Oral Bioavailability of Carbohydrate Mimetics

An *in vitro* and *in vivo* Evaluation

Inauguraldissertation

zur

Erlangung der Würde eines Doktors der Philosophie

vorgelegt der

Philosophisch-Naturwissenschaftlichen Fakultät

der Universität Basel

von

Philipp Dätwyler

aus Staffelbach, AG, Schweiz

Basel, 2020

Genehmigt von der Philosophisch-Naturwissenschaftlichen Fakultät der Universität Basel auf Antrag von

Prof. em. Dr. Beat Ernst, Institut für Molekulare Pharmazie, Universität Basel
Dr. Edmund Hoppe, Grünenthal Gmbh, Aachen, Deutschland

Basel den 27 März 2018

Prof. Dr. Martin Spiess
Dekan

Preface

Philipp Dätwyler performed the work described in this thesis from January 2014 to February 2018 under the supervision of Prof. Dr. Beat Ernst at the Institute of Molecular Pharmacy at the University of Basel. Additionally, work conducted during the master thesis of Philipp Dätwyler under the supervision of Dr. Wojciech Schönemann and Dr. Simon Kleeb is implemented. The work presented in this thesis has been published in peer-reviewed journals, termed as publications or are not yet submitted, marked as manuscripts. The specific contribution of Philipp Dätwyler is mentioned individually before publications and manuscripts:

Manuscript 1: P. Dätwyler, X. Jiang, N. Varga, B. Wagner, T. Mühlethaler, K. Hostettler, B. Ernst. E-selectin antagonist prodrugs with enhanced pharmacokinetic profile.

Manuscript 2: P. Dätwyler, B. Wagner, X. Jiang, T. Mühlethaler, B. Ernst. The potential for oral bioavailability of amidic E-selectin antagonists.

Manuscript 3: N. Varga, X. Jiang, T. Mühletaler, B. Wagner, M. Smiesko, R. Jakob, **P. Dätwyler**, P. Zihlmann, T. Maier, O. Schwardt, B. Ernst. Getting in Shape for Binding – Rational Design of a Non-Conventional Hydrogen Bond in Sialyl Lewisx-Mimetics.

Manuscript 4: C. P. Sager, P. Zihlmann, **P. Dätwyler**, T. Mühlethaler, X. Jiang, J. Cramer, M. Smieško, B. Ernst. Bridging oligosaccharides and glycomimetics with protein receptors — Acca-Bruca, an in-silico tool to categorize carbohydrate–lectin hydrogen bond interactions.

Manuscript 5: T. Mühlethaler, P. Zihlmann, N. Varga, R. P. Jakob, M. Smieško, **P. Dätwyler**, B. Wagner, X. Jiang, T.Maier and B. Ernst. Towards a Nanomolar E-selectin Antagonist: Structures, Thermodynamics and Group Contributions.

Manuscript 6: W. Schönemann, T. Mühlethaler, **P. Dätwyler**, P. Zihlmann, R. P. Jakob, B. Fiege, C. P. Sager, M. Smieško, S. Rabbani, D. Eris, O. Schwardt, T. Maier, B. Ernst. Tyrosine-perfluoroarene interactions lead to picomolar FimH antagonists.

Publication 1: W. Schönemann, S. Kleeb, **P. Dätwyler**, O. Schwardt, B. Ernst. Prodrugability of Carbohydrates – Oral FimH Antagonists. *Can. J. Chem.* **2016**; 94, 909–919

Publication 2: T. Guo, **P. Dätwyler**, E. Demina, M. R. Richards, P. Ge, C. Zou, R. Zheng, A. Fougerat, A. V. Pshezhetsky, B. Ernst, and C. W. Cairo. Selective Inhibitors of Human Neuraminidase 3. *J. Med. Chem.* **2018**. [Epub ahead of print]

Acknowledgements

I would like to thank Prof. Dr. Beat Ernst for letting me be part of this group. He guided me through these years with great enthusiasm, always exploring new ideas and thoughts. I am grateful that he shared his knowledge and experience with me and advised me with the right inputs, but still giving me the freedom to explore my own ideas and to collaborate with other scientists. The interdisciplinarity of the group was a great opportunity to have insights in all important fields of early drug development.

Furthermore, I would like to thank Dr. Edmund Hoppe for his helpful and precise inputs and being my co-referee in the thesis committee.

A special thank you goes to Dr. Simon Kleeb and Dr. Jacqueline Bezençon for introducing me into the field of *in vitro* pharmacokinetic analysis and their help, whenever I had an open question. With great passion and patience, they helped me with the LC-MS, technicalities and analysis.

I further want to thank all current and former members of the team, especially Deniz, Tobias, Pascal, Christoph, Anja, Butrint, Jonathan, Ruben, Rachel, Priska, Marleen, Delphine. The group had a great atmosphere going beyond the work place, which I highly appreciated. In particular, I would like to thank all chemists in our group for letting me test their compounds in our assays, namely Norbert, Bea, Xiaohua, Wojciech, Priska and Blijke.

Many thanks go to the former master students Roman Koch and Katja Hostettler, as well as the internship student Fee Gölitz, for their work and patience, generating additional ideas and results implemented in this thesis.

I would also like to thank Claudia Huber and Marina Kuhn for their support, whenever technical questions or administrative work occurred.

Finally, I would like to thank my family and friends. My parents Heidi and Martin and my brothers Alex and Stefan were very supportive and helped me wherever they could. Especially, I would like to thank my girlfriend Carmela for all her love, joy and support she gave me during this time. Altogether, you all helped me making this thesis not only a scientific, but also a lifetime experience.

Abstract

Carbohydrates play a crucial role in metabolism, cell recognition, cell differentiation, and adhesion processes. Therefore, these molecules represent a potent source for the development of new treatments against many different diseases with an unmet medical need. However, compounds of this class have major inherent drawbacks related to their chemical structure. Carbohydrates are complex and hydrophilic structures with a large polar surface area, which commonly results in poor pharmacokinetic (PK) properties including low oral bioavailability and short plasma half-life due to fast renal clearance. To overcome the poor PK properties of the natural ligands, specific structural modifications have to be implemented in the drug development process from the beginning.

Improvements of newly synthesized compounds have to be constantly monitored with *in vitro* and *in vivo* PK measurements, allowing a direct feedback for further structural modifications. To this purpose, a PADMET platform (**P**hysicochemical properties, **A**bsorption, **D**istribution, **M**etabolism, **E**limination, **T**oxicity) of *in vitro* assays, addressing different aspects that influence the PK properties of a molecule was developed and optimized. The intestinal absorption is a major hurdle to achieve sufficient oral bioavailability of carbohydrate mimetics and therefore,the focus of this work is set on the permeability of carbohydrate mimetic by passive permeation or active transport.

In this thesis, three different targets for the development of potent lead structures starting from natural carbohydrates are discussed:

- E-selectin is a lectin expressed on endothelial cells upon an inflammatory stimulus and is crucial for the recruitment of leukocytes to the side of inflammation. Therefore, E-selectin has been recognized as a potent target for the treatment of various diseases with an inflammatory component. The carbohydrate epitope recognized by E-selectin is the tetrasaccharide sialyl Lewis^x (sLe^x). For the treatment of chronic inflammatory diseases, an oral administration is of interest. However, the development of an orally bioavailable E-selectin antagonist from sLe^x is challenging due to its chemical properties. To overcome the hurdle of insufficient intestinal absorption, an ester prodrug strategy and a bioisosteric replacement were followed and further evaluated by *in vivo* PK studies in mice. The ester prodrug approach resulted in insufficient oral bioavailability

but improvement of the apparent plasma half-life, whereas the bioisosteric approach lead to the first orally bioavailable E-selectin antagonist.

- Uropathogenic *Eschericha coli* (UPEC) are the main cause of urinary tract infections (UTI). UPEC are expressing the virulence factor FimH on the distal tip of type 1 fimbriae, which binds to mannosides on the luminal surface of the bladder to prevent bacteria from being washed out by urine flow. FimH is therefore a promising target for an antiadhesive treatment of UTIs in order to replace current antibiotic treatment strategies. In this work, known biphenyl-α-D-manno-pyranosides were further developed in terms of affinity and *in vitro* PK properties.
- Sialic acids bound or in free circulation are regulated by sialyltransferases and neuraminidases (NEU). Selective inhibitors for the human neuraminidase NEU3 is of interest to study the physiological and pathophysiological role of neuraminidases and further evaluate the potential to develop therapeutics. Here, the development of specific NEU3 inhibitors, as well as the attempt to optimize their PK properties is reported.

Table of contents

Abbreviations

ABC	ATP binding cassette
AIDS	acquired immunodeficiency syndrome
ATP	adenosine-5'-triphosphate
AUC	area under the curve
BDDCS	Biopharmaceutics Drug Distribution and Classification System
BNPP	Bis(4-nitrophenyl) phosphate
C_0	initial concentration at time point zero
Caco-2	colon carcinoma cell line 2
CAM	cell adhesion molecule
CAUTI	catheter induced urinary tract infection
CES	human carboxylesterase
clog*P*	calculated octanol water partition coefficient
CMP-NANA	CMP-N-acetylneuraminic acid
CRD	carbohydrate recognition domain
CYP450	cytochrome P450
D	dose
Da	dalton
DAMP	damage associated molecular pattern molecule
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
e.g.	*exempli gratia,* for example
EGF	epidermal growth factor like domain
ExPEC	extraintestinal pathogenic *Escherichia coli*
F	oral bioavailability
FMO	flavin monooxygenases
Fuc	fucose
g	gram
GABA	gamma-aminobutyric acid
Gal	galactose
Glc*N*Ac	glucosamine
HAS	human serum albumin
HCl	hydrochloric acid
HP-LC	high-pressure liquid chromatography
i.e.	*id est*
IBC	intracellular bacterial community
ICD	intracellular domain
IL	interleukin
Ile	isoleucine
IV	intravenous administration
KDN	deaminoneuraminic acid
L	liter
LC	liquid chromatography
log$D_{7.4}$	octanol water distribution coefficient at pH 7.4
log*P*	octanol water partition coefficient
M	molar

MTT	3-(4, 5-dimethylthiazolyl-2)-2, 5-diphenyltetrazolium bromide
MS	mass spectrometry
MW	molecular weight
NADPH	nicotinamide adenine dinucleotide phosphate
NEU	human neuraminidases
Neu*N*Ac	*N*-acetyl-neuraminic acid
Neu*N*Gc	*N*-glycolyl-neuraminic acid
NFkB	nuclear factor 'kappa-light-chain-enhancer' of activated B-cells
NMR	nuclear magnetic resonance
NRTI	nucleotide reverse-transcriptase inhibitors
OATP	organic anion transporting polypeptide
p	statistical significance of zero hypothesis
P-gp	P-glycoprotein
PADMET	physicochemical properties, absorption, distribution, metabolism, elimination, toxicity
PAMP	pathogen-associated molecular pattern molecule
PAMPA	parallel artificial membrane permeation assay
P_{app}	apparent permeability
PAPS	3'-phosphoadenosine 5'-phospho-sulfate
PD	pharmacodynamics
PDB	protein data bank
P_e	effective permeation
PEPT1	peptide transporter 1
PK	pharmacokinetics
pK_a	acid dissociation constant
PO	*per os,* oral dosage
PPB	plasma protein binding
PPR	pattern recognition receptor
PSA	polar surface area
RLM	rat liver microsomes
S9	supernatant 9000 g
SCD	sickle cell disease
SCR	short census repeat
SGLT2	sodium dependent glucose transporter 2
sLe^x	sialyl Lewis^x
ST	sulfotransferase
TMD	transmembrane domain
TNF	tumor necrosis factor
Tween 80	polyoxyethylene (20) sorbitan monooleate
Tyr	tyrosine
UDPGA	Uridine diphosphate α-D-glucuronic acid
UGT	Uridine diphosphate glucuronyl transferases
UPEC	uropathogenic *Escherichia coli*
UTI	urinary tract infection
V_D	volume of distribution
VOC	vaso-occlusive crisis
α-AGP	α_1-acid glycoprotein

1 Carbohydrate mimetics

Carbohydrates are one of the most important energy sources of the human metabolism. In addition, they are also important as structural element and as cell surface epitopes involved in numerous cell recognition and differentiation processes. Every single human cell is coated with complex carbohydrate structures linked to proteins (→ glycoproteins, proteoglycans) and lipids (→ glycolipids), forming a glycocalyx layer up to 100 nm in thickness (1-4). The structural variability of the glycocalyx is a consequence of different monosaccharides existing in the pyranose or furanose form, which are linked in α- or β-configuration and different regiochemistry. This variability is needed for specific cell-cell interactions and differentiation of cells in various processes (5). To promote specific interactions between cells in physiological and pathological processes, carbohydrate-binding proteins, so-called lectins, recognize specific epitopes of the glycocalyx (6, 7). Furthermore, carbohydrates are recognized by bacteria and viruses during infection cycles, and can lead to autoimmune diseases, when antibodies recognizing endogenous carbohydrate moieties are formed (8-10).

Therefore, the enzymatic metabolism of carbohydrates (i), the active transport of carbohydrates across membrane barriers by transporters (ii), the composition and reduction of glycans by transferases (iii) and hydrolases (iv), and the recognition of carbohydrate epitopes by lectins (v) or antibodies (vi) are all potential targets for the treatment and prevention of carbohydrate-related diseases (11). Although important in numerous pathological processes, only a small fraction of therapeutics on the market are derived from carbohydrates. This is partially due to the poor pharmacokinetic properties inherently linked to the chemical space that must be covered to address carbohydrate targets. Carbohydrates are hydrophilic and contain many hydrogen bond (H-bond) donors, leading to insufficient passive permeability across the lipid membranes of the enterocytes necessary for oral bioavailability. Furthermore, when administered parenterally, carbohydrates suffer from fast renal excretion leading to short plasma half-lives.

The pharmacokinetic parameters required for carbohydrate-derived drugs are highly dependent on the site of action. For example, marketed inhibitors of the α-glycosidase in the brush border of the small intestine (acarbose, miglitol, voglibose) do not require oral bioavailability to reach their therapeutic target (12-14). Additionally, carbohydrate-derived

drugs applied by injection do not need to be permeable if acting on extracellular targets, but have to overcome problems related to short half-life. Thus, Fondaparinux, an antithrombotic agent selectively inhibiting coagulation factor Xa and Rivipansel, a pan-selectin inhibitor used for the treatment of vaso-occlusive crisis in sickle cell patients, are administered by intravenously and have both sufficient long half-lives due to high plasma protein binding (15, 16).

