of the two acylhydrazones. Linear regression of the experimental data obtained using a 6:1 substrate ratio showed a moderate Pearson's correlation coefficient (r) of -0.502 (Figure 4A). For the 4:1 substrate ratio, the correlation coefficient slightly increased to -0.574 (Figure 4B), and for the 2:1 ratio, the highest correlation with an r-value of -0.655 was obtained (Figure 4C).

Overall, the tdDCC experiments delivered results comparable to SPR. Given the narrow range of K_D-values in the compound series, this clearly highlights the sensitivity of tdDCC. Furthermore, the tdDCC approach offers great economy of time: while SPR often requires independent synthesis and purification prior to affinity measurement, tdDCC combines the two steps into a single assay, thus clearly accelerating the process of hit identification. An additional benefit is that the described tdDCC protocol requires only standard laboratory equipment, while SPR requires an elaborate and costly instrument.

Bioisosteric replacement

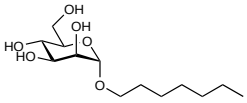
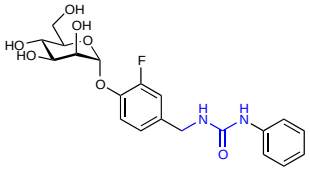
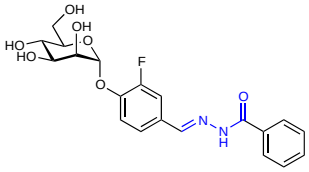
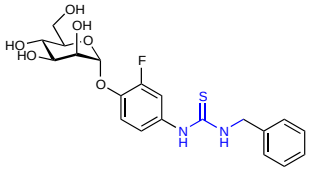
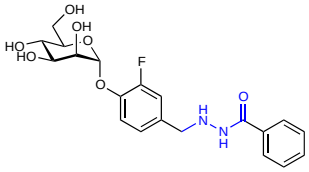
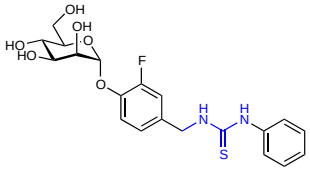
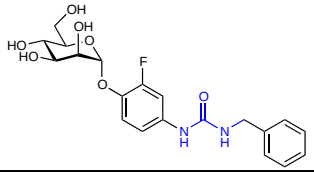
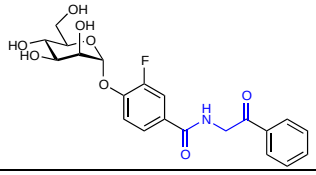
Besides their use in tdDCC,[21,24,27-31] several acylhydrazones have been reported to exhibit therapeutic properties[32] in areas such as cancer,[33-34] viral[35-36] and bacterial[37-38] infection, and pain and inflammation.[39] Furthermore, hydrazone linkages have been exploited for pH-responsive drug delivery.[40] However, both the cytotoxic activity inherently linked to anti-cancer drugs and the instability of the hydrazone moiety which affords its pH-responsiveness give rise to general concerns towards inclusion of the acylhydrazone group in potential FimH antagonists. Jumde *et al.* recently reported on bioisosteric replacement of the acylhydrazone moiety.[41] Therefore, in an effort to improve on stability and reduce toxicity, we generated a small library of bioisosteric analogues.

To explore possible bioisosteres of acylhydrazone, six alternatives to **3f** were synthesized (→ **5**-**10;** Table 2). Conveniently, reduction of **3f** with $NaBH_3CN$ yielded hydrazide **5**. Ureas **6** and **7** were synthesized from the corresponding anilines and amines, which were coupled via intermediates formed from 4-nitrophenyl chloroformate. Thioureas **8** and **9** were generated from the same aniline and amine starting materials through activation with 1,1'-thiocarbonyldiimidazole. Lastly, amide **10** was obtained by first assembling the aglycone from acid and amine precursors using standard peptide coupling, followed by mannosylation.

The affinities of compounds **5**-**10** were evaluated in a competitive fluorescence polarization assay (FPA),[14,17] using a non-biotinylated version of the FimH full-length protein ($FimH_{FL}$).[23] In type 1 pili of UPECs, the FimH subunit is stabilized by the N-terminal donor strand of the adjacent FimG subunit. Because isolated FimH turned out to be unstable, $FimH_{FL-B}$ and $FimH_{FL}$ required stabilization by a peptide consisting of the 15 terminal amino acids of FimG which mimics the donor strand. In the case of $FimH_{FL-B}$, biotin was linked to the pentadecapeptide, which does not alter $FimH_{FL-B}$'s binding properties as compared to $FimH_{FL}$. Hence, affinities determined with either of the constructs should be comparable. In the competitive FPA, the compounds under investigation displace a fluorescently labeled FimH antagonist (see **11** in the Supporting Information)[14] whose fluorescence polarization depends on target binding. By running a ligand dilution series, the dissociation constants could be determined and are summarized in Table 2. For **3f**, a K_D value of 515 nM was obtained, which is in excellent agreement with the affinity measured by SPR (550 nM; Table 1). All bioisosteres except for amide **10** exhibited a diminished affinity for $FimH_{FL}$. Noteworthy, only when the benzoyl moiety of the acylhydrazone was preserved as in hydrazide **5** and amide **10**, a strong loss of affinity could be avoided.

In summary, as evidenced by antagonist **10**, and as described by Jumde *et al.*[41] replacement by an amide provides a good starting point for further optimization of acylhydrazones.

Table 2. Dissociation constant K_D and relative K_D (rK_D) as determined with the fluorescece polarization assay.

Cpd.	Structure	K_D [nM]	rK_D	Cpd.	Structure	K_D [nM]	rK_D
4		3600 [17]	-	**7**		3498	6.79
3f		515.0	1.00	**8**		2828	5.49
5		783.6	1.52	**9**		1673	3.24
6		1097	2.13	**10**		447.5	0.87

Conclusions

In tdDCC experiments, the resulting composition is generally only qualitatively ranked in a "hit/no-hit"-manner. One goal of this communication was to explore whether a tdDCC ranking correlated to an affinity ranking that had been determined using surface plasmon resonance (SPR) experiments. We therefore established a 21-membered acylhydrazone library using aldehyde scaffold **1** and the commercially available hydrazide fragments **2a**-**u** (→ **3a**-**u**). TdDCC acylhydrazone libraries were generated both in the absence and presence of biotinylated target protein $FimH_{FL-B}$, and then analyzed using UV-HPLC. By calculating the *normalized changes of relative peak area (RPA)* between template and blank libraries, the influence of $FimH_{FL-B}$ could be assessed. Surprisingly, the library composition observed post-equilibration was influenced by the acylhydrazones/$FimH_{FL-B}$ ratio (6:1, 4:1, and 2:1), where differences between components became more enhanced with greater relative amounts of $FimH_{FL-B}$ (*i.e.* up- or down-regulation was more pronounced). When the tdDCC results were qualitatively ranked in a "hit/no-hit"-manner, the majority of amplified acylhydrazones indeed also exhibited high affinities in SPR experiments, whereas lower K_Ds correlated with down-regulated compounds. Next, when the *normalized changes of RPA* were plotted against K_D values, a linear correlation was observed. The best alignment was obtained from the libraries with a 2:1 ratio of acylhydrazone to protein target, but the correlation was diminished when libraries with the 4:1 and 6:1 ratios were evaluated. These results suggest that a stoichiometric ratio between library constituents and target protein would be ideal for the generation of libraries in which all members exhibit affinity. However, different ratios are conceivable for

libraries, which cover a wider range of affinities. In a situation where only a few good binders are present, their formation would more efficiently outcompete the others.

In subsequent efforts, replacement of the potentially hazardous acylhydrazone moiety with various bioisosteres was investigated. Whereas urea (→ **6** and **7**) and thiourea (→ **8** and **9**) analogues of parent compound **3f** exhibited decreased affinities, hydrazide derivative **5** retained affinity and amide analogue **10** mildly enhanced the affinity, indicating that the latter two bioisosteres could represent a good starting point for further optimization of acylhydrazone hits from tdDCC.

In summary, applying tdDCC to $FimH_{FL-B}$ using acylhydrazone libraries of structurally related mannosides successfully confirmed the high sensitivity of this approach. Most importantly, a linear association between the *normalized change of RPA* and the K_D values determined by SPR could be observed.

Experimental Section

General. Affinity values were determined using a Biacore T200 system (GE Healthcare). UV-HPLC measurements were made using an Agilent 1100/1200 system (Agilent). FPA was measured on a Synergy H1 hybrid multimode microplate reader (BioTek Instruments Inc., Winooski, VT, USA).
FPA,[14,17] protein production,[21] SPR measurements,[21] and tdDCC experiments[21] were conducted as previously described. UV-HPLC analysis of libraries: Column: Waters Altlantis T3, 3 µm, 2.1 x 150 mm (Waters Corporation), solvent A: H_2O + 0.01% TFA (trifluoroacetic acid); solvent B: MeCN + 0.01% TFA. Detection: UV absorption at 310 nm. Gradient: 5% → 25% B (35 min) → 50% B (65 min), flow rate: 0.3 mL/min, injection volume: 20 µL.

Synthesis

General. NMR spectra were recorded on a Bruker Avance DMX-500 (500.1 MHz) spectrometer. Assignment of ^1H NMR and ^{13}C NMR spectra was achieved using 2D methods (COSY, HSQC). Chemical shifts are expressed in ppm using residual $CHCl_3$, MeOH, or DMSO as references. Optical rotations were measured with a PerkinElmer Polarimeter 341 and infrared spectra were measured on a PerkinElmer Spectrum One FT-IR Spectrometer. Electrospray ionization mass spectrometry (ESI-MS) data were obtained on a Waters Micromass ZQ instrument. High resolution (HR)MS analysis were carried out using an Agilent 1100 LC equipped with a photodiode array detector and a Micromass QTOF I equipped with a 4 GHz digital-time converter. Reactions were monitored by thin layer chromatography (TLC) using glass plates coated with silica gel 60 F_{254} (Merck) and visualized by UV light and/or by charring with a molybdate solution (0.02 M solution of ammonium cerium sulfate dihydrate and ammonium molybdate tetrahydrate in aqueous 10% H_2SO_4). Medium pressure liquid chromatography (MPLC) separations were carried out on a CombiFlash R_f (Teledyne ISCO, Inc.) with RediSep disposable normal-phase or RP-18 (LiChroprepRP18, Merck) reversed-phase flash columns. Commercially available reagents were purchased from Fluka, Aldrich, Alfa Aesar, Fluorochem, and Apollo. Solvents were purchased from Sigma-Aldrich, Acros Organics, or VWR.

Synthesis of the acylhydrazone library

General procedure A for acylhydrazone formation. A flask was charged with a magnetic stirrer, then aldehyde **1**[21] and hydrazide **2a,c-e,g-h,j-t** were dissolved in H_2O/MeCN (2 mL, 7:3). AcOH (100 µL) was added and the mixture was stirred at r.t. for 5-22 h until only product was detected by MS. Then, the mixture was neutralized with 1 M aq. NaOH and the solvents were removed under reduced pressure. The residue was purified by MPLC

on RP-18 (H_2O/MeCN, 5:95 to 20:80) to give the desired products **3a,c-e,g-h,j-t**. For synthesis of **3b**, **3f**, **3i** and **3u**, please see reference 21.

Note: Compounds **3a**, **3d**, and **3l** were obtained as inseparable mixtures of *E*- and *Z*-isomers, with the *E*-isomer most likely representing the bigger fraction due to its sterically more favorable conformation. For further evaluation, we conducted a high temperature NMR measurement (CD_3OD, 60 °C) with compound **3l**, which clearly showed a decreased resolution, suggesting faster conversion of the two isomers. Further, HPLC traces at concentrations similar to the DCC experiments showed only one peak. When the reaction was catalyzed by aniline instead of AcOH, the same *E*/*Z*-ratio was obtained. Thus, the ratio of isomers is expected to be similar in tdDCC and SPR experiments.

***N'*-[3-Fluoro-4-(α-D-mannopyranosyloxy)benzylidene]nicotinohydrazide** (**3a**)**.** Prepared according to general procedure A from aldehyde **1** (10 mg, 33.1 µmol) and nicotinic hydrazide (**2a**, 4.5 mg, 33.1 µmol). Yield: 8.1 mg (58%) as an inseparable mixture of *E*- and *Z*-isomers (approx. 5:1). NMR data are given for the *E*-conformer. $[\alpha]_D^{20}$ +129.0 (*c* 0.35, MeOH); ^{1}H NMR (500 MHz, $(CD_3)_2SO$): δ = 3.39–3.54 (m, 4H, H-4, H-5, H-6a, H-6b), 3.60 (dd, *J* 3.8, 10.6 Hz, 1H), 3.69 (d, *J* 7.8 Hz, 1H, H-3), 3.88 (s, 1H, H-2), 4.47 (s, 1H, OH-4), 4.89 (m, 2H, OH-3, OH-6), 5.14 (s, 1H, OH-2), 5.50 (s, 1H, H-1), 7.45 (t, *J* 8.3 Hz, 1H, Ar-H), 7.50 (d, *J* 8.6 Hz, 1H, Ar-H), 7.52–7.58 (m, 1H, Ar-H), 7.61 (d, *J* 11.9 Hz, 1H, Ar-H), 8.25 (d, *J* 7.7 Hz, 1H, Ar-H), 8.38 (s, 1H, *H*C=N), 8.75 (d, *J* 4.3 Hz, 1H, Ar-H), 9.06 (s, 1H, Ar-H), 12.05 (s, 1H, NH); ^{13}C NMR (125 MHz, $(CD_3)_2SO$): δ 60.9 (C-6), 66.5 (C-4), 69.8 (C-2), 70.5 (C-3), 75.5 (C-5), 99.7 (C-1), 113.9 (d, $J_{C,F}$ 20 Hz), 118.6, 123.6, 124.3, 129.1 (d, $J_{C,F}$ 6 Hz), 129.4, 135.4, 145.4 (d, $J_{C,F}$ 11 Hz; 8C, Ar-C), 146.9 (d, $J_{C,F}$ 2 Hz, H*C*=N), 148.6, 152.1 (2C, Ar-C), 152.5 (d, $J_{C,F}$ 245 Hz, Ar-C), 161.8 (C=O); IR (KBr): ν 3429 (vs, OH, NH), 1652 (vs, C=N-NH-C=O) cm^{-1}; HRMS: *m/z:* Calcd for $C_{19}H_{20}FN_3NaO_7$ $[M+Na]^+$: 444.1183, found: 444.1181.

(*E*)-*N'*-[3-Fluoro-4-(α-D-mannopyranosyloxy)benzylidene]-2-methylthiazole-5-carbohydrazide (**3c**)**.** Prepared according to general procedure A from aldehyde **1** (10 mg, 33.1 µmol) and 2-methyl-thiazole-4-carboxylic acid hydrazide (**2c**, 5.2 mg, 33.1 µmol). Yield: 8.4 mg (57%). $[\alpha]_D^{20}$ +117.7 (*c* 0.24, MeOH); ^{1}H NMR (500 MHz, $(CD_3)_2SO$): δ 2.75 (s, 3H, CH_3), 3.41 (m, 1H, H-5), 3.44–3.54 (m, 2H, H-6a, H-4), 3.60 (dd, *J* 5.7, 11.2 Hz, 1H, H-6b), 3.88 (s, 1H, H-3), 3.69 (s, 1H, H-2), 4.46 (t, *J* 5.7 Hz, 1H, OH-6), 4.82 (d, *J* 5.4 Hz, 1H, OH-3), 4.87 (d, *J* 5.5 Hz, 1H, OH-4), 5.11 (d, *J* 3.3 Hz, 1H, OH-2), 5.49 (s, 1H, H-1), 7.38–7.49 (m, 2H, Ar-H), 7.55 (d, *J* 11.7 Hz, 1H, Ar-H), 8.29 (s, 1H, Ar-H), 8.51 (s, 1H, *H*C=N), 11.79 (s, 1H, NH); ^{13}C NMR (125 MHz, $(CD_3)_2SO$): δ 18.8 (CH_3), 60.9 (C-6), 66.5 (C-4), 69.8 (C-2), 70.6 (C-3), 75.5 (C-5), 99.7 (C-1), 113.7 (d, $J_{C,F}$ 20 Hz), 118.6, 124.4 (d, $J_{C,F}$ 2 Hz), 125.4, 129.3 (d, *J* 7 Hz), 145.3 (d, $J_{C,F}$ 11 Hz; 6C, Ar-C), 147.1 (H*C*=N), 148.3, 152.5 (d, $J_{C,F}$ 245 Hz), 156.9 (3C, Ar-C), 166.5 (C=O); IR (KBr): ν 3413 (vs, OH, NH), 1659 (s, C=N-NH-C=O) cm^{-1}; HRMS: *m/z:* Calcd for $C_{18}H_{20}FN_3NaO_7S$ $[M+Na]^+$: 464.0904, found: 464.0905.

***N'*-[3-Fluoro-4-(α-D-mannopyranosyloxy)benzylidene]-2,4-dimethylthiazole-5-carbohydrazide** (**3d**)**.** Prepared according to general procedure A from aldehyde **1** (10 mg, 33.1 µmol) and 2,4-dimethyl-thiazole-5-carboxylic acid hydrazide (**2d**, 5.7 mg, 33.1 µmol). Yield: 3.2 mg (21%) as a 5:1 mixture of *E*- and *Z*-conformer. NMR data are given for the *E*-conformer. $[\alpha]_D^{20}$ +110.5 (*c* 0.11, MeOH); ^{1}H NMR (500 MHz, $(CD_3)_2SO$): δ 2.51, 2.66 (2 s, 6H, 2 CH_3), 3.41 (d, *J* 7.8 Hz, 1H, H-5), 3.43–3.55 (m, 2H, H-4, H-6a), 3.60 (dd, *J* 5.3, 11.1 Hz, 1H, H-6b), 3.68 (d, *J* 8.4 Hz, 1H, H-3), 3.88 (s, 1H, H-2), 4.46 (t, *J* 5.7 Hz, 1H, OH-6), 4.85 (s, 1H, OH-3), 4.88 (d, *J* 12.0 Hz, 1H, OH-4), 5.13 (s, 1H, OH-2), 5.49 (s, 1H, H-1), 7.45 (m, 1H, Ar-H), 7.50 (d, *J* 7.2 Hz, 1H, Ar-H), 7.57 (d, *J* 11.7 Hz, 1H, Ar-H), 8.02 (s, 1H, *H*C=N), 11.72 (s, 1H, NH); ^{13}C NMR (125 MHz, $(CD_3)_2SO$):δ 18.5 (2C, 2 *C*H_3), 60.9 (C-6), 66.5 (C-4), 69.8 (C-2), 70.6 (C-3), 75.5 (C-5), 99.7 (C-1), 114.2 (d, $J_{C,F}$ 19 Hz), 118.8, 124.2 (d, $J_{C,F}$ 3 Hz), 128.9, 142.3, 145.2 (d, $J_{C,F}$ 11 Hz; 6C, Ar-C), 150.5 (H*C*=N), 152.4 (d, $J_{C,F}$ 245 Hz), 160.3, 162.2 (3C, Ar-C), 170.2 (C=O); IR (KBr): ν 3436 (OH, NH, vs), 1646 (s, C=N-NH-C=O) cm^{-1}; HRMS: *m/z:* Calcd for $C_{19}H_{22}FN_3NaO_7S$ $[M+Na]^+$: 478.1060, found: 478.1061.

(*E*)-6-Chloro-*N*'-[3-fluoro-4-(α-D-mannopyranosyloxy)benzylidene]nicotinohydrazide (**3e**)**.** Prepared according to general procedure A from aldehyde **1** (10 mg, 33.1 µmol) and 6-chloropyridine-3-carbohydrazide (**2e**, 5.7 mg, 33.1 µmol). Yield: 3.8 mg (25%). $[\alpha]_D^{20}$ +74.1 (*c* 0.30, MeOH); ^{1}H NMR (500 MHz, $(CD_3)_2SO$): δ 3.38–3.60 (m, 4H, H-6a, H-6b, H-5, H-4), 3.68 (d, *J* 9.1 Hz, 1H, H-3), 3.88 (s, 1H, H-2), 5.47 (s, 1H, H-1), 7.41 (m, 1H, Ar-H), 7.45 (d, *J* 8.6 Hz, 1H, Ar-H), 7.60 (s, 1H, Ar-H), 7.62 (d, *J* 3.0 Hz, 1H, Ar-H), 8.35 (d, *J* 8.3 Hz, 1H, *H*C=N), 8.37 (d, *J* 9.6 Hz, 1H, Ar-H), 8.93 (s, 1H, Ar-H); ^{13}C NMR (125 MHz, $(CD_3)_2SO$): δ 60.9 (C-6), 66.4 (C-4), 69.8 (C-2), 70.5 (C-3), 75.5 (C-5), 99.8 (C-1), 113.7 (d, $J_{C,F}$ 20 Hz), 118.50, 123.8, 124.1, 138.9 (8C, Ar-C), 147.0 (H*C*=N), 149.3, 151.9, 152.4 (d, $J_{C,F}$ 244 Hz; 3C, Ar-C), 174.0 (C=O); IR (KBr): ν 3436 (vs, OH, NH), 1634 (s, C=N-NH-C=O) cm^{-1}; HRMS: *m/z:* Calcd for $C_{19}H_{19}ClFN_3NaO_7$ $[M+Na]^+$: 478.0793, found: 478.0799.

(*E*)-*N*'-[3-Fluoro-4-(α-D-mannopyranosyloxy)benzylidene]-1-methyl-1*H*-pyrrole-2-carbohydrazide (**3g**)**.** Prepared according to general procedure A from aldehyde **1** (10 mg, 33.1 µmol) and 1-methyl-1*H*-pyrrole-2-carboxylic acid hydrazide (**2g**, 4.6 mg, 33.1 µmol). Yield: 9.5 mg (68%). $[\alpha]_D^{20}$ +132.5 (*c* 0.61, MeOH); ^{1}H NMR (500 MHz, CD_3OD): δ 3.65 (m, 1H, H-5), 3.69–3.81 (m, 3H, H-4, H-6a, H-6b), 3.93 (dd, *J* 3.1, 9.6 Hz, 1H, H-3), 3.96 (s, 3H, CH_3), 4.08 (s, 1H, H-2), 5.56 (s, 1H, H-1), 6.13 (s, 1H, Ar-H), 6.93 (s, 2H, Ar-H), 7.41 (t, *J* 8.1 Hz, 1H, Ar-H), 7.45 (d, *J* 8.6 Hz, 1H, Ar-H), 7.71 (d, *J* 11.9 Hz, 1H, Ar-H), 8.15 (s, 1H, *H*C=N); ^{13}C NMR (125 MHz, CD_3OD): δ 37.0 (CH_3), 62.6 (C-6), 68.2 (C-4), 71.8 (C-2), 72.3 (C-3), 76.0 (C-4), 101.2 (C-1), 108.7, 115.2 (d, $J_{C,F}$ 20 Hz), 119.5, 124.8, 125.7 (d, *J* 4 Hz), 130.7, 131.4 (d, $J_{C,F}$ 7 Hz; 7C, Ar-C), 146.7 (d, $J_{C,F}$ 6 Hz, H*C*=N), 147.0 (d, $J_{C,F}$ 11 Hz, Ar-C), 152.4 (C=O), 154.6 (d, $J_{C,F}$ 246 Hz, Ar-C); IR (KBr): ν 3436 (vs, OH, NH), 1634 (s, C=N-NH-C=O) cm^{-1}; HRMS: *m/z:* Calcd for $C_{19}H_{22}FN_3NaO_7$ $[M+Na]^+$: 446.1339, found: 446.1341.

(*E*)-*N*'-(3-Fluoro-4-(α-D-mannopyranosyloxy)benzylidene)-2-methylfuran-3-carbohydrazide (**3h**)**.** Prepared according to general procedure A from aldehyde **1** (10 mg, 33.1 µmol) and 2-methyl-furan-carboxylic acid hydrazide (**2h**, 4.6 mg, 33.1 µmol). Yield: 9.1 mg (65%). $[\alpha]_D^{20}$ +98.9 (*c* 0.56, MeOH); ^{1}H NMR (500 MHz, CD_3OD): δ 2.63 (s, 3H, CH_3), 3.67 (d, *J* 3.6 Hz, 1H, H-5), 3.72–3.83 (m, 3H, H-4, H-6a, H-6b), 3.95 (dd, *J* 2.8, 9.4 Hz, 1H, H-3), 4.10 (s, 1H, H-2), 5.59 (s, 1H, H-1), 6.85 (s, 1H, Ar-H), 7.42–7.52 (m, 3H, Ar-H), 7.76 (d, *J* 11.9 Hz, 1H, Ar-H), 8.22 (s, 1H, *H*C=N); ^{13}C NMR (125 MHz, CD_3OD): δ 13.7 (CH_3), 62.6 (C-6), 68.2 (C-4), 71.7 (C-2), 72.3 (C-3), 76.0 (C-5), 101.2 (C-1), 109.4, 115.0, 115.3 (d, $J_{C,F}$ = 20 Hz), 119.4, 126.0, 142.1, 147.3 (d, $J_{C,F}$ 11 Hz; 8C, Ar-C), 147.9 (d, $J_{C,F}$ 1 Hz, H*C*=N), 154.6 (d, $J_{C,F}$ 246 Hz), 160.0 (2C, Ar-C), 163.1 (C=O); IR (KBr): ν 3414 (vs, OH, NH), 1619 (vs, C=N-NH-C=O) cm^{-1}; HRMS: *m/z:* Calcd for $C_{19}H_{21}FN_2Na\,O_8$ $[M+Na]^+$: 447.1182, found: 447.1182.

(*E*)-*N*'-[3-Fluoro-4-(α-D-mannopyranosyloxy)benzylidene]-4-methoxybenzhydrazide (**3j**)**.** Prepared according to general procedure A from aldehyde **1** (10 mg, 33.1 µmol) and 4-methoxybenzhydrazide (**2j**, 5.5 mg, 33.1 µmol). Yield: 4.0 mg (27%). $[\alpha]_D^{20}$ +106.6 (*c* 0.26, MeOH); ^{1}H NMR (500 MHz, CD_3OD): δ 3.64 (m, 1H, H-5), 3.69–3.81 (m, 3H, H-4, H-6a, H-6b), 3.88 (s, 3H, CH_3), 3.93 (dd, *J* 3.0, 9.4 Hz, 1H, H-3), 4.08 (s, 1H, H-2), 5.57 (s, 1H, H-1), 7.05 (d, *J* 8.5 Hz, 2H, Ar-H), 7.43 (t, *J* 8.2 Hz, 1H, Ar-H), 7.49 (d, *J* 8.5 Hz, 1H, Ar-H), 7.77 (d, *J* 11.9 Hz, 1H, Ar-H), 7.92 (d, *J* 8.6 Hz, 2H, Ar-H), 8.25 (s, 1H, *H*C=N); ^{13}C NMR (125 MHz, CD_3OD): δ 56.0 (CH_3), 62.6 (C-6), 68.2 (C-4), 71.7 (C-2), 72.3 (C-3), 76.0 (C-5), 101.2 (C-1), 115.0 (2C), 115.4 (d, $J_{C,F}$ 20 Hz), 119.5, 125.96, 126.07 (d, $J_{C,F}$ 2 Hz; 6C, Ar-C), 130.7 (2C), 131.1 (d, $J_{C,F}$ 7 Hz), 147.3 (d, $J_{C,F}$ 10 Hz; 4C, Ar-C), 148.4 (H*C*=N), 154.6 (d, $J_{C,F}$ 246 Hz, Ar-C), 164.5 (C=O), 166.7 (Ar-C); IR (KBr): ν 3439 (vs, OH, NH), 1648 (s, C=N-NH-C=O) cm^{-1}; HRMS: *m/z:* Calcd for $C_{21}H_{23}FN_2NaO_8$ $[M+Na]^+$: 473.1336, found: 473.1336.

(*E*)-*N*'-[3-Fluoro-4-(α-D-mannopyranosyloxy)benzylidene]-1*H*-indole-2-carbohydrazide (**3k**)**.** Prepared according to general procedure A from aldehyde **1** (10 mg, 33.1 µmol) and 1*H*-indole-2-carboxylic acid hydrazide (**2k**, 5.8 mg, 33.1 µmol). Yield: 7.7 mg (51%). $[\alpha]_D^{20}$ +100.9 (*c* 0.47,MeOH); ^{1}H NMR (500 MHz, CD_3OD): δ 3.65 (m, 1H, H-5), 3.70–3.82 (m, 3H, H-4, H-6a, H-6b), 3.93 (dd, *J* 3.0, 9.4 Hz, 1H, H-3), 4.09 (s, 1H, H-2), 5.56 (s, 1H, H-1), 7.15–7.27 (m, 2H, Ar-H), 7.38–7.50 (m, 3H, Ar-H), 7.73 (s, 1H, Ar-H), 8.06 (d, *J* 6.4 Hz, 1H,

Ar-H), 8.17 (s, 1H, *H*C=N), 8.22 (d, *J* 7.1 Hz, 1H, Ar-H); ^{13}C NMR (125 MHz, CD_3OD): δ 62.6 (C-6), 68.2 (C-4), 71.8 (C-2), 72.3 (C-3), 76.0 (C-5), 101.2 (C-1), 112.8, 115.2 (d, $J_{C,F}$ 19 Hz), 119.5, 122.25, 122.31, 123.9, 125.7, 127.6, 127.7, 129.6, 131.5 (d, $J_{C,F}$ 7 Hz), 138.0 (12C, Ar-C), 146.2 (H*C*=N), 147.0 (d, $J_{C,F}$ 11 Hz), 154.6 (d, $J_{C,F}$ 246 Hz; 2C, Ar-C), 157.3 (C=O); IR (KBr): ν 3415 (vs, OH, NH), 1619 (s, C=N-NH-C=O) cm^{-1}; HRMS: *m/z:* Calcd for $C_{22}H_{22}FN_3NaO_7$ $[M+Na]^+$: 482.1339, found: 482.1340.

***N*'-[3-Fluoro-4-(α-D-mannopyranosyloxy)benzylidene]-2-(1*H*-indol-2-yl)acetohydrazide (3l).** Prepared according to general procedure A from aldehyde **1** (10 mg, 33.1 µmol) and indole-3-acetic hydrazide (**2l**, 6.3 mg, 33.1 µmol). Yield: 12.8 mg (82%) as a 2:1 mixture of *E*- and *Z*-conformation. $[\alpha]_D^{20}$ +99.1 (*c* 0.39, MeOH); ^{1}H NMR (500 MHz, CD_3OD; *E*:*Z* = 2:1, normalized to *E* conformation): δ 3.61–3.68 (m, 1.5H, H-5 *E*+*Z*), 3.70–3.81 (m, 6.5H, CH_2 *E*, H-6a *E*+*Z*, H-6b *E*+*Z*, H-4 *E*+*Z*), 3.91–3.96 (m, 1.5H, H-3 *E*+*Z*), 4.08 (dd, *J* 1.9, 3.3 Hz, 1H, H-2 *E*), 4.10 (dd, *J* 1.9, 3.4 Hz, 0.5H, H-3 *Z*), 4.20 (s, 1H, CH_2 *Z*), 5.57 (d, *J* 1.6 Hz, 1H, H-1 *E*), 5.58 (d, *J* 1.6 Hz, 0.5H, H-1 *Z*), 6.99–7.07 (m, 2H, Ar-H *E*+*Z*), 7.08–7.15 (m, 1.5H, Ar-H *E*+*Z*), 7.21 (s, 0.5H, Ar-H *Z*), 7.25 (s, 1H, Ar-H *E*), 7.47–7.33 (m, 4.5H, 3 Ar-H *E*, 3 Ar-H *Z*), 7.57 (d, *J* 12.6 Hz, 0.5H, Ar-H *Z*), 7.61 (d, *J* 7.9 Hz, 1H, Ar-H *E*), 7.65 (d, *J* 8.0 Hz, 0.5H, Ar-H *Z*), 7.70 (dd, *J* 1.6, 12.0 Hz, 1H, Ar-H *E*), 7.91 (s, 0.5H, CH=N *Z*), 8.06 (s, 1H, CH=N *E*); ^{13}C NMR (125 MHz, CD_3OD; only *E* conformer): δ 32.7 (CH_2), 62.6 (C-6), 68.2 (C-4), 71.8 (C-2), 72.3 (C-3), 76.0 (C-5), 101.2 (C-1), 126.1, 124.9, 122.6, 122.4, 120.0, 119.4, 115.5, 112.4, 101.4, 148.1 (16 C, 14 Ar-C, H*C*=N, C=O); IR (KBr): ν 3429 (vs, OH, NH), 1651 (s, C=N-NH-C=O) cm^{-1}; HRMS: *m/z:* Calcd for $C_{23}H_{24}FN_3NaO_7$ $[M+Na]^+$: 496.1496 found: 496.1496.

(*E*)-*N*'-[3-Fluoro-4-(α-D-mannopyranosyloxy)benzylidene]-4-methylbenzohydrazide (3m). Prepared according to general procedure A from aldehyde **1** (10 mg, 33.1 µmol) and *p*-toluic hydrazide (**2m**, 5.0 mg, 33.1 µmol). Yield: 7.7 mg (54%). $[\alpha]_D^{20}$ +104.2 (*c* 0.47, MeOH); ^{1}H NMR (500 MHz, CD_3OD): δ 2.42 (s, 3H, CH_3), 3.65 (d, *J* 5.4 Hz, 1H, H-5), 3.70–3.81 (m, 3H, H-4, H-6a, H-6b), 3.93 (d, *J* 9.3 Hz, 1H, H-3), 4.08 (s, 1H, H-2), 5.57 (s, 1H, H-1), 7.34 (d, *J* 7.8 Hz, 2H, Ar-H), 7.43 (t, *J* 8.2 Hz, 1H, Ar-H), 7.49 (d, *J* 8.5 Hz, 1H, Ar-H), 7.76 (d, *J* 11.9 Hz, 1H, Ar-H), 7.83 (d, *J* 7.7 Hz, 2H, Ar-H), 8.25 (s, 1H, *H*C=N); ^{13}C NMR (125 MHz, CD_3OD): δ 21.5 (CH_3), 62.6 (C-6), 68.2 (C-4), 71.7 (C-2), 72.3 (C-3), 76.0 (C-5), 101.2 (C-1), 115.4 (d, $J_{C,F}$ 21 Hz), 119.4, 126.1, 128.8 (2C), 130.4 (2C), 131.0 (d, $J_{C,F}$ 7 Hz), 131.2, 144.3, 147.4 (d, $J_{C,F}$ 12 Hz; 11C, Ar-C), 148.8 (H*C*=N), 154.6 (d, $J_{C,F}$ 246 Hz, Ar-C), 167.1 (C=O); IR (KBr): ν 3421 (vs, OH, NH), 1651 (s, C=N-NH-C=O) cm^{-1}; HRMS: *m/z:* Calcd for $C_{21}H_{23}FN_2NaO_7$ $[M+Na]^+$: 457.1387, found: 457.1387.

(*E*)-*N*'-[3-Fluoro-4-(α-D-mannopyranosyloxy)benzylidene]-5-methylthiophene-2-carbohydrazide (3n). Prepared according to general procedure A from aldehyde **1** (10 mg, 33.1 µmol) and 5-methyl-thiophene-2-carboxylic acid hydrazide (**2n**, 5.2 mg, 33.1 µmol). Yield: 7.8 mg (54%). $[\alpha]_D^{20}$ +156.0 (*c* 0.47, MeOH); ^{1}H NMR (500 MHz, $(CD_3)_2SO$): δ 2.52 (s, 3H, CH_3), 3.39–3.55 (m, 3H, H-4, H-5, H-6a), 3.60 (dd, *J* 4.9, 10.9 Hz, 1H, H-6b), 3.69 (d, *J* 8.9 Hz, 1H, H-3), 3.88 (s, 1H, H-2), 4.46 (t, *J* 5.8 Hz, 1H, OH-6), 4.83 (d, *J* 3.6 Hz, 1H, OH-3), 4.87 (t, *J* 8.4 Hz, 1H, OH-4), 5.12 (d, *J* 3.0 Hz, 1H, OH-2), 5.50 (s, 1H, H-1), 6.92 (s, 1H, Ar-H), 7.39–7.66 (m, 3H, Ar-H), 7.78 (m, 1H, Ar-H), 8.18 (s, 1H, *H*C=N), 11.75 (s, 1H, NH); ^{13}C NMR (125 MHz, $(CD_3)_2SO$): δ 15.1 (CH_3), 60.9 (C-6), 66.5 (C-5), 69.9 (C-2), 70.6 (C-3), 75.5 (C-4), 99.7 (C-1), 114.03 (d, $J_{C,F}$ 38 Hz), 118.8, 124.2, 125.4, 129.2, 135.1, 142.5, 145.49 (d, $J_{C,F}$ 79 Hz), 152.5 (d, $J_{C,F}$ 245 Hz), 161.2 (11C, 10 Ar-C, H*C*=N), 166.4 (C=O); IR (KBr): ν 3413 (vs, OH, NH), 1619 (vs, C=N-NH-C=O) cm^{-1}; HRMS: *m/z:* Calcd for $C_{19}H_{21}FN_2NaO_7S$ $[M+Na]^+$: 463.0951, found: 463.0954.

(*E*)-4-Chloro-*N*'-[3-fluoro-4-(α-D-mannopyranosyloxy)benzylidene]-carbohydrazide (3o). Prepared according to general procedure A from aldehyde **1** (10 mg, 33.1 µmol) and 4-chlorobenzhydrazide (**2o**, 5.6 mg, 33.1 µmol). Yield: 4.5 mg (30%). $[\alpha]_D^{20}$ +112.4 (*c* 0.24, MeOH); ^{1}H NMR (500 MHz, CD_3OD): δ = 3.63 (ddd, *J* 2.2, 5.3, 9.5 Hz, 1H, H-5), 3.69–3.82 (m, 3H, H-4, H6a, H-6b), 3.93 (dd, *J* 3.4, 9.4 Hz, 1H, H-3), 4.08 (dd, *J* 1.7, 3.1 Hz, 1H, H-2), 5.57 (s, 1H, H-1), 7.44 (t, *J* 8.3 Hz, 1H, Ar-H), 7.50 (d, *J* 9.2 Hz, 1H, Ar-H), 7.54 (d, *J* 8.5 Hz, 2H, Ar-H), 7.76

(m, 1H, Ar-H), 7.92 (d, *J* 8.5 Hz, 2H, Ar-H), 8.26 (s, 1H, *H*C=N); ^{13}C NMR (125 MHz, CD_3OD): δ 62.6 (C-6), 68.2 (C-4), 71.7 (C-2), 72.3 (C-3), 76.0 (C-5), 101.1 (C-1), 115.5 (d, $J_{C,F}$ 20 Hz), 119.4, 126.2, 130.0 (2C), 130.5 (2C), 130.8 (d, $J_{C,F}$ 7 Hz), 132.8, 139.5, 147.5 (d, $J_{C,F}$ 11 Hz; 11C, Ar-C), 149.3 (H*C*=N), 154.6 (d, $J_{C,F}$ 246 Hz, Ar-C), 165.9 (C=O); IR (KBr): ν 3436 (vs, OH, NH), 1651 (s, C=N-NH-C=O) cm^{-1}; HRMS: *m/z:* Calcd for $C_{20}H_{20}ClFN_2NaO_7$ $[M+Na]^+$: 477.0841, found: 477.0841.

(*E*)-3-Chloro-*N*'-[(3-fluoro-4-(α-D-mannopyranosyloxy)benzylidene]-4-methylthiophene-2-carbohydrazide (**3p**)**.** Prepared according to general procedure A from aldehyde **1** (10 mg, 33.1 µmol) and 3-chloro-4-methyl-2-thiophenecarboxylic acid hydrazide (**2p**, 6.3 mg, 33.1 µmol). Yield: 5.4 mg (34%). $[\alpha]_D^{20}$ +90.7 (*c* 0.15, MeOH); ^{1}H NMR (500 MHz, $(CD_3)_2SO$): δ2.20 (s, 3H, CH_3), 3.37–3.55 (m, 3H, H-6a, H-4, H-5), 3.60 (dd, *J* 3.6, 11.2 Hz, 1H, H-6b), 3.68 (dd, *J* 2.9, 9.0 Hz, 1H, H-3), 3.87 (s, 1H, H-2), 4.47 (s, 1H, OH-6), 4.90 (s, 2H, OH-3, OH-4), 5.13 (s, 1H, OH-2), 5.49 (s, 1H, H-1), 7.40–7.48 (m, 2H, Ar-H), 7.56 (d, *J* 11.9 Hz, 1H, Ar-H), 7.65 (m, 1H, Ar-H), 8.06 (m, 1H, Ar-H), 11.77 (s, 1H, NH); ^{13}C NMR (125 MHz, $(CD_3)_2SO$): δ 14.3 (CH_3), 60.9 (C-6), 66.5 (C-4), 69.8 (C-2), 70.5 (C-3), 75.5 (C-5), 99.7 (C-1), 113.9 (d, $J_{C,F}$ 20 Hz), 118.6 (d, $J_{C,F}$ 6 Hz), 124.3, 129.1, 131.2, 136.6, 143.6, 143.8 (d, $J_{C,F}$ 9 Hz), 145.3 (9C, Ar-C), 146.8 (H*C*=N), 152.4 (d, $J_{C,F}$ = 245 Hz, Ar-C), 171.2 (C=O); IR (KBr): ν 3430 (vs, OH, NH), 1642 (s, C=N-NH-C=O) cm^{-1}; HRMS: *m/z:* Calcd for $C_{19}H_{20}ClFN_2NaO_7S$ $[M+Na]^+$: 497.0561, found: 497.0561.

(*E*)-5-Chloro-*N*'-[3-fluoro-4-(α-D-mannopyranosyloxy)benzylidene]-thiophene-2-carbohydrazide (**3q**). Prepared according to general procedure A from aldehyde **1** (10 mg, 33.1 µmol) and 5-chlorothiophene-2-carboxylic acid hydrazide (**2q**, 5.6 mg, 33.1 µmol). Yield: 3.9 mg (26%). $[\alpha]_D^{20}$ +138.3 (*c* 0.15, MeOH); ^{1}H NMR (500 MHz, $(CD_3)_2SO$): δ3.40–3.56 (m, 3H, H-4, H-5, H-6a), 3.62 (dd, *J* 4.1, 11.2 Hz, 1H, H-6b), 3.70 (dd, *J* 3.0, 9.1 Hz, 1H, H-3), 3.90 (s, 1H, H-2), 4.49 (t, *J* 5.6 Hz, 1H, OH-4), 4.89 (m, 2H, OH-3, OH-6), 5.15 (s, 1H, OH-2), 5.53 (s, 1H, H-1), 7.29 (d, *J* 4.1 Hz, 1H, Ar-H), 7.43–7.69 (m, 3H, Ar-H), 8.39 (m, 2H, Ar-H, *H*C=N), 12.00 (s, 1H, NH); ^{13}C NMR (125 MHz, $(CD_3)_2SO$): δ61.0 (C-6), 66.5 (C-5), 69.8 (C-2), 70.5 (C-3), 75.5 (C-4), 99.7 (C-1), 114.4 (d, $J_{C,F}$ 19 Hz), 118.8, 124.4, 126.6, 128.6 (d, $J_{C,F}$ 6 Hz), 130.4, 134.5, 137.3 (8C, Ar-C), 143.5 (H*C*=N), 145.4 (d, $J_{C,F}$ 11 Hz), 152.4 (d, $J_{C,F}$ 245 Hz; 2C, Ar-C), 160.0 (C=O); IR (KBr): ν 3436 (vs, OH, NH), 1651 (s, C=N-NH-C=O) cm^{-1}; HRMS: *m/z:* Calcd for $C_{18}H_{18}ClFN_2NaO_7S$ $[M+Na]^+$: 483.0405, found: 483.0406.

(*E*)-*N*'-[3-Fluoro-4-(α-D-mannopyranosyloxy)benzylidene]-2-naphthohydrazide (**3r**)**.** Prepared according to general procedure A from aldehyde **1** (10 mg, 33.1 µmol) and 2-naphthhydrazide (**2r**, 6.2 mg, 33.1 µmol). Yield: 3.9 mg (25%). $[\alpha]_D^{20}$ +81.6 (*c* 0.16, MeOH); ^{1}H NMR (500 MHz, CD_3OD): δ3.65 (m, 1H, H-5), 3.69–3.82 (m, 3H, H-4, H-6a, H-6b), 3.93 (dd, *J* 2.1, 9.4 Hz, 1H, H-3), 4.09 (s, 1H, H-2), 5.58 (s, 1H, H-1), 7.45 (t, *J* 8.2 Hz, 1H, Ar-H), 7.53 (d, *J* 8.5 Hz, 1H, Ar-H), 7.57–7.66 (m, 2H, Ar-H), 7.80 (d, *J* 11.9 Hz, 1H, Ar-H), 7.94–8.06 (m, 4H, Ar-H), 8.32 (s, 1H, *H*C=N), 8.50 (s, 1H, Ar-H); ^{13}C NMR (125 MHz, CD_3OD): δ62.6 (C-6), 68.2 (C-4), 71.8 (C-2), 72.3 (C-3), 76.1 (C-5), 101.2 (C-1), 115.5 (d, $J_{C,F}$ = 20 Hz), 119.5, 125.0, 126.2, 128.1, 128.9, 129.2, 129.5, 129.6, 130.1, 131.0 (d, $J_{C,F}$ 7 Hz), 131.3, 134.0, 136.6, 147.5 (d, $J_{C,F}$ 11 Hz; 15C, Ar-C), 149.0 (H*C*=N), 154.6 (d, $J_{C,F}$ 246 Hz, Ar-C), 167.1 (C=O); IR (KBr): ν 3422 (vs, OH, NH), 1651 (s, C=N-NH-C=O) cm^{-1}; HRMS: *m/z:* Calcd for $C_{24}H_{23}FN_2NaO_7$ $[M+Na]^+$: 493.1387 found: 493.1388.

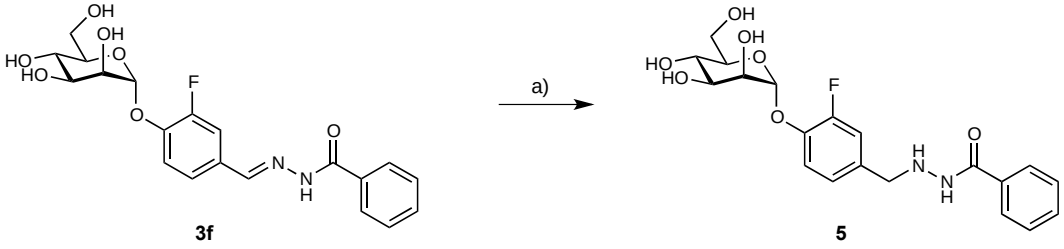
***(E)*-*N*'-[3-Fluoro-4-(α-D-mannopyranosyloxy)benzylidene]-4-(trifluoromethyl)benzo-hydrazide** (**3s**)**.** Prepared according to general procedure A from aldehyde **1** (10 mg, 33.1 µmol) and 4-(trifluoromethyl)benzhydrazide (**2s**, 6.8 mg, 33.1 µmol). Yield: 8.6 mg (53%). $[\alpha]_D^{20}$ +98.8 (*c* 0.39, MeOH); ^{1}H NMR (500 MHz, CD_3OD): δ3.64 (d, *J* 6.2 Hz, 1H, H-5), 3.69–3.81 (m, 3H, H-4, H-6a, H-6b), 3.93 (dd, *J* 2.2, 9.3 Hz, 1H, H-3), 4.08 (s, 1H, H-2), 5.58 (s, 1H, H-1), 7.45 (t, *J* 8.2 Hz, 1H, Ar-H), 7.51 (d, *J* 8.5 Hz, 1H, Ar-H), 7.77 (d, *J* 11.8 Hz, 1H, Ar-H), 7.84 (d, *J* 8.0 Hz, 2H, Ar-H), 8.10 (d, *J* 8.0 Hz, 2H, Ar-H), 8.28 (s, 1H, *H*C=N); ^{13}C NMR (125 MHz, CD_3OD): δ62.6 (C-6), 68.2 (C-4), 71.7 (C-2), 72.3 (C-3), 76.1 (C-5), 101.1 (C-1), 115.5 (d, $J_{C,F}$ 20 Hz), 119.4 (2C, Ar-C), 125.2 (q, $J_{C,F}$ 267 Hz, CF_3), 126.3 (d, *J* 3 Hz), 126.7 (q, *J* 3 Hz, 2C), 129.6 (2C), 130.7 (d, $J_{C,F}$ 7 Hz), 134.6 (q, $J_{C,F}$ 33 Hz), 137.9, 147.6 (d, $J_{C,F}$ 11

Hz; 9C, Ar-C), 149.7 (H*C*=N), 154.6 (d, $J_{C,F}$ 246 Hz, Ar-C), 165.6 (C=O); IR (KBr): ν 3430 (vs, OH, NH), 1663 (vs, C=N-NH-C=O) cm^{-1}; HRMS: *m/z:* Calcd for $C_{21}H_{20}F_4N_2NaO_7$ $[M+Na]^+$: 511.1104 found: 511.1107.

(*E*)-4-*N*'-[3-Fluoro-4-(α-D-mannopyranosyloxy)benzylidene]benzo[*b*]thiophene-2-carbohydrazide (**3t**). Prepared according to general procedure A from aldehyde **1** (10 mg, 33.1 µmol) and 3-chloro-benzo[*b*]thiophene acid hydrazide (**2t**, 7.5 mg, 33.1 µmol). Yield: 4.2 mg (25%). $[\alpha]_D^{20}$ +82.4 (*c* 0.18, MeOH); ^{1}H NMR (500 MHz, $(CD_3)_2SO$): δ 3.38–3.51 (m, 3H, H-6a, H-6b, H-5), 3.59 (s, 1H, H-4), 3.68 (s, 1H, H-3), 3.88 (s, 1H, H-2), 4.47 (s, 1H, OH-4), 4.71–4.96 (m, 2H, OH-3, OH-6), 5.11 (s, 1H, OH-2), 5.49 (s, 1H, H-1), 7.35–7.69 (m, 5H, Ar-H), 8.44–8.84 (m, 3H, 2 Ar-H, *H*C=N), 12.09 (s, 1H, NH); ^{13}C NMR (125 MHz, $(CD_3)_2SO$): δ 60.9 (C-6), 66.5 (C-4), 69.8 (C-2), 70.5 (C-3), 75.5 (C-5), 99.8 (C-1), 114.2 (d, $J_{C,F}$ 20 Hz), 118.6, 122.4, 122.6, 123.4, 124.4, 126.0, 127.4, 127.7, 132.6, 143.5, 145.4, 147.5, 152.5 (d, $J_{C,F}$ 245 Hz; 15C, Ar-C, H*C*=N), 160.2 (C=O); IR (KBr): ν 3460 (vs, OH, NH), 1656 (vs, C=N-NH-C=O) cm^{-1}; HRMS: *m/z:* Calcd for $C_{22}H_{20}ClFN_2NaO_7S$ $[M+Na]^+$: 533.0561 found: 533.0562.

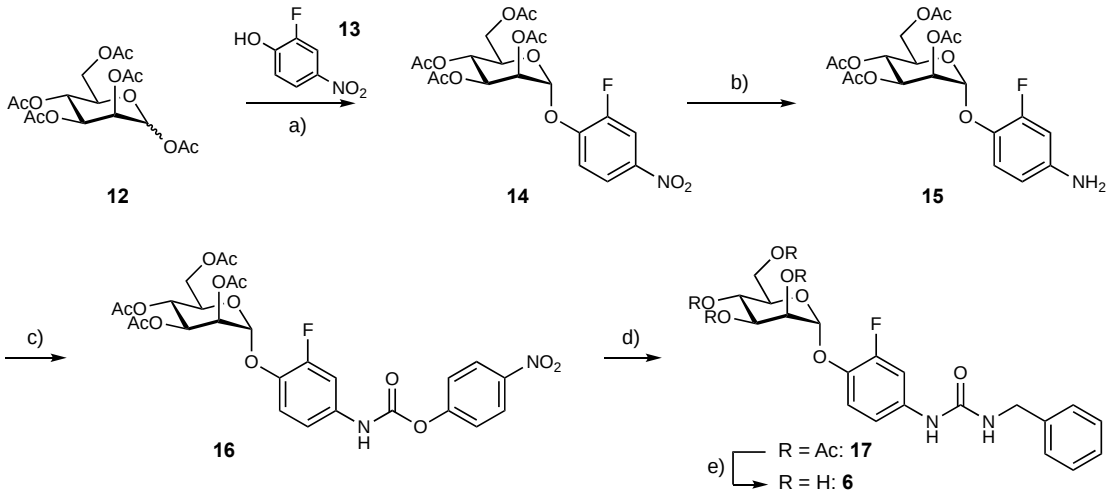
Synthesis of bioisosteres

General procedure B for deprotection of acetylated mannosides. Protected mannosides were dissolved in dry MeOH (2 mL) under argon atmosphere and freshly prepared 1 M NaOMe (100 µL) was added. The mixtures were stirred at r.t. for 30-45 min, until TLC (CH_2Cl_2/MeOH, 9:1) showed no remaining starting material. The mixtures were neutralized with amberlite ion-exchange resin (H^+-form, IR120, Sigma Aldrich), filtered, and concentrated *in vacuo*. The residues were purified by MPLC (RP-18; H_2O/MeCN, 95:5 to 20:80) to yield 60-82% of the desired products.



Scheme 2. Reagents and conditions: a) $NaBH_3CN$, aq. HCl., MeOH, r.t., 23 h, 95%.

***N*'-[3-Fluoro-4-(α-D-mannopyranosyloxy)benzyl]benzohydrazide** (**5**). Compound **3f** (4.5 mg, 10.7 µmol, 1 eq.) was dissolved in MeOH (3 mL) and $NaBH_3CN$ (6.7 mg, 107 µmol, 10 eq.) and five drops of 36% aq. HCl. were added. The mixture was flushed with argon for 5 min and stirred at r.t. After 23 h the mixture was neutralized with solid NaOH and concentrated. The residue was purified by MPLC on RP-18 (H_2O/MeCN, 95:5 to 20:80) to give **5** (4.3 mg, 95%). $[\alpha]_D^{20}$ +66.3 (*c* 0.22, MeOH); ^{1}H NMR (500 MHz, CD_3OD): δ 3.64–3.79 (m, 4H, H-4, H-5, H-6a, H-6b), 3.91 (dd, *J* 3.4, 9.3 Hz, 1H, H-3), 4.00 (s, 2H, CH_2), 4.06 (dd, *J* 1.8, 3.3 Hz, 1H, H-2), 5.46 (d, *J* 1.5 Hz, 1H, H-1), 7.13 (d, *J* 8.4 Hz, 1H, Ar-H), 7.24 (dd, *J* 1.9, 11.9 Hz, 1H, Ar-H), 7.31 (t, *J* 8.4 Hz, 1H, Ar-H), 7.42 (t, *J* 7.5 Hz, 2H, Ar-H), 7.50 (m, 1H, Ar-H), 7.69–7.75 (m, 2H, Ar-H); ^{13}C NMR (125 MHz, CD_3OD): δ = 55.5 (CH_2), 62.6 (C-6), 68.2 (C-4), 71.9 (C-2), 72.3 (C-3), 75.8 (C-5), 101.6 (C-1), 117.9 (d, $J_{C,F}$ = 19 Hz), 120.0, 126.2 (d, $J_{C,F}$ 3 Hz), 128.2, 129.5, 132.6, 134.5, 134.8 (d, $J_{C,F}$ 17 Hz), 144.8 (d, $J_{C,F}$ = 11 Hz), 154.5 (d, $J_{C,F}$ 245 Hz; 12C, Ar-C), 169.0 (C=O); IR (KBr): ν 3438 (vs, OH), 1646 (s, C=O) cm^{-1}; HRMS: *m/z:* Calcd for $C_{20}H_{24}FN_2NaO_7$ $[M+Na]^+$: 445.1387, found: 445.1386.



Scheme 3. *Reagents and conditions:* a) $BF_3{\cdot}Et_2O$, CH_2Cl_2, MS 4Å, 50 °C, 21 h, 39%; b) H_2, $Pd(OH)_2$, THF/ MeOH (2:1), r.t., 2h, 75%; c) 4-nitrophenyl chloroformate, DIPEA, CH_2Cl_2, 2 h; d) benzylamine, DIPEA, THF, 2.5 h, 29% (over two steps); e) NaOMe, MeOH, r.t., 45 min, 70%.