1.1 Orally bioavailable carbohydrate mimetics

To increase passive permeability and intestinal absorption as a first step to achieve oral bioavailability of a carbohydrate-derived drug, a consequent reduction of the polar surface area and addition of lipophilicity is needed. In Figure 1.1, currently approved orally bioavailable carbohydrate mimetics are listed. All compounds are derived from monosaccharides and achieved their selectivity and oral bioavailability primarily via the aglycone.

The iminosugar Miglustat (**1**), a glucosylceramide synthase inhibitor used for the treatment of Gaucher and Niemann Pick disease, shows high oral bioavailability by passive diffusion in mammalians (17). Although not studied, the low molecular weight below 250 Da would allow paracellular diffusion to achieve the high oral bioavailability (18). Zidovudine (**2**) is the first approved representative of nucleotide reverse-transcriptase inhibitors (NRTIs) used in retro-viral therapies against acquired immunodeficiency syndrome (AIDS) (19). By modifying the ribose unit of nucleosides, NRTIs stop the elongation of DNA on the 3’ position of ribose. Zidovudine and other NRTIs show generally moderate to good oral bioavailability due to passive permeation, but also active transport by nucleoside transporters in the small intestine (20, 21). The antiepileptic drug Topiramate (**3**) exceeds its anticonvulsant effects due to different pharmacodynamics properties, i.e. an inhibitory effect on voltage-gated Na^+ and Ca^{2+} channels, glutamate-activated ion channels, and variable modulatory effects on γ-aminobutyric acid (GABA)-activated ion channels (22). High oral bioavailability was achieved by reducing the hydrophilicity of the D-fructose moiety by two acetonide groups, which, in addition, lock the pyranose moiety in a “twist-boat” conformation, that is conducive to the pharmacological effects (23).

In order to treat diabetes mellitus type 2, the reabsorption of glucose in the kidneys can be blocked by specific inhibitors of sodium dependent glucose transporter 2 (SGLT2).

Dapagliflozine (**4**) was the first approved SGLT2 inhibitor, exhibiting a large lipophilic aglycone, responsible for selectivity and intestinal absorption (24, 25). Natural and unnatural O-glucosidic SGLT2 inhibitors showed metabolic instability in the intestinal brush border due to the presence of β-glucosidases. Therefore, the development of potent compounds was focused on metabolically more robust C-aryl glucosides (26, 27).

The neuraminidase inhibitor Oseltamivir (**5**) for the prophylaxis and treatment of influenza virus infections is the carbohydrate mimetic, which underwent the most consequent reduction of the polar surface to achieve oral bioavailability. Starting from the natural substrate sialic acid, the hydrophilic glycerol side chain was replaced by a lipophilic isopentoxy moiety, the hydroxyl group at the 2 position was eliminated and an ethyl ester prodrug was formed (28). The prodrug shows an oral bioavailability of approximately 80% and the bioactive principle is released after hydrolysis by carboxylesterases (29).

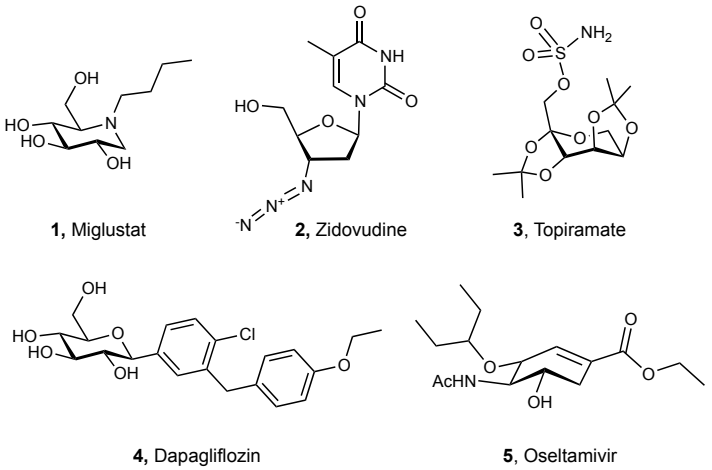


Figure 1.1 Orally bioavailable carbohydrate mimetic drugs currently on the market.

1.2 Projects presented in this thesis

1.2.1 E-selectin antagonists

The lectin E-selectin is expressed on the endothelium upon an inflammatory stimulus (30, 31). By interacting with glycoprotein ligands on leukocytes, E-selectin promotes rolling of these leukocytes on the endothelial surface. This first step initiates the inflammatory cascade to promote firm adhesion and transmigration of leukocytes to the site of inflammation. Thus, the inflammatory cascade is forming a vital defense mechanism in the event of injuries or infections. However, when excessively used, it can turn deleterious in numerous diseases with an inflammatory component, such as stroke, psoriasis or asthma (32). Therefore, blocking selectins with an antagonist and, as a consequence, interrupting the inflammatory cascade has been recognized as a promising therapeutic approach for the treatment of inflammatory diseases (33).

The carbohydrate epitope recognized by E-selectin is the tetrasaccharide sialyl Lewis^x (sLe^x) (34). Carbohydrate mimetics of sLe^x are currently investigated in clinical trials, but due to poor oral bioavailability they have to be administered intravenously (35, 36). For the treatment of chronic inflammatory diseases, an oral administration to reduce the administration burden of patients is more convenient. The development an orally bioavailable E-selectin antagonist is, however, challenging, since carbohydrate mimetics with sufficient pharmacokinetic properties have to be designed from the tetrasaccharide lead structure sLe^x, which possesses most unfavorable physicochemical properties. The evaluation of novel, in-house synthesized antagonists is described in chapter 3.

1.2.2 FimH antagonists

Around 75 % of all urinary tract infections (UTIs) are caused by uropathogenic *E. coli* (UPEC). UPEC are expressing the virulence factor FimH, a bacterial lectin presented on the distal tip of type 1 fimbriae (37). FimH binds to mannosides on the luminal surface of the bladder and therefore prevents the bacteria from being washed out during micturition (38). UTIs are treated with antibiotics, but the increasing antibiotic resistance urges for new treatment options (39). The challenge of developing FimH antagonists is, that after oral absorption, the molecule has to be excreted via the bladder without undergoing metabolism. Previous work

showed the high potential to treat UTIs by FimH antagonists and further progress is presented in chapter 4 (40).

1.2.3 Human neuraminidase 3 inhibitors

The concentration of sialic acids in humans, either bound to residues or in free circulation, is highly regulated, guided by sialyltransferases and neuraminidases (41). Four different human neuraminidases (NEU1-4) have been identified, each with its own specificity, expression pattern and modulations (42). The development of selective inhibitors for human neuraminidases NEU1-4 is of interest, not only for the development of potential therapeutics against different diseases where overexpression of neuraminidases is involved, but also in the study of physiological and pathophysiological role of neuraminidases. To conduct *in vitro* studies in cell lines, as well as *in vivo* studies, antagonists need to emphasize cellular membrane permeation and oral bioavailability. Therefore, the pharmacokinetic properties of neuraminidase inhibitors have to be evaluated. The development of selective neuraminidase 3 inhibitors is presented in chapter 5.

1.3 References

1. Alphonsus CS, Rodseth RN. The endothelial glycocalyx: a review of the vascular barrier. *Anaesthesia*. **2014**;*69(7)*:777-84.
2. Van Teeffelen JW, Brands J, Stroes ES, Vink H. Endothelial glycocalyx: sweet shield of blood vessels. *Trends Cardiovasc Med*. **2007**;*17(3)*:101-5.
3. Nieuwdorp M, Meuwese MC, Mooij HL, Ince C, Broekhuizen LN, Kastelein JJ, et al. Measuring endothelial glycocalyx dimensions in humans: a potential novel tool to monitor vascular vulnerability. *J Appl Physiol (1985)*. **2008**;*104(3)*:845-52.
4. Horiuchi K, Naito I, Nakano K, Nakatani S, Nishida K, Taguchi T, et al. Three-dimensional ultrastructure of the brush border glycocalyx in the mouse small intestine: a high resolution scanning electron microscopic study. *Arch Histol Cytol*. **2005**;*68(1)*:51-6.
5. Weinbaum S, Tarbell JM, Damiano ER. The structure and function of the endothelial glycocalyx layer. *Annu Rev Biomed Eng*. **2007**;*9*:121-67.
6. Turner MW. The role of mannose-binding lectin in health and disease. *Mol Immunol*. **2003**;*40(7)*:423-9.
7. Sharon N, Lis H. Lectins--proteins with a sweet tooth: functions in cell recognition. *Essays Biochem*. **1995**;*30*:59-75.
8. Sharon N. Bacterial lectins, cell-cell recognition and infectious disease. *FEBS Lett*. **1987**;*217(2)*:145-57.
9. Skehel JJ, Wiley DC. Receptor binding and membrane fusion in virus entry: the influenza hemagglutinin. *Annu Rev Biochem*. **2000**;*69*:531-69.
10. Steck AJ, Stalder AK, Renaud S. Anti-myelin-associated glycoprotein neuropathy. *Curr Opin Neurol*. **2006**;*19(5)*:458-63.
11. Ernst B, Magnani JL. From carbohydrate leads to glycomimetic drugs. *Nat Rev Drug Discov*. **2009**;*8(8)*:661-77.
12. Chen X, Zheng Y, Shen Y. Voglibose (Basen, AO-128), one of the most important alpha-glucosidase inhibitors. *Curr Med Chem*. **2006**;*13(1)*:109-16.
13. Campbell LK, Baker DE, Campbell RK. Miglitol: assessment of its role in the treatment of patients with diabetes mellitus. *Ann Pharmacother*. **2000**;*34(11)*:1291-301.
14. Bischoff H. The mechanism of alpha-glucosidase inhibition in the management of diabetes. *Clin Invest Med*. **1995**;*18(4)*:303-11.
15. Petitou M, Duchaussoy P, Herbert JM, Duc G, El Hajji M, Branellec JF, et al. The synthetic pentasaccharide fondaparinux: first in the class of antithrombotic agents that selectively inhibit coagulation factor Xa. *Semin Thromb Hemost*. **2002**;*28(4)*:393-402.
16. Wun T, Styles L, DeCastro L, Telen MJ, Kuypers F, Cheung A, et al. Phase 1 study of the E-selectin inhibitor GMI 1070 in patients with sickle cell anemia. *PLoS One*. **2014**;*9(7)*.
17. van Giersbergen PL, Dingemanse J. Influence of food intake on the pharmacokinetics of miglustat, an inhibitor of glucosylceramide synthase. *J Clin Pharmacol*. **2007**;*47(10)*:1277-82.
18. Knipp GT, Ho NF, Barsuhn CL, Borchardt RT. Paracellular diffusion in Caco-2 cell monolayers: effect of perturbation on the transport of hydrophilic compounds that vary in charge and size. *J Pharm Sci*. **1997**;*86(10)*:1105-10.
19. D'Andrea G, Brisdelli F, Bozzi A. AZT: an old drug with new perspectives. *Curr Clin Pharmacol*. **2008**;*3(1)*:20-37.
20. Blum MR, Liao SH, Good SS, de Miranda P. Pharmacokinetics and bioavailability of zidovudine in humans. *Am J Med*. **1988**;*85(2A)*:189-94.
21. Balimane PV, Sinko PJ. Involvement of multiple transporters in the oral absorption of nucleoside analogues. *Adv Drug Deliv Rev*. **1999**;*39(1-3)*:183-209.
22. Shank RP, Maryanoff BE. Molecular pharmacodynamics, clinical therapeutics, and pharmacokinetics of topiramate. *CNS Neurosci Ther*. **2008**;*14(2)*:120-42.

23. Maryanoff BE, Nortey SO, Gardocki JF, Shank RP, Dodgson SP. Anticonvulsant O-alkyl sulfamates. 2,3:4,5-Bis-O-(1-methylethylidene)-beta-D-fructopyranose sulfamate and related compounds. *J Med Chem*. **1987**;*30(5)*:880-7.
24. Kasichayanula S, Liu X, Lacreta F, Griffen SC, Boulton DW. Clinical pharmacokinetics and pharmacodynamics of dapagliflozin, a selective inhibitor of sodium-glucose co-transporter type 2. *Clin Pharmacokinet*. **2014**;*53(1)*:17-27.
25. Meng W, Ellsworth BA, Nirschl AA, McCann PJ, Patel M, Girotra RN, et al. Discovery of dapagliflozin: a potent, selective renal sodium-dependent glucose cotransporter 2 (SGLT2) inhibitor for the treatment of type 2 diabetes. *J Med Chem*. **2008**;*51(5)*:1145-9.
26. Ehrenkranz JR, Lewis NG, Kahn CR, Roth J. Phlorizin: a review. *Diabetes Metab Res Rev*. **2005**;*21(1)*:31-8.
27. Ellsworth BA, Meng W, Patel M, Girotra RN, Wu G, Sher PM, et al. Aglycone exploration of C-arylglucoside inhibitors of renal sodium-dependent glucose transporter SGLT2. *Bioorg Med Chem Lett*. **2008**;*18(17)*:4770-3.
28. Lew W, Chen X, Kim CU. Discovery and development of GS 4104 (oseltamivir): an orally active influenza neuraminidase inhibitor. *Curr Med Chem*. **2000**;*7(6)*:663-72.
29. He G, Massarella J, Ward P. Clinical pharmacokinetics of the prodrug oseltamivir and its active metabolite Ro 64-0802. *Clin Pharmacokinet*. **1999**;*37(6)*:471-84.
30. Lasky LA. Selectins - Interpreters of Cell-Specific Carbohydrate Information during Inflammation. *Science*. **1992**;*258(5084)*:964-9.
31. Erbe DV, Wolitzky BA, Presta LG, Norton CR, Ramos RJ, Burns DK, et al. Identification of an E-Selectin Region Critical for Carbohydrate Recognition and Cell-Adhesion. *J Cell Biol*. **1992**;*119(1)*:215-27.
32. Ley K. The role of selectins in inflammation and disease. *Trends in Molecular Medicine*. **2003**;*9(6)*:263-8.
33. Rossi B, Constantin G. Anti-selectin therapy for the treatment of inflammatory diseases. *Inflamm Allergy Drug Targets*. **2008**;*7(2)*:85-93.
34. Binder FP, Lemme K, Preston RC, Ernst B. Sialyl Lewis(x): a "pre-organized water oligomer"? *Angew Chem Int Ed Engl*. **2012**;*51(29)*:7327-31.
35. Telen MJ, Wun T, McCavit TL, De Castro LM, Krishnamurti L, Lanzkron S, et al. Randomized phase 2 study of GMI-1070 in SCD: reduction in time to resolution of vaso-occlusive events and decreased opioid use. *Blood*. **2015**;*125(17)*:2656-64.
36. Natoni A, Smith TAG, Keane N, McEllistrim C, Connolly C, Jha A, et al. E-selectin ligands recognised by HECA452 induce drug resistance in myeloma, which is overcome by the E-selectin antagonist, GMI-1271. *Leukemia*. **2017**.
37. Martinez JJ, Mulvey MA, Schilling JD, Pinkner JS, Hultgren SJ. Type 1 pilus-mediated bacterial invasion of bladder epithelial cells. *EMBO J*. **2000**;*19(12)*:2803-12.
38. Klemm P, Christiansen G. Three fim genes required for the regulation of length and mediation of adhesion of Escherichia coli type 1 fimbriae. *Mol Gen Genet*. **1987**;*208(3)*:439-45.
39. Flores-Mireles AL, Walker JN, Caparon M, Hultgren SJ. Urinary tract infections: epidemiology, mechanisms of infection and treatment options. *Nat Rev Microbiol*. **2015**;*13(5)*:269-84.
40. Kleeb S, Pang L, Mayer K, Eris D, Sigl A, Preston RC, et al. FimH antagonists: bioisosteres to improve the in vitro and in vivo PK/PD profile. *J Med Chem*. **2015**;*58(5)*:2221-39.
41. Varki A. Sialic acids in human health and disease. *Trends Mol Med*. **2008**;*14(8)*:351-60.
42. Miyagi T, Yamaguchi K. Mammalian sialidases: physiological and pathological roles in cellular functions. *Glycobiology*. **2012**;*22(7)*:880-96.