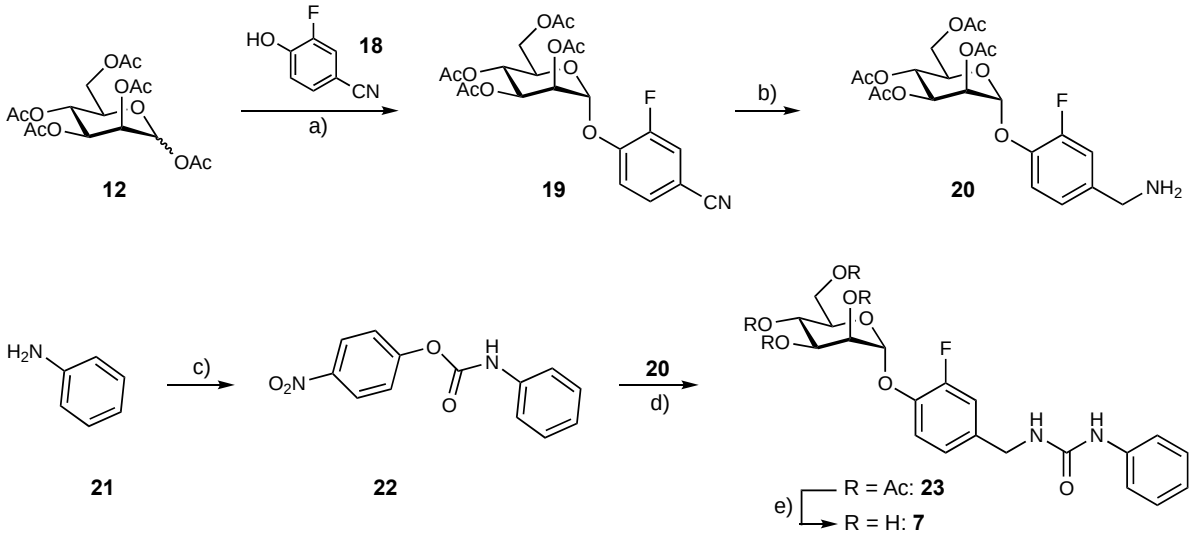
2-Fluoro-4-nitrophenyl 2,3,4,6-tetra-*O*-acetyl-α-D-mannopyranoside (**14**)**.** In a two-necked flask, activated MS 4Å (ca. 500 mg), peracetylated D-mannose (**12**, 500 mg, 1.28 mmol, 1.2 eq.), 3-fluoro-4-nitrophenol (**13**, 168 mg, 1.07 mmol, 1.0 eq.), and dry CH_2Cl_2 (10 mL) were mixed and cooled down in an ice bath. Under argon atmosphere, $BF_3{\cdot}Et_2O$ (395 µL, 3.20 mmol, 3.0 eq.) was added dropwise and the reaction heated to 50 °C. The mixture was refluxed for 21 h. Then it was cooled down to r.t., diluted with EtOAc, and filtered over celite. The filtrate was subsequently washed with satd. aq. $NaHCO_3$ and brine. The organic layer was dried over Na_2SO_4, filtered, and concentrated *in vacuo*. The residue was purified by MPLC on silica (petroleum ether/EtOAc, 1:0 to 4:6) to give **14** (241 mg, 39%). $[\alpha]_D^{20}$ +92.8 (*c* 1.00, $CHCl_3$); ^{1}H NMR (500 MHz, $CDCl_3$): δ 2.03, 2.05, 2.07, 2.22 (4 s, 12H, 4Ac-CH_3), 4.06–4.15 (m, 2H, H-5, H-6a), 4.27 (dd, *J* 5.9, 12.8 Hz, 1H, H-6b), 5.39 (t, *J* 9.9 Hz, 1H, H-4), 5.52 (dd, *J* 1.8, 3.5 Hz, 1H, H-2), 5.55 (dd, *J* 3.5, 9.9 Hz, 1H, H-3), 5.65 (d, *J* 1.5 Hz, 1H, H-1), 7.36 (m, 1H, Ar-H), 8.02–8.08 (m, 2H, Ar-H); ^{13}C NMR (125 MHz, $CDCl_3$): δ 20.63, 20.64, 20.65, 20.8 (4 Ac-CH_3), 61.9 (C-6), 65.6 (C-4), 68.4 (C-3), 68.9 (C-2), 70.2 (C-5), 96.9 (C-1), 113.0 (d, $J_{C,F}$ 23 Hz), 117.1 (d, $J_{C,F}$ 1 Hz), 120.5 (d, $J_{C,F}$ 4 Hz), 143.0 (d, $J_{C,F}$ 7 Hz), 148.8 (d, $J_{C,F}$ 11 Hz), 151.9 (d, $J_{C,F}$ 254 Hz; 6C, Ar-C), 169.6, 169.8, 169.9, 170.4 (4 C=O); ESI-MS: *m/z:* Calcd for $C_{20}H_{22}FNNaO_{12}$ $[M+Na]^+$: 510.10, found: 510.14.

4-Amino-2-fluorophenyl 2,3,4,6-tetra-*O*-acetyl-α-D-mannopyranoside (**15**)**.** A two-necked flask equipped with a magnetic stirrer was charged with **14** (241 mg, 0.527 mmol). Under argon atmosphere, THF/MeOH (2:1, 15 mL) and $Pd(OH)_2$/C (25 mg) were added. The flask was evacuated five times and filled with H_2. Under hydrogen atmosphere, the mixture was stirred at r.t. until TLC (CH_2Cl_2/MeOH, 8:2) indicated completion of the reaction. The mixture was filtered over celite and concentrated. Purification by MPLC (petroleum ether/EtOAc, 1:0 to 0:1) gave **15** (181 mg, 75%). $[\alpha]_D^{20}$ +74.5 (*c* 2.01, $CHCl_3$); ^{1}H NMR (500 MHz, $CDCl_3$): δ 1.97, 2.00, 2.01, 2.12 (4 s, 12H, 4 Ac-CH_3), 3.71 (s, 2H, NH_2), 4.07 (dd, *J* 1.8, 12.1 Hz, 1H, H-6a), 4.18–4.28 (m, 2H, H-5, H-6b), 5.22 (s, 1H, H-1), 5.29 (t, *J* 9.9 Hz, 1H, H-4), 5.44 (dd, *J* 1.7, 3.3 Hz, 1H, H-2), 5.47 (dd, *J* 3.4, 9.9 Hz, 1H, H-3), 6.29 (m, 1H, Ar-H), 6.37 (dd, *J* 2.6, 12.5 Hz, 1H, Ar-H), 6.86 (t, *J* 8.8 Hz, 1H, Ar-H); ^{13}C NMR (125 MHz, $CDCl_3$): δ 20.46, 20.49, 20.6 (4C, 4 Ac-CH_3), 62.1 (C-6), 65.8 (C-4), 68.7 (C-3), 69.2 (2C, C-3, C-5), 98.6 (d, $J_{C,F}$ 1 Hz, C-1), 103.3 (d, $J_{C,F}$ 22 Hz), 110.2 (d, $J_{C,F}$ 3 Hz), 121.4 (d, $J_{C,F}$ 2 Hz), 134.9 (d, $J_{C,F}$ 12 Hz), 144.1 (d, $J_{C,F}$ 10 Hz), 154.3 (d, $J_{C,F}$ 246 Hz; 6C, Ar-C), 169.6, 169.7, 169.8, 170.4 (4 C=O); ESI-MS: *m/z:* Calcd for $C_{20}H_{24}FNNaO_{10}$ $[M+Na]^+$: 480.13, found: 480.13.

4-Nitrophenyl (4'-(2,3,4,6-tetra-*O*-acetyl-α-D-mannopyranosyloxy)-3'-fluorophenyl)carbamate (**16**)**.** In a two-necked flask, 4-nitrophenyl chloroformate (20.7 mg, 0.103 mmol, 1 eq.) was dissolved in dry CH_2Cl_2 (2 mL). Under argon atmosphere, a solution of **15** (47 mg, 0.103 mmol, 1 eq.) and DIPEA (17.6 µL, 0.103 mmol, 1 eq.) in dry CH_2Cl_2 (2 mL) was added dropwise. The mixture was stirred at r.t. for 2 h, until TLC (petroleum ether/EtOAc, 1:1) showed completion of the reaction. The mixture was diluted with EtOAc and washed with 1 M aq. HCl. The organic layer was dried over Na_2SO_4, filtered, and concentrated to give **16** (quant.) which was used without further purification. ESI-MS: *m/z:* Calcd for $C_{27}H_{27}FN_2NaO_{14}$ $[M+Na]^+$: 645.13, found: 645.16.

1-Benzyl-3-[4'-(2,3,4,6-tetra-*O*-acetyl-α-D-mannopyranosyloxy)-3'-fluorophenyl]urea (**17**)**.** Crude **16** was dissolved in dry THF (2 mL) and benzylamine (11.3 mg, 0.103 mmol, 1 eq.) and DIPEA (17.6 µL, 0.103 mmol, 1 eq.) were added. The reaction was stirred for 1.5 h, until TLC (CH_2Cl_2/MeOH, 9:1) indicated full consumption of the starting material. The mixture was concentrated, dissolved in EtOAc, washed with 1 M aq. HCl and the aqueous layer re-extracted with EtOAc (3x). The organic layer was dried over Na_2SO_4, filtered and the solvent removed. Purification by MPLC (petroleum ether/EtOAc, 1:0 to 0:1) gave **17** (17.1 mg, 29% over two steps). $[\alpha]_D^{20}$ +35.8 (*c* 0.89, $CHCl_3$); ^{1}H NMR (500 MHz, $CDCl_3$): δ 2.01, 2.02, 2.06, 2.18 (4 s, 12H, 4 Ac-CH_3), 4.08 (dd, *J* 3.4, 13.2 Hz, 1H, H-6a), 4.21–4.28 (m, 2H, H-6b, H-5), 4.37 (d, *J* 5.7 Hz, 2H, CH_2), 5.32–5.38 (m, 2H, H-4, H-1), 5.41 (t, *J* 5.6 Hz, 1H, NH), 5.49 (dd, *J* 1.8, 3.4 Hz, 1H, H-2), 5.52 (dd, *J* 3.5, 9.9 Hz, 1H, H-3), 6.91 (dd, *J* 1.2, 8.9 Hz, 1H, Ar-H), 6.94 (s, 1H, NH), 7.00 (t, *J* 8.8 Hz, 1H, Ar-H), 7.21–7.27 (m, 4H, Ar-H), 7.27–7.32 (m, 2H, Ar-H); ^{13}C NMR (125 MHz, $CDCl_3$): δ 20.65, 20.66, 20.8 (4C, 4 Ac-CH_3), 44.2 (CH_2), 62.1 (C-6), 65.8 (C-3), 68.9 (C-2), 69.3 (C-4), 69.5 (C-5), 98.1 (C-1), 109.1 (d, $J_{C,F}$ 23 Hz), 115.6 (d, $J_{C,F}$ 3 Hz), 120.1 (d, $J_{C,F}$ 2 Hz), 127.4, 127.5, 128.7, 135.5 (d, $J_{C,F}$ 9 Hz), 138.6, 138.8 (d, $J_{C,F}$ 12 Hz), 153.5 (d, $J_{C,F}$ 247 Hz; 12C, Ar-C), 155.4 (NC=O), 169.8, 170.0, 170.7 (4C, 4 C=O); ESI-MS: *m/z:* Calcd for $C_{28}H_{31}FN_2Na\ O_{11}$ $[M+Na]^+$: 613.18, found: 613.22.

1-Benzyl-3-[3'-fluoro-4'-(α-D-mannopyranosyloxy)phenyl]urea (**6**)**.** Prepared according to general procedure B from **17** (17.7 mg, 30.0 µmol). Yield: 8.8 mg (70%). $[\alpha]_D^{20}$ +109.9 (*c* 0.44, MeOH); ^{1}H NMR (500 MHz, CD_3OD): δ 3.70–3.81 (m, 4H, H-4, H-5, H-6a, H-6b), 3.90 (m, 1H, H-3), 4.06 (dd, *J* 1.8, 3.3 Hz, 1H, H-2), 4.37 (s, 2H, CH_2), 5.35 (d, *J* 1.7 Hz, 1H, H-1), 6.95 (ddd, *J* 1.3, 2.4, 8.9 Hz, 1H, Ar-H), 7.19–7.27 (m, 2H, Ar-H), 7.29–7.34 (m, 4H, Ar-H), 7.38 (dd, *J* 2.5, 13.3 Hz, 1H, Ar-H); ^{13}C NMR (125 MHz, CD_3OD): δ 44.5 (CH_2), 62.7 (C-6), 68.3 (C-4), 71.9 (C-2), 72.3 (C-3), 75.7 (C-5), 102.3 (C-1), 108.8 (d, $J_{C,F}$ 24 Hz), 115.7 (d, $J_{C,F}$ 3 Hz), 121.2 (d, $J_{C,F}$ 2 Hz), 128.1, 128.3, 129.6, 137.1 (d, $J_{C,F}$ 10 Hz), 140.4 (d, $J_{C,F}$ 11 Hz), 140.9, 154.7 (d, $J_{C,F}$ 243 Hz; 12C, Ar-C), 158.0 (C=O); IR (KBr): ν 3347 (vs, OH), 1635 (s, C=O), 1515 (vs, NH) cm^{-1}; HRMS: *m/z:* Calcd for $C_{20}H_{24}FN_2NaO_7$ $[M+Na]^+$: 445.1387, found: 445.1392.



Scheme 4. *Reagents and conditions:* a) $BF_3{\cdot}Et_2O$, CH_2Cl_2, MS 4Å, 50 °C, 29 h, 60%; b) H_2, $Pd(OH)_2$, THF/MeOH (2:1), r.t., 2 h, 64%; c) 4-nitrophenyl chloroformate, DIPEA, CH_2Cl_2, r.t., 22 h, 53%; d) DIPEA, THF, r.t., 2 h, 19%; e)NaOMe/MeOH, r.t., 45 min, 83%.

4'-(2,3,4,6-Tetra-*O*-acetyl-α-D-mannopyranosyloxy)-3'-fluorobenzonitrile (**19**)**.** In a two-necked flask, activated MS 4Å (500 mg), peracetylated D-mannose (**12**, 500 mg, 1.28 mmol, 1.2 eq.), 3-fluoro-4-hydroxybenzonitrile (**18**, 146 mg, 1.07 mmol, 1.0 eq.), and dry CH_2Cl_2 (10 mL) were mixed and cooled down in an ice bath. Under argon atmosphere, $BF_3{\cdot}Et_2O$ (395 µL, 3.20 mmol, 3 eq.) was added slowly and the reaction heated to 50 °C. The mixture was refluxed for 29 h. Then, it was diluted with EtOAc, filtered over celite, and subsequently washed with satd. aq. $NaHCO_3$ and brine. The organic layer was dried over Na_2SO_4, filtered, and concentrated. The residue was purified by MPLC (toluene/EtOAc, 1:0 to 6:4) to give **19** (356 mg, 60%). $[\alpha]_D^{20}$ +88.6 (*c* 1.00, $CHCl_3$); ^{1}H NMR (500 MHz, $CDCl_3$): δ 2.03, 2.04, 2.07, 2.21 (4 s, 12H, 4 Ac-CH_3), 4.06–4.14 (m, 2H, H-5, H-6a), 4.27 (dd, *J* 6.0, 12.9 Hz, 1H, H-6b), 5.38 (t, *J* 9.9 Hz, 1H, H-4), 5.51 (dd, *J* 1.8, 3.4 Hz, 1H, H-2), 5.53 (dd, *J* 3.5, 9.9 Hz, 1H, H-3), 5.63 (d, *J* 1.4 Hz, 1H, H-1), 7.34 (t, *J* 8.3 Hz, 1H, Ar-H), 7.42–7.48 (m, 2H, Ar-H); ^{13}C NMR (125 MHz, $CDCl_3$): δ 20.37, 20.39, 20.41, 20.5 (4 Ac-CH_3), 61.7 (C-6), 65.4 (C-4), 68.2 (C-3), 68.7 (C-2), 69.9 (C-5), 96.6 (C-1), 107.0 (d, $J_{C,F}$ 8 Hz), 117.1 (d, $J_{C,F}$ 2 Hz), 120.3 (d, $J_{C,F}$ 22 Hz), 129.2 (d, $J_{C,F}$ 4 Hz), 147.2 (d, $J_{C,F}$ 11 Hz), 152.2 (d, $J_{C,F}$ 252 Hz; 7C, Ar-C, CN), 169.4, 169.5, 169.6, 170.1 (4 C=O); IR (KBr): ν 2232 (w, CN), 1751 (vs, C=O) cm^{-1}; ESI-MS: *m/z:* Calcd for $C_{21}H_{22}FNNaO_{10}$ $[M+Na]^+$: 490.11, found: 490.06.

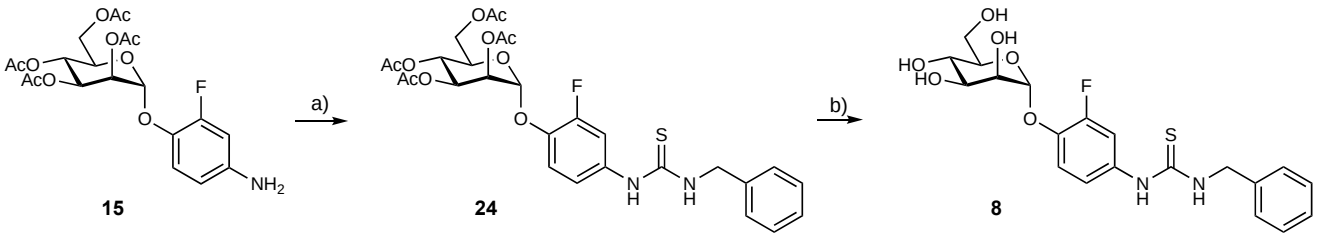
4'-(2,3,4,6-Tetra-*O*-acetyl-α-D-mannopyranosyloxy)-3'-fluorophenylmethanamine (**20**)**.** A two-neckeds flask was charged with **19** (50 mg, 0.107 mmol) and a magnetic stirrer. Under argon atmosphere, THF/MeOH (6 mL, 2:1) and $Pd(OH)_2$/C (15 mg) were added. The flask was evacuated five times and filled with H_2. Under hydrogen atmosphere, the mixture was stirred at r.t. until TLC (CH_2Cl_2/MeOH, 8:2) indicated completion of the reaction. The mixture was filtered over celite, and concentrated. Purification by MPLC (CH_2Cl_2/MeOH, 1:0 to 8:2) gave **20** (32.2 mg, 64%). $[\alpha]_D^{20}$ +93.4 (*c* 1.00, $CHCl_3$); ^{1}H NMR (500 MHz, $CDCl_3$): δ 2.03, 2.04, 2.07, 2.20 (4 s, 12H, 4 Ac-CH_3), 3.83 (s, 2H, CH_2), 4.10 (d, *J* 10.3 Hz, 1H, H-6a), 4.22–4.30 (m, 2H, H-5, H-6b), 5.37 (t, *J* 10.0 Hz, 1H, H-4), 5.46 (d, *J* 1.6 Hz, 1H, H-1), 5.52 (dd, *J* 1.8, 3.4 Hz, 1H, H-2), 5.57 (dd, *J* 3.5, 10.0 Hz, 1H, H-3), 7.02 (d, *J* 8.3 Hz, 1H, Ar-H), 7.09–7.16 (m, 2H, Ar-H); ^{13}C NMR (125 MHz, $CDCl_3$): δ 20.60, 20.63, 20.8 (4C, 4 Ac-CH_3), 45.3 (CH_2), 62.1 (C-6), 65.8 (C-4), 68.7 (C-3), 69.3 (C-2), 69.5 (C-5), 97.7 (C-1), 115.5 (d, $J_{C,F}$ 19 Hz), 119.4 (d, $J_{C,F}$ 1 Hz), 122.9

(d, $J_{C,F}$ 3 Hz), 139.7 (d, $J_{C,F}$ 6 Hz), 141.9 (d, $J_{C,F}$ 11 Hz), 153.4 (d, $J_{C,F}$ 248 Hz; 6C, Ar-C), 160.7, 169.8, 169.8, 170.5 (4 C=O); ESI-MS: *m/z:* Calcd for $C_{21}H_{27}FNO_{10}$ $[M+H]^+$: 472.16, found: 472.13.

4-Nitrophenyl phenylcarbamate (**22**)**.** To a mixture of aniline (**21**, 49.0 µL, 0.537 mmol) and 4-nitrophenyl chloroformate (108 mg, 0.537 mmol) in THF (2 mL) was added DIPEA (91.9 µL, 0.537 mmol). The mixture was stirred at r.t. After 22 h, TLC (toluene/EtOAc, 1:1) indicated no remaining starting materials. The mixture was diluted with EtOAc and washed with 1 M aq. HCl. The aqueous layer was extracted with EtOAc, the combined organic layers were dried over Na_2SO_4, filtered and concentrated. The residue was purified by MPLC (toluene/EtOAc, 1:0 to 1:1) to give **22** (72.8 mg, 52%). ^{1}H NMR (500 MHz, $CDCl_3$): δ 7.06 (s, 1H, NH), 7.16 (t, *J* 7.4 Hz, 1H, Ar-H), 7.34–7.41 (m, 4H, Ar-H), 7.45 (d, *J* 7.8 Hz, 2H, Ar-H), 8.25–8.31 (m, 2H, Ar-H); ^{13}C NMR (125 MHz, $CDCl_3$): δ 119.0 (2C), 122.1 (2C), 124.6, 125.2 (2C), 129.3 (2C), 136.6, 145.1 (12C, Ar-C), 150.1 (C=O), 155.4 (Ar-C).

1-[4'-(2,3,4,6-Tetra-*O*-acetyl-α-D-mannopyranosyloxy)-3'-fluoro]-3-phenylurea (**23**). A mixture of **20** (50 mg, 107 µmol), DIPEA (18.2 µL, 107 µmol) and **22** (27.4 mg, 107 µmol) in dry THF (2 mL) was stirred at r.t. for 2 h, until TLC (CH_2Cl_2/MeOH, 9:1) showed no remaining starting material. The mixture was diluted with EtOAc and washed with 1 M aq. HCl. The aqueous layer was extracted three times with EtOAc, and the combined organic layers were dried over Na_2SO_4, filtered, and concentrated *in vacuo*. The residue was purified by MPLC (CH_2Cl_2/MeOH, 1:0 to 9:1) to give **23** (12.1 mg, 19%). $[\alpha]_D^{20}$ +51.2 (*c* 0.61, $CHCl_3$); ^{1}H NMR (500 MHz, $CDCl_3$): δ = 2.02, 2.03, 2.06, 2.19 (4 s, 12H, 4 Ac-CH_3), 4.08 (dd, *J* 2.0, 12.1 Hz, 1H, H-6a), 4.20 (ddd, *J* 2.0, 5.2, 10.0 Hz, 1H, H-5), 4.25 (dd, *J* 5.2, 12.1 Hz, 1H, H-6b), 4.32 (d, *J* 5.9 Hz, 2H, CH_2), 5.33–5.39 (m, 2H, H-4, NH), 5.42 (d, *J* 1.5 Hz, 1H, H-1), 5.49 (dd, *J* 1.8, 3.4 Hz, 1H, H-2), 5.54 (dd, *J* 3.5, 10.0 Hz, 1H, H-3), 6.80 (s, 1H, NH), 6.96 (d, *J* 8.4 Hz, 1H, Ar-H), 7.03 (dd, *J* 1.8, 11.4 Hz, 1H, Ar-H), 7.05–7.11 (m, 2H, Ar-H), 7.24–7.30 (m, 4H, Ar-H); ^{13}C NMR (125 MHz, $CDCl_3$): δ 20.6, 20.7, 20.8 (4C, 4 Ac-CH_3), 43.2 (CH_2), 62.1 (C-6), 65.8 (C-4), 68.8 (C-3), 69.2 (C-2), 69.6 (C-5), 97.6 (C-1), 115.8 (d, $J_{C,F}$ 19 Hz), 119.4, 121.1, 123.3 (d, $J_{C,F}$ 4 Hz), 124.1, 129.3, 135.9 (d, $J_{C,F}$ 6 Hz), 138.2, 142.3 (d, $J_{C,F}$ 11 Hz), 153.4 (d, $J_{C,F}$ 249 Hz; 12C, Ar-C), 155.8 (C=O(NH)$_2$), 169.7, 169.9, 170.0, 170.6 (4 C=O); ESI-MS: *m/z:* Calcd for $C_{28}H_{31}FN_2NaO_{11}$ $[M+Na]^+$: 613.18, found: 613.31.

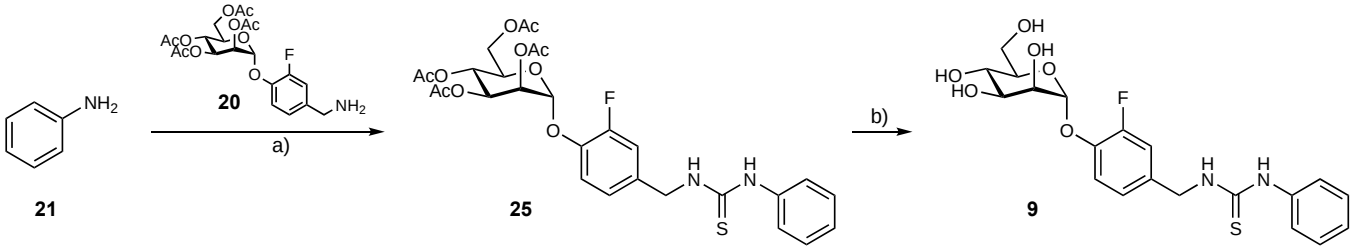
1-(3'-Fluoro-4'-α-D-mannopyranosyloxy)-3-phenylurea (**7**)**.** Prepared according to general procedure B from **23** (12.1 mg, 20.5 µmol). Yield: 7.2 mg (83%). $[\alpha]_D^{20}$ +88.1 (*c* 0.36, MeOH); ^{1}H NMR (500 MHz, CD_3OD): δ 3.66–3.80 (m, 4H, H-4, H-5, H6a, H-6b), 3.91 (dd, *J* 3.4, 9.2 Hz, 1H, H-3), 4.06 (dd, *J* 1.8, 3.3 Hz, 1H, H-2), 5.44 (d, *J* 1.6 Hz, 1H, H-1), 6.97 (t, *J* 7.4 Hz, 1H, Ar-H), 7.07 (d, *J* 8.4 Hz, 1H, Ar-H), 7.11 (dd, *J* 1.9, 11.9 Hz, 1H, Ar-H), 7.21–7.28 (m, 2H, Ar-H), 7.31 (t, *J* 8.4 Hz, 1H, Ar-H), 7.35 (dd, *J* 1.0, 8.5 Hz, 2H, Ar-H); ^{13}C NMR (125 MHz, CD_3OD): δ 43.6 (CH_2), 62.6 (C-6), 68.2 (C-4), 71.9 (C-2), 72.3 (C-3), 75.8 (C-5), 101.7 (C-1), 116.2 (d, $J_{C,F}$ 19 Hz), 120.3, 123.6, 124.3 (d, $J_{C,F}$ 3 Hz), 129.8, 137.0 (d, $J_{C,F}$ 6 Hz), 140.8, 144.4 (d, $J_{C,F}$ 11 Hz), 154.6 (d, $J_{C,F}$ 246 Hz; 12 C, Ar-C), 158.2 (C=O); IR (KBr): ν 3371 (vs, OH), 1651 (s, C=O) cm^{-1}; HRMS: *m/z:* Calcd for $C_{20}H_{24}FN_2NaO_7$ $[M+Na]^+$: 445.1387, found: 445.1385.



Scheme 5. *Reagents and conditions:* a) i. TCDI, CH_2Cl_2, r.t., 17 h; ii. benzylamine, CH_2Cl_2, r.t., 3 h, 52% (over two steps); b) NaOMe, MeOH, 0 °C, 45 min, 71%.

1-Benzyl-3-[4'-(2,3,4,6-tetra-*O*-acetyl-α-D-mannopyranosyloxy)-3'-fluorophenyl]thiourea (**24**). Compound **15** (47.0 mg, 0.103 mmol) and 1,1'-thiocarbonyldiimidazole (TCDI; 18.3 mg, 0.103 mmol.) were dissolved in dry CH_2Cl_2 (2 mL). The mixture was stirred at r.t. for 17 h. Then benzylamine (11.2 µL, 0.103 mmol) was added and the mixture was stirred for 3 h until TLC (petroleum ether/EtOAc, 1:1) showed no remaining starting material. The mixture was diluted with CH_2Cl_2, and washed with 1 M aq. HCl, satd. aq. $NaHCO_3$ and brine. The organic layer was dried over Na_2SO_4, filtered, and concentrated *in vacuo*. The residue was purified by MPLC (petroleum ether/EtOAc, 1:0 to 0:1) to give **24** (32.3 mg, 52%). $[\alpha]_D^{20}$ +51.6 (*c* 1.07, $CHCl_3$); ^{1}H NMR (500 MHz, $CDCl_3$): δ 1.96, 1.99, 2.02, 2.15 (4 s, 12H, 4 Ac-CH_3), 4.04 (dd, *J* 2.4, 12.1 Hz, 1H, H-6a), 4.12 (ddd, *J* 2.3, 5.2, 10.1 Hz, 1H, H-5), 4.20 (dd, *J* 5.2, 12.2 Hz, 1H, H-6b), 4.81 (d, *J* 5.2 Hz, 2H, CH_2), 5.32 (t, *J* 10.1 Hz, 1H, H-4), 5.41 (s, 1H, H-1), 5.43 (dd, *J* 1.8, 3.3 Hz, 1H, H-2), 5.48 (dd, *J* 3.4, 10.0 Hz, 1H, H-3), 6.23 (s, 1H, NH), 6.92 (d, *J* 8.7 Hz, 1H, Ar-H), 7.02 (dd, *J* 2.2, 11.0 Hz, 1H, Ar-H), 7.16 (t, *J* 8.7 Hz, 1H, Ar-H), 7.22–7.27 (m, 3H, Ar-H), 7.27–7.32 (m, 2H, Ar-H), 7.96 (s, 1H, NH); ^{13}C NMR (125 MHz, $CDCl_3$): δ 20.59, 20.61, 20.62, 20.8 (4 Ac-CH_3), 49.4 (CH_2), 62.1 (C-6), 65.7 (C-4), 68.5 (C-3), 69.1 (C-2), 69.7 (C-5), 97.5 (C-1), 114.4 (d, $J_{C,F}$ 20 Hz), 119.9 (d, $J_{C,F}$ 2 Hz), 121.6 (d, $J_{C,F}$ 4 Hz), 127.7, 127.9, 128.8, 137.0, 142.4 (d, $J_{C,F}$ 11 Hz), 153.4 (d, $J_{C,F}$ 252 Hz; 12C, Ar-C), 169.7, 169.8, 169.9, 170.4 (4 C=O), 181.1 (C=S); ESI-MS: *m/z:* Calcd for $C_{28}H_{32}FN_2O_{10}S$ $[M+H]^+$: 607.18, found: 607.30.

1-Benzyl-3-[3-fluoro-4-(α-D-mannopyranosyloxy)phenyl]thiourea (**8**)**.** Prepared according to general procedure B from **24** (32.2 mg, 53.1 µmol). Yield: 16.7 mg (71%). $[\alpha]_D^{20}$ +95.6 (*c* 0.84, MeOH); ^{1}H NMR (500 MHz, CD_3OD): δ 3.65–3.81 (m, 4H, H-4, H-5, H-6a, H-6b), 3.90 (dd, *J* 3.4, 9.1 Hz, 1H, H-3), 4.06 (dd, *J* 1.8, 3.3 Hz, 1H, H-2), 4.80 (s, 2H, CH_2), 5.45 (d, *J* 1.6 Hz, 1H, H-1), 7.03 (ddd, *J* 1.4, 2.3, 8.8 Hz, 1H, Ar-H), 7.25 (m, 1H, Ar-H), 7.27–7.36 (m, 6H, Ar-H); ^{13}C NMR (125 MHz, CD_3OD): δ 49.0 (CH_2), 62.7 (C-6), 68.2 (C-4), 71.8 (C-2), 72.3 (C-3), 75.9 (C-5), 101.7 (C-1), 114.5 (d, $J_{C,F}$ 21 Hz), 120.4, 121.9 (d, $J_{C,F}$ 5 Hz), 128.2, 128.6, 129.5, 135.1 (d, $J_{C,F}$ 9 Hz), 139.9, 143.2 (d, $J_{C,F}$ 11.0 Hz), 154.3 (d, $J_{C,F}$ 246 Hz; 12C, Ar-C), 182.9 (C=S); IR (KBr): ν 3295 (vs, OH), 1563 (vs), 1509 (vs) cm^{-1}; HRMS: *m/z:* Calcd for $C_{20}H_{24}FN_2NaO_6S$ $[M+Na]^+$: 461.1159 found: 461.1161.

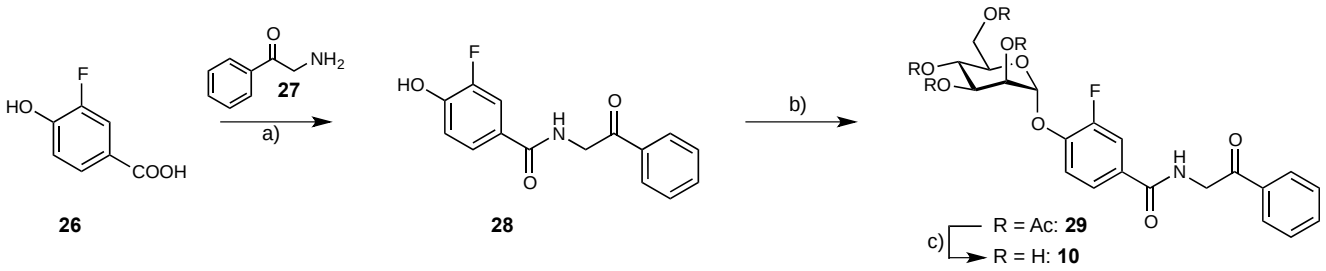


Scheme 6. *Reagents and conditions:* a) i. TCDI, CH_2Cl_2, r.t., 18 h; ii. **20**, CH_2Cl_2, r.t., 6 h, 7% (over two steps); b) NaOMe, MeOH, r.t., 45 min, 66%.

1-[4'-(2,3,4,6-Tetra-*O*-acetyl-α-D-mannopyranosyloxy)-3'-fluorobenzyl]-3-phenylthiourea (**25**)**.** 1,1'-Thiocarbonyldiimidazole (18.9 mg, 0.106 mmol) was dissolved in dry CH_2Cl_2 (2 mL) and **21** (9.7 µL, 0.106 mmol) was added. The mixture was stirred at r.t. for 18 h until TLC (toluene/EtOAc, 1:1) showed no remaining aniline. Then, **20** (50 mg, 0.106 mmol) in dry CH_2Cl_2 (2 mL) was added and the mixture stirred for another 6 h. Then, it was diluted with CH_2Cl_2 and subsequently washed with 1 M aq. HCl, satd. aq. $NaHCO_3$, and brine. The organic layer was dried over Na_2SO_4, filtered, and the solvent removed. The residue was purified by MPLC (toluene/EtOAc, 1:0 to 1:1) to give **25** (4.8 mg, 7 %). $[\alpha]_D^{20}$ +49.8 (*c* 0.24, $CHCl_3$); ^{1}H NMR (500 MHz, $CDCl_3$): δ 2.03, 2.06, 2.19 (3 s, 12H, 4 Ac-CH_3), 4.09 (dd, *J* 1.9, 12.2 Hz, 1H, H-6a), 4.20 (ddd, *J* 1.8, 5.2, 9.9 Hz, 1H, H-5), 4.26 (dd, *J* 5.3, 12.1 Hz, 1H, H-6b), 4.83 (d, *J* 5.7 Hz, 2H, CH_2), 5.36 (t, *J* 10.0 Hz, 1H, H-4), 5.45 (s, 1H, H-1), 5.49 (m, 1H, H-2), 5.55 (dd, *J* 3.5, 10.0 Hz, 1H, H-3), 6.22 (s, 1H, NH), 7.01 (d, *J* 8.4 Hz, 1H, Ar-H), 7.09 (dd, *J* 1.5, 11.3

Hz, 1H, Ar-H), 7.12 (t, *J* 8.2 Hz, 1H, Ar-H), 7.15–7.26 (m, 2H, Ar-H), 7.34 (m, 1H, Ar-H), 7.44 (t, *J* 7.8 Hz, 2H, Ar-H), 7.72 (s, 1H, NH); ^{13}C NMR (125 MHz, $CDCl_3$): δ 20.66, 20.67, 20.68, 20.8 (4 Ac-CH_3), 48.3 (CH_2), 62.1 (C-6), 65.8 (C-4), 68.7 (C-3), 69.3 (C-2), 69.6 (C-5), 97.6 (C-1), 116.1 (d, $J_{C,F}$ 19 Hz), 119.3, 123.6 (d, $J_{C,F}$ 4 Hz), 125.6, 127.8, 128.8, 130.4, 134.0 (d, *J* 6 Hz), 142.8 (d, $J_{C,F}$ 11 Hz), 153.4 (d, $J_{C,F}$ 249 Hz; 12C, Ar-C), 169.7, 169.8, 169.9, 170.5 (4 C=O), 181.3 (C=S); ESI-MS: *m/z:* Calcd for $C_{28}H_{31}FN_2NaO_{10}S$ $[M+H]^+$: 629.16, found: 629.15.

1-(3-Fluoro-4-α-D-mannopyranosyloxybenzyl)-3-phenylthiourea (**9**). Prepared according to general procedure B from **25** (4.6 mg, 7.9 µmol). Yield: 2.3 mg (66%). $[\alpha]_D^{20}$ +61.4 (*c* 0.12, MeOH); ^{1}H NMR (500 MHz, CD_3OD): δ = 3.66–3.79 (m, 4H, H-4, H-5, H-6a, H-6b), 3.91 (dd, *J* 3.4, 9.1 Hz, 1H, H-3), 4.06 (dd, *J* 1.8, 3.4 Hz, 1H, H-2), 4.76 (s, 2H, CH_2), 5.45 (d, *J* 1.7 Hz, 1H, H-1), 7.09 (d, *J* 8.4 Hz, 1H, Ar-H), 7.15 (dd, *J* 2.0, 11.9 Hz, 1H, Ar-H), 7.21 (m, 1H, Ar-H), 7.28–7.39 (m, 5H, Ar-H); ^{13}C NMR (125 MHz, CD_3OD): δ 48.2 (CH_2), 62.6 (C-6), 68.2 (C-4), 71.9 (C-2), 72.3 (C-3), 75.8 (C-5), 101.7 (C-1), 116.6 (d, $J_{C,F}$ 19 Hz), 120.2 (d, $J_{C,F}$ 1 Hz), 124.7 (d, $J_{C,F}$ 3 Hz), 125.9 (d, $J_{C,F}$ 3 Hz), 127.0, 130.3, 139.5, 144.4, 154.5 (d, $J_{C,F}$ 246 Hz; 12C, Ar-C), 182.8 (C=S); IR (KBr): ν 3422 (vs, OH), 1514 (m, NH) cm^{-1}; HRMS: *m/z:* Calcd for $C_{20}H_{24}FN_2NaO_6S$ $[M+Na]^+$: 461.1159 found: 461.1160.



Scheme 7. *Reagents and conditions:* a) HBTU, HOBt, DIPEA, DMF, r.t., 2.5 h, 18%; b) **11**, $BF_3{\cdot}Et_2O$, CH_2Cl_2/ MeCN, MS 4Å, 50-75 °C, 48 h, 12%; c) NaOMe, MeOH, r.t., 45 min, 60%.

3-Fluoro-4-hydroxy-*N*-(2-oxo-2-phenylethyl)benzamide (**28**). 3-Fluoro-4-hydroxybenzoic acid (**26**, 50.0 mg, 0.641 mmol, 1 eq.), HBTU (243 mg, 1.28 mmol, 2 eq.), HOBt hydrate (12% water; 98.4 mg, 0.205 mmol, 2 eq.), and 2-aminoacetophenone hydrochloride (**27**, 35.2 mg, 1.28 mmol, 2 eq.) were dissolved in anhydrous DMF (1.5 mL). Then, DIPEA (110 µL, 2.56 mmol, 4 eq.) was added and the mixture was stirred at r.t. for 2.5 h. Then, it was diluted with EtOAc and subsequently washed with 1 M aq. HCl and brine. The organic layer was dried over Na_2SO_4, filtered and the solvents were removed *in vacuo*. The residue was purified by MPLC (toluene/EtOAc, 1:0 to 1:1) to give **28** (31.8 mg, 18%). ^{1}H NMR (500 MHz, CD_3OD): δ 4.86 (s, 2H, CH_2), 6.99 (t, *J* 8.5 Hz, 1H, Ar-H), 7.53 (t, *J* 7.7 Hz, 2H, Ar-H), 7.60 (dd, *J* 2.1, 8.4 Hz, 1H, Ar-H), 7.62–7.67 (m, 2H, Ar-H), 8.02–8.08 (m, 2H, Ar-H); ^{13}C NMR (125 MHz, CD_3OD): δ 47.7 (CH_2), 116.5 (d, $J_{C,F}$ 20 Hz), 118.5 (d, $J_{C,F}$ 3 Hz), 125.3 (d, $J_{C,F}$ 3 Hz), 126.7 (d, $J_{C,F}$ 6 Hz), 129.0, 129.9, 134.8, 136.5, 150.0 (d, $J_{C,F}$ 13 Hz), 152.4 (d, $J_{C,F}$ 242 Hz; 12C, Ar-C), 169.1 (CONH), 196.4 (C=O); ESI-MS: *m/z:* Calcd for $C_{15}H_{12}FNNaO_3$ $[M+Na]^+$: 296.07, found: 295.49.

4-(2,3,4,6-Tetra-*O*-acetyl-α-D-mannopyranosyloxy)-3-fluoro-*N*-(2-oxo-2-phenylethyl)benzamide (**29**)**.** A two-necked flask was charged with activated MS 4Å (50 mg), peracetylated D-mannose (**12**, 54.5 mg, 0.140 mmol, 1.2 eq.) and dry CH_2Cl_2 (2 mL). Under argon atmosphere, **28** (31.8 mg, 0.116 mmol, 1 eq.) in dry MeCN (2 mL) was added. The mixture was refluxed at 50 °C for 24 h, and another 24 h at 75°C. When TLC (petroleum ether/EtOAc, 1:1) showed no remaining mannose precursor, the mixture was diluted with EtOAc, filtered over celite, and washed with satd. aq. $NaHCO_3$ and brine. The organic layer was dried over Na_2SO_4, filtered and concentrated. The residue was purified by MPLC (petroleum ether/EtOAc, 1:0 to 1:1) to yield **29** (8.3 mg, 12%). Unreacted **28** (15.5 mg, 49%) could be recovered. $[\alpha]_D^{20}$ +60.1 (*c* 0.42, $CHCl_3$); ^{1}H NMR (500 MHz, $CDCl_3$): δ

2.04, 2.04, 2.07, 2.21 (4 s, 12H, 4 Ac-CH_3), 4.09 (dd, *J* 2.1, 12.2 Hz, 1H, H-6a), 4.16 (ddd, *J* 1.9, 5.3, 9.9 Hz, 1H, H-5), 4.28 (dd, *J* 5.4, 12.3 Hz, 1H, H-6b), 4.94 (d, *J* 4.1 Hz, 2H, CH_2), 5.38 (t, *J* 10.1 Hz, 1H, H-4), 5.53 (dd, *J* 1.8, 3.3 Hz, 1H, H-2), 5.55–5.60 (m, 2H, H-3, H-1), 7.22 (s, 1H, NH), 7.28 (t, *J* 8.2 Hz, 1H, Ar-H), 7.53 (t, *J* 7.7 Hz, 2H, Ar-H), 7.62 (d, *J* 8.6 Hz, 1H, Ar-H), 7.65 (t, *J* 7.4 Hz, 1H, Ar-H), 7.70 (dd, *J* 1.9, 11.1 Hz, 1H, Ar-H), 8.03 (d, *J* 7.5 Hz, 2H, Ar-H); ^{13}C NMR (125 MHz, $CDCl_3$): δ 20.64, 20.65, 20.66, 20.8 (4 Ac-CH_3), 46.9 (CH_2), 62.0 (C-6), 65.7 (C-4), 68.6 (C-3), 69.1 (C-2), 69.8 (C-5), 97.0 (C-1), 116.2 (d, $J_{C,F}$ 20 Hz), 118.1, 123.4 (d, $J_{C,F}$ 4 Hz), 128.0, 129.0, 129.9 (d, $J_{C,F}$ 6 Hz), 134.2, 134.4, 146.1 (d, $J_{C,F}$ 11 Hz), 152.8 (d, $J_{C,F}$ 250 Hz; 12C, Ar-C), 165.4, 169.7, 169.8, 169.9, 170.5, 194.0 (6 C=O); ESI-MS: *m/z:* Calcd for $C_{29}H_{30}FNNaO_{12}$ $[M+Na]^+$: 626.17, found: 626.24.

3-Fluoro-4-α-D-mannopyranosyloxy-*N*-(2-oxo-2-phenylethyl)benzamide (**10**)**.** Prepared according to general procedure B from **29**. Yield: 3.6 mg (60%). $[\alpha]_D^{20}$ +88.3 (*c* 0.18, MeOH); ^{1}H NMR (500 MHz, CD_3OD): δ 3.62 (ddd, *J* 2.4, 5.6, 9.8 Hz, 1H, H-5), 3.68–3.81 (m, 3H, H-4, H-6a, H-6b), 3.93 (dd, *J* 3.4, 9.5 Hz, 1H, H-3), 4.09 (dd, *J* 1.8, 3.4 Hz, 1H, H-2), 4.87 (s, 2H, CH_2), 5.62 (d, *J* 1.5 Hz, 1H, H-1), 7.50 (t, *J* 8.5 Hz, 1H, Ar-H), 7.55 (t, *J* 7.7 Hz, 2H, Ar-H), 7.66 (t, *J* 7.4 Hz, 1H, Ar-H), 7.68–7.74 (m, 2H, Ar-H), 8.04–8.08 (m, 2H, Ar-H); ^{13}C NMR (125 MHz, CD_3OD): δ 62.6 (C-6), 68.2 (C-4), 71.7 (C-2), 72.3 (C-3), 76.1 (C-5), 101.0 (C-1), 116.6 (d, $J_{C,F}$ 20 Hz), 119.0, 125.2 (d, $J_{C,F}$ 4 Hz), 129.1, 130.0, 134.9, 136.5, 148.5 (d, $J_{C,F}$ 11 Hz), 153.9 (d, $J_{C,F}$ 246 Hz; 12C, Ar-C), 168.7 (CONH), 196.3 (C=O); IR (KBr): ν 3412 (vs, OH, NH), 1646 (s, C=O) cm^{-1}; HRMS: *m/z:* Calcd for $C_{21}H_{23}FNNaO_8$ $[M+Na]^+$: 458.1227, found: 458.1227.

Supplementary Material

For Surface Plasmon Resonance Experiments and Fluorescence Polarization Assay please refer to the Supporting Information for a detailed description.

References

1. Herrmann, A. *Chem. Soc. Rev.* **2014**, *43*, 1899-1933. https://doi.org/10.1039/C3CS60336A
2. Lehn, J.-M. *Angew. Chem. Int. Ed.* **2015**, *54*, 3276-3289. https://doi.org/10.1002/anie.201409399
3. Ramström, O.; Lehn, J.-M. *Nat. Rev. Drug Discov.* **2002**, *1*, 26-36. https://doi.org/10.1038/nrd704
4. Mondal, M.; Hirsch, A. K. H. *Chem. Soc. Rev.* **2015**, *44*, 2455-2488. https://doi.org/10.1039/C4CS00493K
5. Huang, R.; Leung, I. *Molecules* **2016**, *21*, 910. https://doi.org/10.3390/molecules21070910
6. Frei, P.; Hevey, R.; Ernst, B. *Chem. Eur. J.* **2019**, *25*, 60-73. https://doi.org/10.1002/chem.201803365
7. Nasr, G.; Petit, E.; Supuran, C. T.; Winum, J.-Y.; Barboiu, M. *Bioorg. Med. Chem. Lett.* **2009**, *19*, 6014-6017. https://doi.org/10.1016/j.bmcl.2009.09.047
8. Cheeseman, J. D.; Corbett, A. D.; Shu, R.; Croteau, J.; Gleason, J. L.; Kazlauskas, R. J. *J. Am. Chem. Soc.* **2002**, *124*, 5692-5701. https://doi.org/10.1021/ja017099+

9. Corbett, A. D.; Cheeseman, J. D.; Kazlauskas, R. J.; Gleason, J. L. *Angew. Chem. Int. Ed.* **2004**, *43*, 2432-2436. https://doi.org/10.1002/anie.200453769
10. Foxman, B. *Am. J. Med.* **2002**, *113* (1, Suppl.1), 5-13. https://doi.org/10.1016/S0002-9343(02)01054-9
11. Hooton, T. M.; Roberts, P. L.; Cox, M. E.; Stapleton, A. E. *N. Engl. J. Med.* **2013**, *369*, 1883-1891. https://doi.org/10.1056/NEJMoa1302186
12. Spaulding, C.; Hultgren, S. *Pathogens* **2016**, *5*, 30. https://doi.org/10.3390/pathogens5010030
13. Justice, S. S.; Hung, C.; Theriot, J. A.; Fletcher, D. A.; Anderson, G. G.; Footer, M. J.; Hultgren, S. J. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 1333-1338. https://doi.org/10.1073/pnas.0308125100
14. Kleeb, S.; Pang, L.; Mayer, K.; Eris, D.; Sigl, A.; Preston, R. C.; Zihlmann, P.; Sharpe, T.; Jakob, R. P.; Abgottspon, D.; Hutter, A. S.; Scharenberg, M.; Jiang, X.; Navarra, G.; Rabbani, S.; Smiesko, M.; Lüdin, N.; Bezençon, J.; Schwardt, O.; Maier, T.; Ernst, B. *J. Med. Chem.* **2015**, *58*, 2221-2239. https://doi.org/10.1021/jm501524q
15. Kleeb, S.; Jiang, X.; Frei, P.; Sigl, A.; Bezençon, J.; Bamberger, K.; Schwardt, O.; Ernst, B. *J. Med. Chem.* **2016**, 59, 3163-3182. https://doi.org/10.1021/acs.jmedchem.5b01923
16. Schönemann, W.; Kleeb, S.; Dätwyler, P.; Schwardt, O.; Ernst, B. *Can. J. Chem.* **2016**, *94*, 909-919. https://doi.org/10.1139/cjc-2015-0582
17. Mayer, K.; Eris, D.; Schwardt, O.; Sager, C. P.; Rabbani, S.; Kleeb, S.; Ernst, B. *J. Med. Chem.* **2017**, *60*, 5646-5662. https://doi.org/10.1021/acs.jmedchem.7b00342
18. Mydock-McGrane, L. K.; Cusumano, Z. T.; Janetka, J. W. *Expert Opin. Ther. Pat.* **2015**, *26*, 175-197. https://doi.org/10.1517/13543776.2016.1131266
19. Mydock-McGrane, L.; Cusumano, Z.; Han, Z.; Binkley, J.; Kostakioti, M.; Hannan, T.; Pinkner, J. S.; Klein, R.; Kalas, V.; Crowley, J.; Rath, N. P.; Hultgren, S. J.; Janetka, J. W. *J. Med. Chem.* **2016**, *59*, 9390-9408. https://doi.org/10.1021/acs.jmedchem.6b00948
20. 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. *ChemMedChem* **2016**, *11*, 367-373. https://doi.org/10.1002/cmdc.201600006
21. Frei, P.; Pang, L.; Silbermann, M.; Eriş, D.; Mühlethaler, T.; Schwardt, O.; Ernst, B. *Chem. Eur. J.* **2017**, *23*, 11570-11577. https://doi.org/10.1002/chem.201701601
22. Dirksen, A.; Dirksen, S.; Hackeng, T. M.; Dawson, P. E. *J. Am. Chem. Soc.* **2006**, *128*, 15602-15603. https://doi.org/10.1021/ja067189k
23. Sauer, M. M.; Jakob, R. P.; Eras, J.; Baday, S.; Eris, D.; Navarra, G.; Berneche, S.; Ernst, B.; Maier, T.; Glockshuber, R. *Nat. Commun.* **2016**, *7*, 1-13. https://doi.org/10.1038/ncomms10738
24. Bhat, V. T.; Caniard, A. M.; Luksch, T.; Brenk, R.; Campopiano, D. J.; Greaney, M. F. *Nat. Chem.* **2010**, *2*, 490-497. https://doi.org/10.1038/nchem.658

25. Bouckaert, J.; Berglund, J.; Schembri, M.; De Genst, E.; Cools, L.; Wuhrer, M.; Hung, C.-S.; Pinkner, J.; Slättegård, R.; Zavialov, A.; Choudhury, D.; Langermann, S.; Hultgren, S. J.; Wyns, L.; Klemm, P.; Oscarson, S.; Knight, S. D.; De Greve, H. *Mol. Microbiol.* **2005**, *55*, 441-455. https://doi.org/10.1111/j.1365-2958.2004.04415.x
26. Ladame, S. *Org. Biomol. Chem.* **2008**, *6*, 219-226. https://doi.org/10.1039/B714599C
27. Poulsen, S.-A. *J. Am. Soc. Mass Spectrom.* **2006**, *17*, 1074-1080. https://doi.org/10.1016/j.jasms.2006.03.017
28. Clipson, A. J.; Bhat, V. T.; McNae, I.; Caniard, A. M.; Campopiano, D. J.; Greaney, M. F. *Chem. Eur. J.* **2012**, *18*, 10562-10570. https://doi.org/10.1002/chem.201201507
29. Mondal, M.; Radeva, N.; Köster, H.; Park, A.; Potamitis, C.; Zervou, M.; Klebe, G.; Hirsch, A. K. H. *Angew. Chem. Int. Ed.* **2014**, *53*, 3259-3263. https://doi.org/10.1002/anie.201309682
30. Mondal, M.; Radeva, N.; Fanlo-Virgós, H.; Otto, S.; Klebe, G.; Hirsch, A. K. H. *Angew. Chem. Int. Ed.* **2016**, *55* (32), 9422-9426. https://doi.org/10.1002/anie.201603074
31. Vincent, S.; Fu, J.; Fu, H.; Dieu, M.; Halloumi, I.; Kremer, L.; Xia, Y.; Pan, W. *Chem. Commun.* **2017**, 53 (77), 10632-10635. https://doi.org/10.1039/C7CC05251K
32. Carolina, D. D.; Eliezer, J. B.; Carlos, A. M. F. *Mini-Rev. Med. Chem.* **2007**, *7*, 1108-1119.
33. Yu, X.; Shi, L.; Ke, S. *Bioorg. Med. Chem. Lett.* **2015**, *25*, 5772-5776. https://doi.org/10.1016/j.bmcl.2015.10.069
34. Misra, S.; Ghatak, S.; Patil, N.; Dandawate, P.; Ambike, V.; Adsule, S.; Unni, D.; Venkateswara Swamy, K.; Padhye, S. *Bioorg. Med. Chem.* **2013**, *21*, 2551-2559. https://doi.org/10.1016/j.bmc.2013.02.033
35. Barman, S.; You, L.; Chen, R.; Codrea, V.; Kago, G.; Edupuganti, R.; Robertus, J.; Krug, R. M.; Anslyn, E. V. *Eur. J. Med. Chem.* **2014**, *71* (Suppl. C), 81-90. https://doi.org/10.1016/j.ejmech.2013.10.063
36. Burgeson, J. R.; Gharaibeh, D. N.; Moore, A. L.; Larson, R. A.; Amberg, S. M.; Bolken, T. C.; Hruby, D. E.; Dai, D. *Bioorg. Med. Chem. Lett.* **2013**, *23*, 5840-5843. https://doi.org/10.1016/j.bmcl.2013.08.103
37. Chen, C.; Dolla, N. K.; Casadei, G.; Bremner, J. B.; Lewis, K.; Kelso, M. J. *Bioorg. Med. Chem. Lett.* **2014**, *24*, 595-600. https://doi.org/10.1016/j.bmcl.2013.12.015
38. Pieczonka, A. M.; Strzelczyk, A.; Sadowska, B.; Mlostoń, G.; Stączek, P. *Eur. J. Med. Chem.* **2013**, *64* (Suppl. C), 389-395. https://doi.org/10.1016/j.ejmech.2013.04.023
39. Maia, R. d. C.; Tesch, R.; Fraga, C. A. M. *Expert Opin. Ther. Pat.* **2014**, *24*, 1161-1170. https://doi.org/10.1517/13543776.2014.959491
40. Sonawane, S. J.; Kalhapure, R. S.; Govender, T. *Eur. J. Pharm. Sci.* **2017**, *99* (Suppl. C), 45-65. https://doi.org/10.1016/j.ejps.2016.12.011
41. Jumde, V. R.; Mondal, M.; Gierse, R. M.; Unver, M. Y.; Magari, F.; van Lier, R. C. W.; Heine, A.; Klebe, G.; Hirsch, A. K. H. *ChemMedChem* **2018**, *13*, 2266-2270.

https://doi.org/10.1002/cmdc.201800446

Supporting Information

Comparison of Affinity Ranking by Target-directed Dynamic Combinatorial Chemistry and Surface Plasmon Resonance

Priska Frei, Marleen Silbermann, Tobias Mühlethaler, Xiaohua Jiang, Oliver Schwardt, Rachel Hevey, and Beat Ernst

Supporting information for this article can be found under:

https://www.arkat-usa.org/get-file/66510/

Manuscript 2

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

Pascal Zihlmann,[1)] Xiaohua Jiang,[1)] Christoph P. Sager,[1)] Brigitte Fiege,[1)] Roman P. Jakob,[2)] Stefan Siegrist,[1)] Adam Zalewski,[1)] Said Rabbani,[1)] Deniz Eriş,[1)] Marleen Silbermann,[1)] Lijuan Pang,[1)] Tobias Mühlethaler,[1)] Timothy Sharpe,[3)] Timm Meier,[2)] and Beat Ernst[1)*]

[1)] Institute of Molecular Pharmacy, University of Basel, Klingelbergstr. 50, 4056 Basel, Switzerland

[2)] Structural Biology, University of Basel, Klingelbergstr. 70, 4056 Basel, Switzerland

[3)] Biophysics Facility, University of Basel, Klingelbergstr. 70, 4056 Basel, Switzerland

[*] Corresponding author
Tel.: 0041 61 267 15 51; Fax: 0041 61 207 15 52
E-mail: beat.ernst@unibas.ch

Contributions of Marleen Silbermann

- Kinetic measurement by SPR
- Cloning, expression, purification, and biotinylation of $FimH_{LD}$-AVITag-6His

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.[1-6] 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 α-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 $\Delta\Delta G°$ 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.[7] 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.[8] The high specificity allows carbohydrates to fulfill their broad biological tasks, such as signal transduction,[9,10] cell recognition[11,12] or cell adhesion.[13] Furthermore, considering their fundamental importance in numerous disease-related processes, carbohydrate mimetic drugs offer potential new therapeutic applications.[14] However, the polar character of carbohydrates creates pharmacokinetic challenges related to oral availability, plasma half-live 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"[7] and Hopkins *et al.* regarded the likelihood of modulating a lectin with an orally available small molecule drug to be very low.[15] Therefore, to analyze and solve the structural drawbacks common to carbohydrates is of fundamental importance when therapeutic applications are envisaged.