2 Oral bioavailability of carbohydrate mimetics

Oral dosing of a drug is the most convenient administration route and is enhancing the patient compliance and is reducing the risk of complications. Especially for the treatment of diseases, which do not require acute hospitalization, an oral administration is indicated to improve the patient's quality of life (1, 2). Therefore, we strive for orally bioavailable carbohydrate mimetics in the different projects discussed in this thesis.

Oral bioavailability is defined as the extent of a drug reaching systematic circulation (fraction *F*) after giving an oral dosage form (3). To reach the systemic circulation, the drug has to overcome different physical (e.g. membranes), physicochemical (e.g. solubility, stability) and biological (e.g. metabolism, transporters) barriers summarized in Figure 2.1 (4).

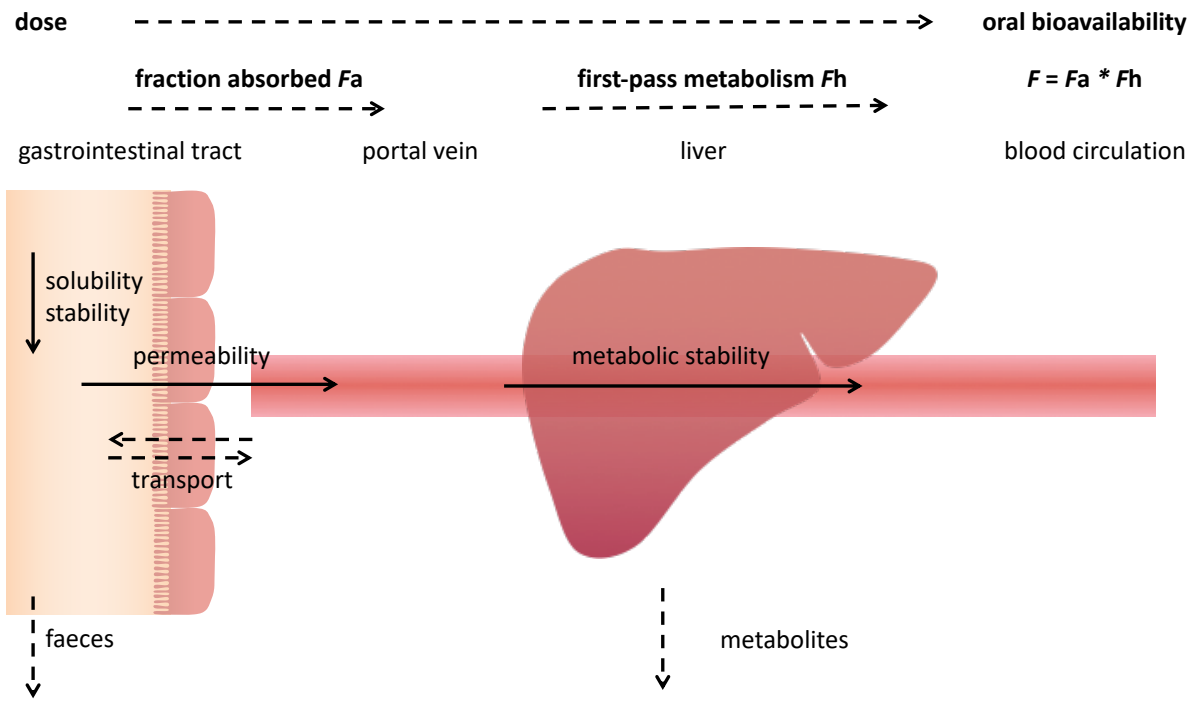


Figure 2.1 Barriers to overcome to achieve oral bioavailability**.**

The impact of different parameters that affect oral bioavailability can be assessed with the Biopharmaceutics Drug Distribution and Classification System (BDDCS) (5, 6). According to the permeability and solubility data, BDDCS divides the drugs in four classes (Table 2.1). As previously discussed, carbohydrate mimetics t are often showing disadvantageous pharmacokinetic properties due to their hydrophilic character, leading to insufficient passive permeability but good aqueous solubility. Therefore, the molecules developed during the course of this thesis predominantly belong to BDDCS class 3, which is characterized by a strong influence of active transport mechanisms in the absorptive direction and possible efflux transport (7). Thus, the permeation over the intestinal membrane either actively or passively is the key challenge to overcome for our goal to achieve orally bioavailable carbohydrate mimetics.

Table 2.1 Biopharmaceutics Drug Distribution and Classification System (BDDCS) adapted from Shugarts *et al.* (7).

	High solubility	**Low solubility**
High permeability	**Class 1** • Extensive metabolism • Minimal transporter effects	**Class 2** • Extensive metabolism • Efflux transporter effects predominate in the gut, while absorptive and efflux transporter effects occur the in liver
Low permeability	**Class 3** • Poor metabolism • Absorptive transporter effects predominate (but may be modulated by efflux transporters)	**Class 4** • Poor metabolism • Absorptive and efflux transporter effects could be important

The oral bioavailability *F* of a drug is measured by comparing the concentration in systemic circulation after oral dosage (*per os,* PO) with intravenous dosing (IV), under the assumption that after IV administration, 100 % of the drug reached circulation (8, 9). After a correction for the applied dose *D*, a determination of the concentration over time (area under the curve, AUC) gives the oral bioavailability *F* of a drug in percent (Equation 1).

$$F_{PO} = 100 \cdot \frac{AUC_{PO} \cdot D_{IV}}{AUC_{IV} \cdot D_{PO}} \qquad \text{(eq. 1)}$$

The oral bioavailability is a key parameter affecting the systemic exposure of a drug. To which extent and time the drug is present at the therapeutic target is dependent on a complex puzzle of different physicochemical, physiological and biological parameters, which are summarized as pharmacokinetic properties. These properties are divided into five parts (PADME). The toxicity of a drug is often included additionally into this set of properties, since the toxicity limits the dosage of the drug (10, 11). In this case, the abbreviation **PADMET** is used:

- **P**hysicochemical properties
- **A**bsorption
- **D**istribution
- **M**etabolism
- **E**limination
- **T**oxicity

Insufficient pharmacokinetic properties causing a lack of sufficient exposure and thus efficacy, were the major cause of failure in clinical trials for a long time. After specific physicochemical and biochemical profiling were implemented in the drug development process, the failures of clinical trials due to pharmacokinetic reasons were reduced to less than 10% (12, 13). To identify possible drawbacks of the molecule and to successfully introduce beneficial structural modifications, the implementation of *in vitro* evaluation methods has to take place in an early phase of drug development. Our group started to establish a PADMET platform with the purpose to profile carbohydrate mimetics and improve their pharmacokinetic properties in 2007. Dr. Mattias Wittwer started this endeavor, which was then further expanded by Dr. Simon Kleeb, Dr. Jacqueline Bezençon, and myself. The current version of the platform is summarized in Table 2.2. The overall aim of our PADMET platform is to predict the pharmacokinetic properties of lead structures and to adapt the standard PADMET evaluation

to a new compound class, the carbohydrate mimetics. Herein, the implemented methods are summarized, and their predictive power and limitations are discussed with a focus on physicochemical properties, absorption and metabolism.

Table 2.2 Assays in the PADMET platform to predict pharmacokinetic properties of carbohydrate mimetics.

Property	**Assay**
Physicochemical properties	
Log$D_{7.4}$	Miniature shake flask method
p*K*a	^{1}H-NMR spectroscopy
Solubility	Kinetic solubility
	Thermodynamic solubility
	Preformulations
Absorption	
Gastrointestinal stability	Simulated gastrointestinal fluids
Permeability and transport	Parallel artificial membrane assay (PAMPA)
	Caco-2 cell-based permeability assay
Distribution	
Plasma protein binding	Equilibrium dialysis
Metabolism	
Phase I metabolism	Human, rat or mouse liver microsomes
Phase II metabolism	Human, rat or mouse liver cytosol and microsomes
Elimination	
No specific assay. Implemented in metabolism and distribution assays	
Toxicity	
Cell toxicity	MTT-assay

2.1 Physicochemical properties

Structural properties of molecules are easily accessible and include valuable information to predict and interpret their pharmacokinetic behavior. Lipinski *et al*. and Veber *et al.* analyzed the structural properties of molecules succeeding or failing in clinical studies and summarized their findings in rules (Table 2.3) (14, 15). Although Lipinski clearly stated that up to two violations of his rule still allow oral availability and many exceptions have been published, the rule of 5 had a large impact on drug development. Many libraries for high throughput screening in industries were adjusted to Lipinski's and Veber's rules as hard cut off criteria, reducing the chemical space used for the initial screening (16). Nevertheless, these rules are an easy to apply guidance indicating possible risk factors during the drug discovery process (17).

Table 2.3 Rules to predict oral bioavailability from structural properties

Lipinski's rule of five		**Veber's rule**	
Molecular weight (MW)	< 500 Da	Rotatable bonds	< 10
Hydrogen bond donors	< 5	Polar surface area (PSA)	< 140 Å^2
Hydrogen bond acceptor	< 10		
clog*P*	< 5		

2.1.1 Lipophilicity

The lipophilicity of a molecule in its uncharged form is quantified by the octanol-water partition coefficient log*P* or, for ionized species, by the distribution coefficient log*D* at a specific pH value (18). Thus, the lipophilicity of a compound at a given pH is dependent on the p*K*a of the molecule. An appropriate lipophilicity is a prerequisite to support the transfer of a molecule from the extracellular aqueous solution into a cell membrane and is therefore needed for transcellular transport. Furthermore, the lipophilicity has an impact on drug binding to plasma or tissue proteins to modify the volume of distribution and plasma half-life (19, 20). However, high lipophilicity can lead to problems regarding solubility and metabolic stability (21, 22). As a result, a majority of orally bioavailable and marketed drugs have a log*P* value between 1 and 4 (23).

The lipophilicity of a molecule can easily be predicted by *in silico* tools (24, 25) using structural parameters as proposed e.g. by Moriguchi (26). However, structural elements specific for carbohydrate mimetics, such as overlapping polar surfaces from neighboring hydroxyl groups and intramolecular hydrogen bonds, are often not implemented in those *in silico* tools. For *in vitro* measurements, different approaches are being used, e.g. high-pressure liquid chromatography (HP-LC) or artificial membranes (27-29). The method of choice for the log*D* determination in the PADMET-platform was the shake flask approach with a water and octanol phase. With this gold standard method, the distribution coefficient log*D* is determined in the solvents of choice by direct measurement and therefore without intrinsic errors. The drawback of the method is the inaccuracy for highly lipophilic (log*D* > 3.5) but also hydrophilic (log*D* < -1.5) compounds due to the high influence of standard measurement errors in these regions. The shake flask method was used to determine the lipophilicity of all molecules presented in this thesis with adapted protocols for each manuscript and publication.

2.1.2 pK_a

The dissociation constant pK_a of acids defines the degree of ionization of a molecule in aqueous solution at a given pH value. When the pH is equal to the pK_a value, 50% of the uncharged and charged species are present (18). Since ionized molecules are more polar than their neutral counterparts, the pK_a is an important characteristic affecting various properties, such as solubility, permeability, plasma protein binding, and excretion (30). The pK_a can either be predicted by *in silico* tools or measured, for example by electropotentiometric methods (31, 32). Dr. Jacqueline Bezençon revisited the determination of pK_a by 1H-NMR spectroscopy (33), which is integrated into the PADMET-platform.

2.1.3 Solubility

The solubility of a compound is a thermodynamic property defined by the equilibrium of a substance in its solid state and its dissolved form in the liquid phase. A compound must be dissolved to create a concentration gradient that drives passive transport through the intestinal membranes. Furthermore, the solubility of a compound determines the upper limit of its concentration to be used in assays. Therefore, the solubility of the compound has to be screened early in the drug development process (9, 34). Numerous factors define solubility, namely the solvent and its composition (addition of co-solvents, ions, proteins), the pH value, and the temperature (35-37). Besides the intrinsic properties of the drug affecting the

solubility (e.g. molecular size, lipophilicity, pK_a), other factors such as crystal packing, particle size of the crystal, melting point and additives influence solubility (38, 39). The solubility required for an oral administration depends on the required dose, but generally should be above 0.5 mg/mL, assuming moderate affinity and permeability (34).

Different techniques to determine the solubility of a compound have been developed and evaluated. For each assay it is important, that the analyte is present in solid state and dissolved form to provide the equilibrium.