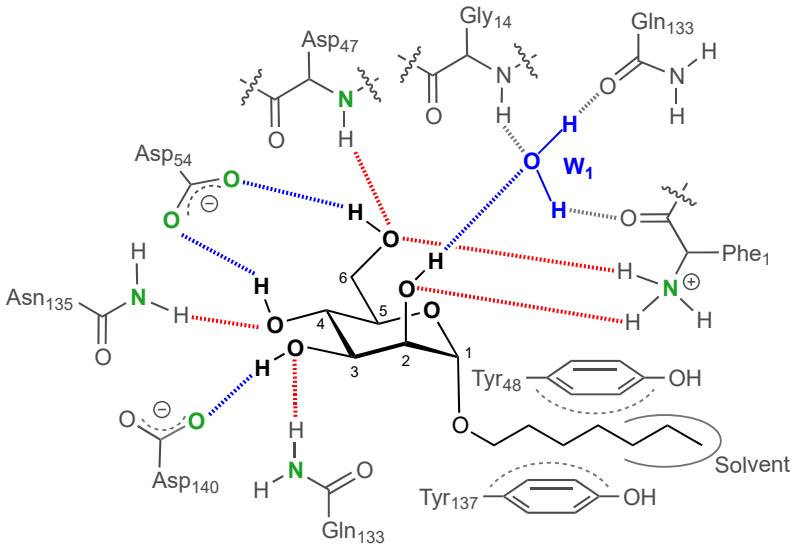
In the present study, the bacterial lectin FimH - one of the rare examples of a lectin undergoing high-affinity interactions with carbohydrates - is analyzed and the role of the individual hydroxyl groups in terms of thermodynamic and kinetic contribution to binding is studied. FimH, a virulence factor of uropathogenic *E. coli* (UPEC), is located at the tip of bacterial type 1 pili.[16,17] 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.[18,19]

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 α-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.[6,20-29]

Results and Discussion

The mannose-binding pocket of FimH (Figure 1). The interaction of α-D-mannosides with the lectin domain of FimH ($FimH_{LD}$) 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,[30] 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 μM.[24] 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.[31-35] This effect leads to a further 100-fold improvement of affinity as documented by *n*-heptyl α-D-mannoside (**1**, K_D: 22 nM).[36]



Acceptor	Donor	Distance [Å]	Angle Donor	Angle Acceptor
O (2-OH)	N (N-terminus)	2.70	161.8	127.2
O (H_2O)	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 α-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 $FimH_{LD}$ in complex with *n*-heptyl α-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 α-D-mannopyranosides with $FimH_{LD}$ (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.[29,37] 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.[38] 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).[39]

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[40] might be lowered by perturbed water molecules in the solvation shell of the interaction surfaces due to their polyamphiphilic character.[29] However, the high desolvation penalty cannot be compensated by the formation of single hydrogen bond.[41] Desolvation costs ($\Delta G°_{solv}$) for the ligand *n*-heptyl α-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 α-D-mannoside was calculated to cost 38.8 kJ/mol, which is almost as much as the entire binding energy between *n*-heptyl α-D-mannoside and $FimH_{LD}$ ($\Delta G°$ = -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’.[42-45] 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.[46,47] 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.[48] 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 $FimH_{LD}$, which accommodates also 7-membered ring analogs of D-mannose without any conformational adaption.[49] In order to demonstrate the rigidity of $FimH_{LD}$, the "pseudo-apo" structure (PDB-Code: 4AUU, containing thioethanol in the binding site) and the crystal structure co-crystallized with *n*-heptyl α-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 α-D-mannoside (**1**) in the bioactive conformation, torsion angles of solution structures obtained by molecular dynamics simulations (ω_{MD}) were compared with its structure ($\omega_{X\text{-}ray}$) in complex with $FimH_{LD}$ (PDB-Code: 4XO8). To classify the energy profile of each hydroxyl group, they were rotated stepwise around their torsion angles (ω_1-ω_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 ($\omega_{1\text{-}3,\,5}$).

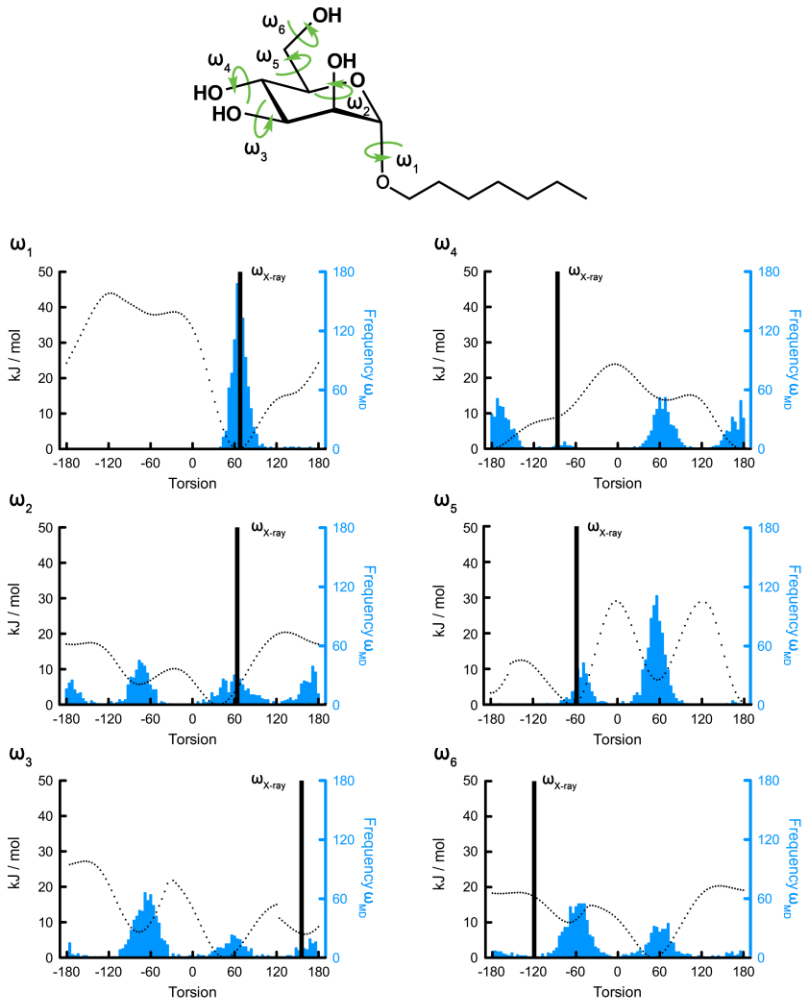
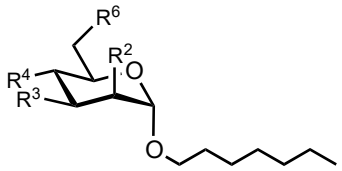


Figure 2. Calculated energy as a function of the torsion angles of methyl α-D-mannopyranoside. Dotted lines correspond to the energy of a given torsion angle. Left y-axis: calculated energy in kJ/mol, right y-axis: frequency of solution conformation from MD simulation, x-axis: torsion angle, conformation of heptyl α-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 $\omega_{1\text{-}6}$ of *n*-heptyl α-D-mannopyranoside (1) in solution and bound to FimH$_{LD}$. Torsion angles ($\omega_{X\text{-}ray}$) 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 ($\omega_{X\text{-}ray}$)	Energy [kJ/mol]	Energy [kJ/mol]
ω_1	1.5	68.5°	0.1	- 1.4
ω_2	9.9	64.7°	4.6	- 5.3
ω_3	8.8	155.4°	6.8	- 2.0
ω_4	6.8	-86.9°	9.0	+ 2.2
ω_5	8.2	-57.4°	0.4	- 7.8
ω_6	9.4	-119.8°	17.3	+ 7.9

Synthesis of deoxy- and deoxy-halogeno-derivatives of *n*-heptyl α-D-mannosides. To study the contribution of the individual hydroxyl groups, a series of deoxy- and deoxy-halogeno-derivatives of *n*-heptyl α-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,[50,51] 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 $FimH_{LD}$ were used (D-Mannose: K_D = 2.3 μM; *n*-heptyl α-D-mannoside (**1**): K_D = 22 nM).[30,36] 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.[36]



1: R^2, R^3, R^4, R^6 = OH
2: **R^2 = F**, R^3, R^4, R^6 = OH
3: **R^2 = Cl**, R^3, R^4, R^6 = OH
4: **R^2 = Br**, R^3, R^4, R^6 = OH
5: **R^2 = H**, R^3, R^4, R^6 = OH
6: **R^3 = F**, R^2, R^4, R^6 = OH
7: **R^3 = H**, R^2, R^4, R^6 = OH
8: **R^4 = F**, R^2, R^3, R^6 = OH
9: **R^4 = H**, R^2, R^3, R^6 = OH
10: **R^6 = F**, R^2, R^3, R^4 = OH
11: **R^6 = H**, R^2, R^3, R^4 = OH

Figure 3. Structure of *n*-heptyl α-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 α-D-mannoside (1) bound to $FimH_{LD}$. To analyze how the substitution of an individual hydroxyl group influences ligand and lectin conformation, the structures **2** & **5** - **9** co-crystallized with $FimH_{LD}$ were determined applying conditions we previously reported for *n*-heptyl α-D-mannoside (**1**).[32,36] 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).

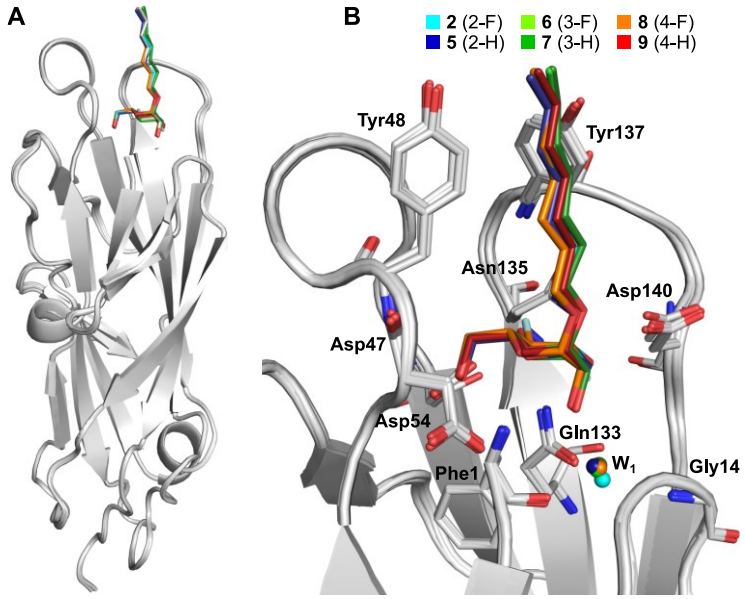


Figure 4. Superimposition of co-crystal structures of FimH$_{LD}$ with *n*-heptyl α-D-mannoside and six deoxy- and deoxy-fluoro derivatives. A) The parent *n*-heptyl α-D-mannoside (**1**) (PDB code 4XO8) and all proteins are shown in grey, while the deoxy- (**5**, **7** & **9**) and the deoxy-fluoro-derivatives (**2**, **6** & **8**) are colored as indicated. B) Important binding pocket residues and the structural water (W_1) 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_1 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 $^{1}H,^{15}N$-HSQC NMR spectra were recorded. Whereas ^{1}H chemical shifts of backbone and side chain amides report on H-bond formations, ^{15}N shifts respond sensitively to changes in the dihedral angles of the protein backbone and side chains.[52] We measured $^{1}H,^{15}N$-HSQC fingerprint spectra of FimH$_{LD}$ 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 α-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 $^{1}H,^{15}N$-HSQC NMR spectrum of **10** (6-F) it was apparent that **10** (6-F) binds very weakly to FimH$_{LD}$, 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^N (W_1 H-bond to 2-OH), Asp47 H^N (to 6-OH), Gln133 $H^{\varepsilon 22}$ (to 3-OH) and Asn135 $H^{\delta 21}$ (to 4-OH) (Figure 5).

The chemical shifts of Gly14 H^N are almost identical for **5** (2-H), **6** (3-F), **8** (4-F) and **9** (4-H), indicating a similar orientation of W_1, even though for **5** (2-H) the interacting hydroxyl group does not exist (Figure 5). The slight proton upfield shift of Gly14 H^N with all ligands indicates that the coordination of mannose weakens the H-bond of this residue to W_1. In agreement with this, the data suggest the strongest Gly14-water interaction for **5** (2-H), for which W_1 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_1 is more buried (Figure 4B) which may contribute to the large chemical shift change of Gly14. The relative chemical shift changes of Asp47 H^N 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 $H^{\varepsilon 22}$ 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 $H^{\varepsilon 22}$ 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 $H^{\delta 21}$ 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 $FimH_{LD}$ as a model to investigate contributions of distinct hydroxyl groups by deoxy- and deoxy-halogeno derivatives.

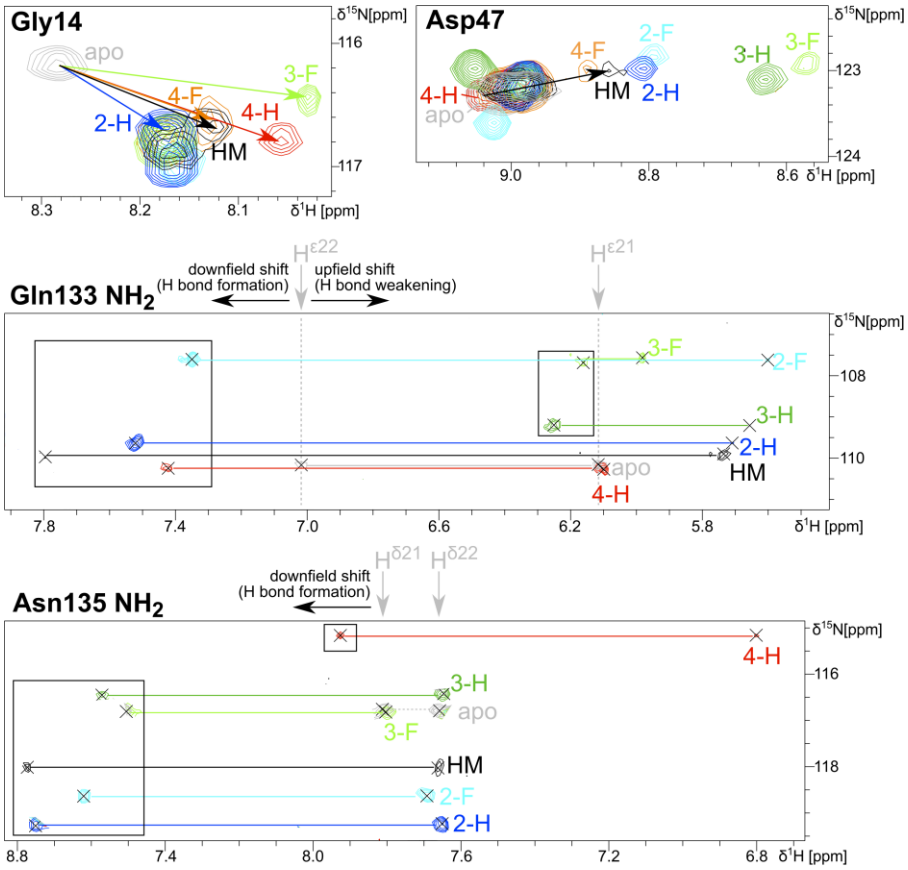


Figure 5. $^{1}H,^{15}N$-HSQC spectral regions of binding pocket residues directly involved in H-bond donation to the mannosyl moiety. Boxes in case of side chain signals of Gln133 and Asn135 indicate relative downfield and upfield proton shifts, respectively. Peaks not of interest were faded-out to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

Thermodynamic analysis of deoxy- and deoxy-halogeno-derivatives of *n*-heptyl α-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 ($\Delta H°$), 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 ($\Delta H°_{ion}$) originating from a partial proton

transfer from the solvent to the N-terminus of the $FimH_{LD}$ protein (Figure S4, Table S3).[53,54] By measuring the enthalpy of *n*-heptyl α-D-mannoside (**1**) binding to $FimH_{LD}$ 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_a 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 ΔΔ-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 $FimH_{LD}$ binding to *n*-heptyl α-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 [μM]	**ITC: K_D** [μM]	**ΔG°** [kJ/mol]	**ΔH°** [kJ/mol]	**-TΔS°** [kJ/mol]
n-heptyl mannoside (**1**)	0.028[32]	0.022	-43.7	-50.5	6.7
			ΔΔG° [kJ/mol]	**ΔΔH°** [kJ/mol]	**-TΔΔS°** [kJ/mol]
2 (2-F)	0.44	0.53	7.9	22.7	-14.8
3 (2-Cl)	2.26	1.34	10.2	25.9	-15.8
4 (2-Br)	4.24	1.83	11.0	29.6	-18.7
5 (2-H)	13.95	9.77	15.1	31.1	-16.0
6 (3-F)	0.45	1.19	9.9	19.1	-9.2
7 (3-H)	23.35	19.88	16.9	25.7	-8.8
8 (4-F)	135.2	103.43	21.0	29.9	-9.0
9 (4-H)	101.2	88.40	20.6	30.6	-10.0
10 (6-F)	n.d.	>1000			
11 (6-H)	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. ΔΔG° = 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** (ΔΔH° = 31.1 kJ/mol), which is partially compensated by a gain in entropy (-TΔΔS° = −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 α-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_1 are lost. The increasing electronegativity of bromine < chlorine < fluorine induces a repulsion of the structural water W_1 (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) ($\Delta\Delta G°$ = 15.1 kJ/mol) and the loss of only the H-bonds of **2** (2-F) ($\Delta\Delta G°$ = 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 α-D-mannoside (**1**) (→ **7**) leads to a substantial loss in enthalpy ($\Delta\Delta H°$ = 25.7 kJ/mol), which is only partially compensated by an entropy gain ($-T\Delta\Delta S°$ = -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 [→ **6** (3-F)] results in an enthalpic gain ($\Delta\Delta H°$ = -6.6 kJ/mol), as well as a small entropic advantage ($-T\Delta\Delta S°$ = -0.4 kJ/mol). Interestingly, crystal structures of *n*-heptyl α-D-mannoside (**1**) and the 3-fluoro derivative **6** co-crystallized with $FimH_{LD}$ (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) ($\Delta\Delta G°$ = 16.9 kJ/mol) and the loss of only the H-bonds to Gln133 and Asp140 of **6** (3-F) ($\Delta\Delta G°$ = 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.[55-57]

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.[58] 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 β-sheet. Crystal structures reveal a change of the orientation between 4-OH and 4-F (dihedral angle HO4-C4-C5-O = -174.9° and F4-C4-C5-O = -167.3°), which supports this assumption. Therefore, the loss of binding energy upon the removal of 4-OH [→ **9** (4-H), $\Delta\Delta G°$ = 20.6 kJ/mol] is almost identical resulting from the loss of the hydrogen bonds [→ **8** (4-F), $\Delta\Delta G°$ = 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 $FimH_{LD}$ 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-α-D-mannosides (Table 3). With kinetic isothermal titration calorimetry (kinITC) kinetic data are accessible from ITC measurements.[59] 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_{off}. The association rate constant k_{on} is calculated from k_{off} and the equilibrium constant ($K_D = k_{off} / k_{on}$) (Table 3).

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

Compound	Method	k_{on} [$M^{-1}s^{-1}$]	$1/rk_{on}$	k_{off} [s^{-1}]	rk_{off}	$t_{1/2}$ [min]
1 (HM)	ITC	$3.32*10^4$	1.0	$7.27*10^{-4}$	1.0	15.90
2 (2-F)	ITC	$2.04*10^4$	1.6	$1.02*10^{-2}$	14.0	1.14
3 (2-Cl)	ITC	$8.90*10^3$	3.7	$1.19*10^{-2}$	16.4	0.97
4 (2-Br)	ITC	$7.90*10^3$	4.2	$1.45*10^{-2}$	19.9	0.80
5 (2-H)	ITC	$5.35*10^3$	6.2	$4.13*10^{-2}$	56.8	0.28
5 (2-H)	SPR	$1.05*10^4$	3.1	$5.08*10^{-2}$	71.1	0.23
6 (3-F)	ITC	$1.66*10^4$	2.0	$1.98*10^{-2}$	27.3	0.58
7 (3-H)	ITC	$6.50*10^3$	5.1	$1.29*10^{-1}$	177.9	0.09
8 (4-F)	ITC	$1.00*10^3$	33.2	$1.03*10^{-1}$	142.3	0.11
9 (4-H)	ITC	$9.56*10^2$	34.7	$8.45*10^{-1}$	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, $FimH_{LD}$ is characterized by unusually slow dissociation rates.[60] The reduced affinity of all deoxy- and deoxy-halogeno derivatives is mainly resulting from increased k_{off} values, leading to substantially reduced complex half-lives ($t_{1/2} = \ln 2/k_{off}$). For example, while $t_{1/2}$ of *n*-heptyl α-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_{on} than k_{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_{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_{on} values to a significantly greater extent than k_{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.[61-64]

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. $FimH_{LD}$ 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 ($\Delta\Delta H°$ = 26 – 31 kJ/mol) and a smaller gain in entropy ($-T\Delta\Delta S°$ = -9 – -16 kJ/mol). Hence, contributions to $\Delta G°$ 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 $\Delta G°$ 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 $\Delta G°$ of *n*-heptyl α-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_{off} than on k_{on}. However, the short-range electrostatic H-bonds were found to be less important for k_{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 $FimH_{LD}$ 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., & Quiocho, 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. Quiocho, 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. Abgottspon, 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. Quiocho, 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^2.

	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
Ligand donor_delta %	**100%**	**100%**	**100%**	**100%**	**100%**	**100%**	**100%**
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
Ligand acceptor_delta %	**99%**	**100%**	**97%**	**100%**	**98%**	**100%**	**100%**
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
Ligand total_delta %	**75%**	**75%**	**75%**	**75%**	**68%**	**76%**	**80%**
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
Receptor total_delta %	**37%**	**37%**	**35%**	**37%**	**42%**	**38%**	**32%**

Table S2. Statistics on diffraction data and refinement of $FimH_{LD}$ and its ligand complexes.

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

*The values in parentheses correspond to the highest resolution shell

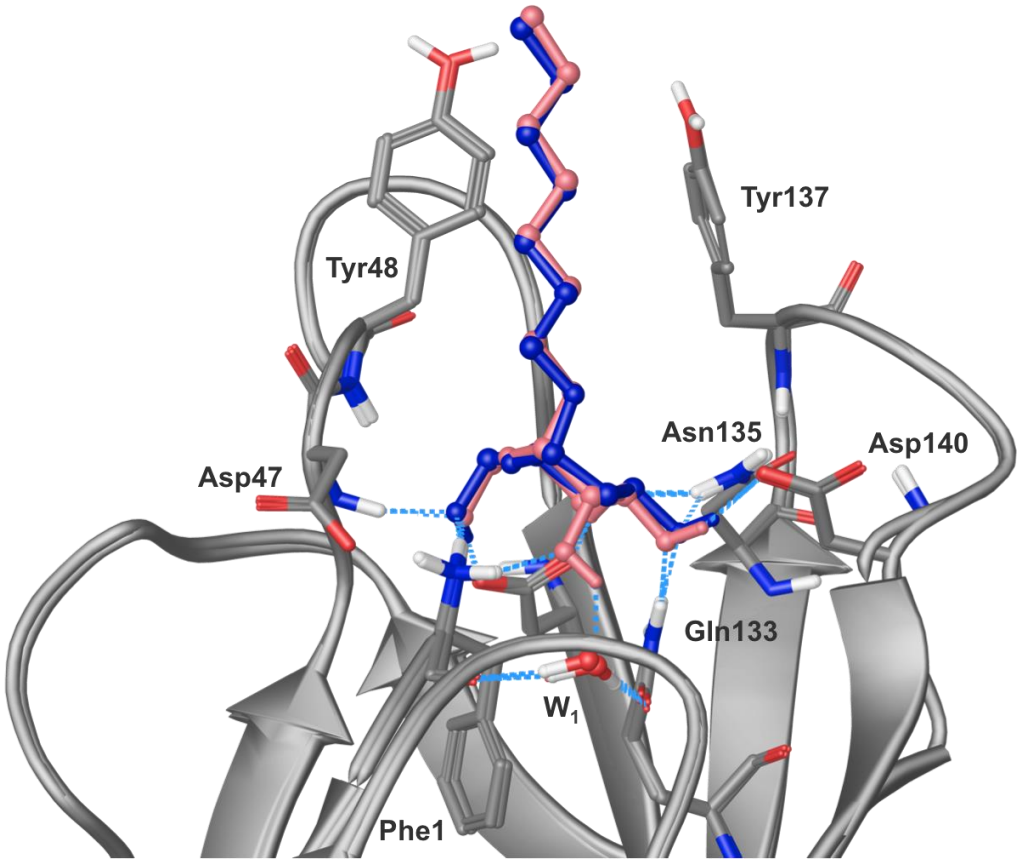


Figure S1. Superimposition of X-ray co-crystal structures of $FimH_{LD}$ with HM and 2-H. Both proteins are shown in grey, the ligands are shown in pink **1** (HM) and blue **5** (2-H). The missing contact of 2-OH to the structural water molecule (W_1) and the positively charged N-terminus of **5** slightly 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.

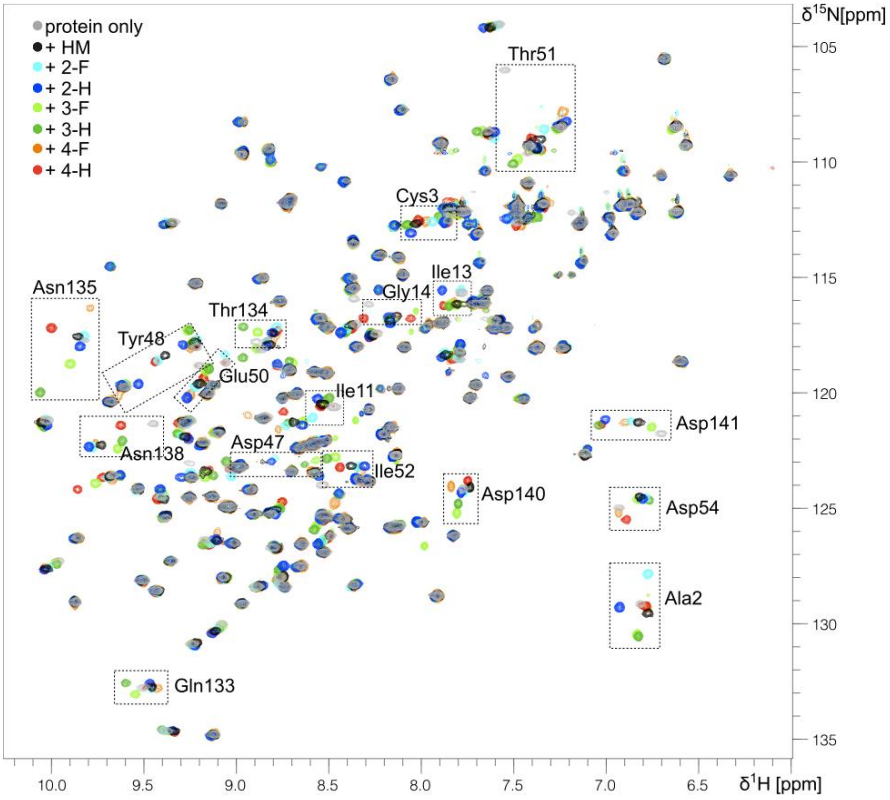


Figure S2. Overlap of $^{1}H,^{15}N$-HSQC spectra of FimH in absence of ligand (grey) and in presence of **1** (HM, black), **2** (2-F, cyan), **5** (2-H, blue), **6** (3-F, light green), **7** (3-H, green), **8** (4-F, orange) and **9** (4-H, red). The spectrum in presence of **10** (6-F) is not shown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

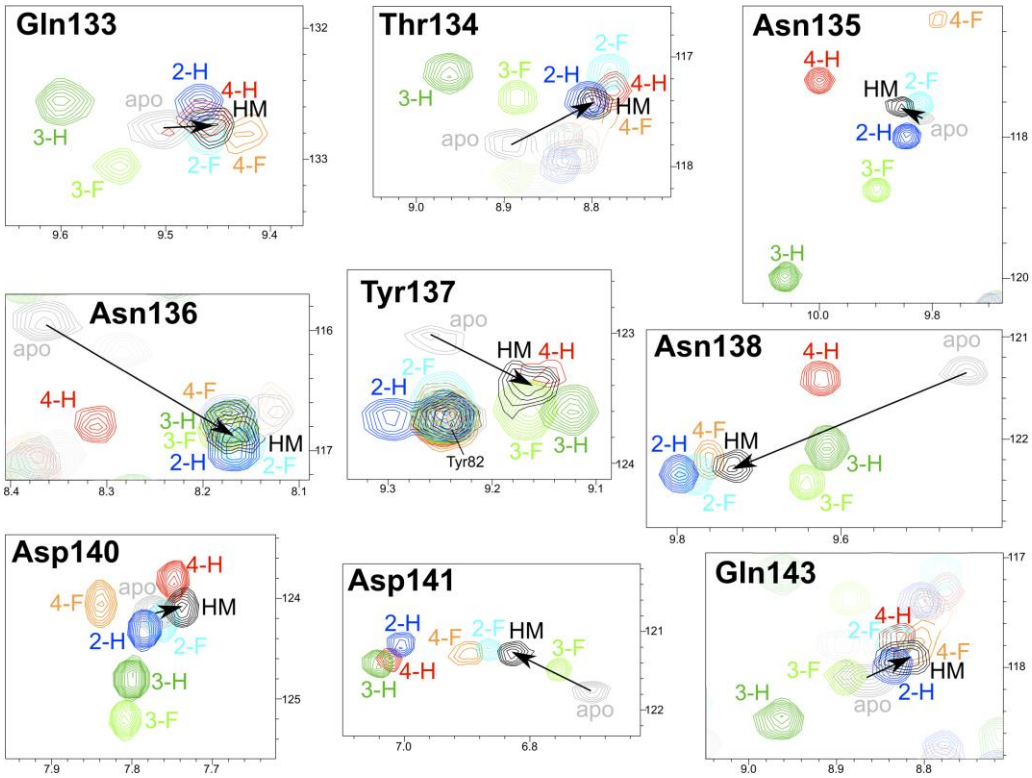


Figure S3. $^{1}H,^{15}N$-HSQC spectral regions of the backbone signals of Gln133 to Gln143.

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 [µM]	Protein [µM]	K_D [nM]	$\Delta G°_{obs}$ [kJ/mol]	$\Delta H°_{obs}$ [kJ/mol]	$-T\Delta S°$ [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

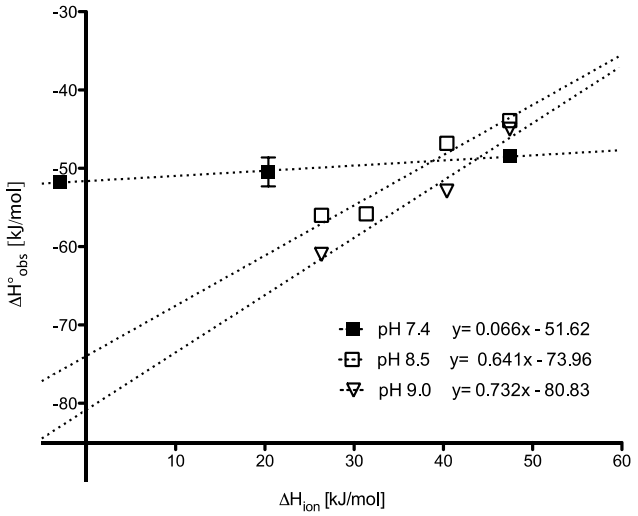


Figure S4. pH-dependent protonation of the FimH N-terminus. Heat of ionization ($\Delta H°_{ion}$) is buffer dependent (Cacodylate, -3 kJ/mol; HEPES, +20.4 kJ/mol; Bicine, +26.34 kJ/mol; Tricine +31.37 kJ/mol; TAPS +40.40 kJ/mol; Tris, +47.45 kJ/mol). The enthalpy ($\Delta H°_{obs}$) of HM binding to 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 $\Delta H°_{obs} = \Delta H°_{int} + nH^+ \cdot \Delta H°_{ion}$, where the steepness of the slope reflects the number of protons taken up or released by this binding event (nH^+) and $\Delta H°_{int}$ 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-Hasselbach 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 α-D-mannopyranoside mannoside (**1**) binding to $FimH_{LD}$ in varying buffers and pH. Measurements were analyzed using AFFINImeter software.

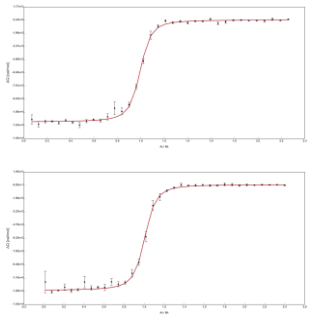
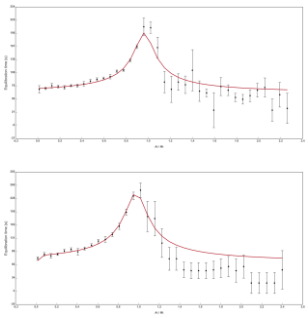
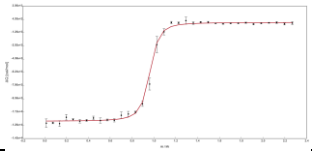
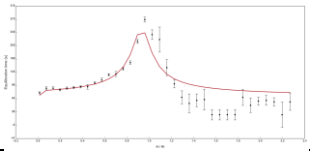
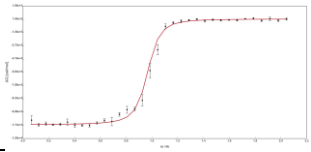
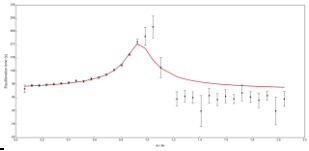
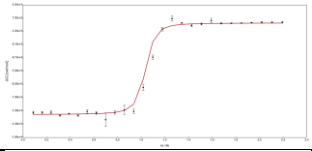
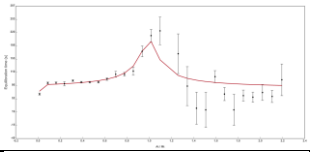
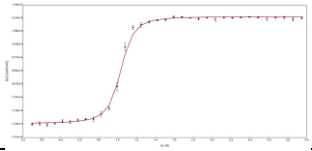
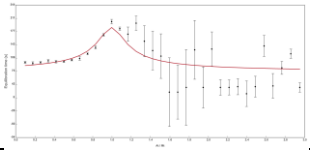
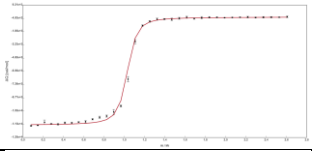
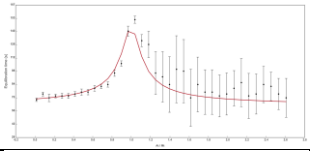
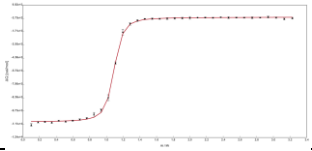
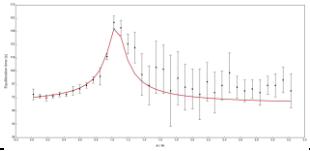
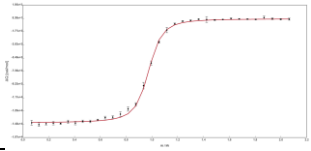
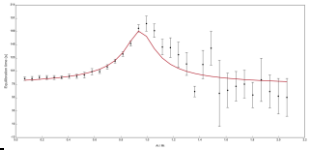
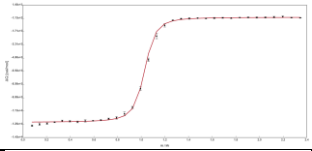
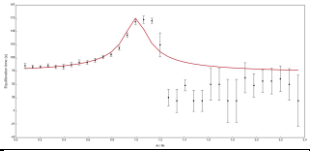
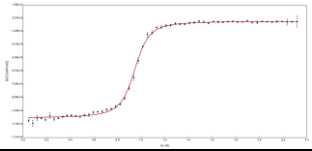
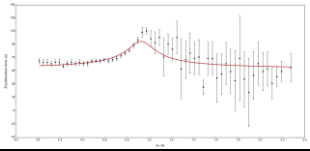
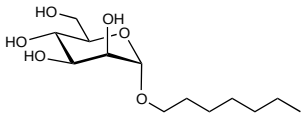
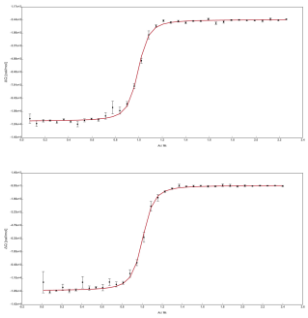
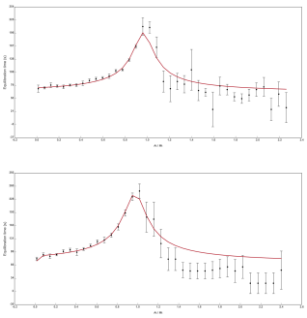
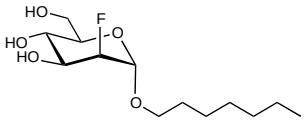
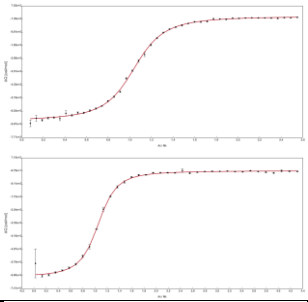
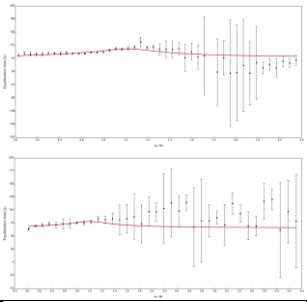
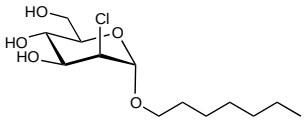
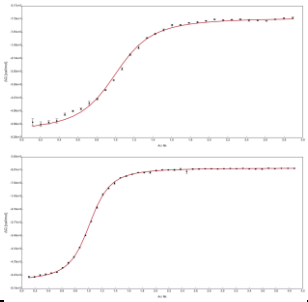
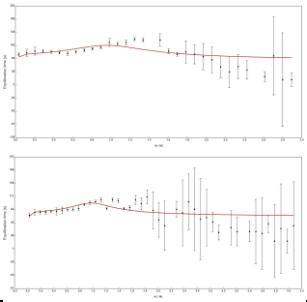
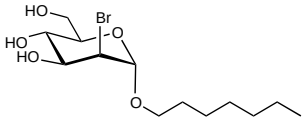
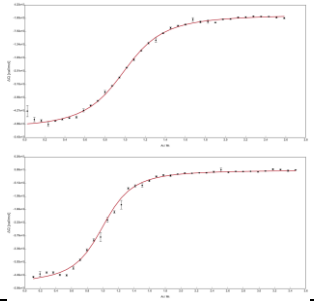
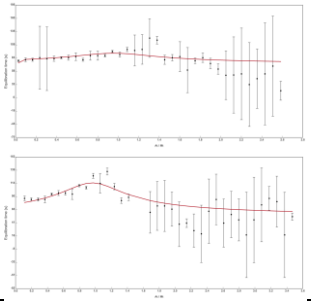
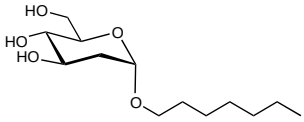
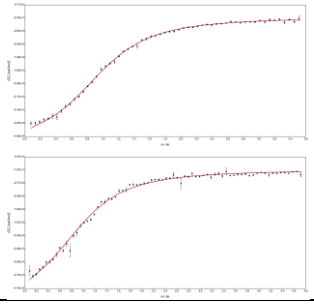
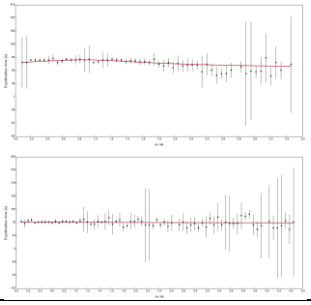
Buffer / pH	Thermodynamic fit	Kinetic fit
Hepes pH 7.4		
Cacodylate pH 7.4		
Tris pH 7.4		
Bicine pH 8.5		
Tricine pH 8.5		
TAPS pH 8.5		
Tris pH 8.5		
Bicine pH 9.0		
TAPS pH 9.0		
Tris pH 9.0		

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 [µM]	Protein [µM]	K_D [µM]	$\Delta G°$ [kJ/mol]	$\Delta H°$ [kJ/mol]	$-T\Delta S°$ [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)			**0.022**	**-43.7**	**-50.5**	**6.7**		
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)			**0.532**	**-35.8**	**-27.7**	**-8.1**		
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)			**1.337**	**-33.5**	**-24.5**	**-9.0**		
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)			**1.830**	**-32.7**	**-20.8**	**-11.9**		
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)			**9.771**	**-28.6**	**-19.3**	**-9.2**		
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)			**1.194**	**-33.8**	**-31.4**	**-2.4**		
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)			**19.878**	**-26.8**	**-24.8**	**-2.1**		
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)			**103.425**	**-22.7**	**-20.5**	**-2.2**		
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)			**88.402**	**-23.1**	**-19.8**	**-3.3**		

Table S6: Structure, and thermodynamic and kinetic fits of *n*-heptyl α-D-mannopyranoside mannoside (**1**) and its deoxy- and deoxy-halogeno derivatives (**2-9)** binding to $FimH_{LD}$. 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			

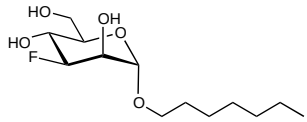
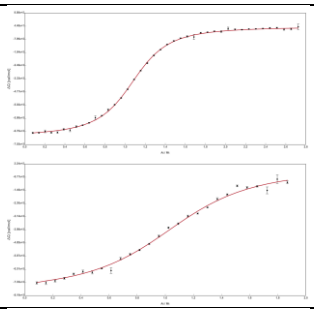
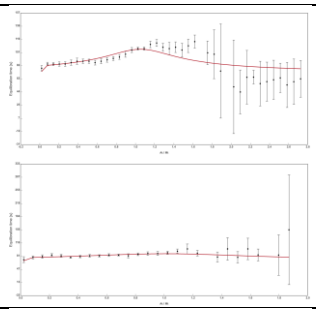
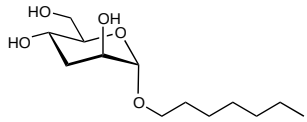
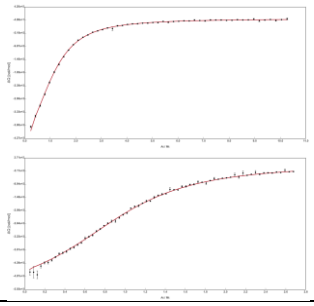
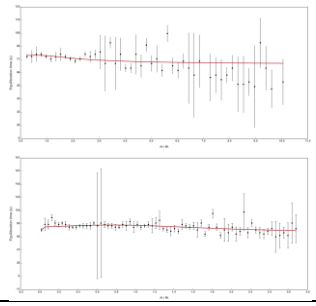
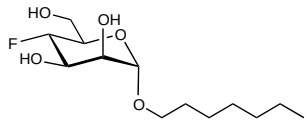
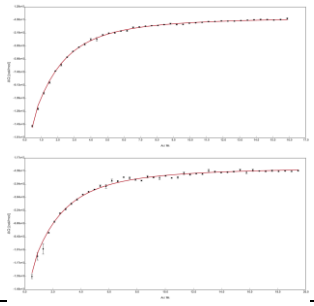
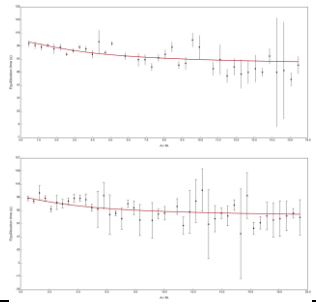
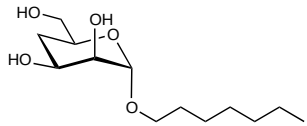
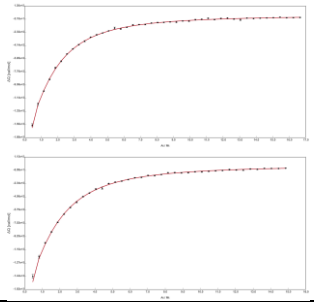
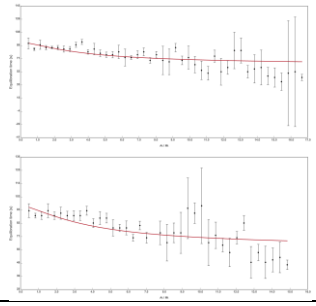
6			
7			
8			
9			

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

Experimental Part

Protein preparation. $FimH_{LD}$ from *E.coli* K-12 strain was expressed with a C-terminal thrombin cleavage site and a 6His-tag ($FimH_{LD}$-Th-6His, 173 residues) following a previously published protocol.[1] Briefly, the clone containing the $FimH_{LD}$ 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_{600} 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 $CaCl_2$. For long time storage the protein was frozen at -80°C. For the production of uniformly ^{15}N-labeled $FimH_{LD}$-Th-6His for NMR experiments, *E.coli* HM125 was cultivated in M9 minimal medium containing 1 g/L $^{15}NH_4Cl$ (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 $FimH_{LD}$ (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_Ds 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[2] and analyzed using Prism (GraphPad Software, U.S.A.). The K_D of **GNFP4** and FimH was defined as constraint and set to 1.7 nM during curve fitting.[3]

Co-crystallization, data processing, and structure refinement. For crystallization, $FimH_{LD}$ (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 $(NH_4)_2SO_4$, 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.[4,5] Structures were solved by molecular replacement with PHASER[6] using the $FimH_{LD}$-*n*-heptyl α-D-mannopyranoside complex (PDB code 4XO8) as search model. The structures were built using the COOT software[7] and periodically refined with the PHENIX and Buster-TNT software.[8,9] Geometric restraints for the ligands were generated with PRODRG[10] and Molprobity[11] 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 $FimH_{LD}$ 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 form 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 α-D-mannoside from a conformational analysis with MacroModel 10.6[12] 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.[13] 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[14-18] 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.[19-21]

NMR experiments. $^{1}H,^{15}N$-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 ^{15}N-labeled $FimH_{LD}$ in 20 mM phosphate buffer pH 7.4 in water with 7% D_2O. Deoxy- and deoxy-fluoro derivatives were solved in H_2O 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).[22] The backbone assignment of $FimH_{LD}$ was available from previous studies.[23] Combined chemical shift differences, $\Delta\delta_{AV}$, between free and ligand-bound protein signals were calculated as in equation 1[24]

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

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 µl, a reference power of 10 µcal/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, $FimH_{LD}$-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 M^{-1} cm^{-1}.[25] 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 $\Delta H°$ (change in enthalpy), the kinetic parameters k_{on} (association rate constant) and k_{off} (dissociation rate constant) as well as N (stoichiometry), were performed using the fully automated analysis software AFFINImeter.[26,27] 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. $\Delta G°$ (free energy of binding) and $\Delta S°$ (change in entropy) were calculated from equation 2

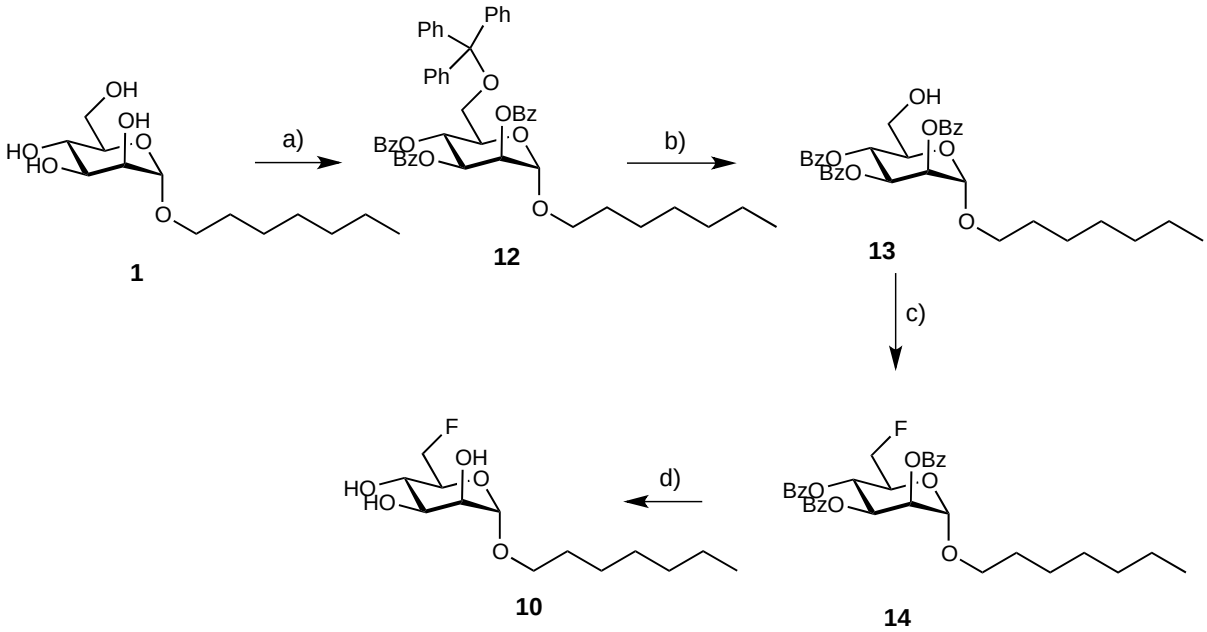
$$\Delta G° = \Delta H° - T\Delta S° = -RT\ln K_A \qquad \text{(eq. 2)}$$

with T being the absolute temperature and R the universal gas constant (8.314 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 (CH_2Cl_2) was dried by filtration through Al_2O_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}H) or 125.76 MHz (^{13}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 (singulet), d (doublet), dd (doublet of a doublet), t (triplet), q (quartet) or m (multiplet). Assignment of the ^{1}H and ^{13}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_{254} 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% H_2SO_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 μm).

Synthesis of 1 (6-F) (*n*-Heptyl 6-Fluoro-6-deoxy-α-D-mannopyranoside)



Scheme S1. a) i. Trityl chloride, pyridine, DMAP, 80°C, overnight, ii. benzoyl chloride, two steps in one pot, (98%); b) $FeCl_3$, DCM, rt, overnight, (87%); c) DAST, DCM, rt, overnight, (45%); d) 0.5 M CH_3ONa/CH_3OH, (80%).

n-Heptyl 2,3,5-tri-O-benzoyl-6-O-triphenylmethyl-α-D-mannopyranoside (12)[28]

To a solution of **1** (310 mg, 1.114 mmol) in pyridine (4.0 mL) was added 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 $NaHCO_3$, water, brine and dried over Na_2SO_4. 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%). ^{1}H-NMR (500 MHz, $CDCl_3$): 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-OC*H*$_2$C$_6$H$_{13}$), 3.58 (dt, *J* = 9.5, 7.0 Hz, 1H, H-OC*H*$_2$C$_6$H$_{13}$), 3.37 (dd, *J* = 10.5, 2.0 Hz, 1H,H-6a), 3.27 (dd, *J* = 10.5, 5.0 Hz, 1H,H-66), 1.71 (m, 2H), 1.43-1.32 (m, 8H), 0.90 (t, *J* = 7.0 Hz, 3H). ^{13}C-NMR (125 MHz, $CDCl_3$): 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 (O*C*H$_2$C$_6$H$_{13}$), 67.08 (C-4), 62.31 (C-6), 31.82, 29.45, 29.14, 26.16, 22.64, 14.12 (OCH$_2$*C*$_6$H$_{13}$). ESI-MS Calcd for $[M+Na]^+$, 855.35, found 855.28.

_n_-Heptyl 2,3,5-tri-O-benzoyl-α-D-mannopyranoside (13)[29]

To a solution of **12** (583 mg, 0.7 mmol) in DCM (10 mL) was added $FeCl_3$ (hydroscopic 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_2SO_4. 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%). $[\alpha]_D^{20}$ = -126.79 (c = 0.60, DCM). ^{1}H-NMR (500 MHz, $CDCl_3$): 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-OC*H*$_2$C$_6$H$_{13}$), 3.54 (dt, *J* = 9.5, 6.5 Hz, 1H, H-OC*H*$_2$C$_6$H$_{13}$), 2.63 (m, 1H, 6-OH), 1.68 (m, 2H), 1.42-1.25 (m, 8H), 0.91 (t, *J* = 7.0 Hz, 3H). ^{13}C-NMR (125 MHz, $CDCl_3$): 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 (O*C*H$_2$C$_6$H$_{13}$), 67.39 (C-4), 61.39 (C-6), 31.77, 29.38, 29.08, 26.08, 22.64, 14.11 (OCH$_2$*C*$_6$H$_{13}$). HR-MS Calcd. for $C_{34}H_{38}O_9$ $[M+Na]^+$, 613.2414, found 613.2420.

_n_-Heptyl 2,3,5-tri-O-benzoyl-6-deoxy-6-fluoro-α-D-mannopyranoside (14)

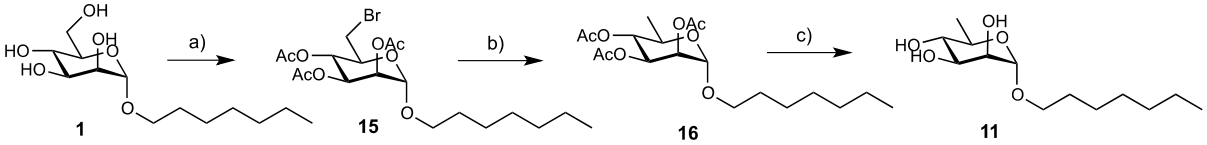
To a solution of **13** (125 mg, 0.2116 mmol) in dry DCM (0.5 mL) was added DAST (0.44 mL) at 0°C, then stirred at rt overnight. After cooling to 0°C, the reaction mixture was quenched with methanol, concentrated, the residue was purified by flash chromatography on silica gel using PE-EE (9:1-6:1) to afford the desired compound as colorless oil (56 mg, 45%). $[\alpha]_D^{20}$ = -126.76 (c = 0.61, DCM). ^{1}H-NMR (500 MHz, $CDCl_3$): 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-OC*H*$_2$C$_6$H$_{13}$), 3.57 (dt, *J* = 9.5, 6.5 Hz, 1H, H-OC*H*$_2$C$_6$H$_{13}$), 1.71 (m, 2H), 1.37 (m, 8H), 0.91 (t, *J* = 7.0 Hz, 3H). ^{13}C-NMR (125 MHz, $CDCl_3$): δ 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 (O*C*H$_2$C$_6$H$_{13}$), 66.36 (d, *J* = 9.0 Hz, C-4), 31.77, 29.36, 29.06, 26.06, 22.63, 14.10 (OCH$_2$*C*$_6$H$_{13}$). HR-MS Calcd for $C_{34}H_{37}FO_8$ [M+Na]$^+$, 615.2370, found 615.2373.

***n*-Heptyl 6-deoxy-6-fluoro-α-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_3ONa/MeOH (80 μL) 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). ^{1}H-NMR (500 MHz, CD_3OD): δ 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-OC*H*$_2$C$_6$H$_{13}$, H-3, H-4, H-5), 3.42 (dt, *J* = 9.5, 6.5 Hz, 1H, H-OC*H*$_2$C$_6$H$_{13}$), 1.59 (m, 2H), 1.32 (m, 8H), 0.91 (t, *J* = 7.0 Hz, 3H, CH_3). ^{13}C-NMR (125 MHz, CD_3OD): δ 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 (O*C*H$_2$C$_6$H$_{13}$), 67.48 (d, *J* = 7.25 Hz, C-4), 32.99, 30.60, 30.21, 27.32, 23.69, 14.42 (OCH$_2$*C*$_6$H$_{13}$). ^{19}F-NMR (376 MHz, CD_3OD): δ -230.46 (m). ESI-MS Calcd for [M+Na]$^+$, 303.16, found 303.07. HR-MS Calcd. for $C_{13}H_{25}FO_5$ [M+Na]$^+$, 303.1578, found 303.1575.

Synthesis of 54 (6-H) (*n*-Heptyl 6-deoxy-α-D-mannopyranoside)



Scheme S2. a) i. CBr_4, Ph_3P, pyr, 0-65°C, 3 h, ii. Ac_2O/pyr, DMAP, two steps (81%); b) Bu_3SnH/toluene, reflux, overnight, (80%); c) CH_3ONa/CH_3OH, (66%).

***n*-Heptyl 2,3,4-tri-O-acetyl-6-bromo-6-deoxy-α-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_4 (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_2O (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, $CHCl_3$); ^{1}H NMR (500 MHz, $CDCl_3$): δ 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, OC*H*$_2$C$_6$H$_{13}$), 3.42 (m, 3H, H-6a, H-6b, OC*H*$_2$C$_6$H$_{13}$), 2.13, 2.06, 1.98 (3xs, 9H, 3xCOC*H*$_3$), 1.59 (m, 2H), 1.30 (m, 8H, OC$_2$H$_4$C$_4$*H*$_8$CH$_3$), 0.87 (t, *J* = 6.5 Hz, 3H, OC$_2$H$_4$C$_4$H$_8$C*H*$_3$); ^{13}C NMR (125 MHz, $CDCl_3$): δ 170.06, 169.84 (3x*C*OCH$_3$), 97.28 (C-1), 69.88 (C-5), 69.67 (C-2), 68.91 (C-3), 68.84 (C-4), 68.41 (O*C*H$_2$C$_6$H$_{13}$), 31.44 (C-6), 31.68, 29.16, 28.98, 25.99, 22.57, 14.05 (OCH$_2$*C*$_6$H$_{13}$), 20.86, 20.74, 20.66 (3xCO*C*H$_3$); ESI-MS Calcd for [M+Na+2]$^+$, 491.13, found 491.11. HR-MS Calcd. for $C_{19}H_{31}BrO_8$ [M+Na]$^+$, 489.1100, found 489.1099.

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

A solution of Bu_3SnH (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); ^{1}H NMR (500 MHz, $CDCl_3$): δ 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, $OCH_2C_6H_{13}$), 3.40 (dt, *J* = 9.5, 6.5 Hz, 1H, $OCH_2C_6H_{13}$), 2.15, 2.05,1.98 (3xs, 9H, 3x$COCH_3$), 1.57 (m, 2H, $OCH_2CH_2C_5H_{11}$), 1.30 (m, 8H, $OC_2H_4C_4H_8CH_3$), 1.21 (d, *J* = 6.0 Hz, 3H, H-6), 0.88 (t, *J* = 7.0 Hz, 3H, $OC_2H_4C_4H_8CH_3$); ^{13}C NMR (125 MHz, $CDCl_3$): δ 170.21, 170.05, 170.00 (3x$COCH_3$), 97.38 (C-1), 71.23 (C-4), 70.02 (C-2), 69.18 (C-3), 68.23 ($OCH_2C_6H_{13}$), 66.16 (C-5), 17.40 (C-6), 31.73, 29.30, 29.01, 26.00, 22.59, 14.06 ($OCH_2C_6H_{13}$), 20.95, 20.81, 20.74 (3x$COCH_3$); HR-MS calcd. for $C_{19}H_{32}O_8$ $[M+Na]^+$, 411.1995, found 411.1992.

_n_-Heptyl 6-deoxy-α-D-mannopyranoside (11)

To a solution of **16** (80 mg, 0.206 mmol) in methanol (2 mL) was added 0.5 M CH_3ONa/MeOH (41 μL) 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); ^{1}H NMR (500 MHz, CD_3OD): δ 4.65 (s, 1H, H-1), 3.78 (m, 1H, H-2), 3.69-3.62 (m, 2H, H-$OCH_2C_6H_{13}$, H-3), 3.57 (m, 1H, H-5), 3.42-3.35 (m, 2H, H-$OCH_2C_6H_{13}$, H-4), 1.58 (m, 2H, $OCH_2CH_2C_5H_{11}$), 1.32 (m, 8H, $OC_2H_4C_4H_8CH_3$), 1.26 (d, *J* = 6.0 Hz, 3H, H-6), 0.91 (t, *J* = 7.0 Hz, 3H, $OC_2H_4C_4H_8CH_3$); ^{13}C NMR (125 MHz, CD_3OD): δ 101.61 (C-1), 73.96 (C-4), 72.44 (C-3), 72.34 (C-2), 69.70 (C-5), 68.52 ($OCH_2C_6H_{13}$), 32.97, 30.64, 30.20, 27.30, 23.68, 14.43 ($OCH_2C_6CH_{13}$), 18.00 (C-6); HR-MS Calcd for $C_{13}H_{26}O_5$ $[M+Na]^+$, 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 3

The tyrosine gate of the bacterial adhesin FimH - An evolutionary remnant paves the way for drug discovery

Deniz Eris,[1)] Rachel Hevey,[1)] Marleen Silbermann,[1)] Said Rabbani,[1)] Christoph P. Sager,[1)] Beat Ernst[1)*]

[1)] Institute of Molecular Pharmacy, University of Basel, Klingelbergstr. 50, 4056 Basel, Switzerland

[*] Corresponding author
Tel.: 0041 61 267 15 51; Fax: 0041 61 207 15 52
E-mail: beat.ernst@unibas.ch

Contributions of Marleen Silbermann

- Thermodynamic profiling by ITC in collaboration with Denis Eris
- Protein expression and purification of $FimH_{FL}$

Abstract

Anti-adhesive glycomimetics have emerged as a promising alternative to conventional antibiotics for the treatment of urinary tract infections (UTIs). Known as FimH antagonists, these glycomimetics avert UTIs by binding and blocking the bacterial lectin FimH, which mediates adhesion to the host urothelium. The natural ligands of FimH are high-mannose N-glycans, which form highly conserved interactions in the binding pocket and additional hydrophobic stacking interactions with a structural element adjacent to the binding pocket, known as the tyrosine gate. It is the prevailing view that the affinity of synthetic FimH antagonists is primarily determined by how well they mimic the hydrophobic interactions between natural ligands and the tyrosine gate. We measured the binding affinities of several di- and oligomannoses with isothermal titration calorimetry (ITC), showing that neither variation of the glycosidic linkages nor variation in the number of mannose residues in oligomannoses affects binding affinities. Previous studies had suggested otherwise, arguing that the tyrosine gate was sensitive to changes in oligomannose structures because of its significant contributions to ligand binding. Partial or complete removal of the tyrosine gate had no effect on the binding affinity of a dimannose, supporting our suspicion that the significance of the tyrosine gate has been overestimated. Molecular dynamics simulations confirmed ITC results, showing that interactions between di- and oligomannoses and the tyrosine gate were limited, especially compared to synthetic FimH antagonists. Indeed, rather than for natural ligand binding, the tyrosine gate seems to be important for synthetic ligands and may be aptly described as an evolutionary remnant.

Introduction

In contrast, pathogens have co-evolved means to exploit human cell-surface glycosides in the form of lectins that can mediate adherence to host cells and the formation of biofilms, both of which are key factors for their survival (13, 14). Examples of such opportunistic bacterial species include *Pseudomonas aeruginosa* with its membrane lectin LecB that has been implicated in biofilm formation (15, 16), and *Burkholderia cenocepacia*, where the characteristic *B. cenocepacia* lectin A (BC2L-A) binds host mannosides with high affinity. BC2L-A is believed to play an important role in the social life of a bacterial cell (17, 18). Another example is the bacterial lectin FimH, which plays a crucial role in urinary tract infections (UTIs) by enabling uropathogenic *Escherichia coli* (UPEC) to adhere to urothelial host cells (19, 20). FimH-mediated bacterial adhesion represents the first and most critical step in an UTI, triggering a cascade of pathogenic processes that ultimately leads to an infection. FimH binds to mannose residues on high-mannose N-glycans (Fig. 1) of the urothelial glycoprotein uroplakin Ia (UPIa) (21). The corresponding mannose binding pocket is located in the N-terminal lectin domain of FimH ($FimH_{LD}$) (22). The prerequisite role of FimH in UTI pathogenesis has turned it into an attractive target to prevent and treat UTIs. Saturation of all FimH receptors with a glycomimetic FimH antagonist can thwart bacterial adhesion, effectively preventing an UTI by removing the requisite on which every subsequent pathogenic mechanism is predicated.

The binding pocket of FimH is well-defined and accommodates a single α-D-mannose (**1**) molecule in an extensive and strong hydrogen bond network (23, 24). Accordingly, any modifications on mannose hydroxyl groups virtually abolish binding affinity (24, 25). Facing outward, (C1)-OH is the only hydroxyl group of **1** susceptible to modifications without simultaneously jeopardizing binding affinity. Early studies of potential FimH antagonists identified several aliphatic groups, *e.g.*, methyl α-D-mannoside (**10**) (26) and *n*-heptyl mannoside (**11**) (23), and aromatic groups (27) as potent (C1)-OH modifications, leading

researchers to speculate about a hydrophobic region adjacent to the binding pocket. This assumption proved to be true when $FimH_{LD}$ was crystallized in complex with various ligands, including the natural glycan known as oligmannose-3 (Table 2) (28). This crystal structure revealed that Manα1-3Manß1-4GlcNAc (Fig. 1, red), hereafter shortened to **7**, is the binding epitope, engaging in several polar and hydrophobic interactions with the FimH lectin. Most notably among these interactions, the trisaccharide epitope stacks onto Tyr48, which together with Tyr137 forms the aptly named tyrosine gate (23). Comparing the crystal structure of the natural glycan with the aliphatic FimH antagonist *n*-butyl mannoside showed how the aliphatic chain was following the α1-3 and ß1-4 glycosidic linkages of oligomannose-3 through the tyrosine gate (28). This lead researchers to conclude that the affinity of synthetic FimH antagonists was dependent on how well the antagonist was mimicking the binding epitope of this natural glycan. Accordingly, high binding affinities of aromatic FimH antagonists, which have been shown to exceed the affinity of **1** by up to three orders of magnitude (29–35), were rationalized solely on the basis of improved hydrophobic stacking interactions with the tyrosine gate.

Intrigued by the concept of FimH antagonists mimicking the binding mode of the natural ligand by interacting with the tyrosine gate, we set out to analyze other mannose-sensitive receptors expecting to find similar structural elements. While the significance of the tyrosine gate for FimH antagonists is unambiguous, our findings made us ponder whether the same degree of certainty is warranted for natural glycans. Consequently, we used isothermal titration calorimetry (ITC) to analyze the binding affinities of several mannose ligands corresponding to sequences that are terminally exposed in natural glycans, which includes the **7** binding epitope (Table 2). In addition, we analyzed the interactions of these ligands with the $FimH_{LD}$ in molecular dynamics (MD) simulations. In this study, we provide an answer to the question: is the tyrosine gate involved in natural ligand binding or is it rather an opportune evolutionary remnant?

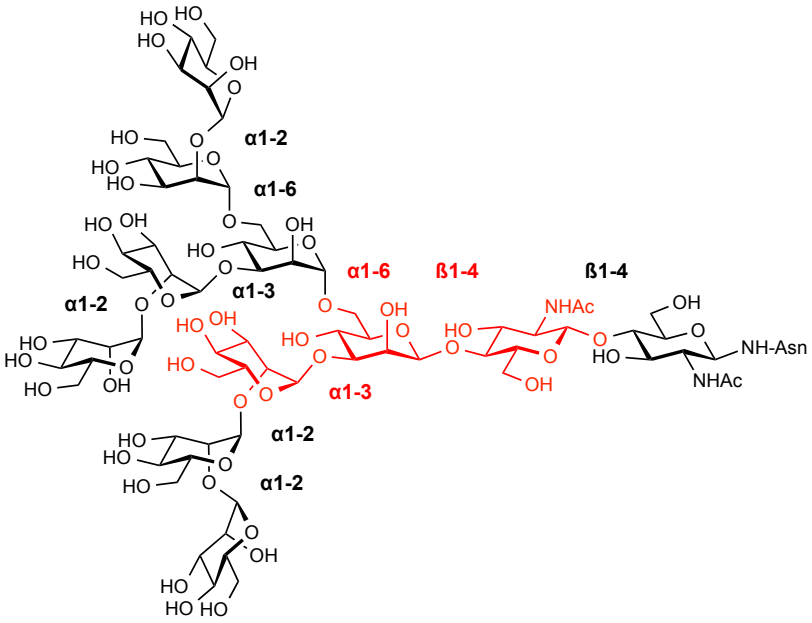
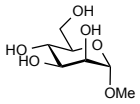
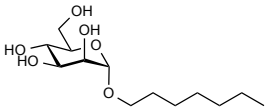
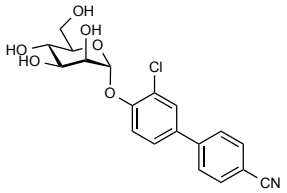


Figure 1. Structure of oligomannose-9, which is the largest biosynthesized high-mannose N-glycan. As a target of α-mannosidases in the endoplasmic reticulum and the Golgi complex, it is usually trimmed down to smaller oligomannoses before being transferred to the appropriate protein. The Manα1-3Manß1-4GlcNAc sequence is the presumed binding epitope of the FimH lectin and highlighted in red. Glycosidic linkages are indicated in bold font.

Results

Binding affinities of mono-, di-, and oligomannose ligands for $FimH_{LD}$. We have determined the dissociation constants (K_D) of $FimH_{LD}$ and **1** as well as several mono-, di-, and oligomannose ligands using ITC (Table 2 and Table 3). As the minimal binding epitope of FimH, **1** was an integral part of the ligand series, as was its congeneric α-anomer, **10**. Additionally, all α-linked dimannoses were measured to assess the influence of the glycosidic linkage on stacking interactions between the second mannose moiety and Tyr48. Measured oligomannoses included **6**, which corresponds to the terminal mannose sequence on the α1-6 branch of high-mannose N-glycans, and **8**, which matches the **6** sequence and additionally includes the α1-6 linked mannose unit of the core sequence of GlcNAcß1 linked N-glycans (Fig. 1 and Table 2). Identified as the highest affinity mannose sequence on high-mannose N-glycans in previous studies (27, 45, 46), **7** represented the most interesting ligand of our experimental series.

Table 1. Mannose ligands of the experimental series in this study.

Experimental mannose ligand series			
D-mannose (**1**)			
Manα1-2Man (**2**)	Manα1-3Man (**3**)	Manα1-4Man (**4**)	Manα1-6Man (**5**)
Manα1-6[Manα1-3]Man (**6**)		Manα1-3Manβ1-4Glc*N*Ac (**7**)	
Manα1-6[Manα1-3]Manα1-6Man (**8**)			
Manα1(6,3)Manß1–4GlcNAcß1–4GlcNAc Oligomannose-3 (**9**)			
 Methyl α-D-mannose (**10**)	 *n*-Heptyl α-D-mannoside (**11**)	 Biphenyl mannoside **12**	
Further relevant oligomannoses			
Manα1(6,3)Manα1(6,3)Manß1–4GlcNAcß1–4GlcNAc Oligomannose-5 (**13**)			
Manα1(6,3)Manα1(6,3)Manα1–2Manα1–3Manß1–4GlcNAcß1–4GlcNAc Oligomannose-6 (**14**)			

ITC experiments with ligands (Table 1) and the lectin domain of FimH ($FimH_{LD}$). Interestingly, the heat signals in the thermograms of all tested mannose ligands were characterized by relatively slow equilibration times. Heat signal curvatures of mannose ligands amenable to kinetic analysis revealed remarkably slow on rate constants in the range of 10^3 M^{-1} s^{-1} (*SI Appendix*, Table S6) as a potential cause for the slowly equilibrating heat signals. The influence of slow kinetics on the heat signal necessitated longer spacing intervals between the injections and higher injection volumes to compensate for the peak broadening (*SI Appendix*, Table S2 and Table S4).

Table 2. Binding affinities

Ligand	$FimH_{FL}$ [µM]	$FimH_{LD}$ [µM]	Y48A mutant of $FimH_{LD}$ [µM]	Y137A mutant of $FimH_{LD}$ [µM]	Y48A/Y137A mutant of $FimH_{LD}$ [µM]
D-Mannose (**1**)	617.3	2.7	-	-	-
Manα1-2Man (**2**)	345.4	2.4	-	-	-
Manα1-3Man (**3**)	152.2	1.1	1.4	2.4	1.5
Manα1-4Man (**4**)	384.9	1.2	-	-	-
Manα1-6Man (**5**)	787.4	3.3	-	-	-
Manα1-6[Manα1-3]Man (**6**)	105.7	1.2	-	-	-
Manα1-3Manβ1-4Glc*N*Ac (**7**)	-	1.2	-	-	-
Manα1-6[Manα1-3]Manα1-6Man (**8**)	114.1	1.5	-	-	-
Oligomannose-3 (**9**)	-	2.7	0.9	6.9	3.9
Methyl α-D-mannose (**10**)	529.5	1.2	-	-	-
n-Heptyl α-D-mannoside (**11**)	-	3×10^{-2}	-	-	-
Biphenylmannoside (**12**)	-	1×10^{-3}	-	-	-

Table 3. Binding thermodynamics of mannose ligands and $FimH_{LD}$ as determined by ITC. Thermodynamic parameters are indicated in kJ mol^{-1}. 68.3% confidence intervals for fitted parameters are given in parentheses. If not stated otherwise, experiments were performed once. The *N* parameter was averaged for ligands that were measured more than once. Individual *N* parameters are given in Table S2.

Ligand	K_D [µM]	LE[d]	$\Delta G°_{obs}$	$\Delta H°_{obs}$	$-T\Delta S_{obs}$	*N*
D-Mannose (**1**)	2.7 (2.2 – 3.2)	0.63	-31.8	-41.6 (-43.5 – -39.8)	9.8	0.91
Manα1-2Man (**2**)	2.4 (1.9 – 3.0)	0.33	-32.1	-33.6 (-35.3 – -32.1)	1.5	1.01
Manα1-3Man (**3**)[a]	1.1 (0.9 – 1.4)	0.32	-34.0	-43.3 (-44.9 – -41.7)	9.2	0.95
Manα1-4Man (**4**)	1.2 (0.9 – 1.5)	0.35	-33.9	-34.0 (-35.7 – -32.5)	0.1	0.98
Manα1-6Man (**5**)[a]	3.3 (2.5 – 4.4)	0.33	-31.3	-28.6 (-30.4 – -27.1)	-2.7	1.08
Manα1-6[Manα1-3]Man (**6**)[a,b]	1.2 (0.9 – 1.7)	0.23	-33.8	-35.3 (-36.8 – -34.0)	1.6	0.93
Manα1-3Manβ1-4Glc*N*Ac (**7**)[a]	1.2 (0.9 – 1.5)	0.22	-33.9	-41.1 (-42.0 – -40.2)	7.2	1.00
Manα1-6[Manα1-3]Manα1-6Man (**8**)[a,b]	1.5 (1.0 – 2.2)	0.18	-33.3	-36.6 (-39.0 – -34.5)	3.3	0.95
Oligomannose-3 (**9**)	2.7 (2.1 – 3.5)	0.12	-31.8	-36.2 (-38.1 – -34.4)	4.4	1.02
Methyl α-D-mannose (**10**)	1.2 (0.9 – 1.6)	0.62	-33.9	-35.3 (-36.8 – -33.9)	1.4	0.93
n-Heptyl α-D-mannoside (**11**)[c]	3×10^{-2}	0.54	-43.0	-50.3	7.3	1.00
Biphenylmannoside (**12**)[c]	1×10^{-3}	0.45	-50.7	-60.9	10.1	1.01

[a]Experiments were performed at least twice. [b]Binding isotherms were fitted to a 2:1 binding model. [c]Values taken from ref. 32. [d]Ligand efficiency.

The inflection point of an ITC binding isotherm indicates the stoichiometric ratio of titrant (here: mannose ligand) to cell material (here: $FimH_{LD}$). Designated simply as *N* in Table 3 and *SI Appendix* Table S2 and Table S4, this parameter is used to infer the binding stoichiometry of the reaction or the protein fraction incompetent of binding. For 1:1 binding,

the *N* parameter is expected to be close to 1, which was the case for most of the mannose ligands in this series. However, **6**, **8**, and **9** were prominent exceptions to this observation and initially showed *N* values around 0.5. Considering that the same $FimH_{LD}$ batch was used throughout the series of ITC experiments, partly inactive $FimH_{LD}$ could be reasonably excluded as possible reason. Instead, an *N* value of 0.5 signified a 2:1 binding stoichiometry of oligomannose to $FimH_{LD}$. As a consequence, we have adjusted the binding model from 1:1 to 2:1 binding (see Materials and Methods for details). If this adjustment is not undertaken, the K_D and $\Delta H°$ of these oligomannoses will be inadvertently overestimated. After applying appropriate binding models to the binding isotherms, the resulting K_D values of all mannose ligands showed a remarkable congruence, only ranging between 1.1 to 3.3 μM (Table 3). Most notably, this includes $Man_2GlcNAc_1$ and $Man_3GlcNAc_2$ with K_D values of 1.2 and 2.7 μM, respectively, which demonstrates that their binding affinities do not exceed those of any other mannose ligands tested in this series.

Binding affinities of $FimH_{LD}$ with a partly or completely defunct tyrosine gate. $FimH_{LD}$ constructs with an incomplete or nonexistent tyrosine gate would be expected to have diminished binding affinities for di- and oligomannose ligands, provided that stacking interactions with the tyrosine gate contribute significantly to the overall binding affinity. To that end, we generated a FimH-Y48A and FimH-Y137A mutant, as well as a double mutant with both tyrosine residues mutated to an alanine residue. ITC experiments were performed with **3** as representative mannose ligand. Heat signals of the $FimH_{LD}$ mutants and **3** were characterized by comparably slow equilibration times as observed for heat signals of the wild type $FimH_{LD}$ measurements. K_D values of **3** and all tyrosine gate mutants were strikingly similar, ranging between 1.1 and 2.4 μM (Table 4). The mutational study demonstrated that partial or complete loss of the tyrosine gate did not affect the binding affinity of a mannose

ligand presumed to engage it in crucial stacking interactions. This result reinforced our findings of the previous ITC experimental series.

Table 4. Comparison of binding thermodynamics of Manα1-3Man (**3**) with $FimH_{LD}$ and several tyrosine gate mutants of $FimH_{LD}$. Thermodynamic parameters are indicated in kJ mol^{-1}. 68.3% confidence intervals for fitted parameters are given in parentheses. If not stated otherwise, experiments were performed at least twice. The *N* parameter was averaged for ligands that were measured more than once. Individual *N* parameters are given in Table S4.

$FimH_{LD}$	**K_D [µM]**	**$\Delta°G_{obs}$**	**$\Delta°H_{obs}$**	**$-T\Delta S°_{obs}$**	***N***
Wild type	1.1 (0.9 – 1.4)	-34.0	-43.3 (-44.9 – -41.7)	9.2	0.95
Y48A mutant	1.4 (1.1 – 1.8)	-33.4	-45.6 (-47.2 – -44.0)	12.1	1.00
Y137A mutant	2.4 (1.9 – 3.0)	-32.1	-39.0 (-40.4 – -37.6)	6.8	1.06
Y48A/Y137A[a] mutant	1.5 (1.2 – 1.9)	-33.2	-41.9 (-43.0 – -40.8)	8.6	1.00

[a]Experiment performed once.

ITC experiments with ligands (Table 1) and full-length FimH ($FimH_{FL}$).

Table 5. Binding thermodynamics of mannose ligands and $FimH_{FL}$ as determined by ITC. Thermodynamic parameters are indicated in kJ mol^{-1}. 68.3% confidence intervals for fitted parameters are given in parentheses. If not stated otherwise, experiments were performed once. The *N* parameter was averaged for ligands that were measured more than once. Individual *N* parameters are given in Table S2.

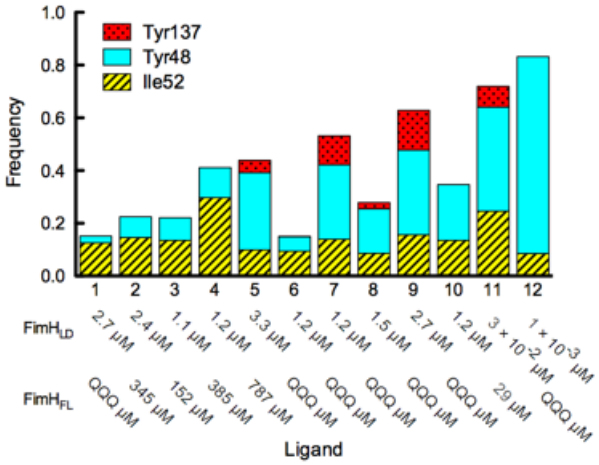
Ligand	K_D [µM]	LE[a]	$\Delta G°_{obs}$	$\Delta H°_{obs}$	$-T\Delta S_{obs}$	*N*
D-Mannose (**1**)	2.7 (2.2 – 3.2)	-	-31.8	-41.6 (-43.5 – -39.8)	9.8	0.91
Manα1-2Man (**2**)	345.4 (334.4 – 356.7)	0.21	-19.8	-62.0 (-63.7 – -60.4)	42.3	1.0[d]
Manα1-3Man (**3**)	152.2 (138.8 – 167.3)	0.23	-21.8	-75.7 (-81.9 – -70.3)	53.9	1.0[d]
Manα1-4Man (**4**)	384.9 (367.5 – 403.0)	0.20	-19.5	-45.9 (-47.6 – -44.3)	26.4	1.0[d]
Manα1-6Man (**5**)	787.4 (757.2 – 818.8)	0.18	-17.7	-63.1 (-65.4 – -60.9)	45.4	1.0[d]
Manα1-6[Manα1-3] Manα1-6Man (**8**)	-	-	-	-	-	-
Oligomannose-3 (**9**)	-	-	-	-	-	-
Methyl α-D-mannose (**10**)	1.2 (0.9 – 1.6)	0.62	-33.9	-35.3 (-36.8 – -33.9)	1.4	0.93
n-Heptyl α-D-mannoside (**11**)[c]	28.9 (25.8 – 32.3)	0.54	-43.0	-50.3 (-50.2 – -50.7)	7.3	1.00
Biphenylmannoside **12**[c]	-	-	-	-	-	-

[a]Ligand efficiency; [b]Binding isotherms were fitted to a 2:1 binding model; [c]Values taken from ref. 32; [d] fixed to 1.

Dynamic and quantum mechanical interaction with tyrosine gate. (Figure 2) Quantitative analysis of ligand–protein interactions of 1000 ns MD simulations revealed that C6 of **1** interacts with the side-chain of Ile52 to a small extent, which is true for all analyzed ligands. The additional methyl group of **10** is able to form further interactions with Tyr48, while the dimannoses (**2**– **5**), trimannose (**6**) and tetramannose (**8**) ligands are rarely in contact with the tyrosine gate residues because of the linked mannose that is less versatile than the methyl group of **10**. The natural binding epitope **7** as well as **9** donate a hydrogen bond with their *N*-

acetylglucosamine (C6)-OH to Thr51 in 63% of the simulation time. This is the reason why additional interactions with Tyr137 can be observed in contrast to the other oligomannoses that lack *N*-acetylglucosamine at the same position. Both synthetic ligands **11** and **12** showed an increased interaction with the tyrosine gate residues, most notably with Tyr48. These results are in good agreement with ITC experiments.

From the MD simulations it was apparent that the first mannose was heavily involved in hydrogen bonding to FimH and hydrophobic contact with Ile52, and that the second mannose moiety or the synthetic aglycones were mainly in contact with the tyrosine gate, especially Tyr48. Therefore, *ab initio* quantum mechanical (QM) calculations were carried out, which revealed that from the tested moieties the second mannose moiety with only -12.1 kJ mol^{-1} interacted the weakest with Tyr48. Replacing this mannose by an exemplary cyclohexane improved the interaction energy to -15.9 kJ mol^{-1}. However, synthetic ligands benefited much more from stacking interactions with Tyr48. The *n*-heptyl aglycone, as seen in **11**, showed an improved interaction energy of -16.8 kJ mol^{-1}, while the biphenyl agylcone of **12** improved even further upon that energy value through optimal π-π stacking, showing an interaction energy of -35.7 kJ mol^{-1}.



Figuge 2. Analysis of the tyrosine gate interaction frequency of different mannose (**1**) and methyl mannoside (**10**), di- (**2-4**), tri- (**6**), and tetramannose (**8**) ligands, as well the natural epitopes (**7** and 9) and synthetic ligands (**11** and **12**). Hydrophobic interactions with Ile52 are colored in yellow, with Tyr48 in cyan, and with Tyr137 in orange. A frequency of 1.0 corresponds to an interaction that is present during the entire simulation time.

Discussion

FimH is not only different in this regard, but also in the fact that from the analyzed mannose binding receptors, it is the only one to feature a hydrophobic structure flanking the mannose binding site. Known as the tyrosine gate and consisting of the two tyrosine residues Tyr48 and Tyr137 (23), this structural element is believed to be engaged in significant interactions with natural high-mannose N-glycans; most notably including hydrophobic stacking interactions between a mannose moiety of the glycan and Tyr48 of FimH (28). To this effect, oligomannose structures with varying numbers of variously linked mannose moieties would be expected to have varying binding affinities depending on how well mannose moieties align and interact with the tyrosine gate. As it happens, binding affinities of various di- and oligomannoses have been the subject of much research over the past decades and were determined in a variety of assays, including agglutination assays (27, 45, 46), an ELISA (49), a bead-based assay (50), a glycan microarray (51), a radioligand displacement assay, and an SPR-based assay (23, 28, 52). Although outcomes of these assays have been partly in conflict with each other, some results have been consistent throughout several studies. For instance, oligomannoses were generally stronger FimH ligands than dimannoses. Among oligomannose structures of high-mannose N-glycans, **7** was found to be the highest affinity carbohydrate sequence (27, 45, 46). In a crystal structure of $FimH_{LD}$ in complex with oligomannose-3 (**9**), the **7** sequence was identified as the principal binding epitope, forming optimal interactions with the tyrosine gate and imparting the oligomannose ligand with its high affinity. Indeed, this seemed to be in line with observations of the higher affinity of small oligomannose ligands such as oligomannose-3 and -5 (52), both of which expose $Man_2GlcNAc_1$ on their non-reducing ends (Table 2). In the bigger congeners oligomannose-6, -7, -8, and -9, the **7** sequence is capped with one or two additional α1-2 linked mannose residues (Fig. 1). Both the α1-2 linkage and the additional mannose unit(s) are thought to disrupt the optimal binding mode of the **7** binding epitope, thereby negatively impacting the binding affinity. However,

when results of a glycan microarray showed preferential binding of FimH to oligomannose-6 instead of oligomannose-3 and -5 (51), the purported significance of the **7** sequence was inadvertently brought into question.

Notably, all mentioned assays are characterized as being competitive binding experiments and/or as having one of the binding partners immobilized. When investigating oligomannose binding where interactions may not be exclusively monovalent, such assay conditions can impose further complexity to a system that already has the potential to be considerably complex. As a consequence, we opted for ITC as our method of choice in determining the binding affinities of several previously characterized mannose ligands. ITC offers the advantage of being a homogeneous assay and the possibility to measure FimH-ligand binding directly. In addition, it indicates the binding stoichiometry of the reaction, which is a crucial piece of information when investigating oligomannose binding. Indeed, our ITC experiments revealed that **6**, **8**, and **9** bind to $FimH_{LD}$ in a 2:1 fashion. By applying the appropriate 2:1 binding model to the isotherm instead of a 1:1 model, fitted K_D values were revised upward, meaning that affinities decreased as a result. In retrospect, K_Ds of oligomannoses determined in previous assays might have been overestimated in much the same way by not accounting for bi- or multivalent binding in the applied binding model.

The K_D values of all mannose ligands tested in this study, including **7** and **9**, were almost identical (Table 3), implying that additional mannose moieties or the glycosidic linkages between them have little to no effect on the binding affinities of di- and oligomannoses. This suggests that the first mannose moiety, which forms an extensive hydrogen bond network in the binding pocket, is the primary factor determining the binding affinity. As a consequence, the concept of the tyrosine gate playing a significant role in binding natural ligands becomes less tenable. Indeed, partial or complete removal of the tyrosine gate from $FimH_{LD}$ had

marginal or no impact on the binding affinity of **3** (Table 4), which was used as representative mannose ligand in this mutational study.

Results of ITC experiments were mirrored in MD simulations, in which we could observe that the interaction of di- and oligomannoses was limited with Tyr48 and not present at all with Tyr137. Furthermore, the interaction energy of mannose with Tyr48 was calculated to be -12.1 kJ mol^{-1}, which was the weakest interaction among all calculated structures. This could explain the measured affinities from ITC experiments, which were the same for the natural binding epitope and the oligomannoses. Due to the hydrogen bond formed with Thr51, the interaction frequency of the natural epitope with the tyrosine gate was elevated in MD simulations - yet the interaction was still weak in comparison to synthetic FimH antagonists that formed stronger and more frequent interactions with Tyr48.

Indeed, in the design of synthetic FimH antagonists researchers have exploited the tyrosine gate for the purpose of π-π stacking by attaching a hydrophobic aryl aglycone to the vital mannose moiety. This has lead to continuously improved affinities of synthetic FimH antagonists (29–35, 53). In contrast, and as demonstrated in this study, di- and oligomannose structures that are terminally exposed in high-mannose N-glycans did not benefit from improved affinities with increasing size and complexity. This is reflected in a decreasing ligand efficiency (LE, Table 3), which is a metric commonly used in drug discovery (54). While **1** and **10** showed good LEs of 0.63 and 0.62, respectively, the introduction of a second mannose (**2**–**5**) halved the LEs of dimannoses to a range between 0.32 and 0.35. The LE of trimannose **6** is reduced further to 0.23, and with an added fourth mannose unit (**8**) the LE drops even lower to 0.17. The natural binding epitope of FimH (**7**) has a ligand efficiency of only 0.20, contrasted by synthetic FimH antagonists **11** and **12**, which exhibit a high LE of 0.54 and 0.45, respectively. Evidently, the synthetically added aglycone moieties establish the intended beneficial interactions with FimH.

As previously mentioned, the prevailing hypothesis has been that these interactions mimic those of the natural **7** binding epitope. In this study, we performed ITC experiments and MD simulations to provide evidence suggesting that **7** does not have an advantage in terms of binding affinity over other di- or oligomannoses. Moreover, the relevance of **7** as natural binding epitope may be overestimated. For instance, one of two major mannosylated glycoproteins involved in UTIs is the transmembrane UPIa glycoprotein (21). Inspection of murine UPIa glycosylation patterns revealed that oligomannose-5 is entirely absent in the glycomoiety (55). Significantly, oligomannose-5 is the only biosynthesized oligomannose structure terminally exposing the **7** sequence. Instead, the major glycoform in UPIa is constituted by oligomannose-8 (43%) (55). As mentioned previously, in structures larger than oligomannose-5, the **7** sequence is buried under one or two α1-2 linked mannose moieties (Fig. 1). Another major glycoprotein with high-mannose N-glycans is Tamm-Horsfall glycoprotein (THP). As the most abundant urinary glycoprotein in mammals, THP carries mostly oligomannose-6 residues (75%) and only marginal amounts of oligomannose-5 (8%) (56). These results attenuate the significance of the **7** sequence as the natural binding epitope, making it rather unlikely that the high affinity of FimH antagonists originates from mimicking the binding mode of a carbohydrate sequence that rarely partakes in mannose ligand binding.

Altogether, these findings have lead us to conclude that the tyrosine gate is not involved in binding of natural oligomannose ligands. Instead, we believe it to be a bygone remnant that is serendipitously located close to the binding site where it represents an important structural target for synthetic ligands.

Materials and Methods

Crystal Structure Preparation. Protein-Ligand complexes from the crystal structures were processed with the Protein Preparation Wizard (57) and further refined using local optimization sampling in Prime 4.2 (58, 59) within 5.0 Å of the ligand using the OPLS3 force-field and VSGB solvation model.

Crystallization and structure determination of FimH/methyl α-D-mannoside. $FimH_{LD}$/**10** was crystallized by sitting-drop vapor diffusion at 4°C. $FimH_{LD}$ (residues 1–158) was used at a final concentration of 10 mg mL^{-1} (ca. 0.5 mM) with a threefold molar excess of **10** in HEPES buffer (pH 7.4, 20 mM). After four weeks equilibration, plate like crystals appeared in 0.2 M $(NH_4)_2SO_4$, 0.1 M HEPES pH 7.0, 30% PEG 3350 (w/v) and were flash-cooled to 100 K by addition of 20% ethylene glycol (Hampton Research, USA). Data were collected with synchrotron radiation at the PXIII beamline of the Swiss Light Source (Paul Scherrer Institute, Switzerland). Data were indexed, integrated, and scaled with XDS (60, 61). Structures were solved by molecular replacement with PHASER (62) with use of the 4X50.pdb (35) as search model. The structures were built in COOT (63, 64) and refined with the PHENIX software (65, 66). Geometric restraints for the ligands were generated with PRODRG (67) and Molprobity (68) was used to validate the available atomic coordinates. The structures are deposited in the Protein Data Bank with PDB codes: 5JCQ and 5JCR.

Molecular dynamics simulations. MD Simulations were carried out using Desmond 4.4 and the OPLS3 force field (69–71). Default parameters were applied unless stated otherwise. TIP3P was the selected water model and a physiological salt concentration (0.15 M) was added to the simulation. An energy barrier of 1 kcal/mol restricted Cα movement during the 1000 ns MD simulations. Energies and trajectories were recorded at 100 ps intervals resulting in 10,000 recorded frames, which were post-processed using the Simulation Interactions Diagram implemented in Desmond.

Mannose ligands. 1 and **10** were purchased from Sigma-Aldrich Chemie (Buchs, Switzerland). All di- and oligomannose ligands with the exception of **7** were purchased from Dextra Laboratories (Reading, UK).

Expression and purification of $FimH_{LD}$. $FimH_{LD}$ corresponding to the *E. coli* laboratory K-12 strain was expressed recombinantly with a C-terminal thrombin cleavage site and His_6-tag in the protease-deficient *E. coli* HM125 strain (72), and purified by affinity chromatography as described previously (73). Appropriate point mutations were inserted by overlap extension PCR (74).

Isothermal titration calorimetry (ITC). ITC experiments were performed with an iTC200 calorimeter (Malvern Instruments, Worcestershire, UK) and carried out in a buffer containing 20 mM HEPES, pH 7.4 and 150 mM NaCl. Before starting with the experiment, FimH solutions were extensively dialyzed against the assay buffer. Mannose ligands were weighed and directly dissolved in the same buffer. Prior to the experiment, FimH concentrations were adjusted *via* the specific absorbance at 280 nm. The molar extinction coefficients used were 24,180, 29,250, 22,900, and 21,620 M^{-1} cm^{-1} for wild type $FimH_{LD}$, $FimH_{LD}$-Y48A, $FimH_{LD}$-Y137A, and $FimH_{LD}$-Y48A/Y137A, respectively. The cell of the calorimeter contained the FimH solution and holds a volume of 203.7 µl. Mannose ligands were injected into the cell in regular intervals. These intervals as well as the injection volumes varied among experiments. Details of individual measurements, including these two parameters, are available in the *SI Appendix*, Table S2 and Table S4. Some parameters were constant throughout the measurements, such as temperature (25 °C), feedback mode (high), stirring speed (750 rpm), and reference power (6 µcal s^{-1}). Baseline correction as well as peak detection and integration was performed automatically with NITPIC (version 1.1.5; University of Texas Southwestern Medical Center, Texas, USA) (75). The resulting binding isotherm was fitted in SEDPHAT (version 12.1c; National Institutes of Health, Maryland, USA) (76) with a 1:1 binding model and the following fitting parameters: change in enthalpy ($\Delta H°$), association constant (K_A),

molar ratio (N), and heat of dilution (ΔH_{dil}). The dissociation constant K_D was obtained through the relationship $K_D = K_A^{-1}$. The K_D was converted into the corresponding Gibbs energy change ($\Delta G°$), from which the change in entropy ($\Delta S°$) was calculated according to:

$$\Delta G° = \Delta H° - T\Delta S° = RT \ln(K_D)$$

Statistical analysis to obtain 68.3% confidence intervals of the fitted parameters was done in SEDPHAT. Isotherms were fitted globally, if multiple data sets were available for a particular ligand. For **6**, **8**, and **9** the binding model was changed from 1:1 to 2:1 binding. Consequently, two fitting parameters were available for each K_A and $\Delta H°$; one parameter set for each binding event. No indications of potential cooperativity were observed in the isotherms of **6**, **8**, and **9**, and binding sites were presumed to be equivalent. As a result, the logarithmic ratio of K_As ("log10(Ka2/Ka1)" in SEDPHAT) was fixed to -0.602. To obtain the microscopic K_A of the interaction, the K_A of the first binding event ("log(Ka1)" in SEDPHAT) was multiplied by 2.

***Ab initio* calculations.** To quantify the van-der-Waals interactions of the tyrosine gate and its interacting moieties, the refined protein-ligand complexes were subjected to energy calculations using the density functional theory (DFT) with the empirical corrected B3LYP-MM functional, which has been specifically parameterized to properly describe non-covalent interactions, and the cc-pVDZ++ basis-set in the gas-phase as implemented in Jaguar 9.0 (77, 78).

ACKNOWLEDGMENTS. We thank Dr. Timothy Sharpe for his help with ITC data analysis.

References

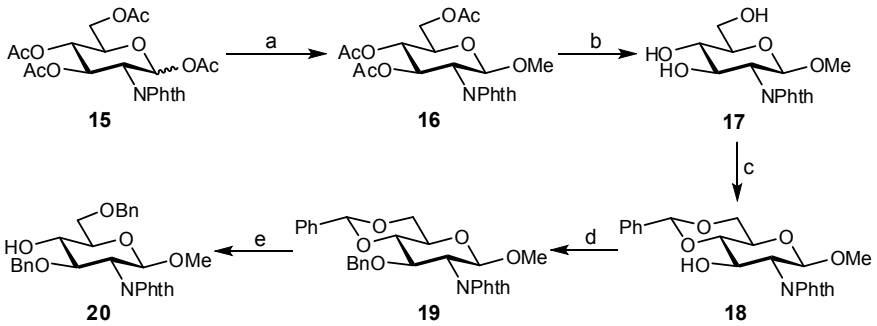
13. Sharon N, Lis H (1989) Lectins as cell recognition molecules. *Science* 246(4927):227–234.
14. Hall-Stoodley L, Costerton JW, Stoodley P (2004) Bacterial biofilms: from the natural environment to infectious diseases. *Nat Rev Microbiol* 2(2):95–108.
15. Loris R, Tielker D, Jaeger KE, Wyns L (2003) Structural basis of carbohydrate recognition by the lectin LecB from Pseudomonas aeruginosa. *J Mol Biol* 331(4):861–870.
16. Tielker D, et al. (2005) Pseudomonas aeruginosa lectin LecB is located in the outer membrane and is involved in biofilm formation. *Microbiology* 151(5):1313–1323.
17. Lameignere E, et al. (2008) Structural basis for mannose recognition by a lectin from opportunistic bacteria Burkholderia cenocepacia. *Biochem J* 411(2):307–18.
18. Marchetti R, et al. (2012) Burkholderia cenocepacia lectin A binding to heptoses from the bacterial lipopolysaccharide. *Glycobiology* 22(10):1387–1398.
19. Mulvey MA, et al. (1998) Induction and Evasion of Host Defenses by Type 1–Piliated Uropathogenic *Escherichia coli*. *Science* 282(5393):1494–1497.
20. Krogfelt KA, Bergmans H, Klemm P (1990) Direct evidence that the FimH protein is the mannose-specific adhesion of Escherichia coli type 1 fimbriae. *Infect Immun* 58(6):1995–1998.
21. Zhou G, et al. (2001) Uroplakin Ia is the urothelial receptor for uropathogenic *Escherichia coli*: evidence from in vitro FimH binding. *J Cell Sci* 114(22):4095–4103.
22. Choudhury D (1999) X-ray Structure of the FimC-FimH Chaperone-Adhesin Complex from Uropathogenic Escherichia coli. *Science* 285(5430):1061–1066.
23. 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.
24. Zihlmann P, et al. High-Affinity Carbohydrate-Lectin Interactions: How Nature Makes it Possible. *unpublished*.
25. Old DC (1972) Inhibition of the Interaction Between Fimbrial Haemagglutinins and Erythrocytes by D-Mannose and Other Carbohydrates. *Microbiology* 71(1):149–157.
26. Aronson M, et al. (1979) Prevention of colonization of the urinary tract of mice with Escherichia coli by blocking of bacterial adherence with methyl alpha-D-mannopyranoside. *J Infect Dis* 139(3):329–332.
27. Firon N, Ofek I, Sharon N (1983) Carbohydrate specificity of the surface lectins of Escherichia coli, Klebsiella pneumoniae, and Salmonella typhimurium. *Carbohydr Res* 120(C):235–249.
28. 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.
29. Han Z, et al. (2010) Structure-based drug design and optimization of mannoside bacterial fimH antagonists. *J Med Chem* 53(12):4779–4792.
30. Jiang X, et al. (2012) Anti-Adhesion Therapy for Urinary Tract Infections - A Balanced PK/PD-Profile Proved to be Key for Success. *J Med Chem* 55(10):4700–13.
31. Pang L, et al. (2012) FimH Antagonists: Structure-Activity and Structure-Property Relationships for Biphenyl α-D-Mannopyranosides. *ChemMedChem* 7(8):1404–1422.
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. Han Z, et al. (2012) Lead optimization studies on FimH antagonists: Discovery of potent and orally bioavailable ortho-substituted biphenyl mannosides. *J Med Chem* 55(8):3945–3959.
34. Jarvis C, et al. (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.
35. 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.
36. Sabin C, et al. (2006) Binding of different monosaccharides by lectin PA-IIL from Pseudomonas aeruginosa: Thermodynamics data correlated with X-ray structures. *FEBS Lett* 580(3):982–987.
37. Jégouzo SA, et al. (2015) A novel mechanism for binding of galactose-terminated glycans by the C-type carbohydrate recognition domain in blood dendritic cell antigen 2. *J Biol Chem* 290(27):16759–16771.
38. Feinberg H, et al. (2013) Common polymorphisms in human langerin change specificity for glycan ligands. *J Biol Chem* 288(52):36762–36771.
39. Crouch E, et al. (2009) Recognition of mannosylated ligands and influenza a virus by human surfactant protein D: Contributions of an extended site and residue 343. *Biochemistry* 48(15):3335–3345.
40. Feinberg H, Mitchell D a, Drickamer K, Weis WI (2001) Structural basis for selective recognition of oligosaccharides by DC-SIGN and DC-SIGNR. *Science* 294(5549):2163–6.
41. Feinberg H, Castelli R, Drickamer K, Seeberger PH, Weis WI (2007) Multiple modes of binding enhance the affinity of DC-SIGN for high mannose N-linked glycans found on viral glycoproteins. *J Biol Chem* 282(6):4202–4209.
42. Ng KKS, et al. (2002) Orientation of bound ligands in mannose-binding proteins. Implications for

multivalent ligand recognition. *J Biol Chem* 277(18):16088–16095.
43. Lee RT, et al. (1991) Ligand-binding characteristics of rat serum-type mannose-binding protein (MBP-A): Homology of binding site architecture with mammalian and chicken hepatic lectins. *J Biol Chem* 266(8):4810–4815.
44. Glusker JP, Kaufmann Katz A, Bock CW (1996) Metal coordination of several divalent cations: Mg2+, Ca2+, Be2+, and Zn2+. *Acta Crystallogr* 52:133.
45. Firon N, Ofek I, Sharon N (1982) Interaction of mannose-containing oligosaccharides with the fimbrial lectin of Escherichia coli. *Biochem Biophys Res Commun* 105(4):1426–1432.
46. Neeser JR, Koellreutter B, Wuersch P (1986) Oligomannoside-type glycopeptides inhibiting adhesion of Escherichia coli strains mediated by type 1 pili: Preparation of potent inhibitors from plant glycoproteins. *Infect Immun* 52(2):428–436.
47. Schmidtke P, Javier Luque F, Murray JB, Barril X (2011) Shielded hydrogen bonds as structural determinants of binding kinetics: Application in drug design. *J Am Chem Soc* 133(46):18903–18910.
48. Quiocho FA (1986) Carbohydrate-binding proteins: tertiary structures and protein-sugar interactions. *Annu Rev Biochem* 55:287–315.
49. Dubber M, Sperling O, Lindhorst TK (2006) Oligomannoside mimetics by glycosylation of "octopus glycosides" and their investigation as inhibitors of type 1 fimbriae-mediated adhesion of Escherichia coli. *Org Biomol Chem* 4(21):3901–3912.
50. Ajisaka K, et al. (2015) Preferential binding of E. coli with type 1 fimbria to d-mannobiose with the Manα1→2Man structure. *Biosci Biotechnol Biochem*:1–7.
51. Taganna J, de Boer AR, Wuhrer M, Bouckaert J (2011) Glycosylation changes as important factors for the susceptibility to urinary tract infection. *Biochem Soc Trans* 39(1):349–54.
52. Bouckaert J, et al. (2006) The affinity of the FimH fimbrial adhesin is receptor-driven and quasi-independent of Escherichia coli pathotypes. *Mol Microbiol* 61(6):1556–1568.
53. Schwardt O, et al. (2011) Design, synthesis and biological evaluation of mannosyl triazoles as FimH antagonists. *Bioorganic Med Chem* 19(21):6454–6473.
54. Hopkins AL, Groom CR, Alex A (2004) Ligand efficiency: A useful metric for lead selection. *Drug Discov Today* 9(10):430–431.
55. Xie B, et al. (2006) Distinct glycan structures of uroplakins Ia and Ib: Structural basis for the selective binding of FimH adhesin to uroplakin Ia. *J Biol Chem* 281(21):14644–14653.
56. Cavallone D, Malagolini N, Monti A, Wu XR, Serafini-Cessi F (2004) Variation of High Mannose Chains of Tamm-Horsfall Glycoprotein Confers Differential Binding to Type 1-fimbriated Escherichia coli. *J Biol Chem* 279(1):216–222.
57. Madhavi Sastry G, 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.
58. Jacobson MP, Friesner RA, 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.
59. Jacobson MP, et al. (2004) A Hierarchical Approach to All-Atom Protein Loop Prediction. *Proteins Struct Funct Genet* 55(2):351–367.
60. Kabsch W (2010) Integration, scaling, space-group assignment and post-refinement. *Acta Crystallogr Sect D Biol Crystallogr* 66(2):133–144.
61. Kabsch W (2010) XDS. *Acta Crystallogr Sect D Biol Crystallogr* 66(2):125–132.
62. McCoy AJ, et al. (2007) Phaser crystallographic software. *J Appl Crystallogr* 40(4):658–674.
63. Emsley P, Cowtan K (2004) Coot: Model-building tools for molecular graphics. *Acta Crystallogr Sect D Biol Crystallogr* 60(12 I):2126–2132.
64. Emsley P, Lohkamp B, Scott WG, Cowtan K (2010) Features and development of Coot. *Acta Crystallogr Sect D Biol Crystallogr* 66(4):486–501.
65. Adams PD, et al. (2002) PHENIX : building new software for automated crystallographic structure determination. *Acta Crystallogr Sect D Biol Crystallogr* 58(11):1948–1954.
66. Adams PD, et al. (2010) PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr Sect D Biol Crystallogr* 66(2):213–221.
67. Schüttelkopf AW, Van Aalten DM (2004) PRODRG: A tool for high-throughput crystallography of protein-ligand complexes. *Acta Crystallogr Sect D Biol Crystallogr* 60(8):1355–1363.
68. Chen VB, et al. (2010) MolProbity: All-atom structure validation for macromolecular crystallography. *Acta Crystallogr Sect D Biol Crystallogr* 66(1):12–21.
69. 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.
70. 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.
71. Bowers K, et al. (2006) Scalable Algorithms for Molecular Dynamics Simulations on Commodity Clusters. *ACM/IEEE SC 2006 Conf* (November):43–43.
72. Meerman HJ, Georgiou G (1994) Construction and characterization of a set of E. coli strains deficient in

all known loci affecting the proteolytic stability of secreted recombinant proteins. *Nat Biotechnol* 12(11):1107–1110.
73. 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.
74. Heckman KL, Pease LR (2007) Gene splicing and mutagenesis by PCR-driven overlap extension. *Nat Protoc* 2(4):924–932.
75. Keller S, et al. (2012) High-precision isothermal titration calorimetry with automated peak-shape analysis. *Anal Chem* 84(11):5066–5073.
76. Houtman JCD, 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.
77. Schneebeli ST, Bochevarov AD, Friesner RA (2011) Parameterization of a B3LYP specific correction for noncovalent interactions and basis set superposition error on a gigantic data set of CCSD(T) quality noncovalent interaction energies. *J Chem Theory Comput* 7(3):658–668.
78. Bochevarov AD, et al. (2013) Jaguar: A high-performance quantum chemistry software program with strengths in life and materials sciences. *Int J Quantum Chem* 113(18):2110–2142.