Kinetic solubility

With the increasing number of compounds moving into drug discovery and development, kinetic solubility measurements have gained importance to meet the high throughput needs (40). These methods use high concentration stock solutions of the compounds dissolved in dimethyl sulfoxide (DMSO). Precipitation is induced by adding the stock solution to a buffer solution, with a subsequent analysis of the supernatant. For high throughput, this approach is very useful, but has some intrinsic limitations, since the obtained kinetic solubility is not taking the solid state and polymorphism of the molecule into account. Additionally, supersaturation may occur and the organic solvent DMSO present in the final solution can enhance the apparent solubility. Therefore, the measured kinetic solubility tends to be increased compared to other measurements and the predicted pharmacokinetic properties of a molecule can therefore be overestimated (41). Kinetic solubility measurements for FimH antagonists are discussed in chapter 4.

Thermodynamic solubility

The thermodynamic solubility is usually determined by the shake flask method, where an excess of solid compound is added to solvent to form a saturated solution. Aliquots are filtered, and the concentration of the supernatant is determined. This technique is considered to be the gold standard for solubility measurements (10). Major drawbacks are the large compound consumption, as well as the time-consuming measurements, since for slowly dissolving compounds, it takes a long time to reach the equilibrium (42). The method was used to determine the solubility of E-selectin and neuraminidase 3 inhibitors, discussed chapters 3 and 5.

Preformulation

If the solubility of a compound is insufficient to reach the required assay concentration, preformulation is needed. Especially for pharmacokinetic studies with intravenous dosing, a complete dissolution of the compound is mandatory, because the injection of a precipitate can lead to embolism (43). A set of different co-solvents (e.g. DMSO), surfactants (e.g. Tween 80) and complexing agents (e.g. cyclodextrin) are available to be used individually or in combination to formulate a solution without precipitate. To minimize the side effects of the injected vehicle, the maximal concentration of each co-solvent used in the pharmacokinetic studies for E-selectin antagonists was chosen according to a study published by Thackaberry et *al.* (44).

2.2 Absorption

After being orally administered, the active ingredient of a drug has to surpass numerous barriers of the gastrointestinal tract to reach the therapeutic target. The gastrointestinal organ system consists of different parts, each with its own physiological and physical environments (9). The majority of drugs are absorbed by transcellular transport across the enterocytes of the small intestines, composed of duodenum, jejunum and ileum. To reach the small intestines, the molecule has to be dissolved in the intestinal lumen and has to be stable in the stomach and duodenum. Therefore, the stability in the gastrointestinal tract, as well as the permeability across the intestinal membranes has to be assessed (4).

2.2.1 Gastrointestinal stability

In the stomach and/or the small intestine, a swallowed tablet disperses, and the drug is dissolved. Fasted or fed state of these organs is largely influencing physiology and residence time (Table 2.4). The effects of food on stability and absorption can be positive or negative. For example, highly lipophilic compounds maybe more likely to be absorbed in fed state, whereas actively transported molecule can suffer from reduced absorption in fed state due to a saturation of the transport system (9, 45, 46). In early drug development, the fasted state is preferred for *in vivo* studies in animals and first in human trials, because this state shows less variability and is easier to standardize. Therefore, in the PADMET-platform, the focus regarding stability of carbohydrate mimetics is on gastrointestinal environments of the fasted state.

Table 2.4 Physiological differences of the gastrointestinal tract in fasted and fed state (18, 45, 47, 48).

	Stomach		**Small intestine**	
	Fasted	**Fed**	**Fasted**	**Fed**
pH	1	2-5	6.5	5
Residence time	20 min	3-4 h	3-4 h	Up to 24 h
Enzymatic activity	low	high	low	high

Simulated gastric fluid

The acidic condition in the stomach is necessary for the function of the gastrointestinal tract. It is needed to activate pepsin from pepsinogen, to denature proteins, and as a natural barrier against bacterial infections (49). Hydrochloric acid, excreted by proton pumps from the stomach glands, is responsible for the acidity of the stomach and can lead to a chemical hydrolysis of carbohydrates(50). Therefore, to test acid stability of compounds, an HCl environment with or without pepsin is used. The results obtained for E-selectin antagonists are discussed in manuscript 1 and 3.

Simulated intestinal fluid

In the first part of the small intestine called duodenum, the gastric acid is neutralized by bicarbonate excretion from Brunner's glands. Furthermore, bile acid for the uptake of fatty acids and pancreatin, a mixture of amylases, proteases and lipases, is added in order to digest food ingredients. However, this enzyme mixture can be disruptive for drugs. Especially ester bonds can be cleaved by lipases, hydrolyzing ester prodrugs before intestinal absorption (51). This effect is studied and discussed in manuscript 1.

2.2.2 Permeability

The majority of nutrients and drugs are absorbed in the small intestine, particular in the jejunum and ileum. The surface of the small intestine is enlarged by villi covered by epithelial cells, which form microvilli on their apical surface, leading to a total surface of more than 100 m^2 over the three to four meter length of small intestine (9, 52). The surface is covered by glycoproteins, leading to a partial negative charge and an unstirred layer mucus coat. For the membrane transfer through the epithelial cells of the small intestines (enterocytes) various pathways have to be considered as described in Figure 2.2. These pathways are used either separately or in combination, making the prediction of the absorption process highly difficult (53).

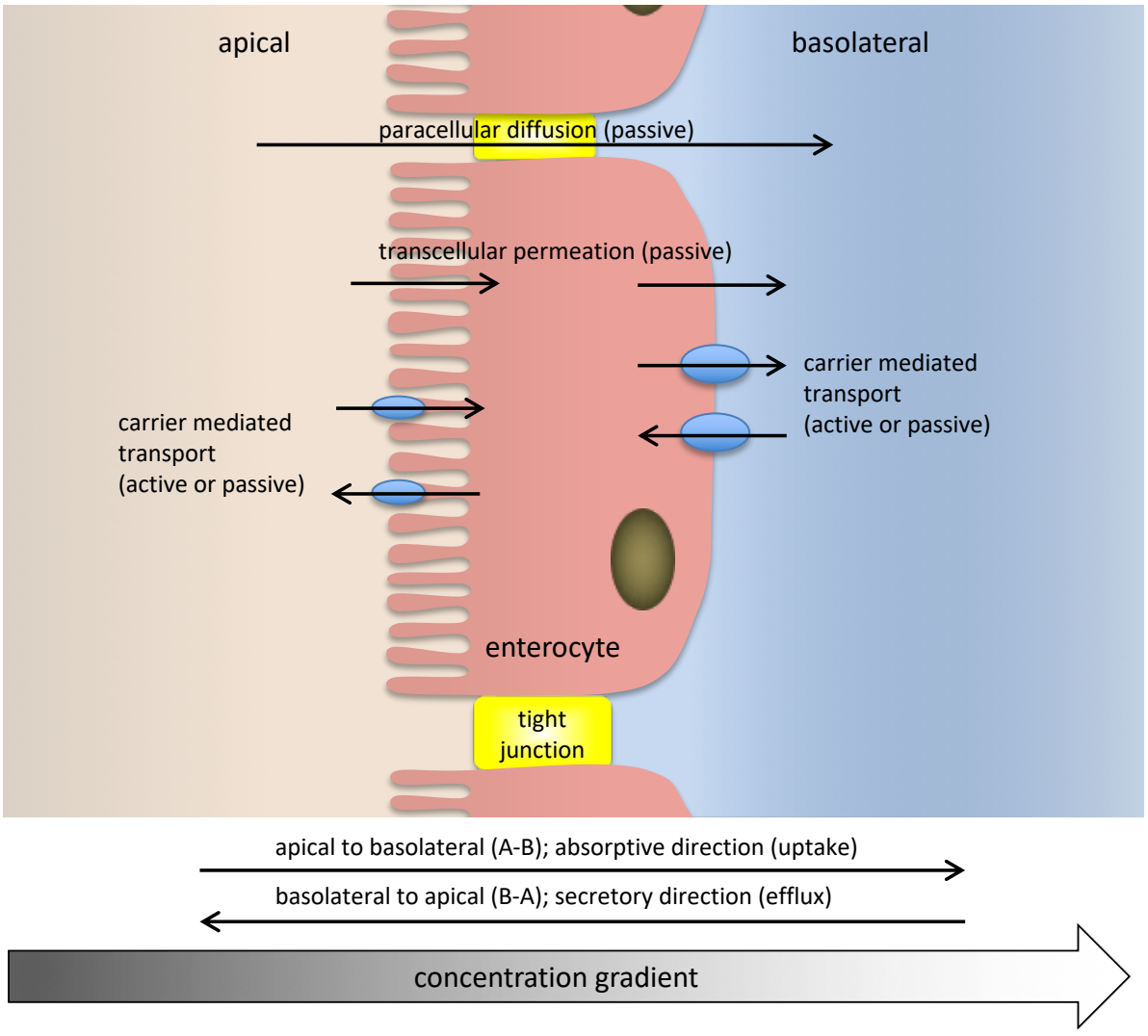


Figure 2.2 Absorption pathways through enterocytes along the concentration gradient (53).

Passive permeability

Passive permeability is driven by a concentration gradient, either paracellularly through tight junctions or transcellularly through epithelial cells. It depends on structural properties and lipophilicity of the molecules. Compared to active transport, it is only slightly influenced by stereochemical variability (18). The key structural properties analyzed by Lipinski and Veber (2.1 Physicochemical properties) are predictive for passive transcellular permeability (14, 15). Paracellular permeation through tight junctions, however, requires different properties, i.e. molecular weight below 250 Da and hydrophilicity (54). For both the trans- as well the paracellular pathway, solubility plays a significant role, since it controls drug concentration in the intestinal lumen. However, unlike for the paracellular route, transcellular permeation can also be influenced by active transport or drug metabolizing enzymes (18).

For the transcellular permeation, drugs have to pass the phospholipid bilayer of the cells. This membrane barrier can be artificially mimicked in a parallel artificial membrane permeation assay (PAMPA) allowing the study of isolated passive permeability and is therefore widely applied in drug development (55). It can be performed in a high throughput format and often correlates well with *in vivo* results (56). Therefore, PAMPA was used to assess the permeability of carbohydrate mimetics in all manuscripts, however with specifically adapted methods for individual projects. Since PAMPA uses an artificial membrane and was developed using drug-like molecules and not carbohydrate mimetics, the results obtained for our molecules have to be handled with care and verification by other methods (e.g. Caco-2 cell assay) is required.

Active Transport

In addition to passive permeation, compounds can also pass membranes via carrier-mediated pathways. These pathway plays a crucial role for compounds not passing membranes passively due to high molecular weight, low lipophilicity or ionization (53). The carrier-mediated transport needs a direct interaction of the compound with a protein embedded in phospholipid bilayer. Therefore, carrier-mediated transport is substrate specific, saturable, and can be inhibited (9). The transport can be divided into facilitated and active transport. Facilitators do not require an energy source and are providing better permeation over the membrane along the concentration gradient. In contrast, active transporters are capable to transport a compound against its concentration gradient and therefore rely on an energy source (57). Symporters and antiporters are using the concentration gradient of another

molecule as driving force, whereas primary active transporter usually use adenosine-5'-triphosphate (ATP) as source of energy (58).

Active transport can therefore be in the absorptive and secretory direction. Efflux transporters, such as the ABC transporter P-glycoprotein (P-gp) exhibit a protective function by excreting xenobiotics back into the intestinal lumen and therefore can have a large effect on compounds with low passive permeability, especially at low doses (59). At high concentrations, P-gp can be saturated, but only for drugs with sufficient solubility. Furthermore, the expression of P-gp increases from the jejunum to the distal ileum. Transporters involved in the active uptake of compounds are mainly peptide transporter 1 (PEPT1) and members of the organic anion transporting polypeptide family (OATP) (57). They are responsible for the uptake of organic anions, di- and tripeptides. In addition, drugs can be specifically designed to target these transporters for enhanced orally bioavailability (60).

For the assessment of active transport an assay based on a colon carcinoma cell line (Caco-2) is used, allowing the measurement of the permeation rate in both directions (61). This assay offers the opportunity to measure permeability in an elaborated system offering good correlation of measured permeability P_{app} and oral absorption in man (62). Since Caco-2 cell express several transport systems simultaneously, other systems or transfected cell lines are preferred to evaluate the effect of a single transporter. Other drawbacks of the Caco-2 cell based model are the expression of enzymes differing from those of human enterocytes (e.g. human esterase CES1 for the study of ester prodrugs) and time-consuming and expensive handling (63). The Caco-2 cell based assay was used for E-selectin and FimH antagonists in chapter 3 and 4.

2.3 Distribution

Once in circulation, the compound distributes into tissues. This process is described by the volume of distribution V_D (Equation 2), where the total amount of the given dose D_{IV} is divided by the extrapolated concentration at time-point zero C_0. This gives an artificial number reflecting the 'apparent volume of distribution' for the compound. Compounds showing pronounced distribution into body tissues and low blood concentrations will also have a large V_D (64).

$$V_D = \frac{D_{IV}}{C_0} \qquad \text{(eq.2)}$$

2.3.1 Plasma Protein Binding (PPB)

Furthermore, compounds in circulation can bind to tissue or plasma components or remain unbound in the aqueous environment. Only the unbound fraction can interact with its target and, therefore, exhibit a therapeutic effect. This unbound fraction, however, is also available for metabolism or elimination. The affinity towards plasma proteins is generally weak and characterized by a fast association and dissociation. High or low PPB does not influence in-vivo efficacy as the effect on clearance and apparent potency by changes in free fraction will cancel out. Therefore, protein binding is not supposed to be an optimization parameter. Thus, structural optimization to improve pharmacokinetic properties in lead development projects should aim at increasing total drug exposure. This can be achieved by improving solubility, permeability and metabolic stability but not by increasing protein binding, i.e. unbound or bound fraction (65-67).

The most important plasma proteins involved in drug binding are human serum albumin (HSA), α_1-acid glycoprotein (α-AGP) and lipoproteins (4). For approx. 50% of 222 analyzed marketed drugs high PPB of more than 90 % was found, i.e. high plasma protein binding is occurring frequently (68).

Several methods for determining *in vitro* PPB are available, e.g. equilibrium dialysis, ultracentrifugation, ultrafiltration, microdialysis and surface plasmon resonance (69, 70). Equilibrium dialysis was chosen for our PADMET-platform, since it is generally accepted as gold standard (71). The principle is a two-chamber system separated by a semipermeable dialysis membrane. In one chamber, plasma proteins are added and after the equilibrium

between bound and unbound state has been established, the compound concentration in the chamber without plasma proteins is measured.

2.4 Metabolism and first pass elimination

Following absorption across the gut lumen, nutrients and drugs are transported along the portal vein to the liver where they encounter the first pass metabolism before reaching systemic blood circulation. Therefore, oral bioavailability of a compounds is also highly dependent on the enzymes recognizing the structure, the abundance of these enzymes and their metabolic rate (9).