Supporting Information

General methods. All commercial reagents were used as supplied unless otherwise stated, and solvents were dried and distilled using standard techniques. Thin layer chromatography was performed on silica-coated glass plates (TLC Silica Gel 60 F_{254}, Merck) with detection by fluorescence, charring with 5% $H_2SO_{4(aq)}$, or staining with a ceric ammonium molybdate solution. Organic solutions were concentrated and/or evaporated to dry under vacuum in a water bath (<50 °C). Molecular sieves were dried at 400 °C under vacuum for 20-30 minutes prior to use. Amberlite IR-120H resin was washed extensively with MeOH and dried under vacuum prior to use. Medium-pressure liquid chromatography (MPLC) was performed using a CombiFlash Companion equipped with RediSep normal-phase flash columns, and solvent gradients refer to sloped gradients with concentrations reported as % *v*/*v*. NMR spectra were recorded on a Bruker Avance DMX-500 (500 MHz) spectrometer, and assignments achieved with the assistance of 1D NOESY, 1D TOCSY, 2D gCOSY, 2D gTOCSY, 2D gHSQC, 2D gHMBC, and 2D gNOESY; chemical shifts are expressed in ppm and referenced to either $Si(CH_3)_4$ (for $CDCl_3$), residual CHD_2OD (for CD_3OD), or MeOH (for D_2O). Electron-spray ionization mass spectrometry (ESI-MS) was performed using a Waters micromass ZQ. High resolution mass spectrometry was performed using an Agilent 1100 LC equipped with a photodiode array detector, and a Mircromass QTOF I equipped with a 4 GHz digital-time converter. Optical rotation was determined in a 10 cm cell at 20 °C using a Perkin-Elmer Model 341 polarimeter. HPLC analysis was performed using an Agilent 1100 LC equipped with an Agilent ZORBAX Eclipse XDB-C8 column and ELSD detection.



(a) MeOH, BF_3•Et_2O, CH_2Cl_2, 71%; (b) NaOMe, MeOH/CH_2Cl_2, 0 °C to rt, quant.; (c) $PhCH(OMe)_2$, CSA, CH_3CN, quant.; (d) BnBr, NaH, DMF, 0 °C to rt, 92%; (e) Et_3SiH, BF_3•Et_2O, CH_2Cl_2, 0 °C to rt, 79%.

Methyl 3,4,6-tri-*O*-acetyl-2-deoxy-2-phthalimido-b-D-glucopyranoside (16). The starting material (**15**; 2.193 g, 4.593 mmol),[ref1] MeOH (0.37 mL, 9.1 mmol), and BF_3•Et_2O (1.73 mL, 13.8 mmol) were added to anhydrous CH_2Cl_2 (20 mL) and left mixing at ambient temperature under Ar. After 48 hours, the reaction mixture was neutralized with Et_3N (to pH 8), diluted with CH_2Cl_2 (20 mL), and then the organic phase was washed with saturated $NaHCO_{3(aq)}$ solution (200 mL), saturated $NaCl_{(aq)}$ solution (200 mL), dried with Na_2SO_4, filtered, and evaporated to dry. The crude material was then purified via MPLC on silica gel using 0→30% EtOAc – toluene to afford the pure product as a white solid (1.465 g, 3.260 mmol, 71% yield). R_f = 0.58 (1:4 acetone : toluene). $[\alpha]_D^{20}$: +42° (*c* 1.0, $CHCl_3$). ^{1}H NMR ($CDCl_3$, 500 MHz): δH 7.88 – 7.84 (m, 2H, Ar), 7.77 – 7.73 (m, 2H, Ar), 5.79 (dd, 1H, *J* = 9.9, 9.9 Hz, H-3), 5.31 (d, 1H, *J* = 8.5 Hz, H-1), 5.19 (dd, 1H, *J* = 9.6, 9.6 Hz, H-4), 4.35 (dd, 1H, *J* = 12.2, 4.4 Hz, H-6^a), 4.34 – 4.29 (m, 1H, H-2), 4.20 (dd, 1H, *J* = 12.1, <2 Hz, H-6^b), 3.90 – 3.87 (m, 1H, H-5), 3.45 (s, 3H, OCH_3), 2.12 (s, 3H, Ac), 2.04 (s, 3H, Ac), 1.86 (s, 3H, Ac). ^{13}C NMR ($CDCl_3$, 125 MHz): δC 170.87 (Ac), 170.30 (Ac), 169.62 (Ac), 167.54 (Phth), 134.45 (Ar), 131.62 (Ar), 123.77 (Ar), 99.20 (C-1), 72.01 (C-5), 70.95 (C-3), 69.15 (C-4), 62.13 (C-6), 57.20 (OCH_3), 54.68 (C-2), 20.91 (Ac), 20.78 (Ac), 20.59 (Ac). ESI-MS *m*/*z* calc'd for $C_{21}H_{23}NO_{10}$ $(M+Na)^+$: 472.12; found: 472.08.

Methyl 2-deoxy-2-phthalimido-b-D-glucopyranoside (17). The starting material (**16**; 1.465 g, 3.260 mmol) was dissolved into anhydrous MeOH/CH_2Cl_2 (3:1; 20 mL), cooled to 0 °C, and then NaOMe solution added drop-wise (1.5 M NaOMe in MeOH; to pH 10). After 1 hour, the reaction flask was warmed to ambient temperature, and after an additional 30 minutes was then neutralized with acidic resin (Amberlite IR-120H; to pH 6), filtered, and evaporated to dry to afford the pure product as a white solid (1.050 g, 3.248 mmol, quantitative yield). R_f = 0.10 (1:19 MeOH : CH_2Cl_2). $[\alpha]_D^{20}$: -10° (*c* 1.0, MeOH). ^{1}H NMR (MeOD, 500 MHz): δH 7.89 – 7.80 (m, 4H, Ar), 5.09 (d, 1H, *J* = 8.5 Hz, H-1), 4.24 (dd, 1H, *J* = 10.7, 8.1 Hz, H-3), 3.96 (dd, 1H, *J* = 10.7, 8.5 Hz, H-2), 3.94 (dd, 1H, *J* = 12.0, 2.1 Hz, H-6^a), 3.76 (dd, 1H, *J* = 11.9, 5.3 Hz, H-6^b), 3.46 – 3.38 (m, 5H, H-5, OCH_3, and H-4). ^{13}C NMR (MeOD, 125 MHz): δC 169.96 (Phth), 135.66 (Ar), 133.20 (Ar), 124.34 (Ar), 100.82 (C-1), 78.45 (C-5), 72.78 (C-3), 72.74 (C-4), 62.86 (C-6), 58.60 (C-2), 57.19 (OCH_3). ESI-MS *m*/*z* calc'd for $C_{15}H_{17}NO_7$ $(M+Na)^+$: 346.09; found: 346.12.

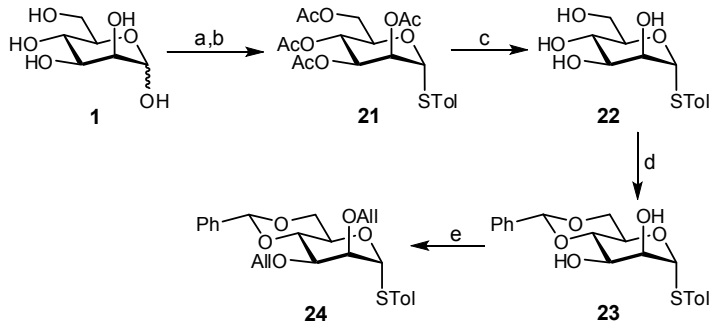
Methyl 4,6-*O*-benzylidene-2-deoxy-2-phthalimido-b-D-glucopyranoside (18). The starting material (**17**; 1.050 g, 3.248 mmol), $PhCH(OMe)_2$ (0.68 mL, 4.5 mmol), and camphorsulfonic acid (to pH 3) were added to anhydrous MeCN (10 mL) and left mixing at ambient temperature under Ar. After 14 hours, additional $PhCH(OMe)_2$ (0.30 mL, 2.0 mmol) was added to react with remaining starting material, and after another 6 hours the mixture was neutralized with Et_3N (to pH 8) and evaporated to dry. The crude material was purified via MPLC on silica gel using 0→15% EtOAc – toluene to afford the pure product as a white solid (1.333 g, 3.240 mmol, quantitative yield). R_f = 0.61 (1:4 acetone : toluene). $[\alpha]_D^{20}$: -26° (*c* 1.0, $CHCl_3$). ^{1}H NMR ($CDCl_3$, 500 MHz): δC 7.82 – 7.79 (m, 2H, Ar), 7.68 – 7.65 (m, 2H, Ar), 7.50 – 7.47 (m, 2H, Ar), 7.36 – 7.33 (m, 3H, Ar), 5.55 (s, 1H, PhC*H*), 5.15 (d, 1H, *J* = 8.5 Hz, H-1), 4.60 – 4.56 (m, 1H, H-3), 4.37 (dd, 1H, *J* = 10.6, 4.6 Hz, H-6^a), 4.21 – 4.17 (m, 1H, H-2), 3.81 (dd, 1H, *J* = 9.8, 9.8 Hz, H-6^b), 3.59 (ddd, 1H, *J* = 9.3, 9.3, 4.6 Hz, H-5), 3.56 (dd, 1H, *J* = 9.2, 9.2 Hz, H-4), 3.41 (s, 3H, OCH_3), 2.95 (d, 1H, *J* = 3.6 Hz, 3-OH). ^{13}C NMR ($CDCl_3$, 125 MHz): δC 168.33 (Phth), 137.18 (Ar), 134.23 (Ar), 131.79 (Ar), 129.42 (Ar), 128.47 (Ar), 126.47 (Ar), 123.62 (Ar), 101.99 (Ph*C*H), 99.95 (C-1), 82.34 (C-4), 68.79 (C-6), 68.65 (C-3), 66.27 (C-5), 57.18 (OCH_3), 56.70 (C-2). ESI-MS *m*/*z* calc'd for $C_{22}H_{21}NO_7$ $(M+Na)^+$: 434.12; found: 434.16.

Methyl 3-*O*-benzyl-4,6-*O*-benzylidene-2-deoxy-2-phthalimido-b-D-gluco-pyranoside (19). The starting material (**18**; 1.323 g, 3.216 mmol) and BnBr (0.50 mL, 4.2 mmol) were added to anhydrous DMF (12 mL), and then NaH was added portion-wise (60% oil dispersion; 193 mg, 4.82 mmol) at 0 °C under Ar. After 30 minutes, the reaction flask was warmed up to ambient temperature and after an additional 6 hours, further BnBr (0.10 mL, 0.84 mmol) and NaH (81 mg, 2.0 mmol) were added. After another 2 hours the reaction was quenched via the slow addition of MeOH (5 mL), evaporated to dry, and then redissolved into EtOAc (60 mL). The organic phase was washed with saturated $NaCl_{(aq)}$ solution (2 x 60 mL), dried with Na_2SO_4, filtered, and evaporated to dry. The crude material was purified via MPLC on silica gel using 0→30% EtOAc – toluene to afford the pure product as a white solid (1.478 g, 2.947 mmol, 92% yield). R_f = 0.86 (1:4 acetone : toluene). $[\alpha]_D^{20}$: +46° (*c* 1.0, $CHCl_3$). ^{1}H NMR ($CDCl_3$, 500 MHz): δH 7.86 – 7.60 (m, 4H, Ar), 7.54 – 7.51 (m, 2H, Ar), 7.42 – 7.35 (m, 3H, Ar), 7.00 – 6.98 (m, 2H, Ar), 6.94 – 6.86 (m, 3H, Ar), 5.63 (s, 1H, PhC*H*), 5.13 (d, 1H, *J* = 8.5 Hz, H-1), 4.80 (d, 1H, *J* = 12.3 Hz, PhC*H*^{a}H^b), 4.50 (d,

1H, *J* = 12.3 Hz, PhCH^{a}*H*^{b}), 4.45 – 4.40 (m, 2H, H-3 and H-6^{a}), 4.23 – 4.19 (m, 1H, H-2), 3.86 (dd, 1H, *J* = 10.3, 10.3 Hz, H-6^{b}), 3.81 (dd, 1H, *J* = 9.2, 9.2 Hz, H-4), 3.64 (ddd, 1H, *J* = 9.7, 9.7, 5.0 Hz, H-5), 3.40 (s, 3H, OCH_3). ^{13}C NMR ($CDCl_3$, 125 MHz): δC 168.01 (Phth), 138.07 (Ar), 137.51 (Ar), 134.00 (Ar), 131.84 (Ar), 129.19 (Ar), 128.47 (Ar), 128.20 (Ar), 128.19 (Ar), 127.55 (Ar), 126.23 (Ar), 123.52 (Ar), 101.49 (Ph*C*H), 100.02 (C-1), 83.30 (C-4), 74.78 (C-3), 74.26 (Ph*C*H_2), 68.96 (C-6), 66.23 (C-5), 57.15 (OCH_3), 55.88 (C-2). ESI-MS *m*/*z* alc'd for $C_{29}H_{27}NO_7$ $(M+Na)^+$: 524.17; found: 524.13.

Methyl 3,6-di-*O*-benzyl-2-deoxy-2-phthalimido-b-D-glucopyranoside (20).

A solution of the starting material (**19**; 1.999 g, 3.986 mmol) in anhydrous CH_2Cl_2 (20 mL) was cooled to 0 °C under Ar, and then Et_3SiH (3.2 mL, 20 mmol) and BF_3•Et_2O (1.0 mL, 8.0 mmol) were slowly added. After 6 hours, the reaction mixture was neutralized with Et_3N (to pH 8), warmed back to ambient temperature, and then quenched via the slow addition of MeOH (5 mL). The crude mixture was evaporated to dry, and then purified via MPLC on silica gel using 0→30% EtOAc – toluene to afford the pure product as a white solid (1.595 g, 3.168 mmol, 79% yield). R_f = 0.13 (1:9 EtOAc : toluene). $[\alpha]_D^{20}$: +42° (*c* 1.0, $CHCl_3$). ^1H NMR ($CDCl_3$, 500 MHz): δH 7.81 – 7.63 (m, 4H, Ar), 7.37 – 7.27 (m, 5H, Ar), 7.04 – 7.02 (m, 2H, Ar), 6.96 – 6.91 (m, 3H, Ar), 5.05 (d, 1H, *J* = 8.4 Hz, H-1), 4.73 (d, 1H, *J* = 12.2 Hz, PhC*H*^{a}H^{b}), 4.64 (d, 1H, *J* = 12.0 Hz, PhC*H*^{a}H^{b}), 4.58 (d, 1H, *J* = 12.0 Hz, PhCH^{a}*H*^{b}), 4.51 (d, 1H, *J* = 12.2 Hz, PhCH^{a}*H*^{b}), 4.22 (dd, 1H, *J* = 10.6, 8.4 Hz, H-3), 4.13 (dd, 1H, *J* = 10.6, 8.4 Hz, H-2), 3.83 (dd, 1H, *J* = 10.1, 4.9 Hz, H-6^{a}), 3.83 – 3.79 (m, 1H, H-5), 3.78 (dd, 1H, *J* = 10.1, 5.1 Hz, H-6^{b}), 3.63 (ddd, 1H, *J* = 9.5, 4.9, 4.9 Hz, H-5), 3.36 (s, 3H, OCH_3), 2.90 (d, 1H, *J* = <1 Hz, 4-OH). ^{13}C NMR ($CDCl_3$, 125 MHz): δC 167.89 (Phth), 138.38 (Ar), 137.82 (Ar), 133.99 (Ar), 128.73 (Ar), 128.36 (Ar), 128.12 (Ar), 128.05 (Ar), 128.02 (Ar), 127.62 (Ar), 99.43 (C-1), 78.92 (C-3), 74.71 (C-4), 74.53 (Ph*C*H_2), 74.01 (Ph*C*H_2), 73.73 (C-5), 70.90 (C-6), 56.87 (OCH_3), 55.46 (C-2). ESI-MS *m*/*z* calc'd for $C_{29}H_{29}NO_7$ $(M+Na)^+$: 526.18; found: 526.18.



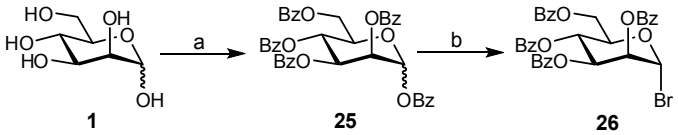
(a) Ac_2O, Py, 60 °C; (b) *p*-TolSH, BF_3•Et_2O, CH_2Cl_2, quant. over 2 steps; (c) NaOMe, MeOH, 50 °C, quant.; (d) $PhCH(OMe)_2$, CSA, CH_3CN/DMF, 60%; (e) AllBr, NaH, DMF, 97%.

***p*-Methylphenyl 2,3,4,6-tetra-*O*-acetyl-1-thio-a-D-mannopyranoside (21).** The starting material (**1**; 8.15 g, 45.2 mmol) was dissolved into pyridine (40 mL) and Ac_2O (40 mL) and left mixing at 65 °C, and after 3 hours was concentrated to a syrup via co-evaporation with toluene (3 x 50 mL). The crude mixture, *p*-methylthiophenol (6.803 g, 54.77 mmol), and BF_3•Et_2O (7.4 mL, 59 mmol) in anhydrous CH_2Cl_2 (150 mL) were then left mixing under Ar at ambient temperature. After 20 hours, additional *p*-methylthiophenol (1.774 g, 14.28 mmol) was added to react with some remaining starting material. After another 24 hours, the reaction mixture was neutralized with the addition of Et_3N (to pH 8), diluted further with CH_2Cl_2 (300 mL), washed with saturated $NaCl_{(aq)}$ solution (2 x 500 mL), dried with Na_2SO_4, filtered, and evaporated to dry. The crude material was then purified via MPLC on silica gel using 0→30% acetone – toluene to afford the pure product as a white solid (20.34 g, 44.75 mmol, quantitative yield over 2 steps). R_f = 0.71 (2:3 acetone : toluene). $[\alpha]_D^{20}$: +103° (*c* 1.0, $CHCl_3$). ^{1}H NMR ($CDCl_3$, 500 MHz): δH 7.39 – 7.36 (m, 2H, Ar), 7.13 – 7.10 (m, 2H, Ar), 5.50 (dd, 1H, *J* = <2, <2 Hz, H-2), 5.42 (d, 1H, *J* = 1.0 Hz, H-1), 5.36 – 5.30 (m, 2H, H-4 and H-3), 4.59 – 4.52 (m, 1H, H-5), 4.30 (dd, 1H, *J* = 12.2, 6.0 Hz, H-6^a), 4.10 (dd, 1H, *J* = 12.2, 2.3 Hz, H-6^b), 2.32 (s, 3H, PhCH_3), 2.13 (s, 3H, Ac), 2.07 (s, 3H, Ac), 2.05 (s, 3H, Ac), 2.01 (s, 3H, Ac). ^{13}C NMR ($CDCl_3$, 125 MHz): δC 170.51 (Ac), 169.90 (Ac), 169.80 (Ac), 169.76 (Ac), 138.46 (Ar), 132.69 (Ar), 130.01 (Ar), 128.86 (Ar), 86.05 (C-1), 70.91 (C-2), 69.50 (C-5), 69.43 (C-3), 66.47 (C-4), 62.54 (C-6), 21.15 (Ph*C*H_3), 20.87 (Ac), 20.73 (Ac), 20.70 (Ac), 20.65 (Ac). ESI-MS *m*/*z* calc'd for $C_{21}H_{26}O_9S$ $(M+Na)^+$: 477.12; found: 477.18.

***p*-Methylphenyl 1-thio-a-D-mannopyranoside (22).** The starting material (**21**; 20.34 g, 44.75 mmol) was dissolved into anhydrous MeOH (200 mL), and then NaOMe solution was added drop-wise (1.5 M NaOMe in MeOH; to pH 10). After 1 hour at ambient temperature, followed by 1 hour at 50 °C, the reaction mixture was neutralized with acidic resin (Amberlite IR-120H; to pH 6), filtered, and then evaporated to dry to afford the pure product as a white solid (13.09 g, 45.71 mmol, quantitative yield). R_f = 0.63 (3:17 MeOH : CH_2Cl_2). $[\alpha]_D^{20}$: +203° (*c* 1.0, MeOH). ^{1}H NMR (MeOD, 500 MHz): δH 7.41 – 7.39 (m, 2H, Ar), 7.14 – 7.11 (m, 2H, Ar), 5.36 (d, 1H, *J* = 1.2 Hz, H-1), 4.08 (dd, 1H, *J* = 3.0, 1.5 Hz, H-2), 4.05 (ddd, 1H, *J* = 9.1, 5.3, 2.5 Hz, H-5), 3.82 (dd, 1H, *J* = 12.0, 2.5 Hz, H-6^a), 3.77 (dd, 1H, *J* = 12.0, 5.3 Hz, H-6^b), 3.73 (dd, 1H, *J* = 9.3, 9.3 Hz, H-4), 3.70 (dd, 1H, *J* = 9.3, 3.0 Hz, H-3), 2.30 (s, 3H, PhC*H*$_3$). ^{13}C NMR (MeOD, 125 MHz): δC 139.02 (Ar), 133.62 (Ar), 132.21 (Ar), 130.88 (Ar), 90.88 (C-1), 75.64 (C-5), 73.85 (C-2), 73.26 (C-3), 68.83 (C-4), 62.74 (C-6), 21.24 (Ph*C*H$_3$). ESI-MS *m*/*z* calc'd for $C_{13}H_{18}O_5S$ $(M+Na)^+$: 309.08; found: 309.05.

***p*-Methylphenyl 4,6-*O*-benzylidene-1-thio-a-D-mannopyranoside (23).**
The starting material (**22**; 13.09 g, 45.71 mmol), $PhCH(OMe)_2$ (8.9 mL, 59 mmol), and camphorsulfonic acid (to pH 3) were added to anhydrous MeCN (140 mL) and anhydrous DMF (50 mL) and left mixing at ambient temperature under Ar. After 16 hours, the mixture was neutralized with Et_3N (to pH 8) and evaporated to dry. The crude material was purified via MPLC on silica gel using 0→20% MeOH – CH_2Cl_2 to afford the pure product as a white solid (10.32 g, 27.56 mmol, 60% yield). R_f = 0.62 (1:9 MeOH : CH_2Cl_2). $[\alpha]_D^{20}$: +276° (*c* 0.50, $CHCl_3$). ^{1}H NMR (MeOD, 500MHz): δH 7.53 – 7.51 (m, 2H, Ar), 7.39 – 7.34 (m, 5H, Ar), 7.18 – 7.16 (m, 2H, Ar), 5.62 (s, 1H, PhC*H*), 5.41 (d, 1H, *J* = <1 Hz, H-1), 4.25 (ddd, 1H, *J* = 9.8, 9.8, 4.9 Hz, H-5), 4.16 (dd, 1H, *J* = 3.2, 1.0 Hz, H-2), 4.13 (dd, 1H, *J* = 10.2, 4.9 Hz, H-6^a), 4.02 (dd, 1H, *J* = 9.7, 9.7 Hz, H-4), 3.94 (dd, 1H, *J* = 9.9, 3.3 Hz, H-3), 3.83 (dd, 1H, *J* = 10.2, 10.2 Hz, H-6^b), 2.33 (s, 3H, CH_3). ^{13}C NMR (MeOD, 125 MHz): δC 139.39 (Ar), 139.35 (Ar), 133.67 (Ar), 131.64 (Ar), 131.04 (Ar), 130.06 (Ar), 129.19 (Ar), 127.68 (Ar), 103.55 (Ph*C*H), 91.51 (C-1), 80.45 (C-4), 74.27 (C-2), 70.14 (C-3), 69.64 (C-6), 66.45 (C-5), 21.25 (CH_3). ESI-MS *m*/*z* calc'd for $C_{20}H_{22}O_5S$ $(M+Na)^+$: 397.11; found: 397.02.

***p*-Methylphenyl 2,3-di-*O*-allyl-4,6-*O*-benzylidene-1-thio-a-D-mannopyranoside (24).** The starting material (**23**; 1.633 g, 4.361 mmol) and allyl bromide (0.91 mL, 11 mmol) were added to anhydrous DMF (15 mL), and then NaH (60% oil dispersion; 531 mg, 13.3 mmol) was slowly added portion-wise under Ar. After 16 hours, the reaction was quenched via the slow addition of MeOH (5 mL), evaporated to dry, and then redissolved into EtOAc (200 mL). The organic phase was washed with saturated $NaCl_{(aq)}$ solution (2 x 200 mL), dried with Na_2SO_4, filtered, and evaporated to dry. The crude material was purified via MPLC on silica gel using 0→10% EtOAc – toluene to afford the pure product as a white solid (1.927 g, 4.239 mmol, 97% yield). R_f = 0.44 (1:19 EtOAc : toluene). $[\alpha]_D^{20}$: +160° (*c* 0.50, $CHCl_3$). ^{1}H NMR ($CDCl_3$, 500 MHz): δH 7.51 – 7.49 (m, 2H, Ar), 7.39 – 7.33 (m, 5H, Ar), 7.14 – 7.12 (m, 2H, Ar), 5.98 – 5.88 (m, 2H, 2x OCH_2CH=CH_2), 5.61 (s, 1H, PhC*H*), 5.47 (d, 1H, *J* = <1 Hz, H-1), 5.35 – 5.28 (m, 2H, 2x OCH_2CH=C*H*^{a}H^b), 5.22 – 5.19 (m, 2H, 2x OCH_2CH=CH^a*H*^{b}), 4.36 – 4.28 (m, 2H, OC*H*^{a}H^{b}CH=CH_2 and H-5), 4.25 – 4.17 (m, 5H, H-6^a, OC*H*^{a}H^{b}CH=CH_2, 2x OCH^a*H*^{b}CH=CH_2, and H-4), 4.03 (dd, 1H, *J* = <2, <2 Hz, H-2), 3.90 – 3.84 (m, 2H, H-3 and H-6^b), 2.34 (s, 3H, PhC*H*$_3$). ^{13}C NMR ($CDCl_3$, 125 MHz): δC 138.15 (Ar), 137.82 (Ar), 135.01 (OCH_2*C*H=CH_2), 134.79 (OCH_2*C*H=CH_2), 132.32 (Ar), 130.28 (Ar), 130.15 (Ar), 129.07 (Ar), 128.41 (Ar), 126.30 (Ar), 118.14 (OCH_2CH=*C*H_2), 117.17 (OCH_2CH=*C*H_2), 101.75 (Ph*C*H), 87.89 (C-1), 79.43 (C-4), 78.35 (C-2), 75.98 (C-3), 72.68 (O*C*H_2CH=CH_2), 72.31 (O*C*H_2CH=CH_2), 68.76 (C-6), 65.43 (C-5), 21.33 (Ph*C*H_3). ESI-MS *m*/*z* calc'd for $C_{26}H_{30}O_5S$ $(M+Na)^+$: 477.17; found: 477.18.



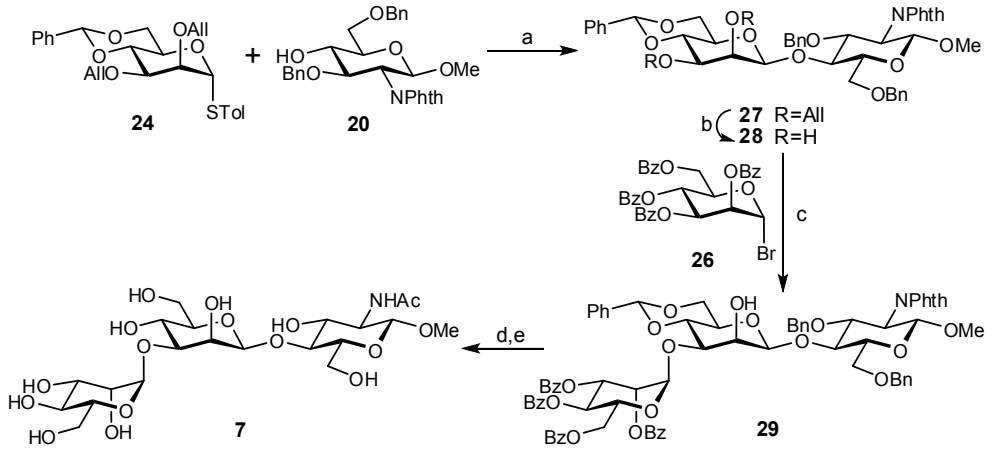
1 **25** **26**

(a) BzCl, Py/CH_2Cl_2, rt to 80 °C, quant.; (b) HBr in AcOH, CH_2Cl_2, 94%.

Benzoyl 2,3,4,6-tetra-*O*-benzoyl-D-mannopyranose (25). The starting material (**1**; 3.70 gm, 20.5 mmol) and benzoyl chloride (14.5 mL, 125 mmol) were added to anhydrous pyridine (40 mL), and after 5 minutes at ambient temperature was begun heating at 60 °C. After 5 hours, the reaction mixture was heated at 100 °C for 15 minutes, and then cooled back to ambient temperature and quenched via the addition of excess ice. The mixture was further diluted with EtOAc (400 mL), then washed with saturated $NaCl_{(aq)}$ solution (400 mL), dried with Na_2SO_4, filtered, and evaporated

to dry. The crude material was then purified via MPLC on silica gel using 0→10% EtOAc – toluene to afford the a-/b-product mixture as a white solid (14.50 g, 20.69 mmol, quantitative yield).[ref2] ESI-MS *m/z* calc'd for $C_{41}H_{32}O_{11}$ $(M+Na)^+$: 723.18; found: 723.33.

2,3,4,6-Tetra-*O*-benzoyl-a-D-mannopyranosyl bromide (26). The starting material (**25**; 14.50 g, 20.69 mmol) was added to anhydrous CH_2Cl_2 (100 mL) and HBr/AcOH solution (33 wt.%; 50 mL) and left mixing at ambient temperature. After 24 hours, the solution was concentrated to a syrup via co-evaporation with toluene, and then purified via MPLC on silica gel using 0→10% EtOAc – toluene to afford the pure product as a white solid (12.83 g, 19.45 mmol, 94% yield). R_f = 0.75 (1:4 EtOAc : toluene). $[\alpha]_D^{20}$: -18° (*c* 1.0, $CHCl_3$). ^{1}H NMR ($CDCl_3$, 500 MHz): δH 8.12 – 8.10 (m, 2H, Ar), 8.04 – 8.02 (m, 2H, Ar), 7.99 – 7.97 (m, 2H, Ar), 7.85 – 7.83 (m, 2H, Ar), 7.61 – 7.56 (m, 2H, Ar), 7.53 – 7.50 (m, 1H, Ar), 7.44 – 7.36 (m, 7H, Ar), 7.28 – 7.24 (m, 2H, Ar), 6.59 (d, 1H, *J* = <1 Hz, H-1), 6.30 (dd, 1H, *J* = 10.2, 2.9 Hz, H-3), 6.26 (dd, 1H, *J* = 9.9, 9.9 Hz, H-4), 5.93 – 5.92 (m, 1H, H-2), 4.75 (dd, 1H, *J* = 12.5, 2.0 Hz, H-6^a), 4.68 – 4.65 (m, 1H, H-5), 4.52 (dd, 1H, *J* = 12.5, 3.6 Hz, H-6^b). ^{13}C NMR ($CDCl_3$, 125 MHz): δC 166.10 (C=O), 165.53 (C=O), 165.46 (C=O), 165.13 (C=O), 133.94 (Ar), 133.84 (Ar), 133.56 (Ar), 133.36 (Ar), 130.06 (Ar), 130.05 (Ar), 129.97 (Ar), 129.94 (Ar), 128.87 (Ar), 128.71 (Ar), 128.67 (Ar), 128.56 (Ar), 83.49 (C-1), 73.31 (C-5), 73.13 (C-2), 69.28 (C-3), 66.11 (C-4), 61.93 (C-6). ESI-MS *m/z* calc'd for $C_{34}H_{27}BrO_9$ $(M+Na)^+$: 681.07; found: 681.12.



(a) 1-benzenesulfinyl piperidine, 2,4,6-tri-*tert*-butylpyrimidine, Tf_2O, CH_2Cl_2, -60 °C to rt, 25%; (b) $Pd(OAc)_2$, PPh_3, 1,3-dimethylbarbituric acid, MeOH/CH_2Cl_2, 40 °C, 74%; (c) AgOTf, CH_2Cl_2, -60 °C to -40 °C, 40%; (d) NH_2NH_2•H_2O, EtOH, 70 °C; then LiOH, 50 °C; then $NaHCO_3$, H_2O, Ac_2O; (e) $Pd(OH)_2$, H_2, MeOH, 99% (2 steps).

Methyl 2,3-di-*O*-allyl-4,6-*O*-benzylidene-b-D-mannopyranosyl-(1→4)-3,6-di-*O*-benzyl-2-deoxy-2-phthalimido-b-D-glucopyranoside (27). The glycosyl donor (**24**; 84 mg, 0.18 mmol), 1-benzenesulfinyl piperidine (41 mg, 0.20 mmol), 2,4,6-tri-*tert*-butylpyrimidine (92 mg, 0.37 mmol), and crushed molecular sieves (4 Å, 213 mg) in anhydrous CH_2Cl_2 (5.0 mL) were left mixing at ambient temperature under Ar [ref3]. After 1 hour, the flask was cooled to –60 °C and then Tf_2O (87 μL, 0.52 mmol) was added. After 5 minutes, a solution of the glycosyl acceptor (**20**; 108 mg, 0.214 mmol) in anhydrous CH_2Cl_2 (1.0 mL) was added, and after another 2 minutes the reaction mixture was warmed to ambient temperature over 15 minutes. The reaction solution was filtered, diluted with CH_2Cl_2 (20 mL), and then washed with saturated $NaHCO_{3(aq)}$ solution (20 mL), saturated $NaCl_{(aq)}$ solution (20 mL), dried with Na_2SO_4, filtered, and evaporated to dry. The crude material was purified via MPLC on silica gel using 0→30% EtOAc – toluene to afford the pure product as a white solid (38 mg, 0.046 mmol, 25% yield). R_f = 0.84 (1:4 acetone : toluene). $[\alpha]_D^{20}$: +20° (*c* 0.53, $CHCl_3$). ^{1}H NMR ($CDCl_3$, 500 MHz): δH 7.79 – 7.61 (m, 4H, Ar), 7.44 – 7.42 (m, 2H, Ar), 7.39 – 7.29 (m, 8H, Ar), 7.00 – 6.98 (m, 2H, Ar), 6.89 – 6.83 (m, 3H, Ar), 5.98 (dddd, 1H, *J* = 17.2, 10.5, 5.5, 5.5 Hz, OCH_2CH=CH_2), 5.90 (dddd, 1H, *J* = 17.2, 10.6, 5.2, 5.2 Hz, OCH_2CH=CH_2), 5.47 (s, 1H, PhC*H*), 5.34 – 5.28 (m, 2H, 2x OCH_2CH=CH_2), 5.18 – 5.15 (m, 2H, 2x OCH_2CH=CH_2), 5.06 (d, 1H, *J* = 8.5 Hz, GlcN_H1), 4.89 (d, 1H, *J* = 12.3 Hz, PhCH^aH^b), 4.77 (d, 1H, *J* = 12.1 Hz, PhCH^aH^b), 4.58 (d, 1H, *J* = <1 Hz, Man_H1), 4.54 (d, 1H, *J* = 12.1 Hz, PhCH^aH^b), 4.46 (d, 1H, *J* = 12.3 Hz, PhCH^aH^b), 4.36 – 4.22 (m, 4H, 3x OCH_2CH=CH_2 and GlcN_H3), 4.18 – 4.14 (m, 2H, GlcN_H2 and Man_H6^a), 4.10 – 4.05 (m, 2H, OCH_2CH=CH_2 and GlcN_H4), 3.95 (dd, 1H, *J* = 9.6, 9.6 Hz, Man_H4), 3.82 – 3.76 (m, 2H, GlcN_H6^a and GlcN_H6^b), 3.69 (dd, 1H, *J* = 2.4, <1 Hz, Man_H2), 3.63 – 3.60 (m, 1H, GlcN_H5), 3.53 (dd, 1H, *J* = 10.2, 10.2 Hz, Man_H6^b), 3.40 (s, 3H, OCH_3), 3.33 (dd, 1H, *J* = 9.9, 2.9 Hz, Man_H3), 3.14 (ddd, 1H, *J* = 9.7, 9.7, 4.9 Hz, Man_H5). ^{13}C NMR ($CDCl_3$, 125 MHz): δC 168.0 (Phth), 138.83 (Ar), 137.89 (Ar), 137.66 (Ar), 135.57 (OCH_2CH=CH_2), 134.97 (OCH_2CH=CH_2), 133.71 (Ar), 131.80 (Ar), 128.79 (Ar), 128.57 (Ar), 128.15 (Ar), 127.99 (Ar), 127.85 (Ar), 127.83 (Ar), 126.98 (Ar), 126.07 (Ar), 123.26 (Ar), 116.79 (OCH_2CH=CH_2), 116.44 (OCH_2CH=CH_2), 101.71 (Man_C1), 101.34 (PhCH), 99.32 (GlcN_C1), 79.10 (GlcN_C4), 78.65 (Man_C4), 78.12 (Man_C3), 77.15 (GlcN_C3), 76.96 (Man_C2), 74.88 (GlcN_C5), 74.61 (PhCH_2), 74.36 (OCH_2CH=CH_2), 73.69 (PhCH_2), 71.72 (OCH_2CH=CH_2), 68.62

(GlcN_C6), 68.57 (Man_C6), 67.26 (Man_C5), 56.60 (OCH_3), 55.60 (GlcN_C2). ESI-MS *m*/*z* calc'd for $C_{48}H_{51}NO_{12}$ $(M+Na)^+$: 856.33; found: 856.40.

Methyl 4,6-*O*-benzylidene-a-D-mannopyranosyl-(1→4)-3,6-di-*O*-benzyl-2-deoxy-2-phthalimido-b-D-glucopyranoside (28). A solution of $Pd(OAc)_2$ (13 mg, 0.060 mmol) and PPh_3 (60 mg, 0.23 mmol) in anhydrous MeOH/CH_2Cl_2 (2:1, 3.0 mL) were left mixing at ambient temperature under Ar. After 20 minutes, the bright yellow solution was transferred to a flask containing the starting material (**27**; 177 mg, 0.212 mmol) and 1,3-dimethylbarbituric acid (134 mg, 0.858 mmol) and left mixing in a sealed flask at 40 °C [ref4]. After 30 hours, the solution was diluted with CH_2Cl_2 (100 mL), washed with saturated $Na_2CO_{3(aq)}$ solution (100 mL), and the aqueous layer re-extracted with CH_2Cl_2 (50 mL). The combined organic phases were then washed with saturated $NaCl_{(aq)}$ solution (2 x 100 mL), dried with Na_2SO_4, filtered, and evaporated to dry. The crude material was purified via MPLC on silica gel using 0→40% acetone – petroleum ether to afford the nearly pure product as a pale yellow solid (119 mg, 0.158 mmol, 74% yield). R_f = 0.25 (2:3 acetone : petroleum ether). $[\alpha]_D^{20}$: +41° (*c* 1.0, $CHCl_3$). ^{1}H NMR ($CDCl_3$, 500 MHz): δH 7.76 – 7.64 (m, 4H, Ar), 7.47 – 7.44 (m, 2H, Ar), 7.41 – 7.34 (m, 8H, Ar), 7.03 – 7.01 (m, 2H, Ar), 6.95 – 6.86 (m, 3H, Ar), 5.46 (s, 1H, PhC*H*), 5.04 (d, 1H, *J* = 8.5 Hz, GlcN_H1), 4.81 (d, 1H, *J* = 12.1 Hz, PhC*H*^{a}H^b), 4.77 (d, 1H, *J* = 12.0 Hz, PhC*H*^{a}H^b), 4.73 (d, 1H, *J* = <1 Hz, Man_H1), 4.53 (d, 1H, *J* = 12.0 Hz, PhCH^a*H*^{b}), 4.43 (d, 1H, *J* = 12.1 Hz, PhCH^a*H*^{b}), 4.38 (dd, 1H, *J* = 10.6, 8.7 Hz, GlcN_H3), 4.18 (dd, 1H, *J* = 10.7, 8.5 Hz, GlcN_H2), 4.16 (dd, 1H, *J* = 10.5, 5.0 Hz, Man_H6^a), 4.14 – 4.10 (m, 1H, GlcN_H4), 3.92 – 3.91 (m, 1H, Man_H2), 3.83 (dd, 1H, *J* = 11.3, 3.1 Hz, GlcN_H6^a), 3.79 – 3.74 (m, 2H, GlcN_H6^b and Man_H4), 3.64 (ddd, 1H, *J* = 9.7, 2.5, 2.5 Hz, GlcN_H5), 3.59 (ddd, 1H, *J* = 9.9, 6.8, 3.4 Hz, Man_H3), 3.56 (dd, 1H, *J* = 10.2, 10.2 Hz, Man_H6^b), 3.39 (s, 3H, OCH_3), 3.13 (ddd, 1H, *J* = 9.7, 9.7, 5.0 Hz, Man_H5), 2.80 (d, 1H, *J* = 2.1 Hz, Man_2OH), 2.66 (d, 1H, *J* = 6.8 Hz, Man_3OH). ^{13}C NMR ($CDCl_3$, 125 MHz): δC 138.49 (Ar), 137.76 (Ar), 137.38 (Ar), 133.98 (Ar), 132.35 (Ar), 132.27 (Ar), 132.17 (Ar), 132.15 (Ar), 131.88 (Ar), 129.37 (Ar), 128.81 (Ar), 128.76 (Ar), 128.67 (Ar), 128.48 (Ar), 128.34 (Ar), 128.31 (Ar), 128.21 (Ar), 127.81 (Ar), 127.43 (Ar), 126.48 (Ar), 123.52 (Ar), 102.26 (Ph*C*H), 100.77 ($^1J_{C,H}$ = 158 Hz, Man_C1), 99.53 ($^1J_{C,H}$ = 162 Hz, GlcN_C1), 79.06 (GlcN_C4), 78.66 (Man_C4), 78.03 (GlcN_C3), 74.93 (Ph*C*H_2), 74.67 (GlcN_C5), 74.01 (Ph*C*H_2), 70.98 (Man_C2), 70.92 (Man_C3), 68.68

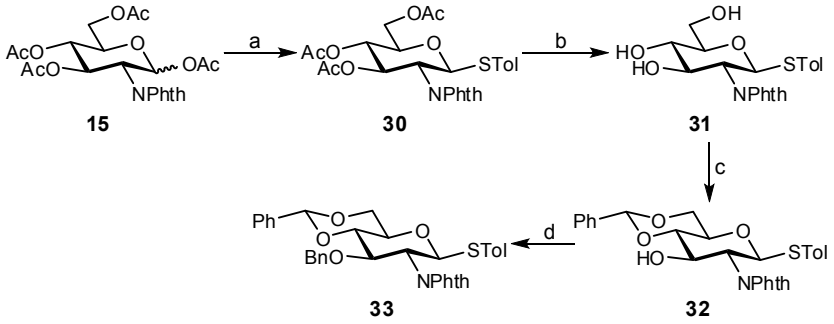
(GlcN_C6), 68.63 (Man_C6), 66.82 (Man_C5), 56.89 (OCH_3), 55.82 (GlcN_C2). ESI-MS *m*/*z* calc'd for $C_{42}H_{43}NO_{12}$ $(M+Na)^+$: 776.27; found: 776.23.

Methyl 2,3,4,6-tetra-*O*-benzoyl-a-D-mannopyranosyl-(1→3)-4,6-*O*-benzylidene-a-D-mannopyranosyl-(1→4)-3,6-di-*O*-benzyl-2-deoxy-2-phthalimido-b-D-glucopyranoside (29). The glycosyl acceptor (**28**; 119 mg, 0.158 mmol), glycosyl donor (**26**; 129 mg, 0.196 mmol), and crushed molecular sieves (3 Å, 207 mg) in anhydrous CH_2Cl_2 (2.0 mL) were left mixing for 20 minutes at ambient temperature under Ar. The foil-covered reaction flask was then cooled to –60 °C, and AgOTf (52 mg, 0.20 mmol) was added over 5 minutes. After 1 hour, the flask was warmed to –40 °C, and after another 1.5 hours additional glycosyl donor (**26**; 50 mg, 0.076 mmol) was added. After 1 hour, the solution was neutralized with Et_3N (to pH 8), warmed to ambient temperature, diluted with CH_2Cl_2 (30 mL), and then filtered over Celite. The organic phase was washed with $NH_4OH_{(aq)}$ solution (0.5 M; 2 x 30 mL), saturated $NaCl_{(aq)}$ solution (30 mL), dried with Na_2SO_4, filtered, and evaporated to dry. The crude material was purified via MPLC on silica gel using 0→20% acetone – toluene to afford the pure product as a white solid (84 mg, 0.063 mmol, 40% yield). R_f = 0.47 (2:3 acetone : petroleum ether). $[\alpha]_D^{20}$: -27° (*c* 1.0, $CHCl_3$). ^{1}H NMR ($CDCl_3$, 500 MHz): δH 8.09 – 8.07 (m, 2H, Ar), 8.06 – 8.04 (m, 2H, Ar), 7.97 – 7.95 (m, 2H, Ar), 7.83 – 7.67 (m, 6H, Ar), 7.64 – 7.58 (m, 2H, Ar), 7.52 – 7.23 (m, 22H, Ar), 7.05 – 7.03 (m, 2H, Ar), 6.98 – 6.90 (m, 3H, Ar), 6.10 (dd, 1H, *J* = 10.1, 10.1 Hz, Man_H4'), 6.01 (dd, 1H, *J* = 10.1, 3.3 Hz, Man_H3'), 5.81 (dd, 1H, *J* = 3.2, 1.9 Hz, Man_H2'), 5.54 (s, 1H, PhC*H*), 5.41 (d, 1H, *J* = 1.7 Hz, Man_H1'), 5.04 (d, 1H, *J* = 8.5 Hz, GlcN_H1), 4.83 (d, 1H, *J* = 12.1 Hz, PhC*H*^{a}H^b), 4.77 (d, 1H, *J* = 12.1 Hz, PhC*H*^{a}H^b), 4.70 – 4.65 (m, 3H, Man_H6$^{a'}$, Man_H5', and Man_H1), 4.51 (d, 1H, *J* = 12.2 Hz, PhCH^a*H*^{b}), 4.51 – 4.47 (m, 1H, Man_H6$^{b'}$), 4.44 (d, 1H, *J* = 12.1 Hz, PhCH^a*H*^{b}), 4.35 (dd, 1H, *J* = 10.7, 8.6 Hz, GlcN_H3), 4.20 (dd, 1H, *J* = 10.7, 8.5 Hz, GlcN_H2), 4.16 (dd, 1H, *J* = 10.4, 4.8 Hz, Man_H6^a), 4.14 – 4.10 (m, 2H, Man_H2 and GlcN_H4), 4.09 (dd, 1H, *J* = 9.6, 9.6 Hz, Man_H4), 3.80 – 3.74 (m, 2H, GlcN_H6^a and GlcN_H6^b), 3.73 (dd, 1H, *J* = 9.9, 3.2 Hz, Man_H3), 3.57 (dd, 1H, *J* = 10.2, 10.2 Hz, Man_H6^b), 3.55 (ddd, 1H, *J* = 9.8, 2.5, 2.5 Hz, GlcN_H5), 3.41 (s, 3H, OCH_3), 3.13 (ddd, 1H, *J* = 9.7, 9.7, 4.9 Hz, Man_H5), 2.56 (d, 1H, *J* = 2.1 Hz, Man_2OH). ^{13}C NMR ($CDCl_3$, 125 MHz): δC 166.44 (C=O), 165.77 (C=O), 165.58 (C=O), 165.30 (C=O), 138.58 (Ar), 137.89 (Ar), 137.30 (Ar), 134.01 (Ar), 133.63

(Ar), 133.49 (Ar), 133.34 (Ar), 131.94 (Ar), 130.07 (Ar), 130.06 (Ar), 129.96 (Ar), 129.95 (Ar), 129.59 (Ar), 129.35 (Ar), 129.22 (Ar), 128.95 (Ar), 128.85 (Ar), 128.79 (Ar), 128.64 (Ar), 128.49 (Ar), 128.34 (Ar), 128.24 (Ar), 128.23 (Ar), 127.77 (Ar), 127.43 (Ar), 126.24 (Ar), 123.54 (Ar), 101.57 (Ph*C*H), 100.35 ($^1J_{C,H}$ = 157 Hz, Man_C1), 99.52 ($^1J_{C,H}$ = 164 Hz, GlcN_C1), 99.13 ($^1J_{C,H}$ = 175 Hz, Man_C1'), 78.73 (GlcN_C4), 77.94 (Man_C3), 77.82 (GlcN_C3), 77.23 (Man_C4), 74.80 (Ph*C*H_2), 74.67 (GlcN_C5), 73.86 (Ph*C*H_2), 71.30 (Man_C2), 70.62 (Man_C2'), 69.99 (Man_C3'), 69.44 (Man_C5'), 68.58 (Man_C6 and GlcN_C6), 67.63 (Man_C4'), 67.04 (Man_C5), 63.40 (Man_C6'), 56.88 (OCH_3), 55.79 (GlcN_C2). ESI-MS *m*/*z* calc'd for $C_{76}H_{69}NO_{21}$ $(M+Na)^+$: 1354.43; found: 1354.73.

Methyl a-D-mannopyranosyl-(1→3)-a-D-mannopyranosyl-(1→4)-2-acetamido-2-deoxy-b-D-glucopyranoside (7). The starting material (**29**; 62 mg, 0.047 mmol) and $NH_2NH_2 \cdot H_2O$ (45 μL, 0.92 mmol) in EtOH (1.5 mL) were heated at 70 °C. After 48 hours, $LiOH_{(aq)}$ solution (1 M; to pH 9) was added and the reaction mixture heated at 50 °C for 30 minutes, and cooled back to ambient temperature; $NaHCO_3$ (396 mg, 4.71 mmol) and H_2O (1.5 mL) were then added, followed by Ac_2O (0.44 mL, 4.7 mmol). After 30 minutes at ambient temperature, the solution was evaporated to dry and then purified via MPLC on silica gel using 0→25% MeOH – CH_2Cl_2 to afford the product as a white solid, which was directly dissolved into MeOH (1.5 mL). $Pd(OH)_2$ (20% w/w; 12 mg, 0.017 mmol) was added, and the flask subjected to positive-pressure $H_{2(g)}$ using a balloon. After 24 hours, the catalyst was removed via filtration, the solution evaporated to dry, and then the crude mixture purified via reverse-phase column chromatography using 0→40% H_2O – MeCN to afford the pure product as a white solid (26 mg, 0.046 mmol, 99% yield). $[\alpha]_D^{20}$: -613° (*c* 0.35, H_2O). ^{1}H NMR (D_2O, 500 MHz): δH 5.10 (d, 1H, *J* = 1.3 Hz, Man_H1'), 4.77 (d, 1H, *J* = <1 Hz, Man_H1), 4.45 (d, 1H, *J* = 8.0 Hz, GlcNAc_H1), 4.22 (dd, 1H, *J* = 3.0, <1 Hz, Man_H2), 4.06 (dd, 1H, *J* = 3.3, 1.6 Hz, Man_H2'), 3.93 – 3.86 (m, 4H, Man_H6^a, Man_H6$^{a'}$, GlcNAc_H6^a, and Man_H3'), 3.79 (ddd, 1H, *J* = 10.0, 6.9, 1.7 Hz, Man_H5'), 3.81 – 3.68 (m, 7H, Man_H3, Man_H6$^{b'}$, GlcNAc_H6^b, Man_H6^b, GlcNAc_H2, GlcNAc_H3, and GlcNAc_H4), 3.68 (dd, 1H, *J* = 9.8, 9.8 Hz, Man_H4), 3.62 (dd, 1H, *J* = 9.8, 9.8 Hz, Man_H4'), 3.56 – 3.53 (m, 1H, GlcNAc_H5), 3.49 (s, 3H, OCH_3), 3.45 (ddd, 1H, *J* = 9.7, 6.4, 2.1 Hz, Man_H5), 2.02 (s, 3H, Ac). ^{13}C NMR (D_2O, 125 MHz): δC 175.32 (C=O), 103.12 ($^1J_{C,H}$ = 172 Hz,

Man_C1'), 102.56 ($^1J_{C,H}$ = 162 Hz, GlcNAc_C1), 100.58 ($^1J_{C,H}$ = 161 Hz, Man_C1), 81.12 (Man_C3), 79.70 (GlcNAc_C4), 76.84 (Man_C5), 75.20 (GlcNAc_C5), 74.07 (Man_C5'), 73.05 (GlcNAc_C3), 70.98 (Man_C3'), 70.92 (Man_C2), 70.64 (Man_C2'), 67.52 (Man_C4'), 66.51 (Man_C4), 61.78 (Man_C6'), 61.47 (Man_C6), 60.80 (GlcNAc_C6), 57.75 (OCH_3), 55.50 (GlcNAc_C2), 22.80 (NHAc). ESI-HRMS *m/z* calc'd for $C_{21}H_{37}NO_{16}$ $(M+Na)^+$: 582.2010; found: 582.2009. HPLC purity analysis: >99.5%, R_t 3.59 minutes, Eclipse XDB-C8 column.



(a) *p*-TolSH, BF_3•Et_2O, CH_2Cl_2, 86%; (b) NaOMe, MeOH/CH_2Cl_2, 97%;
(c) $PhCH(OMe)_2$, CSA, CH_3CN/DMF, 81%; (d) BnBr, NaH, DMF, 98%.

***p*-Methylphenyl 3,4,6-tri-*O*-acetyl-2-deoxy-2-phthalimido-1-thio-b-D-glucopyranoside (30).** The starting material (**15**; 4.999 g, 10.47 mmol), *p*-methylthiophenol (2.092 g, 16.84 mmol), and BF_3•Et_2O (4.0 mL, 32 mmol) in anhydrous CH_2Cl_2 (50 mL) were left mixing under Ar at ambient temperature. After 24 hours, the reaction mixture was neutralized with the addition of Et_3N (to pH 8), diluted further with CH_2Cl_2 (200 mL), washed with saturated $NaCl_{(aq)}$ solution (2 x 250 mL), dried with Na_2SO_4, filtered, and evaporated to dry. The crude material was then purified via MPLC on silica gel using 0→30% EtOAc – toluene to afford the pure product as a white solid (4.885 g, 9.020 mmol, 86% yield). R_f = 0.64 (1:4 acetone : toluene). $[\alpha]_D^{20}$: +37° (*c* 1.0, $CHCl_3$). ^1H NMR ($CDCl_3$, 500 MHz): δH 7.88 – 7.85 (m, 2H, Ar), 7.77 – 7.75 (m, 2H, Ar), 7.32 – 7.29 (m, 2H, Ar), 7.09 – 7.07 (m, 2H, Ar), 5.78 (dd, 1H, *J* = 9.7, 9.7 Hz, H-3), 5.65 (d, 1H, *J* = 10.5 Hz, H-1), 5.12 (dd, 1H, *J* = 9.6, 9.6 Hz, H-4), 4.34 – 4.30 (m, 1H, H-2), 4.28 (dd, 1H, *J* = 12.2, 4.8 Hz, H-6^a), 4.21 (dd, 1H, *J* = 12.1, <2 Hz, H-6^b), 3.90 – 3.86 (m, 1H, H-5), 2.33 (s, 3H, Ph*C*H_3), 2.11 (s, 3H, Ac), 2.02 (s, 3H, Ac), 1.83 (s, 3H, Ac). ^{13}C NMR ($CDCl_3$, 125 MHz): δC 170.78 (Ac), 170.26 (Ac), 169.61 (Ac), 167.99 (Phth), 167.13 (Phth), 138.93 (Ar), 134.61 (Ar), 134.48 (Ar), 134.11 (Ar), 131.79 (Ar), 131.37 (Ar), 129.82

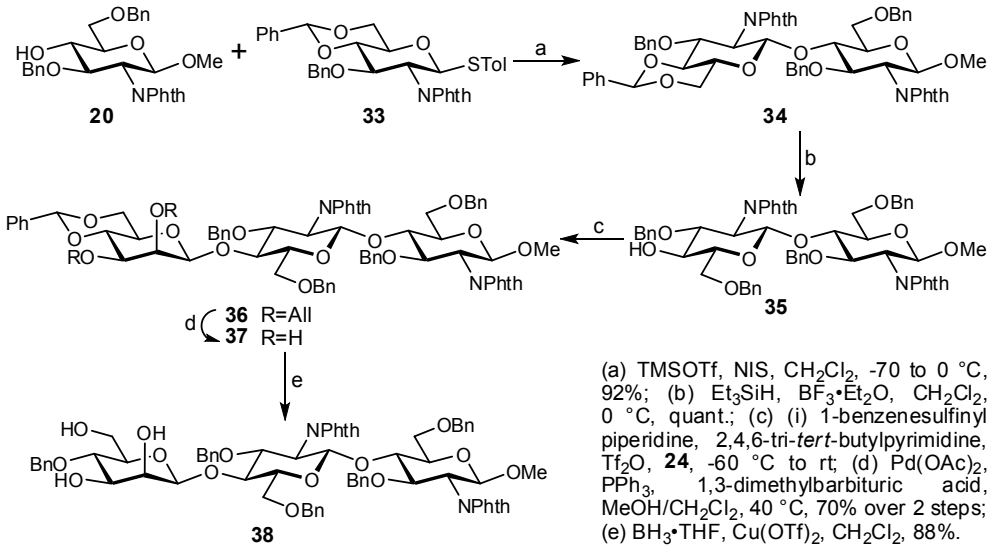
(Ar), 127.14 (Ar), 123.86 (Ar), 83.30 (C-1), 76.03 (C-5), 71.85 (C-3), 68.90 (C-4), 62.38 (C-6), 53.79 (C-2), 21.34 (Ph*C*H_3), 20.93 (Ac), 20.78 (Ac), 20.57 (Ac). ESI-MS *m*/*z* calc'd for $C_{27}H_{27}NO_9S$ $(M+Na)^+$: 564.13; found: 564.15.