The liver is the primary side for metabolism of xenobiotics, playing a central role in pharmacokinetics and toxicity. The liver is organized in lobules, an almost hexagonal cluster of hepatocytes connected to portal triads, which contain small branches of the portal vein, liver artery and bile duct. These lobules receive oxygenated arterial blood from the liver artery, as well as the nutrition rich blood from the portal vein (25% arterial, 75% nutrition rich blood). The bile duct transports bile acids as well as metabolites to the gall bladder and further to the duodenum (72).

Xenobiotics, including drugs and toxins, tend to be lipophilic, enabling passive transport through lipid membranes and compartments. Therefore, distribution and excretion of those molecules are difficult to control. One goal of metabolizing enzymes is to increase the polarity of xenobiotics, improving their water solubility and excretability (9). The metabolic enzymes can be divided into those introducing a new polar functional group by oxidation, reduction or hydrolysis (phase I metabolism) and those introducing new hydrophilic moieties by conjugation to endogenous molecules (phase II metabolism). Which metabolic pathways xenobiotics undergo is highly dependent on their structure; thus direct conjugation by phase II metabolism is especially observed for compounds already containing polar groups such as carbohydrate mimetics (Figure 2.3)(73). Sufficient metabolic stability is a prerequisite to achieve high concentrations in the systemic circulation over an extended time period. In contrast, if the compound is applied as a prodrug, a metabolic transformation is needed for the release of the active principle (74). Therefore, the metabolic profile of each drug candidate has to be carefully evaluated.

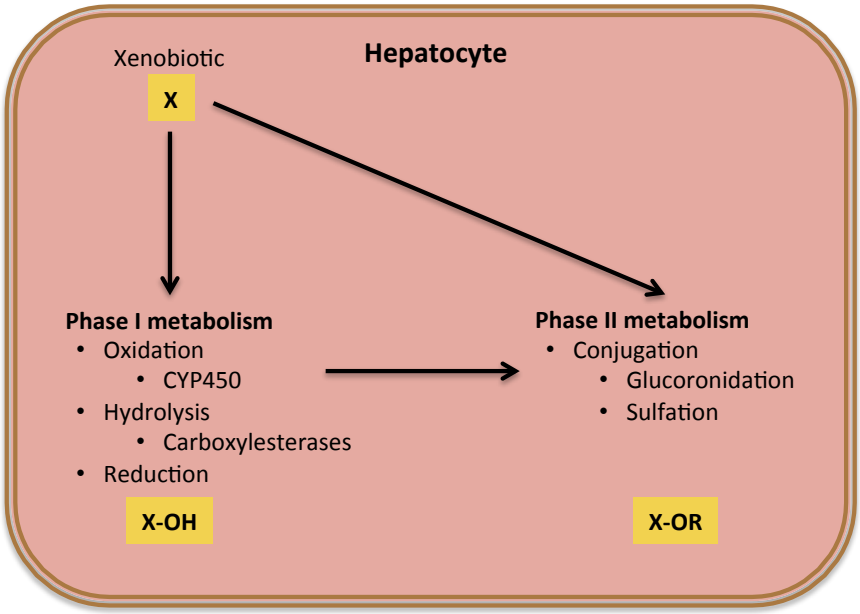


Figure 2.3 Most relevant phase I and II metabolism in hepatocyte.

2.4.1 Phase I metabolism

New functional groups such as hydroxyls, amines or carboxylic acids are introduced by phase I enzymes, leading to higher hydrophilicity compared to the parent compound. Phase I enzymes include cytochrome P450s (CYP450s), flavin monooxygenases (FMOs) dehydrogenases, reductases, and esterases (75).

Cytochrome P450s

The CYP450 superfamily, located on the endoplasmatic reticulum of hepatocytes, makes the largest contribution to phase I metabolism. These proteins are also present in other extrahepatic tissues such as small intestine, colon, kidneys, and lung (76). CYP450 enzymes contain a heme group coordinating an iron ion, which is responsible for the catalytic monooxygenation of substrates with two electrons provided by cytochrome reductases(77). The CYP450 superfamily consists of approximately two thousands of subtypes, expressed in different tissues and species, with just a few isoenzymes mainly responsible for the metabolism of xenobiotics in man (e.g. CYP3A4, CYP2D6, CYP2C9) (78). These isoenzymes are therefore of interest when a new compound is either metabolized by or interacts with one of those enzymes. Furthermore, possible polymorphism of these enzymes makes it crucial to know the metabolic pathway of new compounds (78).

To study the metabolic pathway of new compounds, *in silico* tools were developed (79). These tools often recognize potential *soft spots* of molecules. However, for new substance classes,

including carbohydrate mimetics, the results obtained with *in silico* tools are generally insufficient.

Several *in vitro* tools using liver microsomes, S9 fractions or liver slides from different species are available (80). Liver microsomes are the most common way to evaluate the metabolic behavior and therefore were chosen for the PADMET-platform. By adding the co-factor nicotinamide adenine dinucleotide phosphate (NADPH), cytochrome reductases are able to provide two electrons to activate CYP450 and the metabolic stability of the molecule can be assessed over time. Furthermore, metabolites can be extracted and identified or used for further studies. This assay format was used to measure the metabolic stability of E-selectin antagonists.

Carboxylesterases (CES)

Carboxylesterases are hydrolyzing esters to the corresponding carboxylic acids and alcohols. The multigene family of mammalian carboxylesterases is divided into 5 major subfamilies (CES1-5), of which CES1 and CES2 contribute the most to the metabolic activity (81). The tissue expression pattern and the substrate specificity of CES1 and CES2 show significant differences. CES1 preferably hydrolyzes substrate with a large acyl and small alcohol moiety but is also capable of hydrolyzing esters with both large alcohol and acyl moieties. In contrast, CES2 recognizes substrates with large alcohol group and small acyl moiety (82). Both carboxylesterase families are present in hepatocytes, but only CES2 is present in enterocytes (83, 84).

Carboxylesterases play a crucial role in the bioconversion of ester prodrugs (74). Hydrolysis of ester prodrugs in the enterocytes during absorption into the more hydrophilic carboxylic acid can reduce oral bioavailability of the active principle (84). Therefore, it is crucial to evaluate the hydrolysis of ester prodrug and an enzymatic specificity for CES1 is preffered. In liver microsomes, both isoenzymes CES1 and CES2 are present. Thus, the hydrolysis of ester prodrugs and their enzymatic specificity can be studied with the use of selective inhibitors of CES1 or CES2 (FimH antagonists, Publication 1) The hydrolysis and specificity of E-selectin ester prodrug was tested in recombinant carboxylesterases (Manuscript 1).

2.4.2 Phase II metabolism

Phase II metabolism comprises conjugation reactions of xenobiotics, which involve the covalent transfer of a polar endogenous compound such as glucuronic acid, glutathione, sulphate or acetyl to a suitable functional group on the substrate (9). These conjugations require a nucleophilic group on the xenobiotic, which can but does not have to be formed by phase I metabolism (73). In the majority of the cases, the large hydrophilic conjugate leads to biological inactivation of the parent compound and to increased biliary or renal elimination.

UDP-glucuronyltransferases (UGT)

UDP-glucuronyltransferases (UGT) catalyze a reaction referred to as "glucuronidation" by conjugating glucuronic acid, derived from the cofactor UDP-α-D-glucuronic acid (UDPGA), with a substrate bearing a suitable functional group, such as alcohols, carboxylic acids, amines, thiols and acidic carbon atoms. The conjugation occurs in a second order nucleophilic substitution (S_N2) mechanism leading to β-configuration (85). The majority of UGTs are expressed in the endoplasmatic reticulum of hepatocytes (86). Thus, glucuronidation of xenobiotics can be studied with liver microsomes in the presence of the co-factor UDPGA (87).

Besides CYP450s, UGTs are the most frequently occurring enzymes metabolizing xenobiotics (85). Additionally, carbohydrate mimetics, such as dapagliflozine and other gliflozines are partly glucuronidated without prior phase I metabolism (88). Whether direct glucuronidation plays an important role in the metabolism of E-selectin antagonists has been tested in manuscript 1.

Sulfotransferases (ST)

Sulfotransferases (STs) are responsible for the sulfation of xenobiotics and endogenous substances. STs catalyze the transfer of a hydrophilic sulfuryl group from the co-factor 3'-phosphoadenosine 5'-phospho-sulfate (PAPS) to amine and hydroxyl substrates and can be divided in two subclasses. Cytosolic STs sulfate small compounds, primarily hormones and xenobiotics, whereas membrane bound STs sulfate carbohydrates and proteins (89). To assess the metabolic behavior of carbohydrate mimetics, both classes of STs have to be considered to be active. For example, the natural ligand of E-selectin sLe^x is a substrate of membrane bound ST, but more lipophilic E-selectin antagonists can additionally be a target by cytosolic

STs (90). Therefore, the metabolic stability of E-selectin antagonists by adding PAPS to liver microsomes and liver cytosol was investigated (manuscript 1).

2.5 Elimination

The elimination of xenobiotics, can involve two main routes, i.e. elimination by metabolism/chemical modification or elimination by secretion as unchanged drug. Drug metabolizing enzymes are responsible for the first route, while drug transporters are mediating the latter route. The majority of drugs are inactivated by metabolism, and the metabolites are then excreted. Carbohydrate mimetics, however, are hydrophilic and polar, not prone to metabolism and therefore mostly eliminated as unchanged form by the renal pathway (91, 92). Especially for FimH antagonists (Chapter 8), with the bladder as target, the renal elimination is of major importance (93, 94).

2.5.1 Renal elimination

The kidney's physiological role is not only the excretion of xenobiotics, toxins, and metabolites, but also the control of body fluids, electrolytes and blood pressure as well as the reabsorption of important nutrients. The functional parts of the kidney are nephrons, each consisting of a glomerulus and renal tubuli. Primary urine, which is formed by glomerular filtration, is then further processed by active and passive reabsorption and secretion of water, nutrients, and other compounds (95). The glomerular filtration is size dependent. Larger molecules (> 5 kDa) and molecules bound to albumin or other plasma proteins are not filtered (96, 97). Once the renal tubuli are reached, compounds can be reabsorbed back to circulation either actively or passively. For passive reabsorption, lipophilicity (log$D_{7.4}$) larger than 1 and the rule of 5 parameters by Lipinski are important factors. Membrane bound transporters responsible for active reabsorption are similar to those found in the small intestine (98).

2.6 Toxicity

Toxicity of a substance is quantified by the degree of 'adverse events' it can cause to the organism. Toxic side effects are a major reason for the attrition rate of drugs during clinical phases (13). The field of toxicity is highly complex, since toxic effects are often cell-, organ-, and species specific. To reduce the risk of toxicity during drug development, an integrated implementation of qualified and validated *in silico, in vitro*, and *in vivo* tools is necessary (99). This systematic assessment of the toxic potential of new drug candidates requires a lot of

resources and therefore is rarely executed by academic groups. However, by maximizing the selectivity towards the intended target and by measuring the acute cell toxicity the risk of toxic side effects can be reduced substantially.

2.6.1 Cell Toxicity

Acute cell toxicity is easily assessable and gives primary information on the toxic potential of a compound (100). Several assays are available and the MTT (3-(4, 5-dimethylthiazolyl-2)-2, 5-diphenyltetrazolium bromide) cell toxicity assay was chosen for the PADMET-platform (101). The assay is based on a color reaction of metabolically active HepG2 cells reducing the yellow tetrazolium MTT. As a result, intracellular purple formazan is formed which can be solubilized and quantified by a spectrophotometric readout (102). This assay allows to measure the cell viability at certain time points with or without analyte added.

2.7 References

1. Atherly-John YC, Cunningham SJ, Crain EF. A randomized trial of oral vs intravenous rehydration in a pediatric emergency department. *Arch Pediatr Adolesc Med*. **2002**;*156(12)*:1240-3.
2. Eek D, Krohe M, Mazar I, Horsfield A, Pompilus F, Friebe R, et al. Patient-reported preferences for oral versus intravenous administration for the treatment of cancer: a review of the literature. *Patient Prefer Adherence*. **2016**;*10*:1609-21.
3. Chen ML, Shah V, Patnaik R, Adams W, Hussain A, Conner D, et al. Bioavailability and bioequivalence: an FDA regulatory overview. *Pharm Res*. **2001**;*18(12)*:1645-50.
4. Kerns EH, Di L. Pharmaceutical profiling in drug discovery. *Drug Discov Today*. **2003**;*8(7)*:316-23.
5. Wu CY, Benet LZ. Predicting drug disposition via application of BCS: transport/absorption/ elimination interplay and development of a biopharmaceutics drug disposition classification system. *Pharm Res*. **2005**;*22(1)*:11-23.
6. Amidon GL, Lennernas H, Shah VP, Crison JR. A theoretical basis for a biopharmaceutic drug classification: the correlation of in vitro drug product dissolution and in vivo bioavailability. *Pharm Res.* **1995**;*12(3)*:413-20.
7. Shugarts S, Benet LZ. The role of transporters in the pharmacokinetics of orally administered drugs. *Pharm Res*. **2009**;*26(9)*:2039-54.
8. Shargel LY. Applied Biopharmaceutics and Pharmacokinetics; **1999**.
9. Hu M, Li X. Oral Bioavailability, Wiley; **2011**.
10. Wang J, Urban L, Bojanic D. Maximising use of in vitro ADMET tools to predict in vivo bioavailability and safety. *Expert Opin Drug Metab Toxicol*. **2007**;*3(5)*:641-65.
11. Hodgson J. ADMET--turning chemicals into drugs. *Nat Biotechnol*. **2001**;*19(8)*:722-6.
12. Kola I, Landis J. Can the pharmaceutical industry reduce attrition rates? *Nat Rev Drug Discov*. **2004**;*3(8)*:711-5.
13. Hwang TJ, Carpenter D, Lauffenburger JC, Wang B, Franklin JM, Kesselheim AS. Failure of Investigational Drugs in Late-Stage Clinical Development and Publication of Trial Results. *JAMA Intern Med*. **2016**;*176(12)*:1826-33.
14. Lipinski CA, Lombardo F, Dominy BW, Feeney PJ. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv Drug Deliv Rev*. **2001**;*46(1-3)*:3-26.
15. Veber DF, Johnson SR, Cheng HY, Smith BR, Ward KW, Kopple KD. Molecular properties that influence the oral bioavailability of drug candidates. *J Med Chem*. **2002**;*45(12)*:2615-23.
16. Leeson PD, Springthorpe B. The influence of drug-like concepts on decision-making in medicinal chemistry. *Nat Rev Drug Discov*. **2007**;*6(11)*:881-90.
17. Keseru GM, Makara GM. The influence of lead discovery strategies on the properties of drug candidates. *Nat Rev Drug Discov*. **2009**;*8(3)*:203-12.
18. Avdeef A. Absorption and drug development: Solubility, permeability, and charge state. 2nd. Edition; **2012**.
19. Waring MJ. Lipophilicity in drug discovery. *Expert Opin Drug Discov*. **2010**;*5(3)*:235-48.
20. Valko K, Nunhuck S, Bevan C, Abraham MH, Reynolds DP. Fast gradient HPLC method to determine compounds binding to human serum albumin. Relationships with octanol/water and immobilized artificial membrane lipophilicity. *J Pharm Sci*. **2003**;*92(11)*:2236-48.
21. Lewis DF, Dickins M. Baseline lipophilicity relationships in human cytochromes P450 associated with drug metabolism. *Drug Metab Rev*. **2003**;*35(1)*:1-18.
22. Johnson TW, Dress KR, Edwards M. Using the Golden Triangle to optimize clearance and oral absorption. *Bioorg Med Chem Lett*. **2009**;*19(19)*:5560-4.
23. Bickerton GR, Paolini GV, Besnard J, Muresan S, Hopkins AL. Quantifying the chemical beauty of drugs. *Nat Chem*. **2012**;*4(2)*:90-8.

24. Ghose AK, Crippen GM. Atomic physicochemical parameters for three-dimensional-structure-directed quantitative structure-activity relationships. 2. Modeling dispersive and hydrophobic interactions. *J Chem Inf Comput Sci*. **1987**;*27(1)*:21-35.
25. Tetko IV, Bruneau P. Application of ALOGPS to predict 1-octanol/water distribution coefficients, logP, and logD, of AstraZeneca in-house database. *J Pharm Sci*. **2004**;*93(12)*:3103-10.
26. Moriguchi I, Hirono S, Liu Q, Nakagome I, Matsushita Y. Simple method of calculating octanol water partition-coeffizient. *Chemical and Pharmaceutical Bulletin*. **1992**;*40(1)*:4.
27. Mirrlees MS, Moulton SJ, Murphy CT, Taylor PJ. Direct measurement of octanol-water partition coefficients by high-pressure liquid chromatography. *J Med Chem*. **1976**;*19(5)*:615-9.
28. Ming X, Han SY, Qi ZC, Sheng D, Lian HZ. Chromatographic retention prediction and octanol-water partition coefficient determination of monobasic weak acidic compounds in ion-suppression reversed-phase liquid chromatography using acids as ion-suppressors. *Talanta*. **2009**;*79(3)*:752-61.
29. Wohnsland F, Faller B. High-throughput permeability pH profile and high-throughput alkane/water log P with artificial membranes. *J Med Chem*. **2001**;*44(6)*:923-30.
30. Manallack DT. The pK(a) Distribution of Drugs: Application to Drug Discovery. *Perspect Medicin Chem*. **2007**;*1*:25-38.
31. Clarke FH, Cahoon NM. Ionization constants by curve fitting: determination of partition and distribution coefficients of acids and bases and their ions. *J Pharm Sci*. **1987**;*76(8)*:611-20.
32. Reijenga J, van Hoof A, van Loon A, Teunissen B. Development of Methods for the Determination of pKa Values. *Anal Chem Insights*. **2013**;*8*:53-71.
33. Bezencon J, Wittwer MB, Cutting B, Smiesko M, Wagner B, Kansy M, et al. pKa determination by (1)H NMR spectroscopy - an old methodology revisited. *J Pharm Biomed Anal*. **2014**;*93*:147-55.
34. Lipinski CA. Drug-like properties and the causes of poor solubility and poor permeability. *J Pharmacol Toxicol Methods*. **2000**;*44(1)*:235-49.
35. Streng WH, Hsi SK, Helms PE, Tan HG. General treatment of pH-solubility profiles of weak acids and bases and the effects of different acids on the solubility of a weak base. *J Pharm Sci*. **1984**;*73(12)*:1679-84.
36. Black SN, Collier EA, Davey RJ, Roberts RJ. Structure, solubility, screening, and synthesis of molecular salts. *J Pharm Sci*. **2007**;*96(5)*:1053-68.
37. Mota FL, Carneiro AP, Queimada AJ, Pinho SP, Macedo EA. Temperature and solvent effects in the solubility of some pharmaceutical compounds: Measurements and modeling. *Eur J Pharm Sci*. **2009**;*37(3-4)*:499-507.
38. Kobayashi Y, Ito S, Itai S, Yamamoto K. Physicochemical properties and bioavailability of carbamazepine polymorphs and dihydrate. *Int J Pharm*. **2000**;*193(2)*:137-46.
39. Aguiar AJ, Krc J, Jr., Kinkel AW, Samyn JC. Effect of polymorphism on the absorption of chloramphenicol from chloramphenicol palmitate. *J Pharm Sci*. **1967**;*56(7)*:847-53.
40. Bevan CD, Lloyd RS. A high-throughput screening method for the determination of aqueous drug solubility using laser nephelometry in microtiter plates. *Anal Chem*. **2000**;*72(8)*:1781-7.
41. Box KJ, Volgyi G, Baka E, Stuart M, Takacs-Novak K, Comer JE. Equilibrium versus kinetic measurements of aqueous solubility, and the ability of compounds to supersaturate in solution--a validation study. *J Pharm Sci*. **2006**;*95(6)*:1298-307.
42. Baka E, Comer JE, Takacs-Novak K. Study of equilibrium solubility measurement by saturation shake-flask method using hydrochlorothiazide as model compound. *J Pharm Biomed Anal*. **2008**;*46(2)*:335-41.
43. Product Development Under the Animal Rule - Guidance for Industry [press release]. FDA2015.
44. Thackaberry EA, Wang X, Schweiger M, Messick K, Valle N, Dean B, et al. Solvent-based formulations for intravenous mouse pharmacokinetic studies: tolerability and recommended solvent dose limits. *Xenobiotica*. **2014**;*44(3)*:235-41.

45. Horter D, Dressman JB. Influence of physicochemical properties on dissolution of drugs in the gastrointestinal tract. *Adv Drug Deliv Rev*. **2001**;*46(1-3)*:75-87.
46. Schmidt LE, Dalhoff K. Food-drug interactions. *Drugs*. **2002**;*62(10)*:1481-502.
47. Hopkins A. The pattern of gastric emptying: a new view of old results. *J Physiol*. **1966**;*182(1)*:144-9.
48. Camilleri M, Malagelada JR, Brown ML, Becker G, Zinsmeister AR. Relation between antral motility and gastric emptying of solids and liquids in humans. *Am J Physiol*. **1985**;*249(5 Pt 1)*:G580-5.
49. Dykes CW, Kay J. Conversion of pepsinogen into pepsin is not a one-step process. *Biochem J*. **1976**;*153(1)*:141-4.
50. Jacobson ED. The gastrointestinal circulation. *Annu Rev Physiol*. **1968**;*30*:133-46.
51. Di L, Kerns EH, Carter GT. Drug-like property concepts in pharmaceutical design. *Curr Pharm Des*. **2009**;*15(19)*:2184-94.
52. DeSesso JM, Jacobson CF. Anatomical and physiological parameters affecting gastrointestinal absorption in humans and rats. *Food Chem Toxicol*. **2001**;*39(3)*:209-28.
53. Sugano K, Kansy M, Artursson P, Avdeef A, Bendels S, Di L, et al. Coexistence of passive and carrier-mediated processes in drug transport. *Nat Rev Drug Discov*. **2010**;*9(8)*:597-614.
54. van De Waterbeemd H, Smith DA, Beaumont K, Walker DK. Property-based design: optimization of drug absorption and pharmacokinetics. *J Med Chem*. **2001**;*44(9)*:1313-33.
55. Kansy M, Senner F, Gubernator K. Physicochemical high throughput screening: parallel artificial membrane permeation assay in the description of passive absorption processes. *J Med Chem*. **1998**;*41(7)*:1007-10.
56. Avdeef A, Bendels S, Di L, Faller B, Kansy M, Sugano K, et al. PAMPA--critical factors for better predictions of absorption. *J Pharm Sci*. **2007**;*96(11)*:2893-909.
57. International Transporter C, Giacomini KM, Huang SM, Tweedie DJ, Benet LZ, Brouwer KL, et al. Membrane transporters in drug development. *Nat Rev Drug Discov*. **2010**;*9(3)*:215-36.
58. Amidon GL, W S. Membrane Transporters as Drug Targets. New York: Kluwer Academic/Plenum Publisher; **1999**.
59. Nagy I, Toth B, Gaborik Z, Erdo F, Krajcsi P. Membrane Transporters in Physiological Barriers of Pharmacological Importance. *Curr Pharm Des*. **2016**;*22(35)*:5347-72.
60. Zhang Y, Sun J, Sun Y, Wang Y, He Z. Prodrug design targeting intestinal PepT1 for improved oral absorption: design and performance. *Curr Drug Metab*. **2013**;*14(6)*:675-87.
61. Hubatsch I, Ragnarsson EG, Artursson P. Determination of drug permeability and prediction of drug absorption in Caco-2 monolayers. *Nat Protoc*. **2007**;*2(9)*:2111-9.
62. Hou T, Wang J, Zhang W, Xu X. ADME evaluation in drug discovery. 7. Prediction of oral absorption by correlation and classification. *J Chem Inf Model*. **2007**;*47(1)*:208-18.
63. Imai T, Imoto M, Sakamoto H, Hashimoto M. Identification of esterases expressed in Caco-2 cells and effects of their hydrolyzing activity in predicting human intestinal absorption. *Drug Metab Dispos*. **2005**;*33(8)*:1185-90.
64. Smith DA, Beaumont K, Maurer TS, Di L. Volume of Distribution in Drug Design. *J Med Chem*. **2015**;*58(15)*:5691-8.
65. Benet LZ, Hoener BA. Changes in plasma protein binding have little clinical relevance. *Clin Pharmacol Ther*. **2002**;*71(3)*:115-21.
66. Smith DA, Di L, Kerns EH. The effect of plasma protein binding on in vivo efficacy: misconceptions in drug discovery. *Nat Rev Drug Discov*. **2010**;*9(12)*:929-39.
67. Heuberger J, Schmidt S, Derendorf H. When is protein binding important? *J Pharm Sci*. **2013**;*102(9)*:3458-67.
68. Zhang F, Xue J, Shao J, Jia L. Compilation of 222 drugs' plasma protein binding data and guidance for study designs. *Drug Discov Today*. **2012**;*17(9-10)*:475-85.
69. Ruiz-Garcia A, Bermejo M, Moss A, Casabo VG. Pharmacokinetics in drug discovery. *J Pharm Sci*. **2008**;*97(2)*:654-90.

70. Singh SS, Mehta J. Measurement of drug-protein binding by immobilized human serum albumin-HPLC and comparison with ultrafiltration. *J Chromatogr B Analyt Technol Biomed Life Sci*. **2006**;*834(1-2)*:108-16.
71. Kariv I, Cao H, Oldenburg KR. Development of a high throughput equilibrium dialysis method. *J Pharm Sci*. **2001**;*90(5)*:580-87.
72. Malarkey DE, Johnson K, Ryan L, Boorman G, Maronpot RR. New insights into functional aspects of liver morphology. *Toxicol Pathol*. **2005**;*33(1)*:27-34.
73. David Josephy P, Peter Guengerich F, Miners JO. "Phase I and Phase II" drug metabolism: terminology that we should phase out? *Drug Metab Rev*. **2005**;*37(4)*:575-80.
74. Rautio J, Kumpulainen H, Heimbach T, Oliyai R, Oh D, Jarvinen T, et al. Prodrugs: design and clinical applications. *Nat Rev Drug Discov*. **2008**;*7(3)*:255-70.
75. Nelson SD, Gordon WP. Mammalian drug metabolism. *J Nat Prod*. **1983**;*46(1)*:71-8.
76. Guengerich FP. Cytochrome P450s and other enzymes in drug metabolism and toxicity. *AAPS J*. **2006**;*8(1)*:E101-11.
77. Williams PA, Cosme J, Vinkovic DM, Ward A, Angove HC, Day PJ, et al. Crystal structures of human cytochrome P450 3A4 bound to metyrapone and progesterone. *Science*. **2004**;*305(5684)*:683-6.
78. Zanger UM, Schwab M. Cytochrome P450 enzymes in drug metabolism: regulation of gene expression, enzyme activities, and impact of genetic variation. *Pharmacol Ther*. **2013**;*138(1)*:103-41.
79. Crivori P, Poggesi I. Computational approaches for predicting CYP-related metabolism properties in the screening of new drugs. *Eur J Med Chem*. **2006**;*41(7)*:795-808.
80. Obach RS. Prediction of human clearance of twenty-nine drugs from hepatic microsomal intrinsic clearance data: An examination of in vitro half-life approach and nonspecific binding to microsomes. *Drug Metab Dispos*. **1999**;*27(11)*:1350-9.
81. Hosokawa M. Structure and catalytic properties of carboxylesterase isozymes involved in metabolic activation of prodrugs. *Molecules*. **2008**;*13(2)*:412-31.
82. Satoh T, Taylor P, Bosron WF, Sanghani SP, Hosokawa M, La Du BN. Current progress on esterases: from molecular structure to function. *Drug Metab Dispos*. **2002**;*30(5)*:488-93.
83. Li B, Sedlacek M, Manoharan I, Boopathy R, Duysen EG, Masson P, et al. Butyrylcholinesterase, paraoxonase, and albumin esterase, but not carboxylesterase, are present in human plasma. *Biochem Pharmacol*. **2005**;*70(11)*:1673-84.
84. Imai T. Human carboxylesterase isozymes: catalytic properties and rational drug design. *Drug Metab Pharmacokinet*. **2006**;*21(3)*:173-85.
85. Rowland A, Miners JO, Mackenzie PI. The UDP-glucuronosyltransferases: their role in drug metabolism and detoxification. *Int J Biochem Cell Biol*. **2013**;*45(6)*:1121-32.
86. Dong D, Ako R, Hu M, Wu B. Understanding substrate selectivity of human UDP-glucuronosyltransferases through QSAR modeling and analysis of homologous enzymes. *Xenobiotica*. **2012**;*42(8)*:808-20.
87. Aprile S, Del Grosso E, Grosa G. In vitro and in vivo phase II metabolism of combretastatin A-4: evidence for the formation of a sulphate conjugate metabolite. *Xenobiotica*. **2009**;*39(2)*:148-61.
88. Obermeier M, Yao M, Khanna A, Koplowitz B, Zhu M, Li W, et al. In vitro characterization and pharmacokinetics of dapagliflozin (BMS-512148), a potent sodium-glucose cotransporter type II inhibitor, in animals and humans. *Drug Metab Dispos*. **2010**;*38(3)*:405-14.
89. Chapman E, Best MD, Hanson SR, Wong CH. Sulfotransferases: structure, mechanism, biological activity, inhibition, and synthetic utility. *Angew Chem Int Ed Engl*. **2004**;*43(27)*:3526-48.
90. Imai Y, Lasky LA, Rosen SD. Sulphation requirement for GlyCAM-1, an endothelial ligand for L-selectin. *Nature*. **1993**;*361(6412)*:555-7.
91. Varma MV, Feng B, Obach RS, Troutman MD, Chupka J, Miller HR, et al. Physicochemical determinants of human renal clearance. *J Med Chem*. **2009**;*52(15)*:4844-52.