***p*-Methylphenyl 2-deoxy-2-phthalimido-1-thio-b-D-glucopyranoside (31).** The starting material (**30**; 4.826 g, 8.911 mmol) was dissolved into anhydrous MeOH–CH_2Cl_2 (1:1; 40 mL) and then NaOMe solution was added drop-wise (1.5 M NaOMe in MeOH; to pH 10). After 1 hour, the reaction flask was neutralized with acidic resin (Amberlite IR-120H; to pH 6), filtered, and evaporated to dry to afford the pure product as a white solid (3.603 g, 8.672 mmol, 97% yield). R_f = 0.16 (1:19 MeOH : CH_2Cl_2). $[\alpha]_D^{20}$: +59° (*c* 0.8, MeOH). ^{1}H NMR (MeOD, 500 MHz): δH 7.92 – 7.81 (m, 4H, Ar), 7.27 – 7.25 (m, 2H, Ar), 7.05 – 7.03 (m, 2H, Ar), 5.51 (d, 1H, *J* = 10.4 Hz, H-1), 4.23 (dd, 1H, *J* = 10.2, 8.2 Hz, H-3), 4.07 (dd, 1H, *J* = 10.3, 10.3 Hz, H-2), 3.93 (dd, 1H, *J* = 12.1, 2.0 Hz, H-6^a), 3.74 (dd, 1H, *J* = 12.1, 5.3 Hz, H-6^b), 3.45 (ddd, 1H, *J* = 9.8, 5.3, 2.1 Hz, H-5), 3.42 (dd, 1H, *J* = 9.8, 8.1 Hz, H-4), 2.26 (s, 3H, PhC*H*$_3$). ^{13}C NMR (MeOD, 125 MHz): δC 169.91 (Phth), 169.44 (Phth), 139.30 (Ar), 135.77 (Ar), 135.72 (Ar), 133.64 (Ar), 133.29 (Ar), 133.09 (Ar), 130.79 (Ar), 130.67 (Ar), 124.59 (Ar), 124.28 (Ar), 85.81 (C-1), 82.81 (C-5), 73.99 (C-3), 72.40 (C-4), 62.99 (C-6), 58.11 (C-2), 21.22 (Ph*C*H_3). ESI-MS *m*/*z* calc'd for $C_{21}H_{21}NO_6S$ $(M+Na)^+$: 438.10; found: 438.19.

***p*-Methylphenyl 4,6-*O*-benzylidene-2-deoxy-2-phthalimido-1-thio-b-D-glucopyranoside (32).** The starting material (**31**; 3.603 g, 8.672 mmol), $PhCH(OMe)_2$ (2.4 mL, 16 mmol), and camphorsulfonic acid (to pH 3) were added to anhydrous MeCN–DMF (1:1; 40 mL) and left mixing at ambient temperature under Ar. After 16 hours, more $PhCH(OMe)_2$ (2.0 mL, 13 mmol) was added, and after another 8 hours the mixture was neutralized with Et_3N (to pH 8) and evaporated to dry. The crude material was purified via MPLC on silica gel using 0→20% EtOAc – toluene to afford the pure product as a white solid (3.556 g, 7.062 mmol, 81% yield). R_f = 0.70 (1:4 acetone : toluene). $[\alpha]_D^{20}$: +26° (*c* 1.0, $CHCl_3$). ^{1}H NMR ($CDCl_3$, 500 MHz): δH 7.87 – 7.80 (m, 2H, Ar), 7.72 – 7.67 (m, 2H, Ar), 7.48 – 7.43 (m, 2H, Ar), 7.35 – 7.32 (m, 3H, Ar), 7.28 – 7.25 (m, 2H, Ar), 7.06 – 7.04 (m, 2H, Ar), 5.60 (d, 1H, *J* = 10.5 Hz, H-1), 5.53 (s, 1H, PhC*H*), 4.57 (ddd, 1H, *J* = 9.4, 9.4, <2 Hz, H-3), 4.35 (dd, 1H, *J* = 10.4, 4.8 Hz, H-6^a), 4.27 (dd, 1H, *J* = 10.2, 10.2 Hz, H-2), 3.78 (dd, 1H, *J* = 10.2,

10.2 Hz, H-6^{b}), 3.61 (ddd, 1H, *J* = 9.6, 9.6, 4.9 Hz, H-5), 3.53 (dd, 1H, *J* = 9.2, 9.2 Hz, H-4), 2.76 (d, 1H, *J* = <2 Hz, 3-OH), 2.29 (s, 3H, PhC*H*$_3$). ^{13}C NMR ($CDCl_3$, 125 MHz): δC 168.43 (Phth), 167.72 (Phth), 138.57 (Ar), 137.10 (Ar), 134.36 (Ar), 133.45 (Ar), 131.82 (Ar), 131.74 (Ar), 129.88 (Ar), 129.51 (Ar), 128.52 (Ar), 128.05 (Ar), 126.49 (Ar), 124.01 (Ar), 123.50 (Ar), 102.07 (Ph*C*H), 84.64 (C-1), 82.02 (C-4), 70.44 (C-5), 69.86 (C-3), 68.72 (C-6), 55.84 (C-2), 21.30 (Ph*C*H$_3$). ESI-MS *m*/*z* calc'd for $C_{28}H_{25}NO_6S$ $(M+Na)^+$: 526.13; found: 526.29.

***p*-Methylphenyl 3-*O*-benzyl-4,6-*O*-benzylidene-2-deoxy-2-phthalimido-1-thio-b-D-glucopyranoside (33).** The starting material (**32**; 3.556 g, 7.062 mmol) and BnBr (1.5 mL, 13 mmol) were added to anhydrous DMF (30 mL), and then NaH was added portion-wise (60% oil dispersion; 623 mg, 15.6 mmol) under Ar. After 6 hours the reaction was quenched via the slow addition of MeOH (5 mL), evaporated to dry, and then redissolved into EtOAc (250 mL). The organic phase was washed with saturated $NaCl_{(aq)}$ solution (2 x 250 mL), dried with Na_2SO_4, filtered, and evaporated to dry. The crude material was purified via MPLC on silica gel using 0→10% EtOAc – toluene to afford the pure product as a white solid (4.126 g, 6.950 mmol, 98% yield). R_f = 0.62 (1:9 EtOAc : toluene). $[\alpha]_D^{20}$: +52° (*c* 1.0, $CHCl_3$). ^{1}H NMR ($CDCl_3$, 500 MHz): δH 7.86 – 7.84 (m, 1H, Ar), 7.74 – 7.62 (m, 3H, Ar), 7.52 – 7.50 (m, 2H, Ar), 7.40 – 7.35 (m, 3H, Ar), 7.25 – 7.23 (m, 2H, Ar), 7.05 – 7.03 (m, 2H, Ar), 6.98 – 6.96 (m, 2H, Ar), 6.92 – 6.84 (m, 3H, Ar), 5.61 (s, 1H, PhC *H*), 5.55 (d, 1H, *J* = 10.5 Hz, H-1), 4.76 (d, 1H, *J* = 12.3 Hz, PhC*H*^{a}H^{b}), 4.48 (d, 1H, *J* = 12.3 Hz, PhCH^{a}*H*^{b}), 4.43 – 4.39 (m, 2H, H-3 and H-6^{a}), 4.26 (dd, 1H, *J* = 10.2, 10.2 Hz, H-2), 3.83 (dd, 1H, *J* = 10.2, 10.2 Hz, H-6^{b}), 3.77 (dd, 1H, *J* = 9.1, 9.1 Hz, H-4), 3.69 (ddd, 1H, *J* = 9.6, 9.6, 4.8 Hz, H-5), 2.29 (s, 3H, PhC*H*$_3$). ^{13}C NMR ($CDCl_3$, 125 MHz): δC 167.97 (Phth), 167.40 (Phth), 138.60 (Ar), 137.89 (Ar), 137.49 (Ar), 134.16 (Ar), 134.00 (Ar), 133.63 (Ar), 131.79 (Ar), 129.84 (Ar), 129.21 (Ar), 128.47 (Ar), 128.34 (Ar), 128.22 (Ar), 127.84 (Ar), 127.61 (Ar), 126.24 (Ar), 123.71 (Ar), 123.51 (Ar), 101.48 (Ph*C*H), 84.42 (C-1), 83.00 (C-4), 75.68 (C-3), 74.39 (Ph*C*H$_2$), 70.54 (C-5), 68.86 (C-6), 55.00 (C-2), 21.32 (Ph*C*H$_3$). ESI-MS *m*/*z* calc'd for $C_{35}H_{31}NO_6S$ $(M+Na)^+$: 616.18; found: 616.48.



Methyl 3-*O*-benzyl-4,6-*O*-benzylidene-2-deoxy-2-phthalimido-b-D-gluco-pyranosyl-(1→4)-3,6-di-*O*-benzyl-2-deoxy-2-phthalimido-b-D-glucopyranoside (34). The glycosyl acceptor (**20**; 1.595 g, 3.168 mmol), glycosyl donor (**33**; 2.363 g, 3.980 mmol), and crushed molecular sieves (3 Å, 2.027 g) were added to anhydrous CH_2Cl_2 (30 mL) and left mixing at ambient temperature under Ar. After 2 hours, the reaction flask was cooled to -70 °C, NIS (1.502 g, 6.676 mmol) and TfOH (53 μL, 0.60 mmol) were slowly added, and the flask was warmed to 0 °C over 1 hour. After 18 hours, more glycosyl donor (**33**; 1.880 g, 3.167 mmol) and NIS (1.376 g, 6.116 mmol) were added, and after another 4 hours the solution was neutralized with the addition of Et_3N (to pH 8) and warmed to ambient temperature. The crude reaction mixture was filtered over Celite, diluted with CH_2Cl_2 (200 mL), and then washed with saturated $NaHCO_{3(aq)}$ solution (250 mL), saturated $NaCl_{(aq)}$ solution (250 mL), H_2O (250 mL), dried with Na_2SO_4, filtered, and then evaporated to dry. The crude material was purified via MPLC on silica gel using 0→20% EtOAc – toluene to afford the pure product as a white solid (2.827 g, 2.905 mmol, 92% yield). R_f = 0.32 (1:9 EtOAc : toluene). $[\alpha]_D^{20}$: +23° (*c* 1.0, $CHCl_3$). ^{1}H NMR ($CDCl_3$, 500 MHz): δH 7.92 – 7.59 (m, 8H, Ar), 7.50 – 7.48 (m, 2H, Ar), 7.41 – 7.25 (m, 8H, Ar), 7.06 – 7.04 (m, 2H, Ar), 6.99 – 6.97 (m, 2H, Ar), 6.95 – 6.87 (m, 6H, Ar), 5.51 (s, 1H, Ph*C*H), 5.38 (d, 1H, *J* = 8.4 Hz, GlcN_H1'), 4.88 (d, 1H, *J* = 8.4 Hz, GlcN_H1), 4.82 (d, 1H, *J* = 12.4 Hz, PhC*H*^{a}H^b), 4.78 (d, 1H, *J* = 12.3 Hz, PhC*H*^{a}H^b), 4.50 – 4.47 (m, 3H, 2x PhCH^a*H*^{b} and PhC*H*^{a}H^b), 4.45 – 4.41 (m, 2H, GlcN_H3' and PhCH^a*H*^{b}), 4.24 – 4.15 (m, 4H, GlcN_H6^a', GlcN_H2', GlcN_H3, and GlcN_H4), 4.12 – 4.06 (m, 1H, GlcN_H2),

3.71 (dd, 1H, *J* = 9.2, 9.2 Hz, GlcN_H4'), 3.53 (dd, 1H, *J* = 10.2, 10.2 Hz, GlcN_H6^{b}'), 3.50 (dd, 1H, *J* = 10.9, <1 Hz, GlcN_H6^{a}), 3.39 (ddd, 1H, *J* = 9.7, 9.7, 4.9 Hz, GlcN_H5'), 3.37 (dd, 1H, *J* = 10.9, 3.9 Hz, GlcN_H6^{b}), 3.34 – 3.30 (m, 1H, GlcN_H5), 3.27 (s, 3H, OCH_3). ^{13}C NMR ($CDCl_3$, 125 MHz): δC 168.31 (Phth), 167.68 (Phth), 138.74 (Ar), 138.47 (Ar), 138.13 (Ar), 137.59 (Ar), 134.19 (Ar), 133.85 (Ar), 131.92 (Ar), 129.17 (Ar), 128.45 (Ar), 128.23 (Ar), 128.17 (Ar), 128.15 (Ar), 127.94 (Ar), 127.70 (Ar), 127.58 (Ar), 127.50 (Ar), 127.24 (Ar), 126.28 (Ar), 123.84 (Ar), 123.44 (Ar), 101.41 (Ph*C*H), 99.18 (GlcN_C1), 97.94 (GlcN_C1'), 83.37 (GlcN_C4'), 77.13 and 76.55 (GlcN_C3 and GlcN_C4), 74.73 (GlcN_C5 and GlcN_C3'), 74.56 (Ph*C*H_2), 74.30 (Ph*C*H_2), 72.90 (Ph*C*H_2), 68.93 (GlcN_C6'), 68.18 (GlcN_C6), 65.94 (GlcN_C5'), 56.76 (GlcN_C2), 56.58 (OCH_3), 55.70 (GlcN_C2'). ESI-MS *m*/*z* calc'd for $C_{57}H_{52}N_2O_{13}$ $(M+Na)^+$: 995.37; found: 995.64.

Methyl 3,6-di-*O*-benzyl-2-deoxy-2-phthalimido-b-D-glucopyranosyl-(1→4)-3,6-di-*O*-benzyl-2-deoxy-2-phthalimido-b-D-glucopyranoside (35). A solution of the starting material (**34**; 2.660 g, 2.734 mmol) in anhydrous CH_2Cl_2 (25 mL) was cooled to 0 °C under Ar, and then Et_3SiH (2.2 mL, 14 mmol) and BF_3•Et_2O (0.70 mL, 5.6 mmol) were slowly added. After 3.5 hours, more Et_3SiH (1.0 mL, 6.3 mmol) was added and then after another 2 hours the reaction mixture was neutralized with Et_3N (to pH 8), warmed back to ambient temperature, and then quenched via the slow addition of MeOH (5 mL). The crude mixture was evaporated to dry, and then purified via MPLC on silica gel using 0→20% EtOAc – toluene to afford the pure product as a white solid (2.634 g, 2.701 mmol, quantitative yield). R_f = 0.29 (1:4 EtOAc : toluene). $[\alpha]_D^{20}$: +13° (*c* 1.0, $CHCl_3$). ^1H NMR ($CDCl_3$, 500 MHz): δH 7.89 – 7.56 (m, 8H, Ar), 7.37 – 7.24 (m, 10H, Ar), 7.04 – 6.93 (m, 7H, Ar), 6.86 – 6.81 (m, 3H, Ar), 5.30 (d, 1H, *J* = 8.3 Hz, GlcN_H1'), 4.87 (d, 1H, *J* = 8.0 Hz, GlcN_H1), 4.79 (d, 1H, *J* = 12.5 Hz, PhC*H*^{a}H^{b}), 4.79 (d, 1H, *J* = 12.2 Hz, PhC*H*^{a}H^{b}), 4.53 (d, 1H, *J* = 11.9 Hz, PhC*H*^{a}H^{b}), 4.52 (d, 1H, *J* = 12.2 Hz, PhCH^{a}*H*^{b}), 4.50 – 4.45 (m, 4H, PhC*H*^{a}H^{b} and 3x PhCH^{a}*H*^{b}), 4.25 (dd, 1H, *J* = 10.7, 8.3 Hz, GlcN_H3'), 4.19 – 4.07 (m, 4H, GlcN_H4, GlcN_H2', GlcN_H3, and GlcN_H2), 3.81 (ddd, 1H, *J* = 9.6, 8.2, 2.0 Hz, GlcN_H4'), 3.70 (dd, 1H, *J* = 9.9, 4.4 Hz, GlcN_H6^{a}'), 3.55 (dd, 1H, *J* = 11.1, <1 Hz, GlcN_H6^{a}), 3.53 (dd, 1H, *J* = 9.8, 6.3 Hz, GlcN_H6^{b}'), 3.43 (dd, 1H, *J* = 11.0, 3.7 Hz, GlcN_H6^{b}), 3.38 (ddd, 1H, *J* = 9.4, 6.0, 4.6 Hz, GlcN_H5'), 3.31 (ddd, 1H, *J* = 9.5, 3.6, <1 Hz, GlcN_H5), 3.27 (s, 3H, OCH_3), 3.10 (d, 1H, *J* = 2.1 Hz,

GlcN_4-OH'). ^{13}C NMR ($CDCl_3$, 125 MHz): δC 168.64 (Phth), 167.85 (Phth), 138.83 (Ar), 138.60 (Ar), 138.53 (Ar), 137.72 (Ar), 134.20 (Ar), 134.06 (Ar), 133.81 (Ar), 131.96 (Ar), 131.71 (Ar), 128.74 (Ar), 128.43 (Ar), 128.29 (Ar), 128.08 (Ar), 128.06 (Ar), 128.04 (Ar), 127.97 (Ar), 127.93 (Ar), 127.58 (Ar), 127.52 (Ar), 127.10 (Ar), 123.84 (Ar), 123.37 (Ar), 99.21 (GlcN_C1), 97.18 (GlcN_C1'), 78.52 (GlcN_C3'), 76.95 (GlcN_C3), 75.96 (GlcN_C4), 75.67 (GlcN_C4'), 74.77 (GlcN_C5), 74.53 (Ph*C*H_2), 74.42 (Ph*C*H_2), 73.91 (Ph*C*H_2), 72.99 (GlcN_C5'), 72.88 (Ph*C*H_2), 71.19 (GlcN_C6'), 68.33 (GlcN_C6), 56.60 (OCH_3), 56.31 (GlcN_C2'), 55.74 (GlcN_C2). ESI-MS *m*/*z* calc'd for $C_{57}H_{54}N_2O_{13}$ $(M+Na)^+$: 997.35; found: 997.69.

Methyl 2,3-di-*O*-allyl-4,6-*O*-benzylidene-b-D-mannopyranosyl-(1→4)-3,6-di-*O*-benzyl-2-deoxy-2-phthalimido-b-D-glucopyranosyl-(1→4)-3,6-di-*O*-benzyl-2-deoxy-2-phthalimido-b-D-glucopyranoside (36). The glycosyl donor (**24**; 183 mg, 0.403 mmol), 1-benzenesulfinyl piperidine (85 mg, 0.41 mmol), 2,4,6-tri-*tert*-butylpyrimidine (201 mg, 0.809 mmol), and crushed molecular sieves (3 Å, 453 mg) in anhydrous CH_2Cl_2 (10 mL) were left mixing at ambient temperature under Ar [ref3]. After 2 hours, the flask was cooled to –60 °C and then Tf_2O (75 μL, 0.45 mmol) was added. After 5 minutes, a solution of the glycosyl acceptor (**35**; 585 mg, 0.600 mmol) in anhydrous CH_2Cl_2 (4 mL) was added, and after another 2 minutes the reaction mixture was warmed to ambient temperature over 5 minutes. The reaction solution was diluted with CH_2Cl_2 (50 mL), and then washed with saturated $NaHCO_{3(aq)}$ solution (60 mL), saturated $NaCl_{(aq)}$ solution (60 mL), dried with Na_2SO_4, filtered, and evaporated to dry. The crude material was purified via MPLC on silica gel using 0→30% EtOAc – toluene to afford some pure material, as well as compound that had co-eluted with the undesired a-configured by-product; all fractions were combined and used directly for the subsequent step. R_f = 0.51 (1:4 EtOAc : toluene). $[\alpha]_D^{20}$: +17° (*c* 1.0, $CHCl_3$). ^{1}H NMR ($CDCl_3$, 500 MHz): δH 7.88 – 7.53 (m, 8H, Ar), 7.44 – 7.41 (m, 2H, Ar), 7.35 – 7.24 (m, 13H, Ar), 7.02 – 7.00 (m, 2H, Ar), 6.98 – 6.95 (m, 2H, Ar), 6.93 – 6.86 (m, 3H, Ar), 6.80 – 6.74 (m, 3H, Ar), 5.98 (dddd, 1H, *J* = 17.3, 10.4, 5.7, 5.7 Hz, 2-OCH_2C*H*=CH_2), 5.90 (dddd, 1H, *J* = 17.2, 10.5, 5.3, 5.3 Hz, 3-OCH_2C*H*=CH_2), 5.45 (s, 1H, PhC*H*), 5.32 (dddd, 1H, *J* = 17.3, 1.6, 1.6, 1.6 Hz, 2-OCH_2CH=C*H*^{a}H^b), 5.30 (dddd, 1H, *J* = 17.3, 1.7, 1.7, 1.7 Hz, 3-OCH_2CH=C*H*^{a}H^b), 5.29 (d, 1H, *J* = 8.4 Hz, GlcN_H1'), 5.18 – 5.15 (m, 2H, 2-OCH_2CH=CH^a*H*^{b} and 3-OCH_2CH=CH^a*H*^{b}), 4.92 (d, 1H, *J* = 12.2 Hz, PhC*H*^{a}H^b), 4.87

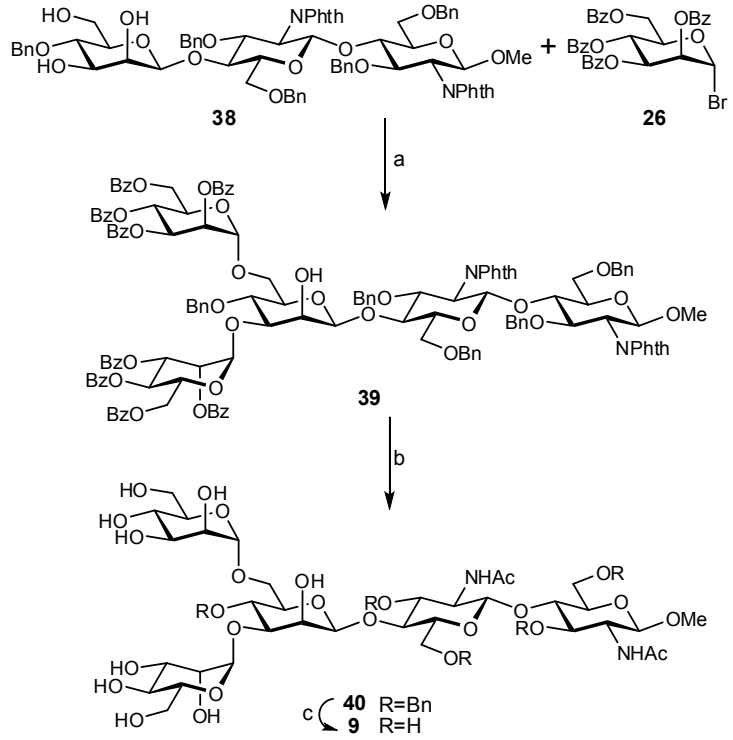
(d, 1H, *J* = 8.2 Hz, GlcN_H1), 4.85 (d, 1H, *J* = 12.8 Hz, PhC*H*^{a}H^{b}), 4.60 (d, 1H, *J* = <1 Hz, Man_H1), 4.59 (d, 1H, *J* = 12.1 Hz, PhC*H*^{a}H^{b}), 4.55 – 4.50 (m, 3H, PhC*H*^{a}H^{b} and 2x PhCH^{a}*H*^{b}), 4.48 – 4.44 (m, 2H, 2x PhCH^{a}*H*^{b}), 4.34 (dddd, 1H, *J* = 12.8, 5.6, 1.4, 1.4 Hz, 2-OC*H*^{a}H^{b}CH=CH$_2$), 4.32 – 4.19 (m, 5H, 2-OCH^{a}*H*^{b}CH=CH$_2$, GlcN_H3', 3-OC*H*^{a}H^{b}CH=CH$_2$, GlcN_H4, and GlcN_H2'), 4.15 – 4.05 (m, 5H, GlcN_H3, Man_H6^{a}, GlcN_H2, GlcN_H4', and 3-OCH^{a}*H*^{b}CH=CH$_2$), 3.94 (dd, 1H, *J* = 9.6, 9.6 Hz, Man_H4), 3.70 (dd, 1H, *J* = 3.1, <1 Hz, Man_H2), 3.67 (dd, 1H, *J* = 11.3, 1.8 Hz, GlcN_H6$^{a'}$), 3.59 – 3.56 (m, 2H, GlcN_H6^{a} and GlcN_H6$^{b'}$), 3.50 (dd, 1H, *J* = 10.3, 10.3 Hz, Man_H6^{b}), 3.44 (dd, 1H, *J* = 11.1, 3.8 Hz, GlcN_H6^{b}), 3.34 (dd, 1H, *J* = 9.9, 3.1 Hz, Man_H3), 3.32 (ddd, 1H, *J* = 9.8, 3.7, 1.3 Hz, GlcN_H5), 3.28 – 3.24 (m, 4H, OCH$_3$ and GlcN_H5'), 3.11 (ddd, 1H, *J* = 9.7, 9.7, 4.8 Hz, Man_H5). ^{13}C NMR ($CDCl_3$, 125 MHz): δC 168.73 (Phth), 167.80 (Phth), 139.13 (Ar), 138.84 (Ar), 138.71 (Ar), 138.15 (Ar), 137.85 (Ar), 135.77 (2-OCH$_2$*C*H=CH$_2$), 135.16 (3-OCH$_2$*C*H=CH$_2$), 134.23 (Ar), 134.01 (Ar), 133.76 (Ar), 132.02 (Ar), 131.96 (Ar), 131.70 (Ar), 128.98 (Ar), 128.75 (Ar), 128.40 (Ar), 128.34 (Ar), 128.22 (Ar), 128.05 (Ar), 128.02 (Ar), 128.01 (Ar), 127.94 (Ar), 127.89 (Ar), 127.55 (Ar), 127.52 (Ar), 127.17 (Ar), 127.06 (Ar), 126.25 (Ar), 123.87 (Ar), 123.33 (Ar), 117.05 (2-OCH$_2$CH=*C*H$_2$), 116.69 (3-OCH$_2$CH=*C*H$_2$), 101.93 ($^{1}J_{C,H}$ = 158 Hz, Man_C1), 101.52 (Ph*C*H), 99.28 ($^{1}J_{C,H}$ = 163 Hz, GlcN_C1), 97.23 ($^{1}J_{C,H}$ = 167 Hz, GlcN_C1'), 79.25 (GlcN_C4'), 78.86 (Man_C4), 78.29 (Man_C3), 77.15 (GlcN_C3' and Man_C2), 77.00 (GlcN_C3), 75.96 (GlcN_C4), 74.90 (GlcN_C5'), 74.86 (Ph*C*H$_2$), 74.76 (GlcN_C5), 74.60 and 74.57 (Ph*C*H$_2$ and 2-O*C*H$_2$CH=CH$_2$), 73.58 (Ph*C*H$_2$), 72.86 (Ph*C*H$_2$), 71.91 (3-O*C*H$_2$CH=CH$_2$), 68.76 (Man_C6), 68.41 (GlcN_C6), 68.26 (GlcN_C6'), 67.44 (Man_C5), 56.81 (GlcN_C2'), 56.65 (OCH$_3$), 55.81 (GlcN_C2). ESI-MS *m*/*z* calc'd for $C_{76}H_{76}N_2O_{18}$ $(M+Na)^+$: 1327.50; found: 1327.96.

Methyl 4,6-*O*-benzylidene-b-D-mannopyranosyl-(1→4)-3,6-di-*O*-benzyl-2-deoxy-2-phthalimido-b-D-glucopyranosyl-(1→4)-3,6-di-*O*-benzyl-2-deoxy-2-phthalimido-b-D-glucopyranoside (37). A solution of $Pd(OAc)_2$ (33 mg, 0.15 mmol) and PPh_3 (146 mg, 0.557 mmol) in anhydrous MeOH/CH_2Cl_2 (2:1, 9 mL) were left mixing at ambient temperature under Ar. After 20 minutes, the solution was transferred to a flask containing an a-/b-mixture of the substrate (**36**; 945 mg, 0.724 mmol) and 1,3-dimethylbarbituric acid (452 mg, 2.89 mmol) and left mixing at 40 °C [ref4]. After 30 hours, the solution was diluted with CH_2Cl_2 (300 mL), washed with

saturated $Na_2CO_{3(aq)}$ solution (300 mL), and the aqueous layer re-extracted with CH_2Cl_2 (300 mL). The combined organic phases were then washed with saturated $NaCl_{(aq)}$ solution (2 x 300 mL), dried with Na_2SO_4, filtered, and evaporated to dry. The crude material was purified via MPLC on silica gel using 0→50% acetone – toluene to afford the pure product as a white solid (619 mg, 0.505 mmol, 70% yield over 2 steps) together with the a-anomeric byproduct (226 mg, 0.184 mmol, 25% yield over 2 steps). R_f = 0.64 (2:3 acetone : toluene). $[\alpha]_D^{20}$: +18° (*c* 1.0, $CHCl_3$). ^{1}H NMR ($CDCl_3$, 500 MHz): δH 7.87 – 7.54 (m, 8H, Ar), 7.46 – 7.44 (m, 2H, Ar), 7.37 – 7.24 (m, 13H, Ar), 7.03 – 7.01 (m, 2H, Ar), 6.98 – 6.89 (m, 5H, Ar), 6.80 – 6.76 (m, 3H, Ar), 5.45 (s, 1H, PhC*H*), 5.28 (d, 1H, *J* = 8.4 Hz, GlcN_H1'), 4.87 (d, 1H, *J* = 8.3 Hz, GlcN_H1), 4.86 (d, 1H, *J* = 12.6 Hz, PhC*H*^{a}H^b), 4.82 (d, 1H, *J* = 12.1 Hz, PhC*H*^{a}H^b), 4.74 (d, 1H, *J* = <1 Hz, Man_H1), 4.59 (d, 1H, *J* = 12.0 Hz, PhC*H*^{a}H^b), 4.54 (d, 1H, *J* = 11.9 Hz, PhC*H*^{a}H^b), 4.53 – 4.50 (m, 2H, 2x PhCH^a*H*^{b}), 4.46 (d, 1H, *J* = 12.0 Hz, PhCH^a*H*^{b}), 4.42 (d, 1H, *J* = 12.1 Hz, PhCH^a*H*^{b}), 4.39 (dd, 1H, *J* = 10.7, 8.6 Hz, GlcN_H3'), 4.28 – 4.23 (m, 1H, GlcN_H4), 4.23 (dd, 1H, *J* = 10.7, 8.4 Hz, GlcN_H2'), 4.15 – 4.11 (m, 4H, GlcN_H4', Man_H6^a, GlcN_H2, and GlcN_H3), 3.94 – 3.92 (m, 1H, Man_H2), 3.75 (dd, 1H, *J* = 9.5, 9.5 Hz, Man_H4), 3.65 (dd, 1H, *J* = 11.4, 1.7 Hz, GlcN_H6$^{a'}$), 3.62 – 3.56 (m, 3H, GlcN_H6^a, Man_H3, and GlcN_H6$^{b'}$), 3.53 (dd, 1H, *J* = 10.3, 10.3 Hz, Man_H6^b), 3.44 (dd, 1H, *J* = 11.2, 3.6 Hz, GlcN_H6^b), 3.32 – 3.28 (m, 5H, GlcN_H5, GlcN_H5', and OCH_3), 3.12 (ddd, 1H, *J* = 9.8, 9.8, 5.0 Hz, Man_H5), 2.69 (d, 1H, *J* = 2.0 Hz, Man_2-OH), 2.54 (d, 1H, *J* = 6.8 Hz, Man_3-OH). ^{13}C NMR ($CDCl_3$, 125 MHz): δC 168.72 (Phth), 167.74 (Phth), 138.80 (Ar), 138.62 (Ar), 138.58 (Ar), 137.79 (Ar), 137.36 (Ar), 134.26 (Ar), 134.10 (Ar), 133.78 (Ar), 131.98 (Ar), 131.64 (Ar), 129.37 (Ar), 128.81 (Ar), 128.48 (Ar), 128.45 (Ar), 128.22 (Ar), 128.21 (Ar), 128.19 (Ar), 128.06 (Ar), 127.74 (Ar), 127.60 (Ar), 127.49 (Ar), 127.45 (Ar), 127.12 (Ar), 126.47 (Ar), 123.92 (Ar), 123.40 (Ar), 102.26 (Ph*C*H), 100.62 ($^1J_{C,H}$ = 163 Hz, Man_C1), 99.28 ($^1J_{C,H}$ = 166 Hz, GlcN_C1), 97.15 ($^1J_{C,H}$ = 166 Hz, GlcN_C1'), 78.93 (GlcN_C4'), 78.70 (Man_C4), 77.89 (GlcN_C3'), 76.91 (GlcN_C3), 75.81 (GlcN_C4), 74.94 (Ph*C*H$_2$), 74.72 (GlcN_C5'), 74.60 (GlcN_C5), 74.58 (Ph*C*H$_2$), 73.71 (Ph*C*H$_2$), 72.86 (Ph*C*H$_2$), 70.96, 70.93 (Man_C2 and Man_C3), 68.63 (Man_C6), 68.34 (GlcN_C6), 68.02 (GlcN_C6'), 66.78 (Man_C5), 56.77 (GlcN_C2'), 56.63 (OCH_3), 55.79 (GlcN_C2). ESI-MS *m*/*z* calc'd for $C_{70}H_{68}N_2O_{18}$ $(M+Na)^+$: 1247.44; found: 1247.65.

Methyl 4-*O*-benzyl-b-D-mannopyranosyl-(1→4)-3,6-di-*O*-benzyl-2-deoxy-2-phthalimido-b-D-glucopyranosyl-(1→4)-3,6-di-*O*-benzyl-2-deoxy-2-phthalimido-b-D-glucopyranoside (38). The starting material (**37**; 684 mg, 0.558 mmol) and BH_3•THF (1 M solution; 2.8 mL, 2.8 mmol) in anhydrous CH_2Cl_2 (6.0 mL) were left mixing for 5 minutes at ambient temperature under Ar, and then $Cu(OTf)_2$ (40 mg, 0.11 mmol) which had immediately prior been co-evaporated with toluene (3x) was added to the solution [ref5]. After 6 hours, the reaction mixture was neutralized with Et_3N (to pH 8) and then quenched via the slow addition of MeOH (2 mL). The crude mixture was evaporated to dry and then purified via MPLC on silica gel using 0→40% EtOAc – toluene to afford the pure product as a white solid (603 mg, 0.491 mmol, 88% yield). R_f = 0.55 (2:3 acetone : toluene). $[\alpha]_D^{20}$: +23° (*c* 1.0, $CHCl_3$). ^{1}H NMR ($CDCl_3$, 500 MHz): δH 7.87 – 7.53 (m, 8H, Ar), 7.36 – 7.23 (m, 15H, Ar), 7.04 – 7.02 (m, 2H, Ar), 6.98 – 6.90 (m, 5H, Ar), 6.80 – 6.75 (m, 3H, Ar), 5.28 (d, 1H, *J* = 8.4 Hz, GlcN_H1'), 4.89 – 4.84 (m, 3H, GlcN_H1 and 2x PhC*H*^{a}H^b), 4.78 (d, 1H, *J* = 11.4 Hz, PhC*H*^{a}H^b), 4.67 (d, 1H, *J* = <1 Hz, Man_H1), 4.61 (d, 1H, *J* = 11.4 Hz, PhCH^a*H*^{b}), 4.58 – 4.50 (m, 3H, 2x PhC*H*^{a}H^b and PhCH^a*H*^{b}), 4.50 (d, 1H, *J* = 11.8 Hz, PhCH^a*H*^{b}), 4.48 (d, 1H, *J* = 12.1 Hz, PhCH^a*H*^{b}), 4.44 (d, 1H, *J* = 12.1 Hz, PhCH^a*H*^{b}), 4.40 (dd, 1H, *J* = 10.7, 8.7 Hz, GlcN_H3'), 4.27 – 4.21 (m, 2H, GlcN_H4 and GlcN_H2'), 4.15 – 4.09 (m, 3H, GlcN_H3, GlcN_H4', and GlcN_H2), 3.83 – 3.82 (m, 1H, Man_H2), 3.72 (ddd, 1H, *J* = 11.9, 6.6, 2.5 Hz, Man_H6^a), 3.64 (dd, 1H, *J* = 11.5, 1.9 Hz, GlcN_H6^a'), 3.60 – 3.55 (m, 2H, GlcN_H6^b' and GlcN_H6^a), 3.54 – 3.47 (m, 3H, Man_H4, Man_H6^b, and Man_H3), 3.43 (dd, 1H, *J* = 11.1, 3.7 Hz, GlcN_H6^b), 3.32 – 3.29 (m, 2H, GlcN_H5 and GlcN_H5'), 3.27 (s, 3H, OCH_3), 3.11 (ddd, 1H, *J* = 8.9, 5.0, 2.5 Hz, Man_H5), 2.72 (d, 1H, *J* = 3.2 Hz, ManC2_OH), 2.47 (d, 1H, *J* = 7.4 Hz, ManC3_OH), 2.07 (dd, 1H, *J* = 6.8, 6.8 Hz, ManC6_OH). ^{13}C NMR ($CDCl_3$, 125 MHz): δC 168.72 (Phth), 167.74 (Phth), 138.77 (Ar), 138.58 (Ar), 138.36 (Ar), 138.32 (Ar), 137.78 (Ar), 134.29 (Ar), 134.13 (Ar), 133.75 (Ar), 131.94 (Ar), 131.58 (Ar), 128.78 (Ar), 128.71 (Ar), 128.43 (Ar), 128.37 (Ar), 128.23 (Ar), 128.15 (Ar), 128.11 (Ar), 128.08 (Ar), 128.02 (Ar), 127.59 (Ar), 127.57 (Ar), 127.46 (Ar), 127.44 (Ar), 127.08 (Ar), 123.92 (Ar), 123.38 (Ar), 99.89 ($^1J_{C,H}$ = 159 Hz, Man_C1), 99.25 ($^1J_{C,H}$ = 164 Hz, GlcN_C1), 97.15 ($^1J_{C,H}$ = 167 Hz, GlcN_C1'), 78.15 (GlcN_C4'), 77.98 (GlcN_C3'), 76.93 (GlcN_C3), 75.87 (GlcN_C4), 75.80 (Man_C4), 75.52 (Man_C5), 75.00 (Ph*C*H$_2$), 74.91 (Ph*C*H$_2$), 74.79, 74.71 (GlcN_C5 and GlcN_C5'), 74.65 (Ph*C*H$_2$), 74.36 (Man_C3), 73.67 (Ph*C*H$_2$), 72.84 (Ph*C*H$_2$),

70.98 (Man_C2), 68.31 (GlcN_C6), 67.72 (GlcN_C6'), 62.23 (Man_C6), 56.75 (GlcN_C2'), 56.61 (OCH_3), 55.78 (GlcN_C2). ESI-MS *m/z* calc'd for $C_{70}H_{70}N_2O_{18}$ $(M+Na)^+$: 1249.45; found: 1249.65.



(a) AgOTf, CH_2Cl_2, -60 °C to -50 °C, 24%; (b) (i) $NH_2NH_2 \cdot H_2O$, EtOH, 70 °C; (ii) $NaHCO_3$, Ac_2O, MeOH, 76% over 2 steps; (c) Na, NH_3, KO^tBu, -70 °C, 65%.

Methyl 2,3,4,6-tetra-*O*-benzoyl-a-D-mannopyranosyl-(1→3)-[2,3,4,6-tetra-*O*-benzoyl-a-D-mannopyranosyl-(1→6)]-4-*O*-benzyl-b-D-mannopyranosyl-(1→4)-3,6-di-*O*-benzyl-2-deoxy-2-phthalimido-b-D-glucopyranosyl-(1→4)-3,6-di-*O*-benzyl-2-deoxy-2-phthalimido-b-D-glucopyranoside (39). The glycosyl acceptor (**38**; 586 mg, 0.477 mmol), glycosyl donor (**26**; 1.015 g, 1.539 mmol), and crushed molecular sieves (3 Å, 428 mg) in anhydrous CH_2Cl_2 (6.0 mL) were left mixing for 30 minutes at ambient temperature under Ar. The foil-covered reaction flask was cooled to –60 °C, and then AgOTf (373 mg, 1.45 mmol) was added. After 1.5 hours, additional glycosyl donor (**26**; 186 mg, 0.282 mmol) was added to the reaction mixture, followed by a second additional portion of glycosyl donor (**26**; 169 mg, 0.256 mmol) after another hour had passed. The reaction flask was warmed to –50 °C, and then after 3 hours was neutralized with Et_3N (to pH 8), warmed to ambient temperature, diluted further with CH_2Cl_2 (200 mL), and filtered over Celite. The

organic phase was washed with $NH_4OH_{(aq)}$ solution (0.5 M; 2 x 250 mL), saturated $NaCl_{(aq)}$ solution (250 mL), dried with Na_2SO_4, filtered, and evaporated to dry. The crude material was purified via MPLC on silica gel using 0→50% EtOAc – toluene to afford the pure product as a white solid (269 mg, 0.113 mmol, 24% yield) as a well as a mixture of mono-mannosylated intermediates (543 mg, 0.301 mmol, 63% yield). R_f = 0.41 (1:4 EtOAc : toluene). $[\alpha]_D^{20}$: -11° (*c* 1.0, $CHCl_3$). ^{1}H NMR ($CDCl_3$, 500 MHz): δH 8.15 – 8.13 (m, 2H, Ar), 8.10 – 8.08 (m, 2H, Ar), 8.04 – 8.02 (m, 2H, Ar), 8.00 – 7.98 (m, 2H, Ar), 7.90 – 7.85 (m, 6H, Ar), 7.79 – 7.77 (m, 2H, Ar), 7.74 – 7.14 (m, 47H, Ar), 6.93 – 6.91 (m, 2H, Ar), 6.82 – 6.79 (m, 2H, Ar), 6.75 – 6.69 (m, 3H, Ar), 6.57 – 6.51 (m, 3H, Ar), 6.23 (dd, 1H, *J* = 10.2, 10.2 Hz, Man_H4'^{3}), 6.09 (dd, 1H, *J* = 10.2, 3.1 Hz, Man_H3'^{3}), 6.00 (dd, 1H, *J* = 9.9, 9.9 Hz, Man_H4'^{6}), 5.88 (dd, 1H, *J* = 10.1, 3.3 Hz, Man_H3'^{6}), 5.79 – 5.78 (m, 1H, Man_H2'^{3}), 5.68 (dd, 1H, *J* = 3.2, 1.7 Hz, Man_H2'^{6}), 5.38 (d, 1H, *J* = 1.4 Hz, Man_H1'^{6}), 5.23 (d, 1H, *J* = 8.4 Hz, GlcN_H1'), 5.22 (d, 1H, *J* = 11.6 Hz, PhC*H*^{a}H^b), 5.21 (d, 1H, *J* = <2 Hz, Man_H1'^{3}), 5.05 (ddd, 1H, *J* = 10.2, 3.2, 3.2 Hz, Man_H5'^{3}), 4.90 (d, 1H, *J* = 11.4 Hz, PhCH^a*H*^{b}), 4.84 (d, 1H, *J* = 12.7 Hz, PhC*H*^{a}H^b), 4.83 (d, 1H, *J* = 8.3 Hz, GlcN_H1), 4.80 (d, 1H, *J* = 12.6 Hz, PhC*H*^{a}H^b), 4.75 (dd, 1H, *J* = 12.2, 2.3 Hz, Man_H6^{a,3}), 4.71 (d, 1H, *J* = <1 Hz, Man_H1), 4.56 (d, 1H, *J* = 12.2 Hz, PhC*H*^{a}H^b), 4.56 (dd, 1H, *J* = 12.2, 4.1 Hz, Man_H6^{b,3}), 4.51 – 4.44 (m, 5H, Man_H6^{a,6}, PhC*H*^{a}H^b, and 3x PhCH^a*H*^{b}), 4.39 (d, 1H, *J* = 12.6 Hz, PhCH^a*H*^{b}), 4.35 – 4.29 (m, 2H, Man_H6^{b,6} and Man_H5'^{6}), 4.28 (dd, 1H, *J* = 10.5, 8.6 Hz, GlcN_H3'), 4.23 – 4.12 (m, 5H, Man_H2, Man_H4, GlcN_H4, GlcN_H2', and GlcN_H4'), 4.09 (dd, 1H, *J* = 10.6, 8.0 Hz, GlcN_H2), 4.08 (dd, 1H, *J* = 10.6, 8.0 Hz, GlcN_H3), 4.03 (dd, 1H, *J* = 12.1, 2.8 Hz, Man_H6^a), 3.94 (dd, 1H, *J* = 12.2, <2 Hz, Man_H6^b), 3.67 (dd, 1H, *J* = 11.2, <2 Hz, GlcN_H6^a'), 3.58 – 3.55 (m, 2H, GlcN_H6^b' and Man_H3), 3.49 (dd, 1H, *J* = 10.9, <2 Hz, GlcN_H6^a), 3.37 (dd, 1H, *J* = 11.1, 3.8 Hz, GlcN_H6^b), 3.34 – 3.31 (m, 1H, Man_H5), 3.26 – 3.23 (m, 4H, OCH_3 and GlcN_H5), 3.21 (ddd, 1H, *J* = 9.9, <2, <2 Hz, GlcN_H5'), 2.78 (broad s, 1H, ManC2_OH). ^{13}C NMR ($CDCl_3$, 125 MHz): δC 168.29 (Phth), 167.59 (Phth), 166.40 (Bz), 166.24 (Bz), 165.85 (Bz), 165.64 (Bz), 165.51 (Bz), 165.50 (Bz), 165.45 (Bz), 138.82 (Ar), 138.54 (Ar), 138.51 (Ar), 138.35 (Ar), 138.03 (Ar), 133.86 (Ar), 133.66 (Ar), 133.50 (Ar), 133.47 (Ar), 133.40 (Ar), 133.37 (Ar), 133.24 (Ar), 133.16 (Ar), 133.04 (Ar), 131.97 (Ar), 131.66 (Ar), 130.21 (Ar), 130.17 (Ar), 130.10 (Ar), 130.06 (Ar), 130.00 (Ar), 129.95 (Ar), 129.94 (Ar), 129.90 (Ar), 129.61 (Ar), 129.46 (Ar), 129.39 (Ar), 129.37 (Ar), 129.29 (Ar), 129.22

(Ar), 128.82 (Ar), 128.78 (Ar), 128.74 (Ar), 128.69 (Ar), 128.59 (Ar), 128.56 (Ar), 128.45 (Ar), 128.36 (Ar), 128.22 (Ar), 128.19 (Ar), 128.06 (Ar), 127.96 (Ar), 127.90 (Ar), 127.75 (Ar), 127.61 (Ar), 127.45 (Ar), 126.98 (Ar), 123.64 (Ar), 123.28 (Ar), 100.99 ($^1J_{C,H}$ = 160 Hz, Man_C1), 99.83 ($^1J_{C,H}$ = 171 Hz, Man_C1'^{3}), 99.16 ($^1J_{C,H}$ = 165 Hz, GlcN_C1), 97.92 ($^1J_{C,H}$ = 173 Hz, Man_C1'^{6}), 97.11 ($^1J_{C,H}$ = 169 Hz, GlcN_C1'), 85.95 (Man_C3), 79.85 (GlcN_C4'), 77.17 (GlcN_C3'), 77.03 (GlcN_C3), 75.93 (GlcN_C4), 75.90 (Ph*C*H$_2$), 75.66 (Man_C5), 74.78 (Ph*C*H$_2$), 74.76 (GlcN_C5), 74.59 (Ph*C*H$_2$), 74.40 (GlcN_C5'), 73.34 (Ph*C*H$_2$), 73.07 (Man_C4), 72.88 (Ph*C*H$_2$), 71.26 (Man_C2), 70.96 (Man_C2'^{3}), 70.88 (Man_C2'^{6}), 70.50 (Man_C3'^{3}), 70.00 (Man_C3'^{6}), 69.76 (Man_C5'^{3}), 69.08 (Man_C5'^{6}), 68.20 (GlcN_C6), 68.07 (GlcN_C6'), 67.06 (Man_C4'^{6}), 66.94 (Man_C4'^{3}), 66.33 (Man_C6), 63.24 (Man_C6'^{3}), 62.97 (Man_C6'^{6}), 56.78 (GlcN_C2'), 56.51 (OCH_3), 55.73 (GlcN_C2). ESI-MS *m*/*z* calc'd for $C_{138}H_{122}N_2O_{36}$ $(M+Na)^+$: 2406.77; found: 2406.80.

Methyl a-D-mannopyranosyl-(1→3)-[a-D-mannopyranosyl-(1→6)]-4-*O*-benzyl-b-D-mannopyranosyl-(1→4)-2-acetamido-3,6-di-*O*-benzyl-2-deoxy-b-D-glucopyranosyl-(1→4)-2-acetamido-3,6-di-*O*-benzyl-2-deoxy-b-D-glucopyranoside (40). The starting material (**39**; 170 mg, 0.0713 mmol) and $NH_2NH_2 \cdot H_2O$ (87 μL, 1.8 mmol) in EtOH (1.5 mL) were heated at 70 °C for 24 hours. Additional $NH_2NH_2 \cdot H_2O$ (17 μL, 0.35 mmol) was added and the solution left for another 16 hours, and then the reaction mixture was evaporated to dry. The crude material, $NaHCO_3$ (237 mg, 2.82 mmol), and Ac_2O (145 μL, 1.53 mmol) in dry MeOH (1.0 mL) were left mixing at ambient temperature for 1.5 hours. The reaction mixture was evaporated to dry and purified via MPLC on silica gel using 0→35% MeOH – CH_2Cl_2 to afford the pure product as a white solid (74 mg, 0.054 mmol, 76% yield over 2 steps). R_f = 0.59 (1:10:39 H_2O : MeOH : CH_2Cl_2). $[\alpha]_D^{20}$: +11° (*c* 1.0, MeOH). ^{1}H NMR (MeOD, 500 MHz): δH 7.38 – 7.16 (m, 25H, Ar), 5.02 – 4.99 (m, 3H, 2x PhC*H*^{a}H^b and Man_H1'^{3}), 4.80 (d, 1H, *J* = 1.6 Hz, Man_H1'^{6}), 4.79 (d, 1H, *J* = 11.1 Hz, PhC*H*^{a}H^b), 4.75 (d, 1H, *J* = 12.3 Hz, PhCH^a*H*^{b}), 4.64 – 4.55 (m, 6H, GlcNAc_H1', PhC*H*^{a}H^b, 3x PhCH^a*H*^{b}, and Man_H1), 4.44 (d, 1H, *J* = 12.0 Hz, PhC*H*^{a}H^b), 4.41 (d, 1H, *J* = 12.0 Hz, PhCH^a*H*^{b}), 4.30 (d, 1H, *J* = 8.1 Hz, GlcNAc_H1), 4.12 (dd, 1H, *J* = 3.1, <1 Hz, Man_H2), 4.00 (dd, 1H, *J* = 9.2, 9.2 Hz, GlcNAc_H4'), 3.98 (dd, 1H, *J* = 8.8, 8.8 Hz, GlcNAc_H4), 3.92 – 3.86 (m, 3H,

Man_H3'^{3}, Man_H6^{a}'^{3}, and Man_H2'^{3}), 3.85 (dd, 1H, *J* = 9.5, 8.3 Hz, GlcNAc_H2), 3.83 – 3.69 (m, 9H, Man_H2'^{6}, GlcNAc_H6^{a}, GlcNAc_H6^{b}, Man_H4, GlcNAc_H2', Man_H6^{a}, Man_H4'^{6}, Man_H4'^{3}, and Man_H6^{b}'^{3}), 3.68 – 3.54 (m, 10H, GlcNAc_H6^{a}', GlcNAc_H3', Man_H5'^{3}, Man_H5'^{6}, Man_H6^{a}'^{6}, Man_H6^{b}'^{6}, Man_H6^{b}, GlcNAc_H3, GlcNAc_H6^{b}', and Man_H3), 3.52 – 3.46 (m, 2H, Man_H3'^{6} and GlcNAc_H5), 3.43 (s, 3H, OCH_3), 3.25 (ddd, 1H, *J* = 9.5, 3.6, 2.1 Hz, GlcNAc_H5'), 3.19 (ddd, 1H, *J* = 9.8, 5.2, 1.8 Hz, Man_H5), 1.83 (s, 3H, Ac), 1.82 (s, 3H, Ac). ^{13}C NMR (MeOD, 125 MHz): δC 173.52 (Ac), 173.40 (Ac), 140.79 (Ar), 140.67 (Ar), 139.91 (Ar), 139.86 (Ar), 139.76 (Ar), 129.69 (Ar), 129.59 (Ar), 129.59 (Ar), 129.49 (Ar), 129.45 (Ar), 129.32 (Ar), 129.11 (Ar), 129.08 (Ar), 128.93 (Ar), 128.88 (Ar), 128.83 (Ar), 128.66 (Ar), 128.35 (Ar), 104.16 ($^{1}J_{C,H}$ = 169 Hz, Man_C1'^{3}), 103.63 ($^{1}J_{C,H}$ = 164 Hz, GlcNAc_C1), 101.79 (GlcNAc_C1'), 101.69 ($^{1}J_{C,H}$ = 169 Hz, Man_C1'^{6}), 101.47 ($^{1}J_{C,H}$ = 159 Hz, Man_C1), 84.24 (Man_C3), 82.54 (GlcNAc_C3), 81.47 (GlcNAc_C3'), 78.43 (GlcNAc_C4'), 77.19 (GlcNAc_C4), 76.59 (GlcNAc_C5), 76.27 (GlcNAc_C5'), 76.09 (Man_C5), 76.06 (Ph*C*H$_2$), 75.57 (Man_C4), 75.41 (Ph*C*H$_2$), 75.30 (Man_C3'^{3}), 75.29 (Ph*C*H$_2$), 74.58 (Man_C3'^{6}), 74.46 (Ph*C*H$_2$), 74.45 (Ph*C*H$_2$), 72.85, 72.73 (Man_C4'^{3} and Man_C4'^{6}), 72.69 (Man_C2'^{3}), 72.13 (Man_C2'^{6}), 71.95 (Man_C2), 70.02 (GlcNAc_C6), 69.81 (GlcNAc_C6'), 68.91 (Man_C5'^{3}), 68.75 (Man_C5'^{6}), 67.30 (Man_C6), 63.19 (Man_C6'^{3}), 62.81 (Man_C6'^{6}), 57.30 (GlcNAc_C2'), 57.08 (OCH_3), 55.89 (GlcNAc_C2), 23.38 (Ac), 23.11 (Ac). ESI-MS *m*/*z* calc'd for $C_{70}H_{90}N_2O_{26}$ $(M+Na)^+$: 1397.57; found: 1397.91. ESI-HRMS *m*/*z* calc'd for $C_{70}H_{90}N_2O_{26}$ $(M+Na)^+$: 1397.5680; found: 1397.5682.

Methyl a-D-mannopyranosyl-(1→3)-[a-D-mannopyranosyl-(1→6)]-b-D-manno-pyranosyl-(1→4)-2-acetamido-2-deoxy-b-D-glucopyranosyl-(1→4)-2-acetamido-2-deoxy-b-D-glucopyranoside (9). The starting material (**40**; 64 mg, 0.047 mmol) and *t*-BuOH (1.0 mL) were added to condensed NH_3 (40 mL) at –70 °C, and then $Na_{(s)}$ (0.68 g, 30 mmol) was added over 4 hours. The reaction was quenched via the slow addition of MeOH (10 mL), warmed back to ambient temperature, and the bulk NH_3 left to evaporate overnight. The remaining mixture was co-evaporated with MeOH (2 x 10 mL), redissolved into H_2O (15 mL), and then adjusted to pH 8 using dilute $AcOH_{(aq)}$ (25% *v*/*v*). The solution was dialyzed, concentrated, and then purified via size-exclusion chromatography (P2 Bio-Gel) to afford the pure product as a white

solid (28 mg, 0.030 mmol, 65% yield). $[\alpha]_D^{20}$: -21° (*c* 1.0, H_2O). ^1H NMR (D_2O, 500 MHz): δH 5.13 (d, 1H, *J* = 1.4 Hz, Man_H1'^{3}), 4.94 (d, 1H, *J* = 1.6 Hz, Man_H1'^{6}), 4.81 – 4.80 (m, 1H, Man_H1), 4.63 (d, 1H, *J* = 7.9 Hz, GlcNAc_H1'), 4.46 (d, 1H, *J* = 8.0 Hz, GlcNAc_H1), 4.28 (dd, 1H, *J* = 1.5, <1 Hz, Man_H2), 4.09 (dd, 1H, *J* = 3.3, 1.7 Hz, Man_H2'^{3}), 4.00 (dd, 1H, *J* = 3.4, 1.7 Hz, Man_H2'^{6}), 3.97 – 3.88 (m, 7H, Man_H6^a, Man_H6^{a,6}, Man_H6^{a,3}, GlcNAc_H6$^{a'}$, Man_H3'^{3}, Man_H3'^{6}, and GlcNAc_H6^a), 3.85 – 3.62 (m, 19H, Man_H6^b, GlcNAc_H2', Man_H5'^{3}, GlcNAc_H3', Man_H3, Man_H4, Man_H6^{b,3}, Man_H6^{b,6}, GlcNAc_H6$^{b'}$, GlcNAc_H4', GlcNAc_H2, GlcNAc_H3, GlcNAc_H6^b, Man_H5'^{6}, Man_H5, Man_H4'^{3}, Man_H4'^{6}, GlcNAc_H4, and GlcNAc_H5'), 3.54 (ddd, 1H, *J* = 9.7, 5.4, 2.0 Hz, GlcNAc_H5), 3.52 (m, 3H, OCH_3), 2.11 (s, 3H, Ac), 2.06 (s, 3H, Ac). ^{13}C NMR (D_2O, 125 MHz): δC 175.34 (Ac), 175.32 (Ac), 103.16 ($^1J_{C,H}$ = 171 Hz, Man_C1'^{3}), 102.48 ($^1J_{C,H}$ = 162 Hz, GlcNAc_C1), 102.00 ($^1J_{C,H}$ = 162 Hz, GlcNAc_C1'), 101.01 ($^1J_{C,H}$ = 162 Hz, Man_C1), 100.26 ($^1J_{C,H}$ = 172 Hz, Man_C1'^{6}), 81.14 (Man_C3), 80.27 (GlcNAc_C4'), 79.97 (GlcNAc_C4), 75.16 (GlcNAc_C5), 75.02 (GlcNAc_C5'), 74.82 (Man_C5), 74.08 (Man_C5'^{3}), 73.32 (Man_C5'^{6}), 73.18 (GlcNAc_C3), 72.61 (GlcNAc_C3'), 71.05 (Man_C3'^{6}), 70.97 (Man_C3'^{3}), 70.79 (Man_C2), 70.64 (Man_C2'^{3}), 70.51 (Man_C2'^{6}), 67.50 (Man_C4'^{3}), 67.43 (Man_C4'^{6}), 66.48 (Man_C6), 66.46 (Man_C4), 61.78 (Man_C6'^{3}), 61.60 (Man_C6'^{6}), 60.73 (GlcNAc_C6), 60.63 (GlcNAc_C6'), 57.75 (OCH_3), 55.51 (GlcNAc_C2 and GlcNAc_C2'), 22.83 (Ac), 22.79 (Ac). ESI-MS *m*/*z* calc'd for $C_{35}H_{60}N_2O_{26}$ $(M+Na)^+$: 947.33; found: 947.52. ESI-HRMS *m*/*z* calc'd for $C_{35}H_{60}N_2O_{26}$ $(M+Na)^+$: 947.3332; found: 947.3336. HPLC purity analysis: >99.5%, R_t 5.79 minutes, Eclipse XDB-C8 column.

References

[1] McGeary, R. P.; Wright, K.; Toth, I. *J. Org. Chem.* **2001**, *66*, 5102-5105.

[2] D'Accorso, N. B.; Thiel, I. M. E. *Carbohydr. Res.* **1983**, *124*, 177-184.

[3] Crich, D.; Smith, M. *J. Am. Chem. Soc.* **2001**, *123*, 9015-9020.

[4] Tsukamoto, H.; Suzuki, T.; Kondo, Y. *Synlett* **2007**, *20*, 3131-3136.

[5] Orwenyo, J.; Huang, W.; Wang, L. X. *Bioorg. Med. Chem.* **2013**, *21*, 4768-4777.

ChemMedChem, 2019, *14*(7), 749-757
doi: 10.1002/cmdc.201900051

Publication 6

Improvement of Aglycone π-Stacking Yields Nanomolar to Sub-nanomolar FimH Antagonists

Wojciech Schönemann,[1)] Jonathan Cramer,[1)] Tobias Mühlethaler,[1)] Brigitte Fiege,[1)] Marleen Silbermann,[1)] Said Rabbani,[1)] Philipp Dätwyler,[1)] Pascal Zihlmann,[1)] Roman P. Jakob,[2)] Christoph P. Sager,[1)] Martin Smieško,[1)] Oliver Schwardt,[1)] Timm Maier,[2)] Beat Ernst[1)*]

[1)] Institute of Molecular Pharmacy, University of Basel,
Klingelbergstr. 50, 4056 Basel, Switzerland

[2)] Structural Biology, Biozentrum, University of Basel,
Klingelbergstr. 70, 4056 Basel, Switzerland

[*] Corresponding author
Tel.: 0041 61 267 15 51; Fax: 0041 61 207 15 52
E-mail: beat.ernst@unibas.ch

Contributions of Marleen Silbermann:

- Thermodynamic profiling by ITC in collaboration with Jonathan Cramer
- Protein expression and purification of $FimH_{LD}$ and $FimH_{FL}$
- Proofreading of the manuscript

ChemPubSoc Europe
DOI: 10.1002/cmdc.201900051
CHEMMEDCHEM Full Papers

VIP *Very Important Paper*

Improvement of Aglycone π-Stacking Yields Nanomolar to Sub-nanomolar FimH Antagonists

Wojciech Schönemann,[a] Jonathan Cramer,[a] Tobias Mühlethaler,[a] Brigitte Fiege,[a] Marleen Silbermann,[a] Said Rabbani,[a] Philipp Dätwyler,[a] Pascal Zihlmann,[a] Roman P. Jakob,[b] Christoph P. Sager,[a] Martin Smieško,[a] Oliver Schwardt,[a] Timm Maier,[b] and Beat Ernst*[a]

Antimicrobial resistance has become a serious concern for the treatment of urinary tract infections. In this context, an anti-adhesive approach targeting FimH, a bacterial lectin enabling the attachment of *E. coli* to host cells, has attracted considerable interest. FimH can adopt a low/medium-affinity state in the absence and a high-affinity state in the presence of shear forces. Until recently, mostly the high-affinity state has been investigated, despite the fact that a therapeutic antagonist should bind predominantly to the low-affinity state. In this communication, we demonstrate that fluorination of biphenyl α-D-mannosides leads to compounds with perfect π–π stacking interactions with the tyrosine gate of FimH, yielding low nanomolar to sub-nanomolar K_D values for the low- and high-affinity states, respectively. The face-to-face alignment of the perfluorinated biphenyl group of FimH ligands and Tyr48 was confirmed by crystal structures as well as $^{1}H,^{15}N$-HSQC NMR analysis. Finally, fluorination improves pharmacokinetic parameters predictive for oral availability.

Introduction

Urinary tract infections (UTI) caused by uropathogenic *E. coli* (UPEC) are among the most common infections and account for more than 70% of the reported UTI cases.[1] Particularly often affected are women, who have a 40–50% risk of experiencing at least one symptomatic UTI episode during their life.[2] Because of the high morbidity and the resulting associated costs, UTI pose a serious burden on healthcare systems.[3] To date, patients with acute uncomplicated lower UTI are treated with antibiotics to relieve them from associated symptoms and to prevent the infection from aggravating into life-threatening pyelonephritis or urosepsis.[4] However, the repeated use of antibiotics as a first-line treatment for UTI increasingly provokes antibiotic resistance,[5] highlighting the need for new strategies for the prevention and treatment of UTI.[6]

The adhesin FimH, located on the tip of type 1 pili of UPEC, is a bacterial virulence factor.[7] Its adherence to the highly mannosylated glycoprotein uroplakin 1a on urothelial cells of the host not only represents the initial step of the infection, but also prevents clearance of UPEC by the urine bulk flow.[8] FimH consists of two domains: an N-terminal lectin domain ($FimH_{LD}$), acting as a mannose-specific carbohydrate recognition domain (CRD), and a C-terminal pilin domain ($FimH_{PD}$) anchoring the adhesin FimH into the pilus core.[9] To successfully colonize urothelial cells, bacteria only weakly interact with the host cell surfaces, still allowing their exploration for optimal nutrition supply. However, when shear forces arise, strong adherence to host cells is necessary to avoid clearance from the bladder. To fulfill these two opposing tasks, FimH relies on a sophisticated allosteric mechanism, fine-tuning its mannose binding affinity through conformational adaption. In the medium-affinity state, a close interaction between $FimH_{PD}$ and $FimH_{LD}$ allosterically attenuates the mannose binding capacity. However, when shear forces induced by micturition provoke the separation of the two domains, the CRD adopts the high-affinity state. This conformational switch is by no means a one-step mechanism. Intermediate states and a high variability of the conformational ensemble have been described.[10,11] Although all conformational states of FimH play a relevant role for the pathogenicity of UPEC, efforts for the development of therapeutic FimH antagonists have almost exclusively been focused on the high-affinity state of the protein.[12]

Within the last three decades, a large number of monovalent α-D-mannosides with lipophilic aglycones, such as alkyl,[13] phenyl,[14] diarylamides,[15] dioxocyclo-butenylaminophenyl,[16] umbelliferyl,[14] biphenyl,[17–20] indol(in)ylphenyl,[21] triazolyl,[22] and thiazolylamino[23] have been synthesized. Moreover, a wide range of multivalent ligands has also been explored.[24] Generally, FimH antagonists consist of a mannose moiety establishing an extended hydrogen bond network within the CRD and a lipophilic aglycone with α configuration that interacts with

[a] *Dr. W. Schönemann, Dr. J. Cramer, T. Mühlethaler, Dr. B. Fiege, M. Silbermann, Dr. S. Rabbani, Dr. P. Dätwyler, Dr. P. Zihlmann, Dr. C. P. Sager, Dr. M. Smieško, Dr. O. Schwardt, Prof. Dr. B. Ernst*
Department of Pharmaceutical Sciences, University of Basel, Klingelbergstrasse 50, 4056 Basel (Switzerland)
E-mail: beat.ernst@unibas.ch

[b] *Dr. R. P. Jakob, Prof. Dr. T. Maier*
Department Biozentrum, Focal Area Structural Biology, University of Basel, Klingelbergstrasse 70, 4056 Basel (Switzerland)

Supporting information and the ORCID identification number(s) for the author(s) of this article can be found under: https://doi.org/10.1002/cmdc.201900051.

the hydrophobic amino acids Tyr48, Ile52, and Tyr137 lining the entrance to the CRD.[25]

To exploit electrostatic complementarity to the electron-rich tyrosine side chain, a series of biphenyl aglycones were designed with the goal of placing an electron-deficient terminal phenyl moiety within interaction distance to Tyr48. Therefore, we explored substitution with fluorine atoms having a van der Waals radius similar to that of hydrogen to avoid unfavorable steric clashes as a consequence of an altered docking mode.[26] In case of hexafluorobenzene, the high electronegativity of fluorine causes an inductive polarization leading to a positive potential above and below the ring plane, resulting in a quadrupole moment of 31.7×10^{-40} C m^2. Because of the inverse electron density distribution, the quadrupole moment of benzene exhibits the same magnitude but with opposite sign (-33.3×10^{-40} C m^2).[27]

Such a complementary charge distribution results in a particularly strong electrostatic attraction as experimentally demonstrated on the rotational barrier in 1,8-diarylnaphthalenes.[28,29] Fluorine substitution led to an increase of the energy barrier of approximately 2.1 kJ mol^{-1} per fluorine atom.[30] Finally, the influence of fluoroarenes on protein binding affinities was demonstrated for various targets, including carbonic anhydrase II and the cysteine protease cathepsin L.[31]

Based on a series of biphenyl α-D-mannosides (**1**–**3**, Figure 1), we aimed to improve the affinity for all relevant conformational states of FimH by introducing different fluorination

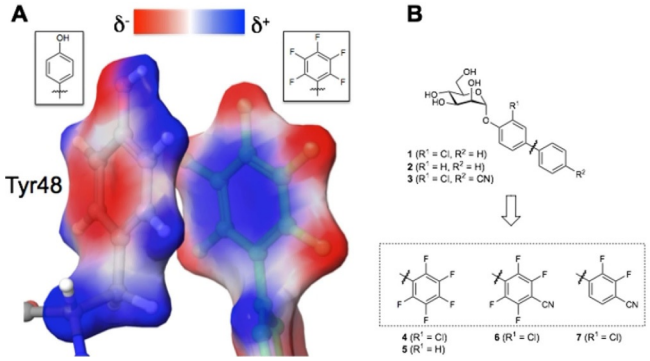


Figure 1. A) Electrostatic potential visualized on the interaction of the perfluorinated aglycone of mannoside **5** and Tyr48 of FimH. Regions of negative potential (red) of the perfluorinated aglycone correspond with regions of positive potential (blue) of the Tyr48 side chain and vice versa. B) The biphenyl α-D-mannosides **4**–**7** exhibiting different fluorination patterns were synthesized.

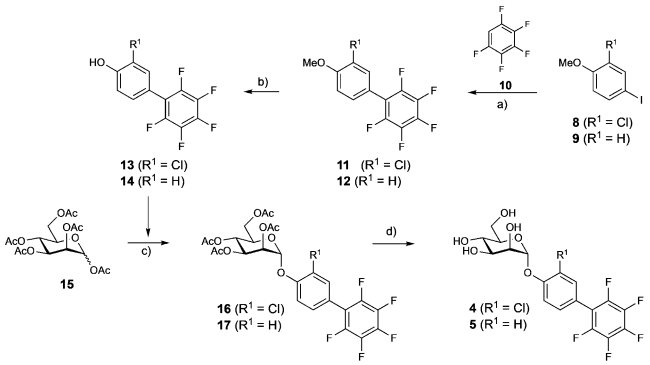
patterns on the terminal aromatic ring. In addition to perfluorination, as in compounds **4** and **5**, a cyano substituent was introduced in the *para* position (→**6**). This modification is marked by an increase in the Hammett constant,[32] which is an indicator for the electrostatic properties of aromatic groups and signifies a decrease in electron density, further increasing the electrostatic attraction to the Tyr48 side chain. Finally, the difluorinated compound **7** was designed to enable favorable σ–π interactions in a T-shape with the tyrosine gate residues.

Results and Discussion

For exploitation of π–π stacking interactions between the outer aromatic ring of biphenyl α-D-mannopyranosides and Tyr48, four representatives with different fluorination patterns (→**4**–**7**) were synthesized.

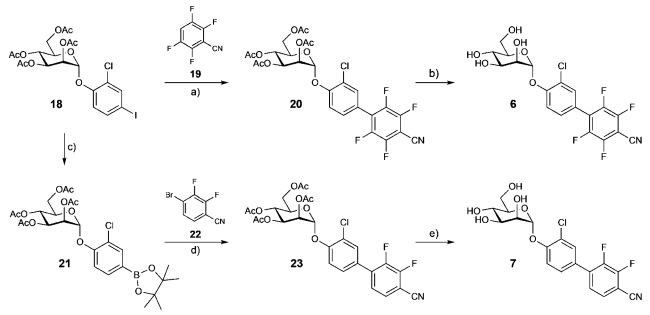
Synthesis

The synthesis of test compounds **4** and **5** is depicted in Scheme 1. The pentafluoro-substituted biphenyls **13** and **14** were obtained from the anisole derivatives **8** and **9** in a copper-catalyzed arylation of arene **10** followed by demethylation using boron tribromide.[33] $BF_3{\cdot}Et_2O$-promoted mannosylation (→**16** & **17**) followed by deprotection under Zemplén conditions afforded the mannosides **4** and **5**.



Scheme 1. a) CuI, 1,10-phenantroline, K_3PO_4, DMF/*p*-xylene, 130 °C, 2–10 h, 36% for **11**, 46% for **12**; b) BBr_3, CH_2Cl_2, −78 °C→RT, 5–19 h, 88% for **13**, 90% for **14**; c) $BF_3{\cdot}Et_2O$, CH_2Cl_2, 40 °C, 3–24 h, 40% for **16**, 42% for **17**; d) MeONa/MeOH, RT, 2 h, 74% for **4**, 89% for **5**.

For the synthesis of the tetrafluoro-substituted derivative **6** (Scheme 2), mannoside **18**[21] was coupled in the presence of CuI and 1,10-phenanthroline with commercially available 2,3,5,6-tetrafluorobenzonitrile (**19**) in a microwave-assisted arylation reaction to yield intermediate **20**. Because of the high electrophilicity of the cyano group, cleavage of the acetyl



Scheme 2. a) CuI, 1,10-phenantroline, K_3PO_4, *p*-xylene, microwave, 150→160 °C, 13 h, 66%; b) 2.0 M aq. NH_3, *i*PrOH, THF, RT, 24 h, 39%; c) bis(pinacolato)diborone, $PdCl_2(dppf){\cdot}CH_2Cl_2$, KOAc, DMF, microwave, 120 °C, 2 h, 61%; d) $PdCl_2(dppf){\cdot}CH_2Cl_2$, K_3PO_4, DMF, 90 °C, 2.5 h, 58%; e) 1.3 M aq. NH_3, *i*PrOH, THF, RT, 4.5 days, 73%.

groups under Zemplén conditions (NaOMe, MeOH) almost exclusively led to the formation of by-products. Although this could not be completely avoided under milder deacetylation conditions (aq. NH_3 in isopropanol/THF), the test compound **6** could be isolated, however, only with moderate yield.

For the synthesis of the difluoro-substituted derivative **7** (Scheme 2), mannoside **18** was transformed into the boronic acid pinacol ester **21** followed by Suzuki coupling (→**23**).[34] Finally, by deprotection with aq. NH_3, by-product formation could be avoided in this case, yielding test compound **7** in 73% yield.

FimH constructs

Two different recombinant FimH constructs were used to determine binding affinities. The isolated lectin domain of FimH ($FimH_{LD}$) locked in the high-affinity conformation was expressed as previously described.[17] The full-length FimH variant ($FimH_{FL}$), representing the low- and, after ligand binding, the medium-affinity state, consists of the tethered lectin and pilin ($FimH_{PD}$) domain, the latter stabilized by a synthetic oligopeptide corresponding to residues 1–14 of the natural FimG donor strand (Figure 2 A).[10]

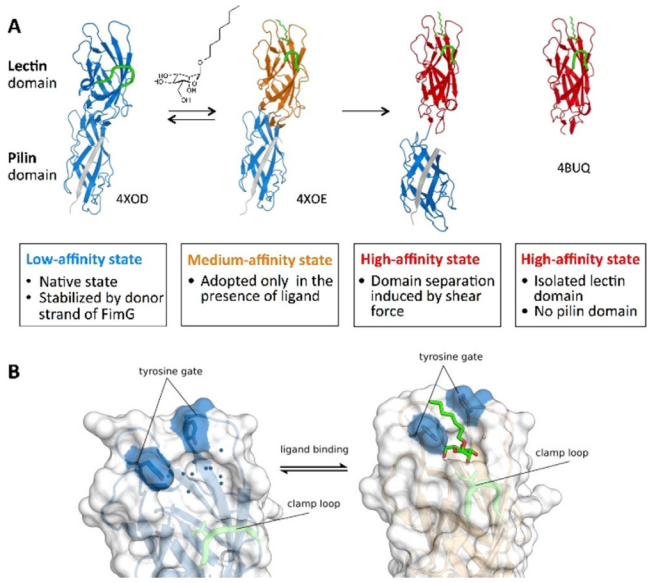


Figure 2. A) Structural models of different FimH conformations. Except for the first structure (low-affinity state), the ligand *n*-heptyl-α-D-mannoside is bound to the active site. B) Ligand-induced conformational adjustments of the mannose binding site of low- (PDB ID: 4XOD) to medium- (PDB ID: 4XOE) affinity FimH states.

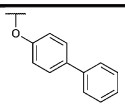
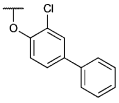
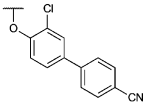
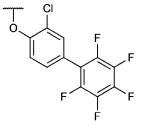
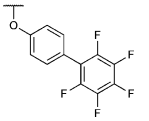
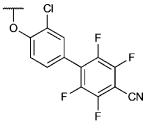
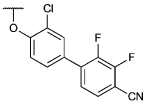
Competitive fluorescence polarization assay (FP assay)

In this assay, the antagonist of interest displaces a fluorescently labeled competitor from the binding site, thereby causing a decrease in fluorescence polarization.[19] As a surrogate for the high-affinity conformation of FimH, $FimH_{LD}$ was used in most published studies for affinity determination. However, because the low/medium-affinity conformation most likely represents the relevant therapeutic state of the protein,[12,35,36] we also investigated affinities of $FimH_{FL}$ for selected compounds (→**4**–**7**) with the FP assay (Figure 2 A).

Affinity to $FimH_{FL}$

Relative to their non-fluorinated counterparts, all fluoro-substituted antagonists exhibited higher affinities to full-length FimH (Table 1). While pentafluorination led to a roughly three-fold improvement in affinity (**1**→**5**, K_D 1100 nM vs. 346 nM), antagonist **4,** exhibiting an additional *o*-chloro substituent on the phenyl ring adjacent to the anomeric position, showed an even 4.5-fold improvement (**2**→**4**, K_D 458 nM vs. 103 nM). Apart from the electron-withdrawing effect of fluorine, which enhances π–π stacking interactions to Tyr48, the new antagonists can also benefit from a slightly increased surface area leading to improved van der Waals contacts. Because the Hammett constants (σ_{para}) revealed that the introduction of a cyano group in the *para* position of the terminal phenyl ring impli-

Table 1. Affinity of FimH antagonists to $FimH_{FL}$ and $FimH_{LD}$ determined by competitive fluorescence polarization assay.

Compd		K_D [nM][a]	
		$FimH_{FL}$	$FimH_{LD}$
1[19]		1100	15.1
2[35]		458	2.8
3[19]		59 (47–75)	1.06 (0.85–1.32)
4		103 (85–123)	4.48 (3.75–5.35)
5		346 (296–403)	4.01 (3.30–4.87)
6		22 (15–32)	0.35 (0.24–0.51)
7		23 (17–30)	0.58 (0.45–0.75)

[a] Dissociation constants were determined by FP assay using the lectin receptors $FimH_{FL}$, $FimH_{LD}$, and the fluorescent reference compound 3′-chloro-*N*-(2-(3-(3′,6′-dihydroxy-3-oxo-3*H*-spiro-[isobenzofuran-1,9′-xanthen]-5-yl)thioureido)ethyl)-4′-(α-D-mannopyranosyloxy)biphenyl-4-carboxamide.[19] For all new compounds, 95% confidence intervals from curve fitting of data for three independent experiments are given as an estimation of experimental error.

cates a significantly higher electron-withdrawing potential than a *para*-fluoro substituent (σ_{para} 0.660 vs. 0.062), the increase in the binding affinity (**4**→**6**, K_D 103 nM vs. 22 nM) is not surprising. The cyano-difluoro derivative **7** was designed to capitalize on σ–π interactions resulting from T-shaped orientation to Tyr48, which has been shown to be energetically more favorable than the face-to-face conformation in the benzene homodimer.[29] However, the observed affinity is practically identical to the cyano-tetrafluoro derivative **6**, possibly because the more beneficial σ–π interaction is compensated by a higher desolvation penalty, originating from the strong polarization of the aromatic ring in **7** and the related larger dipole moment.

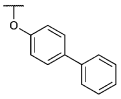
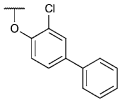
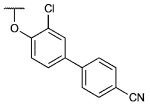
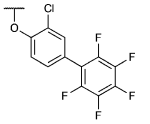
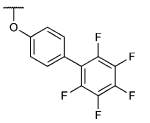
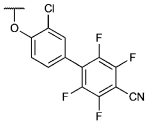
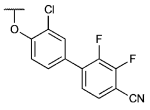
Affinity to $FimH_{LD}$ (locked in the high-affinity state)

The fluorination of the terminal phenyl moiety in the biphenyl aglycone resulted in an almost four-fold gain in affinity (**1**→**5**; **3**→**6**). A similar effect was found for the difluoro-substituted derivative **7**. In contrast to earlier observations,[20] the additional *ortho*-chloro substituent in mannoside **4** unexpectedly decreased binding affinity.

Ab initio calculation of π–π stacking interactions

As an estimate for the π–π stacking interaction energies of different aglycones and Tyr48, quantum mechanical calculations were performed based on X-ray structural data. Analysis of the interaction energies in Table 2 suggest that the π–π stacking is strongest between Tyr48 and the tetrafluoro-cyano derivative **6** ($\Delta E_{gas} = -42.9$ kJ mol^{-1}) and weakest for compounds **1** and **2** featuring unsubstituted terminal aromatic rings ($\Delta E_{gas} = -19.8$ and -19.6 kJ mol^{-1}, respectively). Introduction of the *o*-chloro substituent to the biphenyl derivatives **1** (→**2**) and **5** (→**4**) does not influence the π–π stacking. Addition of a *p*-cyano group or perfluorination of the terminal phenyl ring equally improves the π–π stacking energies as observed for compounds **3**, **4** and **5** (ΔE_{gas} between -34.3 and -35.2 kJ mol^{-1}). However, substitution of the *p*-fluoro with a *p*-cyano group (→**6**) induces further polarization and accordingly a further en-

Table 2. Physicochemical and pharmacokinetic parameters (solubility, distribution coefficient, and permeability), and quantum mechanical calculation of the π–π stacking energies of Tyr48 with different biphenyl derivatives in the gas phase.

Compd		Solubility [μg mL^{-1}]	log$D_{7.4}$[c]	PAMPA logP_e [cm s^{-1}][d]	ΔE_{gas} [kJ mol^{-1}]
1[19]		21 ± 1[a]	2.1 ± 0.1	−4.7 ± 0.1	−19.8
2[35]		5.5 ± 0.2[a]	2.6 ± 0.1	n.d.	−19.6
3[19]		195 ± 5[b]	2.1 ± 0.0	−5.2 ± 0.0	−34.3
4		47 ± 7[b]	2.7 ± 0.2	−4.7 ± 0.2	−35.2
5		131 ± 10[b]	2.3 ± 0.1	−4.7 ± 0.2	−34.4
6		111 ± 7[b]	2.2 ± 0.1	−4.8 ± 0.1	−42.9
7		97 ± 15[b]	2.2 ± 0.1	−5.0 ± 0.1	−39.8[e]/ −30.7[f]

[a] Kinetic solubility was measured in a 96-well format using the μSOL Explorer solubility analyzer in triplicate. The experimental error for all experiments is given as standard deviation from the mean. [b] Thermodynamic solubility was measured in 0.1 M Tris·HCl buffer pH 7.4 in duplicate. [c] Octanol/water distribution coefficients (log$D_{7.4}$) were determined at pH 7.4 by a miniaturized shake-flask procedure in sextuplicate. [d] P_e = effective permeability: diffusion through an artificial membrane was determined by the parallel artificial membrane permeability assay (PAMPA) in quadruplicate at pH 7.4. [e] ΔE_{gas} obtained with fluorine atoms on the same side as the chlorine atom. [f] ΔE_{gas} obtained with fluorine atoms on opposing side of the chlorine.

hancement of the π–π stacking energy ($\Delta E_{gas} = -42.9$ kJ mol^{-1}). Interestingly, the difluoro-cyano derivative **7** exhibits similar interaction energy when the two fluorine atoms are on the solvent-exposed, coinciding side with the chlorine atom ($\Delta E_{gas} = -39.8$ kJ mol^{-1}). See the Supporting Information for experimental details.

X-ray crystallography

To get better insight into the effect of fluorination on the binding mode, antagonists **5** and **6** were co-crystallized with $FimH_{LD}$ (Figure 3; see the Supporting Information for experimental details). The structures were obtained at a resolution of 2.2 Å and 2.1 Å, respectively. The binding mode of the mannose moieties is identical to previously reported non-fluorinated analogues **1** (PDB ID: 4X50)[15] and **3** (PDB ID: 4CST).[19] In all crystal structures, the biphenyl aglycones interact with Tyr48 and Tyr137. Although side chain orientation of Tyr137, which is in contact with the inner aromatic ring, is not affected, substantial effects can be seen for the orientation of Tyr48. With the non-fluorinated antagonist **1** (Figure 3A), the Tyr48 side chain adopts a slightly tilted orientation with a distance of 4.4 Å between the ring centers, implicating a suboptimal arrangement for π–π stacking interactions with the aglycone. Although the introduction of the electron-withdrawing cyano group (→**3**, Figure 3B) is supposed to improve π–π stacking, a face-to-face conformation with Tyr48 is not observed. Instead, the Tyr48 side chain displays a certain degree of flexibility implicated from the observation of two distinct rotamers (3.6 Å and 3.8 Å distance between ring centers).[19]

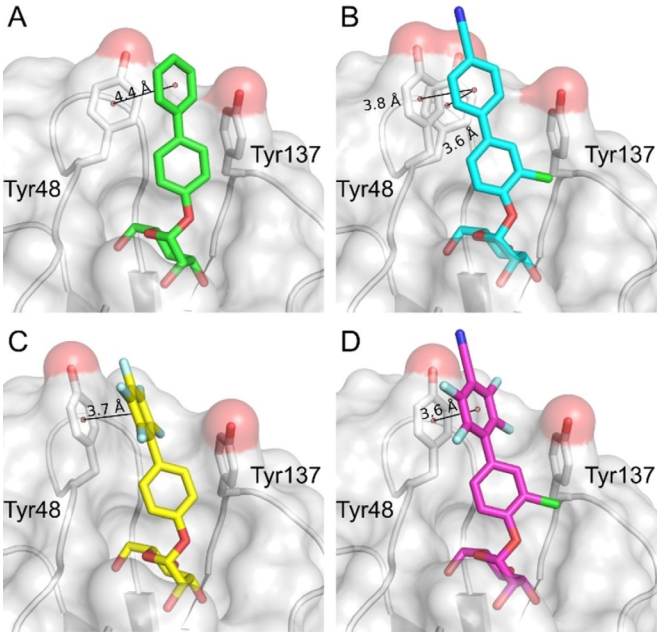


Figure 3. Comparison of the crystal structures of $FimH_{LD}$ co-crystallized with fluorinated antagonists **5** and **6**. A) Binding mode of non-fluorinated analogue **1** (green sticks, PDB ID: 4X50). B) Binding mode of non-fluorinated analogue **3** (cyan sticks, PDB ID: 4CST), Tyr48 exhibits two alternate conformations). C) Binding mode of **5** (yellow sticks, PDB ID: 6G2S). D) Binding mode of **6** (magenta sticks, PDB ID: 6G2R).

By contrast, the cyano-tetrafluoro derivative **6** (Figure 3D) is in an optimal parallel alignment with Tyr48 (3.6 Å distance), most likely due to a markedly enhanced π–π stacking resulting from the electron deficiency of the fluorinated ring. A similar orientation is found for the pentafluorinated derivative **5** (3.7 Å distance, Figure 3C). Because the binding sites are solvent exposed, it should be mentioned that the aglycones of the co-crystallized antagonists **5** and **6** are involved in crystal packing contacts potentially influencing the binding modes.

NMR chemical shift perturbation (CSP)

The structural information obtained from X-ray crystallography was complemented with data from NMR CSP experiments of $FimH_{LD}$ in solution. $^{1}H,^{15}N$-HSQC spectra of ^{15}N-labeled $FimH_{LD}$[15] were measured upon addition of the fluorinated compounds **4**, **5** and **6** as well as their non-fluorinated counterparts **1** and **3**. Residues in the mannose binding pocket exhibit nearly identical chemical shifts for all five antagonists (Figure 4A) confirming a similar orientation and hydrogen bond network of their mannosyl moieties as observed in the X-ray structures shown in Figure 3. In contrast, chemical shift differences were observed for residues in the binding loop around Tyr48. Previously, the residues Glu50 and Thr51 have been identified as sensitive reporters for the conformation of the Tyr48 residue.[15] The formation of a T-shaped NH–π interaction of Thr51 H^N and the Tyr48 ring in complex with the non-fluorinated biphenyl mannosides **1** and **3** leads to a strong upfield shift of Thr51 H^N while Glu50 is shifted downfield, corresponding to shielding and deshielding effects by the aromatic ring current from Tyr48, respectively (Figure 4B). The $^{1}H,^{15}N$-HSQC spectra of $FimH_{LD}$ with the fluorinated counterparts **4**, **5** and **6** revealed the presence of this Tyr48 conformation in agreement with the X-ray structural data (Figures 3C, D). It is tempting to

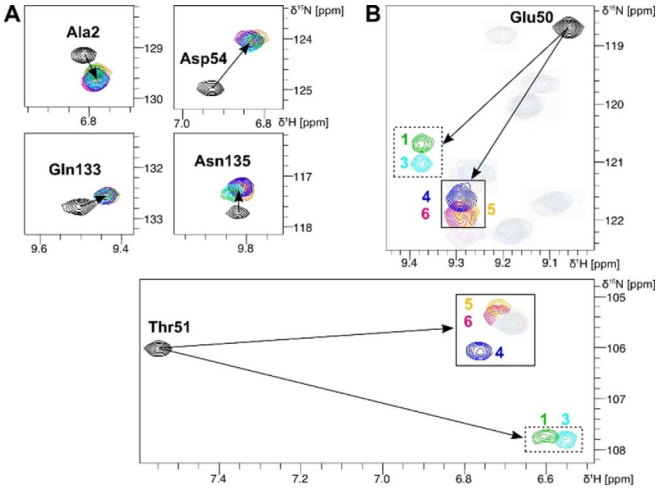


Figure 4. $^{1}H,^{15}N$-HSQC spectral regions of $FimH_{LD}$ with fluorinated compounds **4**, **5** and **6** as well as their non-fluorinated counterparts **1** and **3**. A) Signal shifts of four representative amino acids in the mannose binding site. B) Signal shifts of Glu50 and Thr51.

speculate that the observed conformation of the tyrosine gate and the aglycone in the X-ray structures of these three compounds represents the optimal stacking arrangement of Tyr48 and aglycone.

Physicochemical and in vitro pharmacokinetic characterization

The introduction of fluorine in pharmacologically active molecules can have a significant impact on their physicochemical and pharmacokinetic characteristics.[37] In general, non-fluorinated FimH antagonists, such as **3**, are known to exhibit acceptable pharmacokinetic properties.[19] To exclude disadvantageous effects from fluorination, we determined the aqueous solubility[38] and lipophilicity[39] and passive permeability in a parallel artificial membrane permeation assay (PAMPA)[40] of the new antagonists. The results are summarized in Table 2.

Whereas the introduction of an *ortho*-chloro substituent (**1**→**2** or **5**→**4**) substantially decreased solubility, fluorination of the terminal aromatic ring (**1**→**5**, **2**→**4**) led to a six- to eight-fold improvement. When the *para*-fluoro substituent was replaced by a *para*-cyano group (**5**→**6**), solubility remained unchanged. However, against expectations, tetrafluorination of cyanide **3** (→**6**) decreased solubility two-fold. Interestingly, despite a higher predicted polarization of the terminal phenyl ring, the difluorinated analogue **7** had essentially the same solubility as **6**. For the two sub-nanomolar antagonists **6** and **7**, solubility is above the 52 $\mu g\,mL^{-1}$ threshold required for complete absorption of a 1 $mg\,kg^{-1}$ dose of moderately permeable compounds.[41]

Non-fluorinated and fluorinated antagonists show similar lipophilicity with $\log D_{7.4}$ between 2.2 and 2.7 and are expected to be well absorbed by means of passive diffusion upon oral administration.[42] Moreover, a high lipophilicity supports tubular re-absorption from the glomerular filtrate favoring slower renal clearance and thus a prolonged therapeutic effect in the bladder upon application of a single dose.[43] The $\log P_e$ values determined in the PAMPA for both non-fluorinated and fluorinated antagonists are also a clear indication of oral permeability, as $\log P_e$ values above −5.7 indicate a high potential for intestinal absorption.[44]

In summary, fluorination of the aromatic aglycones showed little effect on lipophilicity and passive permeability. Interestingly, solubility is much more influenced by the presence of an *ortho*-chloro substituent on the aromatic ring adjacent to the anomeric center and by a *para*-cyano substituent on the terminal aromatic ring. Based on the results of prior pharmacokinetic evaluations[17,19,20,35] and in vivo studies[17,19] of biphenyl α-D-mannosides, favorable in vivo profiles for the high-affinity FimH antagonists **6** and **7** can be reasonably expected.

Conclusions

Over the years, the high-affinity conformation of FimH present in its isolated lectin domain ($FimH_{LD}$) was mainly used for the development of anti-adhesives for prevention and therapy of UTI, while the more relevant low-affinity state ($FimH_{PD}$) displayed by the full-length FimH was neglected. In a recent study, we addressed this issue, demonstrating an offset of approximately two orders of magnitude between the affinity of a compound for $FimH_{FL}$ relative to $FimH_{LD}$.[35] We therefore designed and synthesized a series of FimH antagonists with fluorinated aglycones with the goal to increase affinity for both states by enhancing the π–π stacking with the tyrosine gate. Pentafluorination (Table 1, **1**→**5** or **2**→**6**) gave a three- to four-fold improvement of binding to $FimH_{FL}$. When the *para*-fluoro substituent was replaced by a *para*-cyano group (Table 1, **4**→**6**), a further increase in affinity to a K_D of 22.1 nM was obtained. To our knowledge, this represents the best reported affinity for a $FimH_{FL}$ antagonist to date. In the case of $FimH_{LD}$, a similar improvement was observed, with **6** and **7** reaching even sub-nanomolar affinities. Structural models of the interaction of $FimH_{LD}$ with fluorinated antagonists confirm strong π–π stacking interactions. In summary, the fluorination of the terminal phenyl rings of biphenyl aglycone results in improved π–π stacking interactions leading to low nanomolar antagonists for $FimH_{FL}$ and sub-nanomolar antagonists for $FimH_{LD}$. In addition, the two best antagonists **6** and **7** display physicochemical properties supporting oral bioavailability.

Experimental Section

Synthesis: The synthesis of compounds **4**–**7**, including compound characterization data, can be found in the Supporting Information.

Protein production and purification: $FimH_{LD}$ and $FimH_{FL}$ were produced and purified as reported earlier.[10,45]

Fluorescence polarization (FP) assay: FP assays were essentially performed as previously described.[19] A detailed experimental protocol is given in the Supporting Information.

FimH crystallization and structure refinement: For crystallization, $FimH_{LD}$ (residues 1–158) at a final concentration of 15 $mg\,mL^{-1}$ (ca. 0.9 mM) with a three-fold molar excess of ligand **5** (ligand dissolved in DMSO) in 20 mM HEPES pH 7.4, 2% DMSO was used. $FimH_{LD}$/**5** crystals were grown in sitting-drop vapor diffusion at 19 °C in 0.1 M BisTrisPropane, propionic acid, cacodylate (PCPT buffer) pH 8.0 and 25% PEG1500. Crystals appeared after two months and grew to their final size within three months. They were cryopreserved by the addition of 20% ethylene glycol (*v*/*v*) and flash-cooled with liquid nitrogen. $FimH_{LD}$/**6** was crystallized in 0.1 M NaOAc pH 5.0, 1.5 M $(NH_4)_2SO_4$ at 19 °C and flash-cooled after a quick soak in 2.5 M Li_2SO_4.[46] Data were collected at the X06SA beamline of the Swiss Light Source (Paul Scherrer Institute, Switzerland) and indexed, integrated, and scaled with XDS.[47] The structures were solved by molecular replacement with PHASER[48] using the $FimH_{LD}$/biphenyl α-D-mannopyranoside (**1**) complex (PDB ID: 4X50)[15] as search model. The structures were built using COOT software[49] and periodically refined with PHENIX software.[50] Geometric restraints for the ligands were generated with PRODRG.[51] The final protein models were validated with Molprobity[52] and deposited in the Protein Data Bank with PDB IDs 6G2S and 6G2R.

NMR chemical shift perturbation (CSP): NMR samples for backbone CSP experiments contained 250 μM uniformly ^{15}N-labeled $FimH_{LD}$ and an excess (1.3–1.7 mM nominal due to solubility limitations) of fluorinated compounds **4**, **5** and **6**, respectively, in 20 mM sodium phosphate buffer pH 7.0 with 7% D_2O; 0.1 mM [D_4]TSP (3-

(trimethylsilyl)-2,2′,3,3′-tetradeuteropropionic acid, Armar Chemicals, Switzerland) was added as an internal reference. All spectra were acquired on a 500 MHz Bruker Avance III NMR spectrometer equipped with a 5 mm TXI RT probe head at a temperature of 298 K. All spectra were acquired and processed with Topspin 3.2 (Bruker BioSpin, Switzerland) and analyzed with CcpNmr Analysis 2.4[53] The backbone resonance assignment as well as CSP data with the non-fluorinated reference compounds **1** and **3** were available from a previous study.[15] As all tested mannoside ligands bound in the slow-exchange regime, peak assignment was performed based on chemical shift proximity.

Physicochemical and in vitro pharmacokinetics studies

Materials: Tris·HCl, 1-octanol, NaOH, formic acid, and 1-propranolol were purchased from Sigma–Aldrich Chemie GmbH (Steinheim, Germany). PRISMA HT universal buffer, GIT-0 Lipid Solution, and Acceptor Sink Buffer were purchased from pIon (Billerica, USA). MeCN and MeOH were ordered from Acros Organics (Geel, Belgium). Solubility, permeability, and lipophilicity of **3** was reported previously.[19]

Log$D_{7.4}$ determination (shake-flask method): Equal amounts of 0.1 M Tris·HCl buffer pH 7.4 and 1-octanol were mixed and vigorously shaken for 5 min to saturate the phases. The mixture was left until complete separation of the phases occurred and the buffer was retrieved. The test compound was diluted with buffer to a concentration of 10 μM. The buffer was transferred to a 96-well plate and saturated 1-octanol was added, resulting in a 3:180 and 4:180 1-octanol/water ratio, respectively. Each ratio was measured in triplicate, and simultaneous measurements were conducted with 1-propanolol as positive control. 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 liquid chromatography–mass spectrometry (LC–MS, see below).

The log$D_{7.4}$ coefficients were calculated from the 1-octanol/buffer ratio (*o*:*b*), the initial concentration of the analyte in buffer (10 μM), and the concentration of the analyte in buffer (c_b) with Equation (1).

$$logD_{7.4} = log(\frac{10\,\mu M - c_b}{c_b} \times \frac{1}{o:b}) \quad (1)$$

Aqueous solubility: Lyophilized compound was added to 200 μL 0.1 M Tris·HCl buffer pH 7.4 in duplicate until a precipitate was formed. The vials were put into a super-sonication bath (Branson 2510, Danbury, USA) for 30 min at 25 °C and then left at 25 °C for 24 h to equilibrate. The dispersion was filtered (0.2 μm), diluted in Tris·HCl buffer, and analyzed by LC–MS (see below).

Parallel artificial membrane permeability assay (PAMPA): Effective permeability (logP_e) was determined in a 96-well format with PAMPA.[40,44] For each compound, measurements were performed at pH 7.4 in quadruplicate. Four wells of a deep-well plate were filled with 650 μL of PRISMA HT universal buffer (pIon, Billerica, USA) and adjusted to pH 7.4 by adding the requested amount of NaOH (0.5 M). Samples (150 μL) were withdrawn from each well to determine the blank spectra by UV/Vis spectroscopy (190 to 500 nm, SpectraMax 190, Molecular Devices, Silicon Valley, CA, USA). Then, analyte dissolved in DMSO (10 mM) was added to the remaining buffer to yield 50 μM solutions. Afterward, samples (150 μL) were withdrawn to determine the reference spectra. Further 200 μL 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 μL of GIT-0 Lipid Solution and 200 μL of Acceptor Sink Buffer was filled into each acceptor well. The sandwich was assembled, placed in the GutBox™ (pION), and left undisturbed for 16 h. Thereafter, it was disassembled and samples (150 μL) were transferred from each donor and acceptor well to UV plates for determination of the UV/Vis spectra. Effective permeability (logP_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).

Liquid chromatography–mass spectrometry measurements (LC–MS): Analyses were performed using a 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.03.01). The column used was an Atlantis® T3 C_{18} column (2.1× 50 mm) with a 3 μm particle size (Waters Corp., Milford, MA, USA). The mobile phase consisted of eluent A: H_2O containing 0.1% formic acid; and eluent B: MeCN containing 0.1% formic acid. The flow rate was maintained at 0.6 mL min^{-1}. 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. Subsequently, the system was brought back to 95% A/5% B, resulting in a total duration of 4 min. Fragmentor voltage and collision energy were optimized for the analysis of compounds in multiple reaction monitoring mode in positive mode. The concentrations of the analytes were quantified by the Agilent Mass Hunter Quantitative Analysis software (version B.04.00).

Interference compounds: Tested compounds do not belong to the known classes of assay interference compounds and are therefore not expected to interfere unselectively with the applied assays.

PDB IDs

The crystal structures of $FimH_{LD}$ co-crystallized with the fluorinated antagonists **5** (PDB ID: 6G2S) and **6** (PDB ID: 6G2R) are shown in Figure 3. The authors will release the atomic coordinates and experimental data upon article publication.

Supporting Information

Experimental and analytical data for compound syntheses, experimental details for affinity data from fluorescence polarization measurements, data collection and refinement statistics from X-ray diffraction experiments, method for ab initio calculations, and HPLC as well as NMR spectra for the compounds can be found in the Supporting Information.

Acknowledgements

Financial support from the Swiss National Science Foundation (P.Z., SNF 200020_146202/1; C.P.S., SNF 200020_146202/1) is gratefully acknowledged. B.F. thanks the German Academic Exchange Service (DAAD) for a scholarship.

Conflict of Interest

The authors declare no competing financial interest.

Keywords: antibiotics · drug design · FimH antagonists · urinary tract infections · uropathogenic *Escherichia coli*

[1] T. J. Wiles, R. R. Kulesus, M. A. Mulvey, *Exp. Mol. Pathol.* **2008**, *85*, 11.
[2] a) B. Foxman, R. Barlow, H. D'Arcy, B. Gillespie, J. D. Sobel, *Ann. Epidemiol.* **2000**, *10*, 509; b) A. Ronald, *Am. J. Med.* **2002**, *113*, 14S–19S.
[3] a) L. Cegelski, G. R. Marshall, G. R. Eldridge, S. J. Hultgren, *Nat. Rev. Microbiol.* **2008**, *6*, 17; b) A. L. Flores-Mireles, J. N. Walker, M. Caparon, S. J. Hultgren, *Nat. Rev. Microbiol.* **2015**, *13*, 269.
[4] a) S. D. Fihn, *N. Engl. J. Med.* **2003**, *349*, 259; b) T. M. Hooton, R. Besser, B. Foxman, T. R. Fritsche, L. E. Nicolle, *Clin. Infec. Dis.* **2004**, *39*, 75.
[5] a) T. M. Hooton, *N. Engl. J. Med.* **2012**, *366*, 1028; b) G. V. Sanchez, R. N. Master, J. A. Karlowsky, J. M. Bordon, *Antimicrob. Agents Chemother.* **2012**, *56*, 2181.
[6] A. E. Clatworthy, E. Pierson, D. T. Hung, *Nat. Chem. Biol.* **2007**, *3*, 541.
[7] a) M. A. Mulvey, J. D. Schilling, J. J. Martinez, S. J. Hultgren, *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 8829; b) J. D. Schilling, M. A. Mulvey, S. J. Hultgren, *J. Infect. Dis.* **2001**, *183*, S36–S40; c) C. H. Jones, J. S. Pinkner, R. Roth, J. Heuser, A. V. Nicholes, S. N. Abraham, S. J. Hultgren, *Proc. Natl. Acad. Sci. USA* **1995**, *92*, 2081.
[8] a) J. Pak, Y. Pu, Z.-T. Zhang, D. L. Hasty, X.-R. Wu, *J. Biol. Chem.* **2001**, *276*, 9924; b) 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.
[9] a) D. Choudhury, A. Thompson, V. Stojanoff, S. Langermann, J. Pinkner, S. J. Hultgren, S. D. Knight, *Science* **1999**, *285*, 1061; b) C.-S. Hung, J. Bouckaert, D. Hung, J. Pinkner, C. Widberg, A. DeFusco, C. G. Auguste, R. Strouse, S. Langermann, G. Waksman et al., *Mol. Microbiol.* **2002**, *44*, 903; c) I. Le Trong, P. Aprikian, B. A. Kidd, M. Forero-Shelton, V. Tchesnokova, P. Rajagopal, V. Rodriguez, G. Interlandi, R. Klevit, V. Vogel et al., *Cell* **2010**, *141*, 645.
[10] M. M. Sauer, R. P. Jakob, J. Eras, S. Baday, D. Eriş, G. Navarra, S. Bernèche, B. Ernst, T. Maier, R. Glockshuber, *Nat. Commun.* **2016**, *7*, 10738.
[11] V. Kalas, J. S. Pinkner, T. J. Hannan, M. E. Hibbing, K. W. Dodson, A. S. Holehouse, H. Zhang, N. H. Tolia, M. L. Gross, R. V. Pappu et al., *Sci. Adv.* **2017**, *3*, e1601944.
[12] D. Eris, R. C. Preston, M. Scharenberg, F. Hulliger, D. Abgottspon, L. Pang, X. Jiang, O. Schwardt, B. Ernst, *ChemBioChem* **2016**, *17*, 1012.
[13] J. Bouckaert, J. Berglund, M. Schembri, E. de Genst, L. Cools, M. Wuhrer, C.-S. Hung, J. Pinkner, R. Slättegård, A. Zavialov et al., *Mol. Microbiol.* **2005**, *55*, 441.
[14] N. Firon, S. Ashkenazi, D. Mirelman, I. Ofek, N. Sharon, *Infect. Immun.* **1987**, *55*, 472.
[15] B. Fiege, S. Rabbani, R. C. Preston, R. P. Jakob, P. Zihlmann, O. Schwardt, X. Jiang, T. Maier, B. Ernst, *ChemBioChem* **2015**, *16*, 1235.
[16] C. Grabosch, M. Hartmann, J. Schmidt-Lassen, T. K. Lindhorst, *ChemBioChem* **2011**, *12*, 1066.
[17] T. Klein, D. Abgottspon, M. Wittwer, S. Rabbani, J. Herold, X. Jiang, S. Kleeb, C. Lüthi, M. Scharenberg, J. Bezençon et al., *J. Med. Chem.* **2010**, *53*, 8627.
[18] a) 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; b) Z. Han, J. S. Pinkner, B. Ford, R. Obermann, W. Nolan, S. A. Wildman, D. Hobbs, T. Ellenberger, C. K. Cusumano, S. J. Hultgren et al., *J. Med. Chem.* **2010**, *53*, 4779; c) C. K. Cusumano, J. S. Pinkner, Z. Han, S. E. Greene, B. A. Ford, J. R. Crowley, J. P. Henderson, J. W. Janetka, S. J. Hultgren, *Sci. Transl. Med.* **2011**, *3*, 109ra115.
[19] S. Kleeb, L. Pang, K. Mayer, D. Eris, A. Sigl, R. C. Preston, P. Zihlmann, T. Sharpe, R. P. Jakob, D. Abgottspon et al., *J. Med. Chem.* **2015**, *58*, 2221.
[20] L. Pang, S. Kleeb, K. Lemme, S. Rabbani, M. Scharenberg, A. Zalewski, F. Schadler, O. Schwardt, B. Ernst, *ChemMedChem* **2012**, *7*, 1404.
[21] X. Jiang, D. Abgottspon, S. Kleeb, S. Rabbani, M. Scharenberg, M. Wittwer, M. Haug, O. Schwardt, B. Ernst, *J. Med. Chem.* **2012**, *55*, 4700.
[22] O. Schwardt, S. Rabbani, M. Hartmann, D. Abgottspon, M. Wittwer, S. Kleeb, A. Zalewski, M. Smieško, B. Cutting, B. Ernst, *Bioorg. Med. Chem.* **2011**, *19*, 6454.
[23] S. Brument, A. Sivignon, T. I. Dumych, N. Moreau, G. Roos, Y. Guérardel, T. Chalopin, D. Deniaud, R. O. Bilyy, A. Darfeuille-Michaud et al., *J. Med. Chem.* **2013**, *56*, 5395.
[24] a) T. K. Lindhorst, C. Kieburg, U. Krallmann-Wenzel, *Glycoconjugate J.* **1998**, *15*, 605; b) N. Nagahori, R. T. Lee, S.-I. Nishimura, D. Pagé, R. Roy, Y. C. Lee, *ChemBioChem* **2002**, *3*, 836; c) C. C. M. Appeldoorn, J. A. F. Joosten, F. Ait el Maate, U. Dobrindt, J. Hacker, R. M. J. Liskamp, A. S. Khan, R. J. Pieters, *Tetrahedron: Asymmetry* **2005**, *16*, 361; d) A. Patel, T. K. Lindhorst, *J. Org. Chem.* **2001**, *66*, 2674; e) M. Touaibia, A. Wellens, T. C. Shiao, Q. Wang, S. Sirois, J. Bouckaert, R. Roy, *ChemMedChem* **2007**, *2*, 1190; f) M. Durka, K. Buffet, J. Iehl, M. Holler, J.-F. Nierengarten, J. Taganna, J. Bouckaert, S. P. Vincent, *Chem. Commun.* **2011**, *47*, 1321; g) J. Bouckaert, Z. Li, C. Xavier, M. Almant, V. Caveliers, T. Lahoutte, S. D. Weeks, J. Kovensky, S. G. Gouin, *Chem. Eur. J.* **2013**, *19*, 7847.
[25] A. Wellens, M. Lahmann, M. Touaibia, J. Vaucher, S. Oscarson, R. Roy, H. Remaut, J. Bouckaert, *Biochemistry* **2012**, *51*, 4790.
[26] A. Bondi, *J. Phys. Chem.* **1964**, *68*, 441.
[27] J. Vrbancich, G. L. D. Ritchie, *J. Chem. Soc. Faraday Trans. 2* **1980**, *76*, 648.
[28] a) M. R. Battaglia, A. D. Buckingham, J. H. Williams, *Chem. Phys. Lett.* **1981**, *78*, 421; b) C. A. Hunter, K. R. Lawson, J. Perkins, C. J. Urch, *J. Chem. Soc. Perkin Trans. 2* **2001**, 651; c) S. L. Cockroft, C. A. Hunter, K. R. Lawson, J. Perkins, C. J. Urch, *J. Am. Chem. Soc.* **2005**, *127*, 8594; d) L.-J. Riwar, N. Trapp, B. Kuhn, F. Diederich, *Angew. Chem. Int. Ed.* **2017**, *56*, 11252; *Angew. Chem.* **2017**, *129*, 11405.
[29] C. A. Hunter, J. K. M. Sanders, *J. Am. Chem. Soc.* **1990**, *112*, 5525.
[30] F. Cozzi, F. Ponzini, R. Annunziata, M. Cinquini, J. S. Siegel, *Angew. Chem. Int. Ed. Engl.* **1995**, *34*, 1019; *Angew. Chem.* **1995**, *107*, 1092.
[31] a) C.-Y. Kim, J. S. Chang, J. B. Doyon, T. T. Baird, C. A. Fierke, A. Jain, D. W. Christianson, *J. Am. Chem. Soc.* **2000**, *122*, 12125; b) M. Giroud, M. Harder, B. Kuhn, W. Haap, N. Trapp, W. B. Schweizer, T. Schirmeister, F. Diederich, *ChemMedChem* **2016**, *11*, 1042; c) E. A. Meyer, R. K. Castellano, F. Diederich, *Angew. Chem. Int. Ed.* **2003**, *42*, 1210; *Angew. Chem.* **2003**, *115*, 1244; d) L. M. Salonen, M. Ellermann, F. Diederich, *Angew. Chem. Int. Ed.* **2011**, *50*, 4808; *Angew. Chem.* **2011**, *123*, 4908.
[32] C. Hansch, A. Leo, R. W. Taft, *Chem. Rev.* **1991**, *91*, 165.
[33] H.-Q. Do, O. Daugulis, *J. Am. Chem. Soc.* **2007**, *129*, 12404.
[34] M. Prieto, E. Zurita, E. Rosa, L. Muñoz, P. Lloyd-Williams, E. Giralt, *J. Org. Chem.* **2004**, *69*, 6812.
[35] K. Mayer, D. Eris, O. Schwardt, C. P. Sager, S. Rabbani, S. Kleeb, B. Ernst, *J. Med. Chem.* **2017**, *60*, 5646.
[36] a) S. Rabbani, B. Fiege, D. Eris, M. Silbermann, R. P. Jakob, G. Navarra, T. Maier, B. Ernst, *J. Biol. Chem.* **2018**, *293*, 1835; b) V. B. Rodriguez, B. A. Kidd, G. Interlandi, V. Tchesnokova, E. V. Sokurenko, W. E. Thomas, *J. Biol. Chem.* **2013**, *288*, 24128; c) D. I. Kisiela, H. Avagyan, D. Friend, A. Jalan, S. Gupta, G. Interlandi, Y. Liu, V. Tchesnokova, V. B. Rodriguez, J. P. Sumida et al., *PLoS Pathog.* **2015**, *11*, e1004857.
[37] a) E. P. Gillis, K. J. Eastman, M. D. Hill, D. J. Donnelly, N. A. Meanwell, *J. Med. Chem.* **2015**, *58*, 8315; b) S. Swallow, *Prog. Med. Chem.* **2015**, *54*, 65.
[38] J. Alsenz, M. Kansy, *Adv. Drug Delivery Rev.* **2007**, *59*, 546.
[39] J. C. Dearden, G. M. Bresnen, *Quant. Struct. Relationships* **1988**, *7*, 133.
[40] M. Kansy, F. Senner, K. Gubernator, *J. Med. Chem.* **1998**, *41*, 1007.
[41] C. A. Lipinski, *J. Pharmacol. Toxicol. Methods* **2000**, *44*, 235.
[42] M. J. Waring, *Expert Opin. Drug Discovery* **2010**, *5*, 235.
[43] B. Feng, J. L. LaPerle, G. Chang, M. V. S. Varma, *Expert Opin. Drug Metab. Toxicol.* **2010**, *6*, 939.
[44] A. Avdeef, S. Bendels, L. Di, B. Faller, M. Kansy, K. Sugano, Y. Yamauchi, *J. Pharm. Sci.* **2007**, *96*, 2893.
[45] S. Rabbani, X. Jiang, O. Schwardt, B. Ernst, *Anal. Biochem.* **2010**, *407*, 188.
[46] K. A. Rubinson, J. E. Ladner, M. Tordova, G. L. Gilliland, *Acta Crystallogr. Sect. D* **2000**, *56*, 996.
[47] a) W. Kabsch, *Acta Crystallogr. Sect. D* **2010**, *66*, 125; b) W. Kabsch, *Acta Crystallogr. Sect. D* **2010**, *66*, 133.
[48] A. J. McCoy, *Acta Crystallogr. Sect. D* **2007**, *63*, 32.
[49] P. Emsley, K. Cowtan, *Acta Crystallogr. Sect. D* **2004**, *60*, 2126.