92. Ernst B, Magnani JL. From carbohydrate leads to glycomimetic drugs. *Nat Rev Drug Discov*. **2009**;*8(8)*:661-77.
93. Aronson M, Medalia O, Schori L, Mirelman D, Sharon N, Ofek I. 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*. **1979**;*139(3)*:329-32.
94. Kleeb S, Pang L, Mayer K, Eris D, Sigl A, Preston RC, et al. FimH antagonists: bioisosteres to improve the in vitro and in vivo PK/PD profile. *J Med Chem*. **2015**;*58(5)*:2221-39.
95. Lote CJ. Principles of Renal Physiology: Springer, New York, NY; **2012**.
96. Chang RL, Ueki IF, Troy JL, Deen WM, Robertson CR, Brenner BM. Permselectivity of the glomerular capillary wall to macromolecules. II. Experimental studies in rats using neutral dextran. *Biophys J*. **1975**;*15(9)*:887-906.
97. Tencer J, Frick IM, Oquist BW, Alm P, Rippe B. Size-selectivity of the glomerular barrier to high molecular weight proteins: upper size limitations of shunt pathways. *Kidney Int*. **1998**;*53(3)*:709-15.
98. Yin J, Wang J. Renal drug transporters and their significance in drug-drug interactions. *Acta Pharm Sin B*. **2016**;*6(5)*:363-73.
99. Dambach DM, Misner D, Brock M, Fullerton A, Proctor W, Maher J, et al. Safety Lead Optimization and Candidate Identification: Integrating New Technologies into Decision-Making. *Chem Res Toxicol*. **2016**;*29(4)*:452-72.
100. Ekwall B. Screening of toxic compounds in mammalian cell cultures. *Ann N Y Acad Sci*. **1983**;*407*:64-77.
101. Hayon T, Dvilansky A, Shpilberg O, Nathan I. Appraisal of the MTT-based assay as a useful tool for predicting drug chemosensitivity in leukemia. *Leuk Lymphoma*. **2003**;*44(11)*:1957-62.
102. Huang R, Southall N, Cho MH, Xia M, Inglese J, Austin CP. Characterization of diversity in toxicity mechanism using in vitro cytotoxicity assays in quantitative high throughput screening. *Chem Res Toxicol*. **2008**;*21(3)*:659-67.

3 E-selectin antagonists

3.1 Leukocytes adhesion cascade

The recruitment of leukocytes to the side of inflammation is crucial for host defence and repair mechanisms. This process is highly regulated due to its importance and the potential of self-damage, triggered by an excessive reaction of the immune system. In Figure 3.1, the leukocyte adhesion cascade is illustrated (1). Initial inflammatory stimuli are expressed by pathogens during infection or by cell damages. Infective pathogens are recognised by their expression of exogenous *pathogen-associated molecular pattern molecules* (PAMP's). Non-infectious cell damage leads to the leak of intracellular *damage associated molecular pattern molecules* (DAMP's), such as DNA and RNA fragments (2). PAMP's and DAMP's are sensed by innate pattern recognition receptors (PRRs), located on sentinel cells of the immune system, such as macrophages and dendritic cells (3). The interaction with PRRs leads to an inflammatory response orchestrated by proinflammatory cytokines such as tumour necrosis factor (TNF), interleukins (IL)-1, and IL-6 causing redness, swelling, heat, pain and loss of tissue function as symptoms of the inflammatory response (4).

Furthermore, the presence of these inflammatory stimuli in a tissue upregulate cell adhesion molecules (CAMs) in the endothelial cells of the blood vessels. A member of CAMs are selectins, which are presented on the luminal membrane in order to recruit leukocytes to the site of inflammation (5).

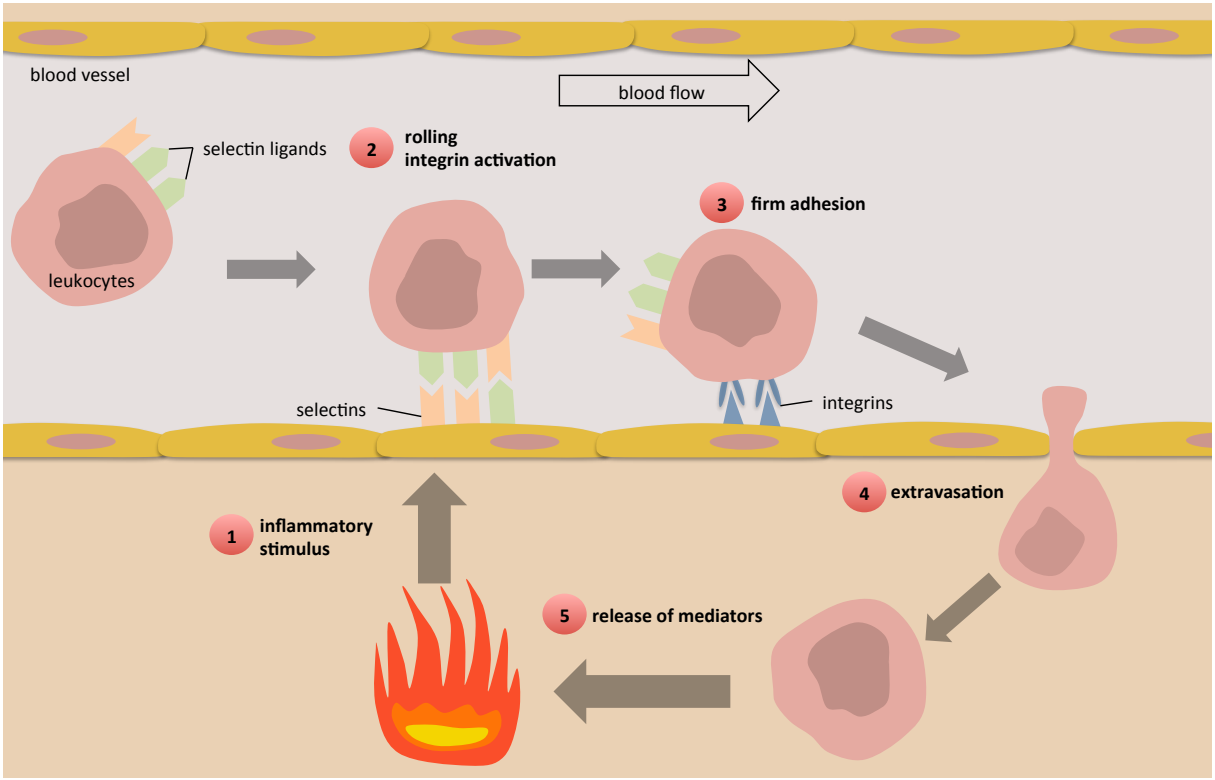


Figure 3.1 Schematic illustration of inflammation induced leukocyte extravasation to affected tissues, initiated by selectins. An inflammatory stimulus leads to the upregulation of selectins on the endothelial surface of blood vessels adjacent to the inflammatory stimulus (1). Free flowing leukocytes adhere via their selectin ligands, resulting in their rolling along the vessel wall. This rolling is associated with a decrease of their flowing velocity and integrin activation (2). This leads to firm attachment of leukocytes (3), initiating extravasation (4) to the adjacent tissue. Once present in the inflammatory tissue, leukocytes produce new inflammatory mediators (5). Adapted from Springer *et al.* (1).

3.2 Selectins

Selectins are C-type lectins (calcium dependent carbohydrate binding proteins), belonging to the type 1 transmembrane protein family. The selectin family consists of three subtypes, E-, P- and L-selectin, which are involved in cell adhesion and cell trafficking to specific tissues (6, 7). E-, P- and L-selectin share a common overall structure shown in

Figure 3.2. The calcium dependent N-terminal lectin domain is responsible for binding the selectin ligands and is located extracellular, pointing towards the vascular lumen. For P- and L-selectin, the lectin domain is composed of 100 amino acid residues whereas in E-selectin, consists of 120 amino acid residues. The lectin domain is followed by an epidermal growth factor like domain (EGF). Together, the two domains form the minimum requirement for selectin binding activity (8). The binding domain is connected via a varying number of short consensus repeats (SCR) to the membrane anchor (9). The intracellular C-terminus forms a 17-

35 amino acid long cytosolic tail responsible for cell signalling and forms interactions with the cytoskeleton (10, 11).

In a first phase of inflammation, P-selectin is responsible for the initiation of the leukocyte diapedesis. P-selectin is stored in granules (Weibel-Palade bodies) of vascular cells, allowing a quick release after stimulation (12). In a later phase, E-selectin is newly synthesized via the NFκB-pathway (13) upon induction by TNFα or IL-1 and then released to the cell surface. In contrast to P-selectin, E-selectin is not stored in advance and hence, its expression on the cell surface is delayed approximately 4-6 hours after inflammatory stimulation (14). The third member of the selectin family, L-selectin, is exclusively presented on lymphocytes and is involved in lymphocyte migration to secondary lymphoid organs by binding to the corresponding ligand on high endothelial cells (15).

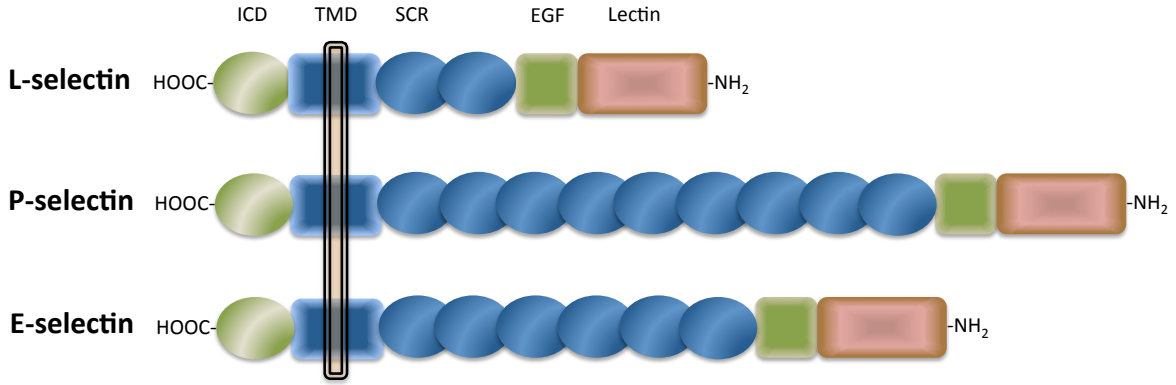


Figure 3.2 L-, P- and E-selectin. Each member of the selectin family consists of a intracellular domain (ICD), transmembrane domain (TMD), a different number of short census repeats (SCR; 2 in L-selectin, 9 in P-selectin and 6 in E-selectin), an epidermal growth factor like domain (EGF) and the lectin domain, which is responsible for the recognition of sLe^x containing structures.

3.3 The role of E-selectin in inflammatory diseases

Inflammation and the resulting leukocyte extravasation are crucial to protect the tissues from pathogenic infection or for the healing process after mechanical, thermic or chemical tissue damage. Nevertheless, excessive and faulty regulated leukocyte extravasation can lead to a massive destruction of healthy tissue (9).

Therefore, E-selectin is involved in different pathophysiological processes with an inflammatory component, such as asthma bronchiale or rheumatoid arthritis due excessive leukocyte extravasation (9, 16, 17). E-selectin overexpression is furthermore involved in the pathology of cardiovascular diseases like arteriosclerosis, myocardial infarction, reperfusion injury and hypertension (18-21). In various cancers, E-selectin is involved in the formation of metastasis of cancer cells by inducing the extravasation of metastatic cancer cells in other tissues, for example in dormant breast cancer micrometastases (22-24). Selectins are also associated with sickle cell adhesion in sickle cell disease (SCD) patients leading to vaso-occlusive crisis (VOC), a disseminated, extremely painful thrombosis in small vessels (25).

3.4 E-selectin antagonists

The carbohydrate epitope recognized by E-selectin is the tetrasaccharide sialyl Lewis^x (sLe^x, **1**) as part of glycan structures of selectin ligands present on leukocytes (26). sLe^x binds to the carbohydrate recognition domain (CRD) of E-selectin by forming a highly coordinated hydrogen bond network, typical for carbohydrate lectin interactions. Mainly the hydroxyl groups of the fucose and galactose moiety, as well as the carboxylic acid of sialic acid are involved in the binding network as shown in Figure 3.3 (27). The binding of sLe^x is predominantly entropically driven, since the binding of the rigid sLe^x leads to the release of numerous water molecules into bulk water. The low binding affinity (800 µM) is due to an enthalpy penalty caused by desolvation costs which is not compensated by the enthalpy gain from the formation of the hydrogen bond network (28).