[50] P. D. Adams, P. V. Afonine, G. Bunkóczi, V. B. Chen, I. W. Davis, N. Echols, J. J. Headd, L.-W. Hung, G. J. Kapral, R. W. Grosse-Kunstleve et al., *Acta Crystallogr. Sect. D* **2010**, *66*, 213.
[51] D. M. van Aalten, R. Bywater, J. B. Findlay, M. Hendlich, R. W. Hooft, G. Vriend, *J. Comput. Aided Mol. Des.* **1996**, *10*, 255.
[52] V. B. Chen, W. B. Arendall, J. J. Headd, D. A. Keedy, R. M. Immormino, G. J. Kapral, L. W. Murray, J. S. Richardson, D. C. Richardson, *Acta Crystallogr. Sect. D* **2010**, *66*, 12.
[53] 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* **2005**, *59*, 687.

Manuscript received: January 23, 2019
Accepted manuscript online: February 1, 2019
Version of record online: February 22, 2019

RSC Chem. Biol., 2020, *1*, 281-287
doi: 10.1039/D0CB00108B

Publication 7

Enhancing the enthalpic contribution of hydrogen bonds by solvent shielding

Jonathan Cramer,[1)] Xiaohua Jiang,[1)] Wojciech Schönemann,[1)] Marleen Silbermann,[1)] Pascal Zihlmann,[1)] Stefan Siegrist,[1)] Brigitte Fiege,[1)] Roman Peter Jakob,[2)] Said Rabbani,[1)] Timm Maier,[2)] and Beat Ernst[1)*]

[1)] Institute of Molecular Pharmacy, University of Basel, Klingelbergstr. 50, 4056 Basel, Switzerland

[2)] Structural Biology, Biozentrum, University of Basel, Klingelbergstr. 70, 4056 Basel, Switzerland

[*] Corresponding author
Tel.: 0041 61 267 15 51; Fax: 0041 61 207 15 52
E-mail: beat.ernst@unibas.ch

Contributions of Marleen Silbermann:

- Protein expression and purification of $FimH_{LD}$ and $FimH_{FL}$
- Proofreading of the manuscript

RSC Chemical Biology

ROYAL SOCIETY OF CHEMISTRY

PAPER

View Article Online
View Journal | View Issue

Cite this: *RSC Chem. Biol.*, 2020, **1**, 281

Enhancing the enthalpic contribution of hydrogen bonds by solvent shielding†

Jonathan Cramer, [a] Xiaohua Jiang,[a] Wojciech Schönemann,[a] Marleen Silbermann,[a] Pascal Zihlmann,[a] Stefan Siegrist,[a] Brigitte Fiege,[a] Roman Peter Jakob, [b] Said Rabbani,[a] Timm Maier [b] and Beat Ernst *[a]

Received 26th June 2020,
Accepted 19th August 2020

DOI: 10.1039/d0cb00108b

rsc.li/rsc-chembio

In biological systems, polar interactions are heavily burdened by high desolvation penalties resulting from strong solute–solvent interactions. As a consequence thereof, enthalpic contributions of hydrogen bonds to the free energy of binding are severely diminished. However, this effect is strongly attenuated for interactions within solvent-shielded areas of proteins. In microcalorimetric experiments, we show that the bacterial lectin FimH utilizes conformational adaptions to effectively shield its binding site from solvent. The transition into a lower dielectric environment results in an enthalpic benefit of approximately −13 kJ mol^{-1} for mannoside binding. However, this effect can be abrogated, if the hydrogen bond network within the binding site is disturbed by deoxygenation of the ligand. Conformational adaption leading to reduced local dielectric constants could represent a general mechanism for proteins to enable enthalpy-driven recognition of polar ligands.

Introduction

The energetics of hydrogen bonds in biomolecular recognition are highly affected by severe desolvation penalties imposed by strong interactions with the aqueous solvent in the unbound state. A common conception is that strongly directional hydrogen bond interactions provide selectivity, but do not necessarily augment the overall binding affinity of a ligand because favorable contributions are often canceled out by significant desolvation penalties.[1–3] The free energy penalty for the transfer of a hydroxyl group from aqueous solvent to the gas phase was estimated as 26 kJ mol^{-1}, with an associated enthalpy penalty of 36 kJ mol^{-1} and the transfer of an alcohol function from bulk solvent into a hydrophobic pocket was associated with a free energy penalty of 18–21 kJ mol^{-1}.[4,5] Furthermore, in an aqueous system, the contribution of a single hydrogen bond to the free energy of binding has been estimated to equally amount to 18–21 kJ mol^{-1}.[6–10] In summary, a hydrogen bond hardly contributes to the binding process, unless additional conditions are fulfilled. It has been proposed that hydrophobic occlusion and solvent shielding are able to advance the strength of hydrogen bonds in biological systems. Thus, enthalpies of formation for solvent-exposed hydrogen bonds on protein surfaces amount to −6 kJ mol^{-1}, whereas buried interactions contribute with up to −25 kJ mol^{-1}.[11,12] For the arabinose-binding protein, removal of buried hydrogen bond interactions in deoxy D-galactose derivatives was associated with an enthalpy-driven loss in binding free energy of 30 kJ mol^{-1}.[13] Similarly, hydrophobic solvent-shielding has been shown to impact the activity of enzymes, as well as the pK_a of buried titratable groups.[14–18] The exclusion of polar water molecules effectively creates a low dielectric environment that supports electrostatic interactions and accelerates the reaction of charged intermediates. Thus, solvent shielding of hydrophilic binding sites may contribute to enthalpy-driven recognition of highly polar molecules such as carbohydrates. In this study, we aim to demonstrate that proteins actively harness the modulation of local dielectric properties by conformational transitions and to quantify the thermodynamic consequences on polar hydrogen bond interactions.

As a model system to study the impact of the dielectric constant ε_r on the thermodynamics of binding, we employed FimH, a bacterial adhesin located on type I pili of uropathogenic *E. coli*.[19–21] The functional lectin domain of FimH features a carbohydrate recognition site tailor-made to accommodate mannose epitopes. The mannose binding affinity of the lectin domain is allosterically controlled by the adjacent regulatory pilin domain.[22,23] When mechanical forces, such as shear forces during micturition, cause the separation of lectin and pilin domain,[23,24] a switch from the low-affinity conformation to the high-affinity conformation of FimH is induced and mannoside ligands are bound with ~100-fold stronger affinity

[a] *Institute of Molecular Pharmacy, University of Basel, Klingelbergstrasse 50, 4056 Basel, Switzerland. E-mail: beat.ernst@unibas.ch*
[b] *Institute of Structural Biology, University of Basel, Klingelbergstrasse 70, 4056 Basel, Switzerland*

† Electronic supplementary information (ESI) available. See DOI: 10.1039/d0cb00108b

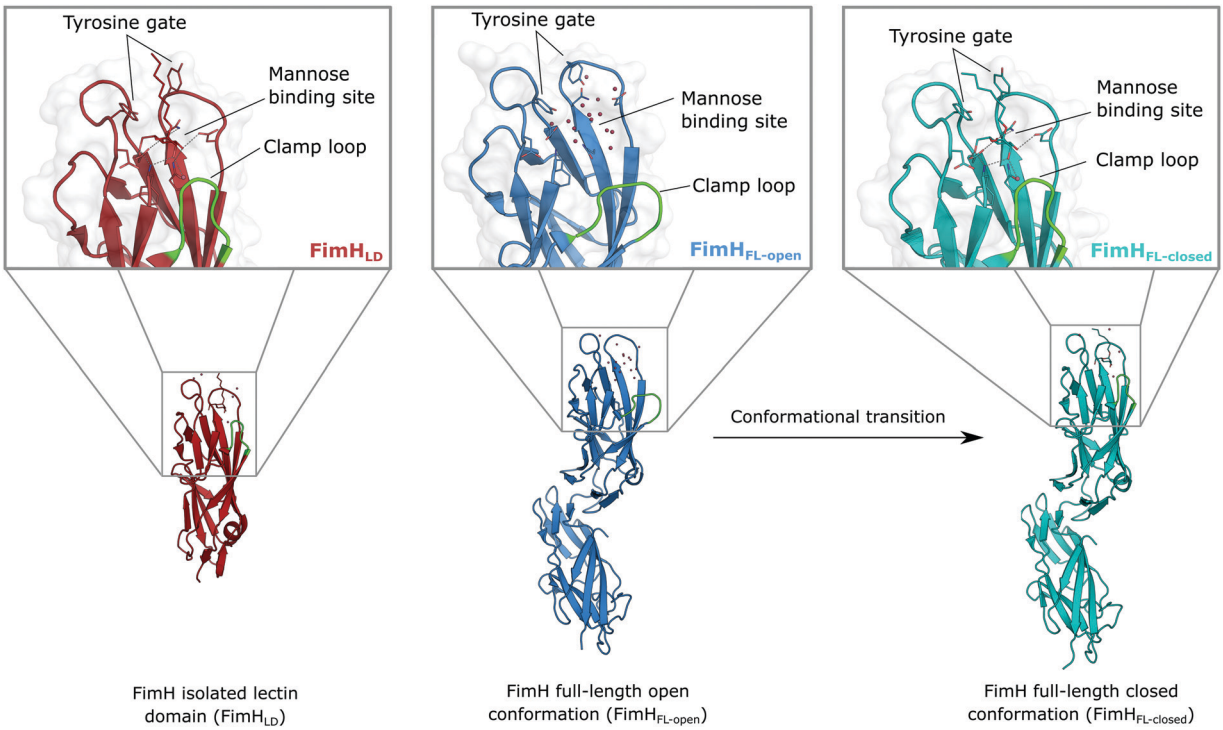


Fig. 1 Molecular structure of the bacterial lectin FimH. The isolated lectin domain (FimH$_{LD}$, PDB 4BUQ) is fairly rigid and pre-organized in the closed conformation even in the absence of a ligand. The recombinant full-length FimH construct (FimH$_{FL}$) undergoes a conformational transition from an open (FimH$_{FL-open}$, PDB 4XOD) to a closed (FimH$_{FL-closed}$, PDB 4XOE) state when a ligand binds to the mannose binding site.

compared to the domain-associated low affinity state. Here, we compare the properties of the two affinity states by means of two different recombinant FimH constructs.

Results and discussion

The isolated FimH lectin domain (FimH$_{LD}$) is a common *in vitro* model for the domain-separated, high-affinity conformation of the protein.[25] FimH$_{LD}$ features a rigid structure and a well-defined carbohydrate binding site (Fig. 1). Mannoside ligands engage in an impressive network of ten hydrogen bonds. In the absence of a ligand, water molecules assume the positions of mannose hydroxyl groups, pre-organizing the binding site for carbohydrate recognition (RMSD$_{apo\text{-}bound}$ = 0.35 Å). The binding event is a single-step mechanism that proceeds without any major conformational adjustment of the residues in the mannose binding site.[26,27] On the other hand, the domain associated full-length FimH construct (FimH$_{FL}$) has been shown to bind mannoside ligands by a more complex mechanism caused by conformational distortions of the protein due to the presence of the regulatory subunit.[22,23] In the absence of a ligand, the protein exists in an ensemble of conformational states that are characterized by a shallow, solvent-exposed binding site, a low affinity to carbohydrate ligands, and a high degree of flexibility (FimH$_{FL-open}$, Fig. 1).[23,28] In the presence of

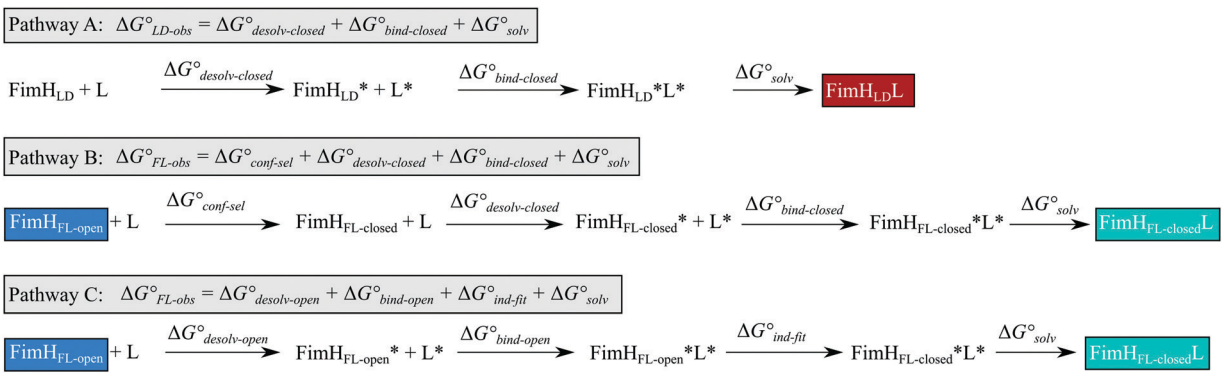


Fig. 2 Thermodynamic model of the interaction of FimH with mannoside ligands. FimH states and complexes are colored analogous to available X-ray structures in Fig. 1. Desolvated states are indicated with an asterisk.

a ligand, the conformational equilibrium is shifted toward a closed state ($FimH_{FL\text{-}closed}$), in which the clamp loop is accommodated close to the carbohydrate binding site and shields the mannose moiety from the solvent. The atomistic features of this mechanism, namely clamp loop mobility in the open and solvent shielding of the binding site in the closed conformation, have been thoroughly characterized in MD simulations.[29]

Although the overall structure of this state differs from the high-affinity conformation, the binding site regions are virtually identical ($RMSD_{lectin\_domain}$ = 2.0 Å, $RMSD_{binding\_site}$ = 0.2 Å). Fig. 2 depicts a thermodynamic model of ligand binding to $FimH_{LD}$ and $FimH_{FL}$. The experimentally observed free energy of binding to $FimH_{LD}$ ($\Delta G^{\circ}_{LD\text{-}obs}$) is composed of terms for desolvation ($\Delta G^{\circ}_{desolv\text{-}closed}$), protein–ligand interaction ($\Delta G^{\circ}_{bind\text{-}closed}$), and solvation of the assembled complex (ΔG°_{solv}) (Fig. 2, pathway A). For $FimH_{FL}$, the conformational rearrangement associated with ligand binding theoretically proceeds *via* a conformational selection (Fig. 2, pathway B) or an induced-fit (Fig. 2, pathway C) mechanism. In case of conformational selection, the experimentally observed thermodynamics ($\Delta G^{\circ}_{FL\text{-}obs}$) contain contributions from a conformational transition to the closed state ($\Delta G^{\circ}_{conf\text{-}sel}$), desolvation ($\Delta G^{\circ}_{desolv\text{-}closed}$), protein–ligand interaction ($\Delta G^{\circ}_{bind\text{-}closed}$), and solvation of the assembled complex (ΔG°_{solv}). In the induced-fit pathway C, $\Delta G^{\circ}_{FL\text{-}obs}$ is composed of desolvation ($\Delta G^{\circ}_{desolv\text{-}open}$), low-affinity binding ($\Delta G^{\circ}_{bind\text{-}open}$), induced-fit transition ($\Delta G^{\circ}_{ind\text{-}fit}$), and solvation of the assembled complex (ΔG°_{solv}). Thus, the difference in the thermodynamic profiles of ligand binding to $FimH_{FL}$ and $FimH_{LD}$ ($\Delta\Delta G^{\circ}_{FL\text{-}LD}$) gives access to an experimental approximation of clamp loop rearrangement and hydrophobic occlusion of the polar mannose binding site. In the conformational selection model (Fig. 2, pathway B), $\Delta G^{\circ}_{conf\text{-}sel}$ is a constant term, and ligand binding affinity scales entirely with binding affinity to the closed state. In the induced-fit model (Fig. 2, pathway C), however, differences between ligands can scale with the affinity of a low-affinity encounter complex ($\Delta G^{\circ}_{bind\text{-}open}$) and the effect of the conformational transition ($\Delta G^{\circ}_{ind\text{-}fit}$).

As a result of an extensive network of hydrogen bonds between the pre-organized binding site and the mannose moiety, the interaction between mannosides **1–4** and $FimH_{LD}$ (high-affinity state) is characterized by a strongly enthalpic contribution (Fig. 3A). The weak entropic penalty has been linked to the rigidification of the tyrosine gate, an ensemble of hydrophobic residues that control access to the mannose binding site and interact with the aglycones of ligands **1–4**.[26,27] An improved π–π-stacking and hydrophobic interaction between aglycone and tyrosine gate further accentuates their thermodynamic profile, as the electrostatic nature of the aglycones gives rise to progressively tighter, enthalpy-driven binding from **4** → **3** → **2** → **1**.

The thermodynamic profiles of mannoside interactions with the $FimH_{FL}$ construct (low-affinity state) show similar characteristics (Fig. 3B). The binding process is equally driven by enthalpy, with a counteracting entropic penalty. However, the magnitude of this enthalpy–entropy compensation is much more substantial. Overall, the binding affinity of the mannosides **1–4** is reduced roughly by 12 kJ mol^{-1}, approximately two orders of magnitude in K_D, compared with the same ligands binding to $FimH_{LD}$. This value represents the experimental approximation of the thermodynamics related to the conformational rearrangement from the open ($FimH_{FL\text{-}open}$) to the closed ($FimH_{FL\text{-}closed}$) state.

The difference in the thermodynamics (Fig. 3C) shows that this fairly constant affinity reduction mainly originates from a

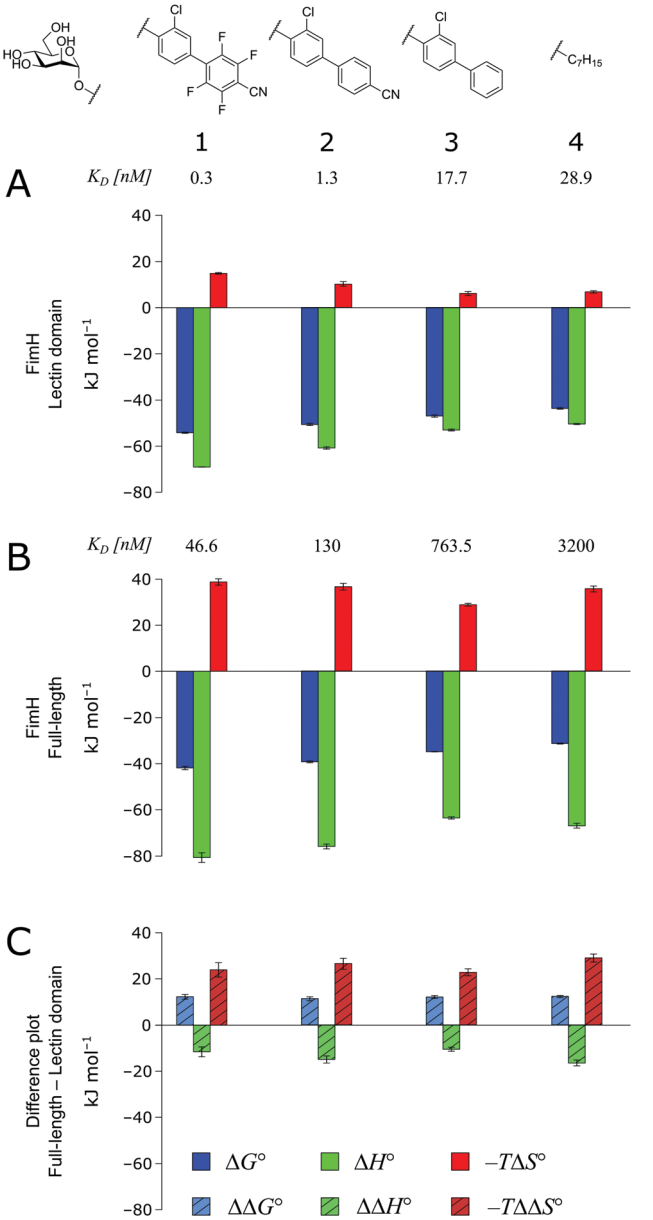


Fig. 3 Thermodynamic profiles (see Fig. S1–S4, ESI†) for the interaction of mannosides **1–4** with $FimH_{LD}$ (A) and $FimH_{FL}$ (B). Data for the interactions $FimH_{LD}$-**2**/**3**/**4** and $FimH_{FL}$-**2**/**4** have been published before.[24,30,31] (C) Difference in the thermodynamic profiles between ligand binding to $FimH_{FL}$ and $FimH_{LD}$. Error bars represent (propagated) 68% confidence intervals. Numeric data and further information can be found in the ESI.†

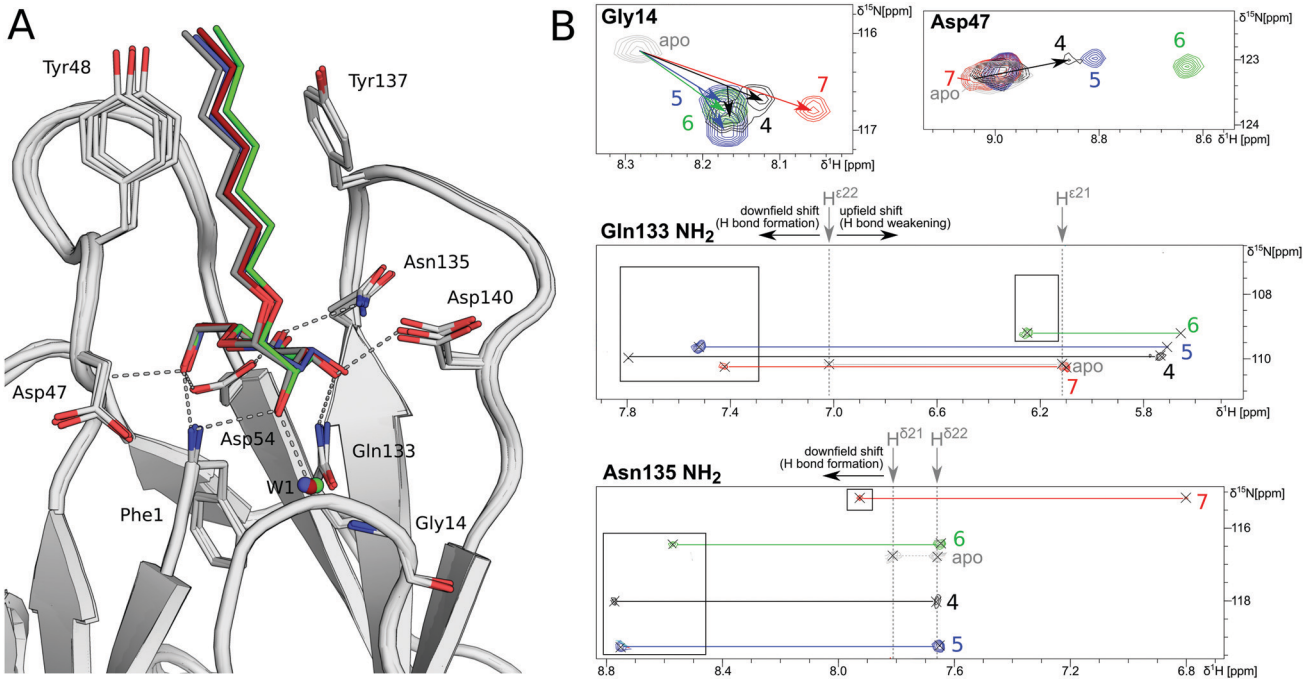


Fig. 4 Structural characterization of interactions between mannoside **4** and the deoxygenated compounds **5–7** and FimHLD. (A) X-ray crystal structures of mannoside **4** (black, PDB 4BUQ), 2-deoxy-mannoside **5** (blue, PDB 5L4T), 3-deoxy-mannoside **6** (green, PDB 5L4V), and 4-deoxy-mannoside **7** (red, PDB 5L4X). Hydrogen bond interactions of **4** are indicated as grey dashes. (B) ^{1}H,^{15}N-HSQC NMR experiments reveal chemical shift perturbations for residues in the mannose binding site in the presence of ligands. Color coding according to panel A. Nomenclature of Asn and Gln δ and ε protons according to PDB guidelines. Full spectrum is shown in Fig. S20 (ESI†).

severe entropy penalty of roughly 26 kJ mol^{-1}, which can be correlated to a shift of the conformational equilibrium of $FimH_{FL}$ toward a closed state, in which the clamp loop is fixed near the mannose moiety. The high mannose binding affinity of $FimH_{LD}$ thus results from pre-organization of the mannose binding site. This concept has been thoroughly characterized in the study of enzyme catalysis.[32,33] Less intuitively, the thermodynamics are also characterized by a favorable enthalpic contribution of approximately −13 kJ mol^{-1}. Since the mannose hydrogen bond networks and any other electrostatic contacts between protein and ligand are identical in the bound states of $FimH_{LD}$ and $FimH_{FL}$, this difference has to originate from the conformational transition of the clamp loop region upon ligand binding. While the local dielectric constant on protein surfaces, as in the open form of $FimH_{FL}$, is commonly estimated to be 20, it is reduced to approximately 5 in the shielded binding site of $FimH_{FL\text{-}closed}$.[17] Based on these simplified qualitative estimations,[34,35] the shift between the open and closed protein states is accompanied by hydrophobic shielding of polar hydrogen bond interactions in the binding site. This leads to reduced local permittivity ε_r in $FimH_{FL\text{-}closed}$, which is associated with an enhanced enthalpic contribution to binding $\left(\Delta H^\circ \propto F_e = \frac{q_1 q_2}{4\pi\varepsilon_0\varepsilon_r r^2}\right)$.

In order to investigate the outcome of a perturbation of the hydrogen bond network, several deoxygenated derivatives of *n*-heptyl α-D-mannoside (**4**) were subjected to thermodynamic and structural analyses. The X-ray crystal structures of **4–7** in complex with $FimH_{LD}$ reveal very little variation (RMSDs ≈ 0.2 Å) upon sequential removal of mannose hydroxyl groups (Fig. 4A), indicating that the loss of hydrogen bonds is not associated with any major conformational adjustments. Next, NMR chemical shift perturbations (CSP) of residues involved in hydrogen bond interactions to mannose hydroxyl groups were studied (Fig. 4B). Gly14, engaged in a water-mediated hydrogen bond to 2-OH, shows a minor CSP upon ligand binding. Variations between different mannosides are negligible, even when the 2-OH group is removed (→**5**). This indicates that the position of the interacting water molecule **W1** is not affected. The backbone amide NH of Asp47 forms a hydrogen bond with 6-OH, leading only to slight CSP upon binding to the ligands **4** and **5** (2-deoxy). In case of the 3-deoxy derivative **6**, the CSP is more pronounced, whereas the signal is superimposed to the apo protein for **7** (4-deoxy). CSP of the amide side chain protons of Gln133 report on hydrogen bond formation with 3-OH. For **4**, the signal of the amide proton (PDB nomenclature $H^{\varepsilon 22}$) is shifted downfield by 0.8 ppm, indicating hydrogen bond formation. A smaller downfield shift of 0.3–0.5 ppm signifies the slightly weaker interaction for **5** and **7**, whereas the absence of the interaction for 3-deoxy derivative **6** is obvious from a strong upfield shift of −1.6 ppm.

Similarly, CSP of the side chain signals of Asn135 indicate the interaction with 4-OH. The signal is shifted downfield for **4**, **5**, and **6**, whereas a relative upfield shift for the 4-deoxy derivative **7** signifies the absence of the interaction. Importantly, none of the observed residues show a marked shift in the

^{15}N-dimension of the spectra, reporting on changes in dihedral angles of protein backbone and side chains.

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. In both experiments, low binding affinity prevented a characterization of compound **8**.

The thermodynamics of **5–7** binding to $FimH_{LD}$ (Fig. 5A) illustrate that removal of the hydroxyl groups in the 2-, 3-, and 4-position of *n*-heptyl α-D-mannoside **4** coincide with a sizeable reduction of the enthalpic contribution to binding.

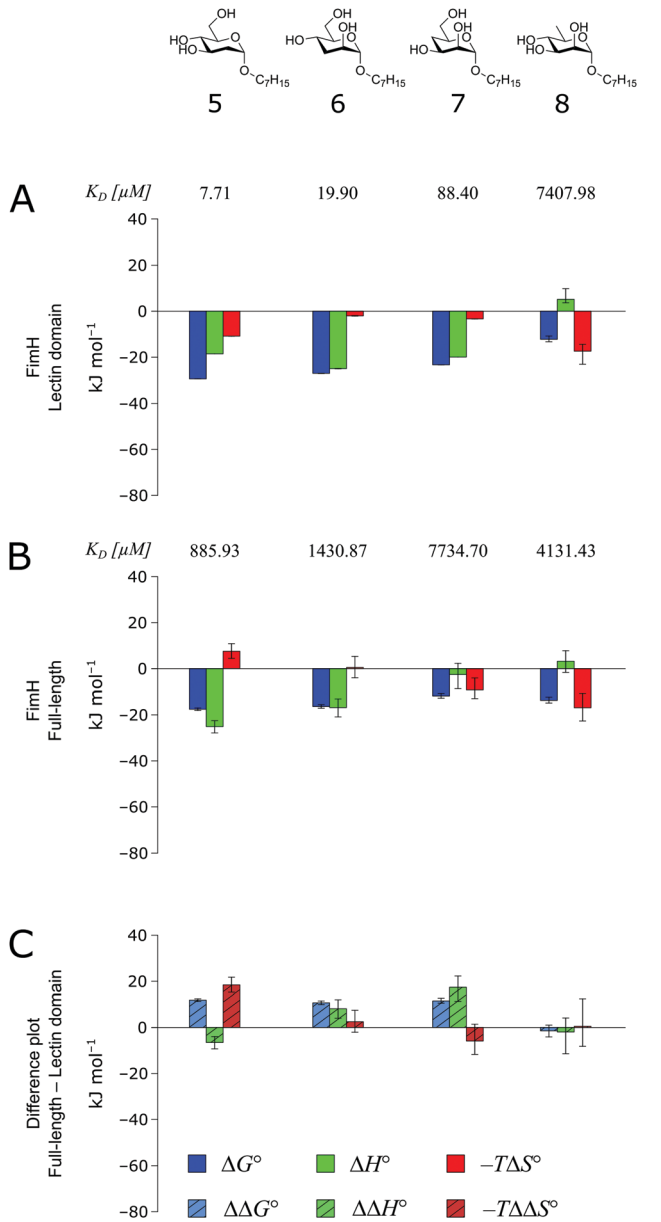


Fig. 5 Thermodynamic profiles (see Fig. S5–S8, ESI†) for the interaction of deoxygenated *n*-heptyl α-D-mannoside derivatives **5–8** with $FimH_{LD}$ (A) and $FimH_{FL}$ (B). Data for the interactions of $FimH_{LD}$ with compounds **5–7** have been published elsewhere.[30] (C) Difference in the thermodynamic profiles between ligand binding to $FimH_{FL}$ and $FimH_{LD}$. Error bars represent (propagated) 68% confidence intervals.

In addition, the negative value for $-T\Delta S^\circ$ indicates an increased degree of disorder, probably as a consequence of the impaired hydrogen bond network. This is in line with the reduced downfield shifts in NMR experiments. The interaction of **5–7** with $FimH_{FL}$ shows an equivalent trend in the thermodynamic profiles (Fig. 5B). Yet, the magnitude of the enthalpy–entropy compensation effect differs substantially. As a consequence, the fairly constant trend we observed in the differential profiles for **1–4** (Fig. 3C) is lost for **5–7** (Fig. 5C). Here, the enthalpic contribution that was attributed to the modulation of binding site permittivity ε_r in $FimH_{FL}$ is considerably diminished and even reversed for **6** and **7**. Instead, a progressive shift toward entropic binding is observed.

When the thermodynamic profiles of **5–7** are compared with the parent compound **4**, notable dissimilarities between the two protein constructs, $FimH_{LD}$ (Fig. 6A) and $FimH_{FL}$ (Fig. 6B) can be observed. For $FimH_{LD}$, the enthalpic penalty for removal of a hydroxyl group amounts to 26–32 kJ mol^{-1}. The same modification in $FimH_{FL}$, however, results in an enthalpic penalty that ranges from 42 kJ mol^{-1} up to 65 kJ mol^{-1}. The discrepancy suggests that the loss of a hydrogen bond in the interaction of **5–7** and $FimH_{FL}$ is superimposed by global

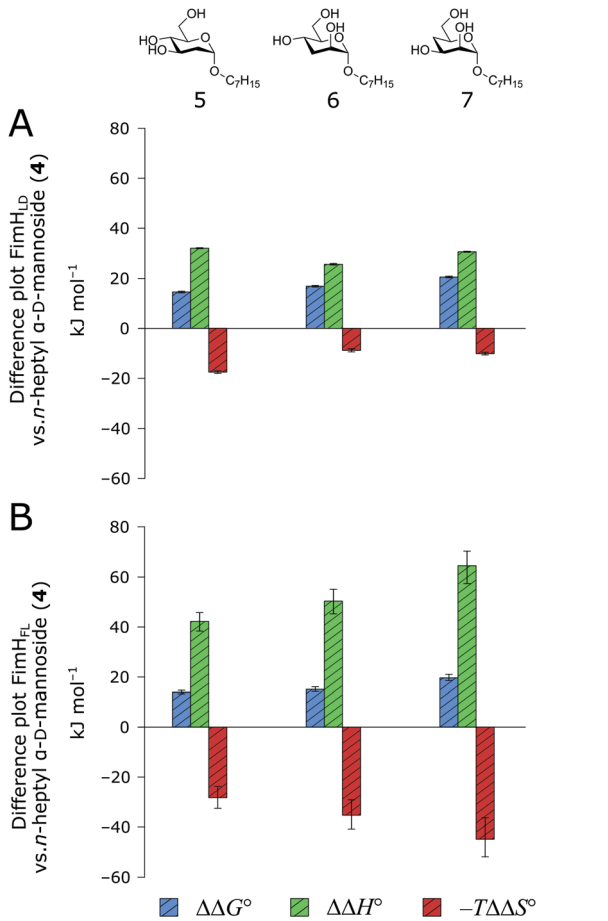


Fig. 6 Difference in thermodynamic profiles between **4** and deoxygenated *n*-heptyl α-D-mannoside derivatives **5–7** with $FimH_{LD}$ (A) or $FimH_{FL}$ (B). Error bars represent propagated 68% confidence intervals.

effects resulting from a modulation of clamp loop dynamics and the associated solvent shielding mechanism.

Thus, the observed thermodynamics are not indicative of the protein–ligand interaction itself, but instead reflect a shift in the conformational equilibrium of the protein. This mechanism has been termed entropy–enthalpy transduction.[36] The existence of distinctive enthalpy–entropy compensation profiles for different ligands might also indicate that binding does not proceed *via* selection of a preexisting $FimH_{FL}$ conformation (Fig. 2, pathway B), but through a ligand-induced fit (Fig. 2, pathway C). The conformational equilibrium in a conformational selection model is an inherent property of the protein and does not depend on the nature of a ligand. Thus, $\Delta G^{\circ}_{\text{conf-sel}}$ must be a constant term, and differences in ligand thermodynamics originate solely from interaction with the binding-competent conformation $\left(\Delta G^{\circ}_{\text{bind-closed}}\right)$. In an induced-fit model, however, formation of an encounter complex $\left(\Delta G^{\circ}_{\text{bind-copen}}\right)$ and contrasting consequences of the conformational transition $\Delta G^{\circ}_{\text{ind-fit}}$ (*e.g.* loss of enthalpic benefit from solvent shielding) can account for differences between ligands. For ligands with an intact mannoside core (**1**–**4**), the induced-fit transition reaches its full effect so that the differential thermodynamic profiles (Fig. 3C) appear constant.

Intriguingly, X-ray and NMR experiments yielded no information about the binding mode of **8** to $FimH_{LD}$. ITC data suggests that this compound actually binds with very low, but similar affinity to the high- and low-affinity construct of the protein (Fig. 5). These observations could signify a different binding mode that does not proceed *via* a conformational rearrangement in $FimH_{FL}$.

Conclusions

In conclusion, we have investigated the binding thermodynamics of two FimH constructs that share a similar binding mode but explore a very different conformational landscape. We could demonstrate that an observed enthalpic contribution of roughly −13 kJ mol^{-1} for mannoside ligands is a direct consequence of solvent shielding enabled by the rearrangement of the clamp loop in $FimH_{FL}$ and the associated modulation of local dielectric properties. When the hydrogen bond network in the binding site is disrupted, the enthalpic benefit is completely abolished (Fig. 5C). In the case of $FimH_{FL}$, the effect of the solvent shielding mechanism is superimposed by an entropic penalty for the arrest of the clamp loop dynamics.

A similar transition of the binding site into a lower dielectric environment was observed for a number of other carbohydrate binding proteins, such as arabinose-binding protein (ABP), glucose/galactose-binding protein (GGBP) und sialic acid-binding periplasmic protein (SiaP). Their conformational transition upon ligand binding involves a twist of rigid protein domains around a fixed axis (''hinge-bending'' motions).[37–40] These transitions do not alter the flexibility of individual protein domains to an extent as observed in FimH and can be considered entropically neutral. Thus, modulation of binding site permittivity could represent a general mechanism for carbohydrate binding proteins to enable enthalpy-driven recognition of polar ligands in aqueous solution.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

B. E. gratefully acknowledges funding from the Swiss National Science Foundation (PZ) and German Academic Exchange Service (BF).

Notes and references

1 H.-J. Böhm and G. Klebe, *Angew. Chem., Int. Ed. Engl.*, 1996, **35**, 2588–2614.
2 R. E. Babine and S. L. Bender, *Chem. Rev.*, 1997, **97**, 1359–1472.
3 G. M. Keserü and D. C. Swinney, *Thermodynamics and Kinetics of Drug Binding*, Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, Germany, 2015.
4 S. Cabani, P. Gianni, V. Mollica and L. Lepori, *J. Solution Chem.*, 1981, **10**, 563–595.
5 E. Barratt, A. Bronowska, J. Vondrášek, J. Černý, R. Bingham, S. Phillips and S. W. Homans, *J. Mol. Biol.*, 2006, **362**, 994–1003.
6 G. A. Jeffrey, *An Introduction to Hydrogen Bonding*, Oxford University Press, Oxford, 1997.
7 T. Steiner, *Angew. Chem., Int. Ed.*, 2002, **41**, 48–76.
8 C. P. Sager, D. Eriş, M. Smieško, R. Hevey and B. Ernst, *Beilstein J. Org. Chem.*, 2017, **13**, 2584–2595.
9 A. Vedani and J. D. Dunitz, *J. Am. Chem. Soc.*, 1985, **107**, 7653–7658.
10 J. Cramer, C. P. Sager and B. Ernst, *J. Med. Chem.*, 2019, **62**, 8915–8930.
11 A. V. Efimov and E. V. Brazhnikov, *FEBS Lett.*, 2003, **554**, 389–393.
12 G. Némethy, I. Z. Steinberg and H. A. Scheraga, *Biopolymers*, 1963, **1**, 43–69.
13 A. H. Daranas, H. Shimizu and S. W. Homans, *J. Am. Chem. Soc.*, 2004, **126**, 11870–11876.
14 S.-O. Shan and D. Herschlag, *Proc. Natl. Acad. Sci. U. S. A.*, 1996, **93**, 14474–14479.
15 M. M. Malabanan, T. L. Amyes and J. P. Richard, *Curr. Opin. Struct. Biol.*, 2010, **20**, 702–710.
16 D. G. Isom, C. A. Castañeda, B. R. Cannon, P. D. Velu and E. B. García-Moreno, *Proc. Natl. Acad. Sci. U. S. A.*, 2010, **107**, 16096–16100.
17 C. A. Fitch, D. A. Karp, K. K. Lee, W. E. Stites, E. E. Lattman and E. B. García-Moreno, *Biophys. J.*, 2002, **82**, 3289–3304.
18 E. L. Mehler, M. Fuxreiter, I. Simon and E. B. Garcia-Moreno, *Proteins: Struct., Funct., Genet.*, 2002, **48**, 283–292.
19 M. A. Mulvey, J. D. Schilling, J. J. Martinez and S. J. Hultgren, *Proc. Natl. Acad. Sci. U. S. A.*, 2000, **97**, 8829–8835.

20 J. D. Schilling, M. A. Mulvey and S. J. Hultgren, *J. Infect. Dis.*, 2001, **183**, S36–S40.
21 C. H. Jones, J. S. Pinkner, R. Roth, J. Heuser, A. V. Nicholes, S. N. Abraham and S. J. Hultgren, *Proc. Natl. Acad. Sci. U. S. A.*, 1995, **92**, 2081–2085.
22 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 and W. E. Thomas, *Cell*, 2010, **141**, 645–655.
23 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.
24 K. Mayer, D. Eris, O. Schwardt, C. P. Sager, S. Rabbani, S. Kleeb and B. Ernst, *J. Med. Chem.*, 2017, **60**, 5646–5662.
25 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.*, 2004, **55**, 441–455.
26 B. Fiege, S. Rabbani, R. C. Preston, R. P. Jakob, P. Zihlmann, O. Schwardt, X. Jiang, T. Maier and B. Ernst, *ChemBioChem*, 2015, **16**, 1235–1246.
27 W. Schönemann, J. Cramer, T. Mühlethaler, B. Fiege, M. Silbermann, S. Rabbani, P. Dätwyler, P. Zihlmann, R. P. Jakob, C. P. Sager, M. Smieško, O. Schwardt, T. Maier and B. Ernst, *ChemMedChem*, 2019, **14**, 749–757.
28 V. Kalas, J. S. Pinkner, T. J. Hannan, M. E. Hibbing, K. W. Dodson, A. S. Holehouse, H. Zhang, N. H. Tolia, M. L. Gross, R. V. Pappu, J. Janetka and S. J. Hultgren, *Sci. Adv.*, 2017, **3**, e1601944.
29 G. Interlandi and W. E. Thomas, *Proteins: Struct., Funct., Bioinf.*, 2016, **84**, 990–1008.
30 P. Zihlmann, M. Silbermann, T. Sharpe, X. Jiang, T. Mühlethaler, R. P. Jakob, S. Rabbani, C. P. Sager, P. Frei, L. Pang, T. Maier and B. Ernst, *Chem. – Eur. J.*, 2018, **24**, 13049–13057.
31 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. Lüdin, J. Bezençon, O. Schwardt, T. Maier and B. Ernst, *J. Med. Chem.*, 2015, **58**, 2221–2239.
32 J. Lameira, R. P. Bora, Z. T. Chu and A. Warshel, *Proteins: Struct., Funct., Bioinf.*, 2015, **83**, 318–330.
33 G. Jindal and A. Warshel, *Proteins: Struct., Funct., Bioinf.*, 2017, **85**, 2157–2161.
34 C. N. Schutz and A. Warshel, *Proteins: Struct., Funct., Genet.*, 2001, **44**, 400–417.
35 A. Warshel, P. K. Sharma, M. Kato and W. W. Parson, *Biochim. Biophys. Acta, Proteins Proteomics*, 2006, **1764**, 1647–1676.
36 A. T. Fenley, H. S. Muddana and M. K. Gilson, *Proc. Natl. Acad. Sci. U. S. A.*, 2012, **109**, 20006–20011.
37 M. J. Borrok, L. L. Kiessling and K. T. Forest, *Protein Sci.*, 2007, **16**, 1032–1041.
38 L. Unione, G. Ortega, A. Mallagaray, F. Corzana, J. Pérez-Castells, A. Canales, J. Jiménez-Barbero and O. Millet, *ACS Chem. Biol.*, 2016, **11**, 2149–2157.
39 A. Müller, E. Severi, C. Mulligan, A. G. Watts, D. J. Kelly, K. S. Wilson, A. J. Wilkinson and G. H. Thomas, *J. Biol. Chem.*, 2006, **281**, 22212–22222.
40 J. F. Darby, A. P. Hopkins, S. Shimizu, S. M. Roberts, J. A. Brannigan, J. P. Turkenburg, G. H. Thomas, R. E. Hubbard and M. Fischer, *J. Am. Chem. Soc.*, 2019, **141**, 15818–15826.

Section III
Curriculum Vitae

Marleen Silbermann • Curriculum Vitae

PERSONAL DETAILS

Address:	Magnolienpark 22, 4052 Basel
Mobile:	+41 79 857 13 57
E-Mail:	marleen.silbermann@gmail.com
Date and place of birth:	7th April 1986, Worms (Germany)
LinkedIn:	https://www.linkedin.com/in/marleen-silbermann-173982a6/

PROFESSIONAL EXPERIENCE

Since Aug. 2019 — **Consultant (Compliance)** at **Arcondis**, Reinach

- Coordination of a qualification project for a biotech company
- Preparation / performance of risk analysis, URS, DQ, IQ, OQ, PQ
- Giving advice in / improving the qualification concept
- Representation of the qualification department in weekly PM meetings
- Supervision and mentoring of junior consultants

EDUCATION

Mar. 2014 – Mar. 2018 — **PhD student** at the **Institute of Molecular Pharmacy**, **University of Basel**

- Kinetic and thermodynamic evaluation of the bacterial lectin FimH

June 2016 – Feb. 2017 — **CAS Pharmaceuticals – From Research to Market, ETH Zurich**

Completed modules: Quality and GMP, Clinical Development and Pharmaceutical Development and Production

Oct. 2008 – Feb. 2013 — **Masters Student** of **Molecular Biotechnology**, **University of Heidelberg**

Main subject area: Drug design

Master's thesis, COS (Dept. Molecular Biology of Plants), Heidelberg

- Regulatory impact of precursor availability on Cys biosynthesis in *A. thaliana*

Oct. 2005 – Aug. 2008 — **Bachelors Student** of **Molecular Biotechnology**, **University of Heidelberg**

Main subject areas: Drug design, biophysical chemistry, bioinformatics

Bachelor's thesis, Merck KGaA, (Dept. Molecular Toxicology), Darmstadt

- Hybridization of antibodies on formalin-fixed paraffin embedded sections

Sept. 1996 – June 2005 — Ludwig-Frank-Gymnasium, Mannheim

- Prize: Best student in Arts (2003 – 2005)

PRACTICAL EXPERIENCES AND STUDY-RELATED JOBS

Research assistantships

- Centre for Organismal Studies (Dept. Molecular Biology of Plants), 2013, Heidelberg (Germany)
- University Hospital Heidelberg, (Inst. Anatomy and Cell Biology), 2011, Heidelberg (Germany)

Academic internships

- Centre for Organismal Studies (Dept. Molecular Biology of Plants), 2012, Heidelberg (Germany)
- German Cancer Research Center (Dept. Biophysics of Macromolecules), 2009, Heidelberg (Germany)
- Max Planck Institute for Medical Research (Dept. Molecular Neurology), 2009, Heidelberg (Germany)
- University of Calgary (Cell and Developmental Biology), 2008, Calgary (Canada)
- Centro de Invest. Biomédica de Occidente (Dept. Molecular Medicine), 2008, Guadalajara (Mexico)
- German Cancer Research Center (Dept. Functional Genome Analysis), 2005, Heidelberg (Germany)

Industrial internships

- Merck KGaA (Inst. Toxicology), 2007, Darmstadt (Germany)
- ORPEGEN Pharma, 2003, Heidelberg (Germany)

PRACTICAL SKILLS

Protein analytics / characterization	Spectroscopy (UV/VIS, CD), chromatographic methods (HPLC: *RP*, *SEC*, *CEX*, *AEX*), electrophoretic methods (SDS-PAGE, CEC), immunological methods (ELISA, Western Blot), microcalorimetry (ITC, DSF), protein-ligand interactions (SPR, FP Assay), microscopy (confocal microscopy, AFM, TIRFM), thermal stability (DSF), enzyme activity assay
Protein Purification	Chromatographic methods (FPLC: SEC, CEX, AEX, affinity), dialysis
Nucleic acid analytics	Polymerase chain reaction (PCR, real-time PCR), electrophoretic methods (agarose gel electrophoresis), cDNA microarrays
Molecular genetics:	Cloning, mRNA isolation, cDNA synthesis, quantitative real-time PCR, hybridization and evaluation of cDNA microarrays
Cell- and tissue-based work:	FACS, MACS, handling of mammalian cells and FFPE-tissue
Teaching:	Supervision of practical courses for undergraduate pharmacy students

COMPUTING SKILLS

Operating systems	Microsoft Windows (advanced)
Software	Microsoft Office (very good), Biacore T200 Software (very good), Empower Pro (advanced), Chromeleon (advanced), NITPIC (advanced), SEDPHAT (advanced), GraphPad Prism (advanced), Photoshop (advanced)

LANGUAGE SKILLS

German (native), English (fluent), French (basic)

COURSES AND CONFERENCES

- QA/GMP representative in the pharmaceutical industry (Concept Heidelberg)
- Quality management in the pharmaceutical industry, lecture (University Basel)
- Good Clinical Practice (GCP) Training for Investigators and Study Teams (Swiss TPH, Basel)
- Career and Mentoring Program: Antelope@Novartis (Novartis, Basel)
- Time Management; Project Management for Researches; Project Management II (Graduate Center and Advanced Studies University of Basel)
- Conflict Management (Graduate Center, University of Basel)
- Leadership Practice (Advanced Studies, University of Basel)
- Logic and argumentation (Institute for Argumentation Competence, Berlin)
- Turning ideas into innovations - The role of intellectual property rights in life sciences (Pharmazentrum, Basel)
- Conference - Biophysics in Drug Discovery (NovAliX, Strasbourg)
- Practical Course: Biophysical Methods (Biozentrum, Basel)
- 11th Swiss Course on Medicinal Chemistry (Swiss Chemical Society, Leysin)
- Workshop: Rhetoric and Communication (Richter, Heidelberg)

SOCIAL RESPONSIBILITY

2010 – 2014; since May 2017	Volunteer at the animal shelter Viernheim; animal protection beider Basel
Sept. 2002 – June 2003	Remedial teaching in German

INTERESTS

Mountaineering, drawing, meditation

Basel, February 2021

PUBLICATIONS AND ARTICLES

Cramer J, Jiang X, Schönemann W, Silbermann M, Zihlmann P, Siegrist S, Fiege B, Jakob R P, Rabbani S, Maier T and Ernst B **(2020)** Enhancing the enthalpic contribution of hydrogen bonds by solvent shielding. *RSC Chemical Biology*, *1*, 281-287.

Schönemann W, Cramer J, T, Fiege B, Silbermann M, Rabbani S, Dätwyler P, Zihlmann P, Jakob R P, Sager C P, Smiesko M, Schwardt O, Maier T und Ernst B **(2019)** Improvement of Aglycone π-Stacking Yields Nano- to Subnanomolar FimH Antagonists, *ChemMedChem*. Wiley, 14(7), S. 749-757. doi: 10.1002/cmdc.201900051.

Frei P, Silbermann M, Mühlethaler T, Jiang X, Schwardt O, Rachel H, und Ernst B **(2019)** Comparison of Affinity Ranking by Target-directed Dynamic Combinatorial Chemistry and Surface Plasmon Resonance, *Arkivoc*, part iv, S. 143-167.

Zihlmann P*****, Silbermann M*****, Sharpe T, Jiang X, Mühlethaler T, Jakob RP, Rabbani S, Sager CP, Frei P, Maier T, and Ernst B **(2018)** KinITC – One Method Supports both Thermodynamic and Kinetic SARs as Exemplified on FimH Antagonists. *Chemistry, 24*(49), 13049-13057. ***Co-First authorship**

Speiser A, Silbermann M, Dong Y, Haberland S, Uslu VV, Wang S, Bangash SAK, Reichelt M, Meyer AJ, Wirtz M, and Hell R **(2018)** Sulfur Partitioning between Glutathione and Protein Synthesis Determines Plant Growth. *Plant Physiology*, *177*(3), 927-937.

Rabbani S, Fiege B, Eris D, Silbermann M, Jakob RP, Navarra G, Maier T, and Ernst B **(2018)** Conformational Switch of the Bacterial Adhesin FimH in the Absence of the Regulatory Domain - Engineering a Minimalistic Allosteric System. *The Journal of Biological Chemistry*, *293*, 1835-1849.

Dong Y., Silbermann M, Speiser A, Forieri I, Linster E, Poschet G, Samami AA, Wanatabe M, Sticht C, Teleman AA, Deragon J-M, Saito K, Hell R, and Wirtz M **(2017)** Sulfur Availability Regulates Plant Growth via Glucose-TOR Signaling. *Nature Communications*, *8*(1), 1174.

Frei P, Pang L, Silbermann M, Eriş D, Mühlethaler T, Schwardt O, and Ernst B **(2017)** Target-directed Dynamic Combinatorial Chemistry: A Study on Potentials and Pitfalls as Exemplified on a Bacterial Target. *Chemistry*, *23*(48), 11570-11577.

Frei P, Navarra G, Sager C, Silbermann M, Varga N, and Wamhoff E-C **(2015)** Drug Discovery Summit: 11th Swiss Course on Medicinal Chemistry. *ChemMedChem*, *10*(3), 575-579.

Silbermann M and Panic G **(2015)** antelope@novartis - connecting female PhD students with industry. https://www.linkedin.com/pulse/antelopenovartis-connecting-female-phd-students-marleen-silbermann/

Winheim S, Hieb AR, Silbermann M, Surmann E-M, Wedig T, Hermann H, Langowski J, and Mücke N **(2011)** Deconstructing the Late Phase of Vimentin Assembly by Total Internal Reflection Fluorescence Microscopy (TIRFM). *PLoS ONE*, *6*(4), e19202.

Perotti C, Wiedl T, Florin L, Reuter H, Moffat S, Silbermann M, Hahn M, Angel P, and Shemanko CS **(2009)** Characterization of Mammary Epithelial Cell Line HC11 using the NIA 15k Gene Array Reveals Potential Regulators of the Undifferentiated and Differentiated Phenotypes. *Differentiation*, *78*(5), 269-82.