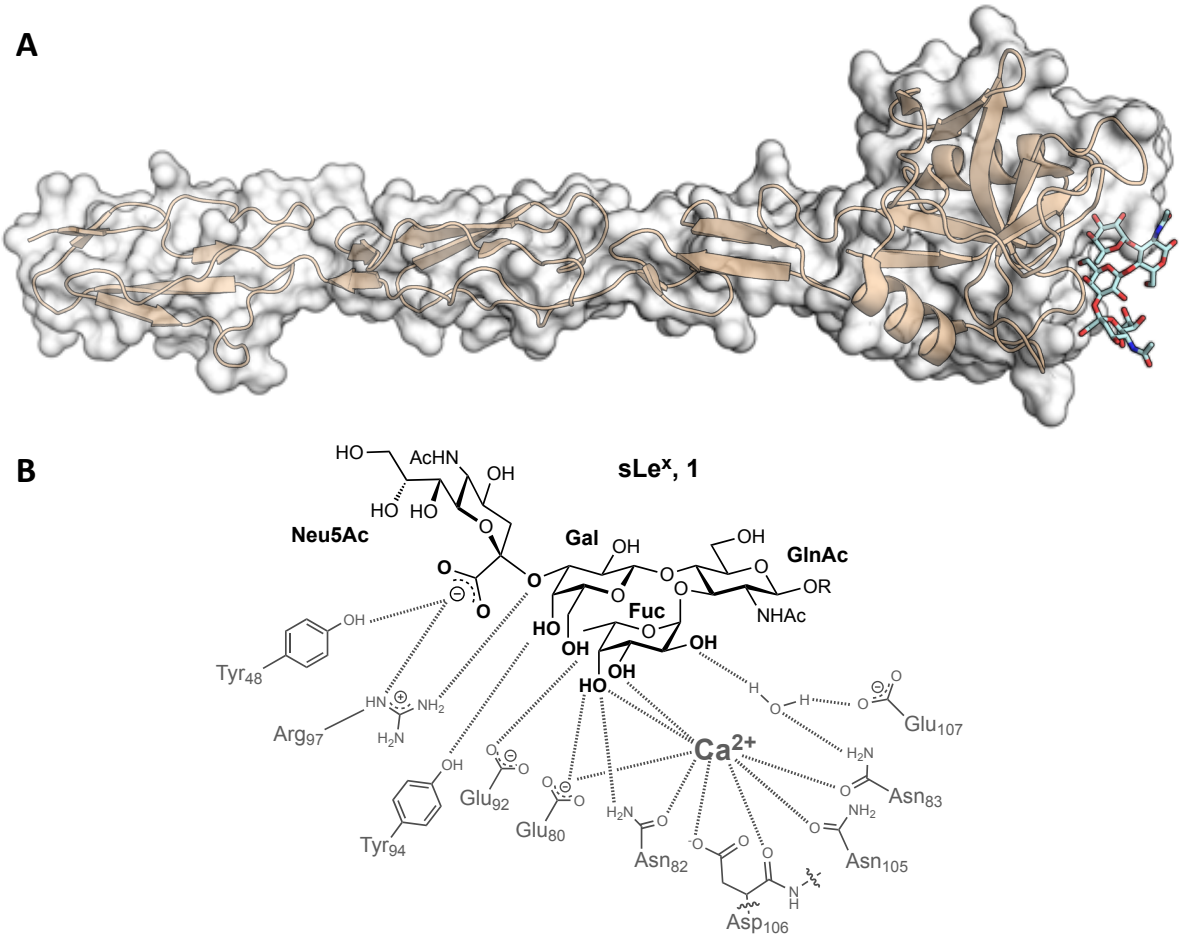


Figure 3.3 A Crystal structure of sLe^x binding to E-selectin with EGF and two SCR domains (PDB: 4CSY). **B** The binding mode of sLe^x (**1**) and E-selectin. sLe^x consists of *N*-acetyl-neuraminic acid (sialic acid, Neu5Ac), galactose (Gal), glucosamine (Glc*N*Ac) and fucose (Fuc). Beneficial hydrogen bonds between E-selectin and the ligand are represented in dotted lines. Adapted from Binder *et al*.(27).

E-selectin antagonists are designed starting from the natural ligand sLe^x (**1**). In order to achieve drug-like properties, not only increased affinity is required but additionally sufficient plasma half-life and if possible, oral bioavailability is aspired. sLe^x (**1**) has very poor pharmacokinetic properties, failing Lipinski's rule of five and Veber's rule in every respect (29, 30). Therefore, for the development of E-selectin antagonists, the focus is on reduction of the complexity and hydrophilicity without losing the pre-organization of the core structure. In the case of the pan-selectin inhibitor Rivipansel (Figure 3.4, **2**), sufficient plasma half-life was achieved by adding a suramin fragment (31). Rivipansel does not only have a good plasma half-life, but the suramin fragment improves binding to P-selectin and thus leads to pan-selectin antagonism. It successfully completed phase 2 clinical trials for the treatment of vaso-occlusive crisis (VOC) in sickle cell disease patients (32). Whereas Rivipansel is administered intravenously, an oral treatment would be preferable for chronic inflammatory diseases with a strong correlation to E-selectin overexpression. Reducing the hydrophilicity and complexity of the core structure led

to CGP69669A (**3**) (33). Compound **3** exhibits an improved affinity (60 µM for E-selectin), a reduced molecular weight and polarity, and is therefore a good starting point for further lead optimisation processes.

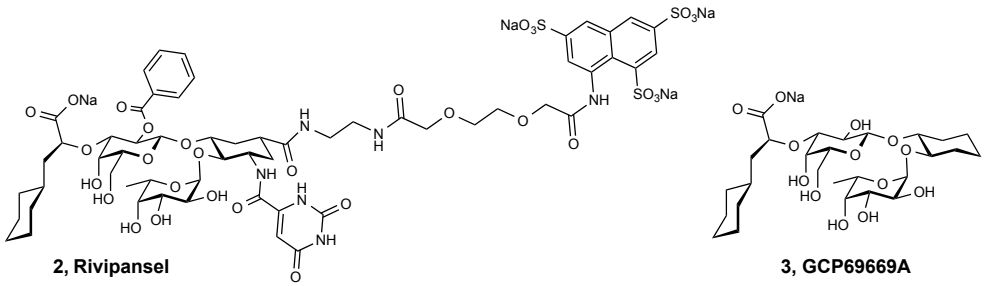


Figure 3.4 E-selectin antagonists Rivipansel (**2**) and GCP69669A (**3**).

3.5 Results and discussion: E-selectin antagonists

A large variety of E-selectin antagonists were synthesized with the goal to improve their physicochemical and pharmacokinetic properties and analysed in the PADMET-Platform (Figure 3.5). The starting point of the modifications is compound **3**, a well-known mimic of the tetrasaccharide sLe^x (**1**) (33). The C4- and C5-hydroxyl groups of galactose and the three hydroxyl groups of fucose are strongly involved in the hydrogen bond network upon E-selectin binding and their modification does not lead to higher affinity (27). Therefore, most modifications were done in order to replace the hydrophilic carboxylic acid by either ester prodrugs (manuscript 1) or amides (manuscript 2). Other modifications of the lactic acid moiety, the galactose 2-position, the lipophilic linker and the fucose moiety addressed specific properties such as affinity, stability or lipophilicity issues.

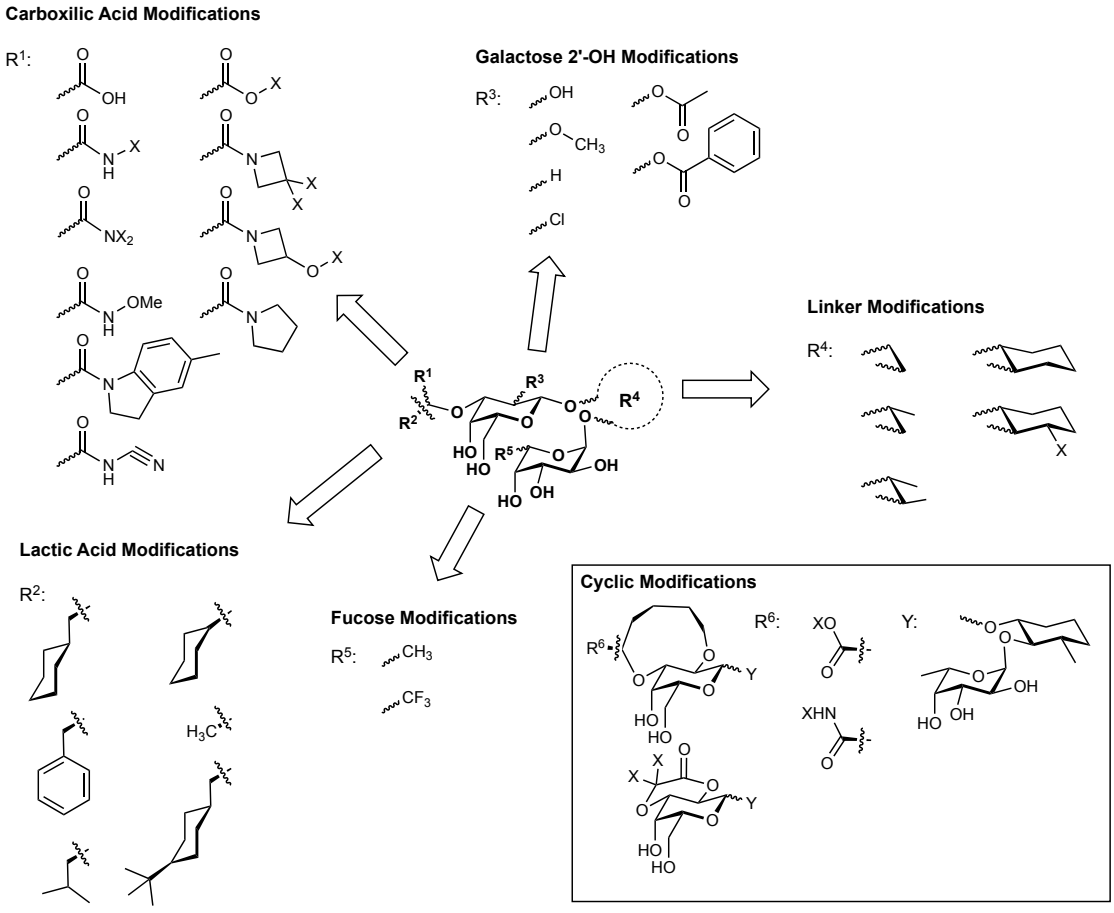


Figure 3.5 Various modifications of E-selectin antagonist analyzed for physicochemical and pharmacokinetic parameters in the PADMET-Platform. The best antagonists were obtained when the various modifications were combined in one molecule. X is either H or an aliphatic or aromatic residue.

3.5.1 Physicochemical properties

The influence of the polar surface area

The polar surface areas (PSA) of carbohydrates are large due to the numerous consecutive H-bond donating hydroxyl-groups. Combining the PSA of the galactose and fucose moiety already results in a PSA of 167.5 Å^2, which is above the threshold allowing oral bioavailability according to Veber *et al*. (29). Additional PSA is added from the carboxylic acid moiety or bioisosteres thereof and results in a total PSA of 180 to 200 Å^2 for most selectin antagonists. Since the hydrophilic core structure is similar for all compounds and the additional moieties are of lipophilic character, a correlation between PSA and clog*P* does not exist (Figure 3.6)

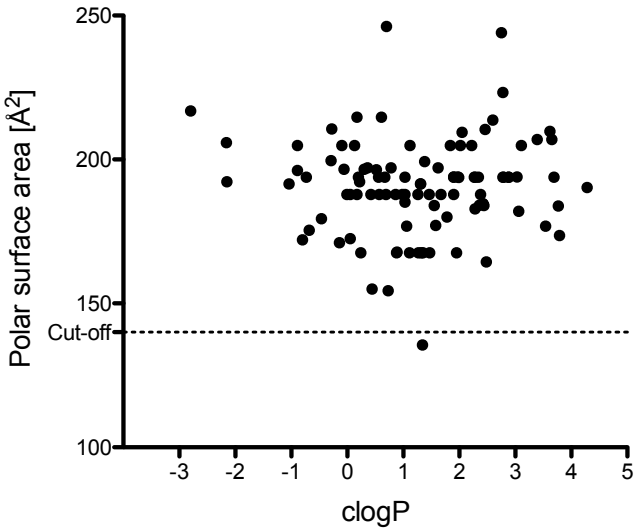
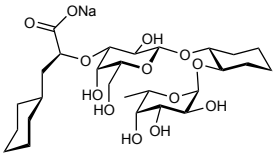
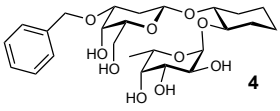


Figure 3.6 Plot of PSA [Å^2] versus clog*P* of E-selectin antagonists. PSA and clog*P* were calculated using MarvinSketch Version 16.10.10.0. Polar surface area was calculated using a fragment approach according to Ertl *et al.* and clog*P* was calculated with the consensus method. The dashed line indicates the cut-off value of 140 Å^2 used by Veber *et al.* to predict oral bioavailability (29, 34).

Consequent minimization of the PSA led to compound **4**. Even though this compound has a PSA below the threshold of 140 Å^2, the compound does not show sufficient permeability in PAMPA and Caco-2 cells as shown in comparison with compound **3** (Table 3.1). Furthermore, the affinity is strongly reduced due to the missing lactic acid residue and therefore, this approach was no longer followed.

Table 3.1 Reduction of PSA of E-selectin antagonists. PSA was calculated using MarvinSketch version 16.10.10.0 and using a fragment approach according to Ertl *et al.* (34). Log$D_{7.4}$ (octanol-water distribution coefficient) values were determined at pH 7.4 by a miniaturized shake flask procedure in sextuplicate. Effective permeability (P_e), i.e. the diffusion through an artificial membrane was determined by the parallel artificial membrane permeability assay (PAMPA) in quadruplicate at pH 7.4 and thresholds for moderate and high absorption potential are -6.3 and -5.7, respectively (35). Errors are given as the standart deviation.

Compound	MW [g/mol]	PSA [Å^2]	log$D_{7.4}$	PAMPA logP_e [log cm/s]
 **GCP69669A, 3**	600.6	196.4	< 1.5	< -10
 **4**	512.6	135.6	1.5 ± 0.2	-7.1 ± 0.1

The influence of lipophilicity on permeability

The starting point for the lead optimization process is **3**, exhibiting a molecular weight (MW) of 600.68 Da and an experimental log$D_{7.4}$ below -1.5. Due to the hydrophilic core structure formed by the galactose and fucose moieties, an increased lipophilicity can mainly be realized by the addition of peripheral lipophilic modifications. This trend can be visualized by plotting the experimental log$D_{7.4}$ against the MW of the molecules (Figure 3.7). For orally bioavailable, marketed drugs, the MW threshold of 500 Da has been the most often violated rule of 5 (30), indicating that slightly increased MWs still allow sufficient permeability (36). Nevertheless, lipophilic modifications have to be carefully evaluated.

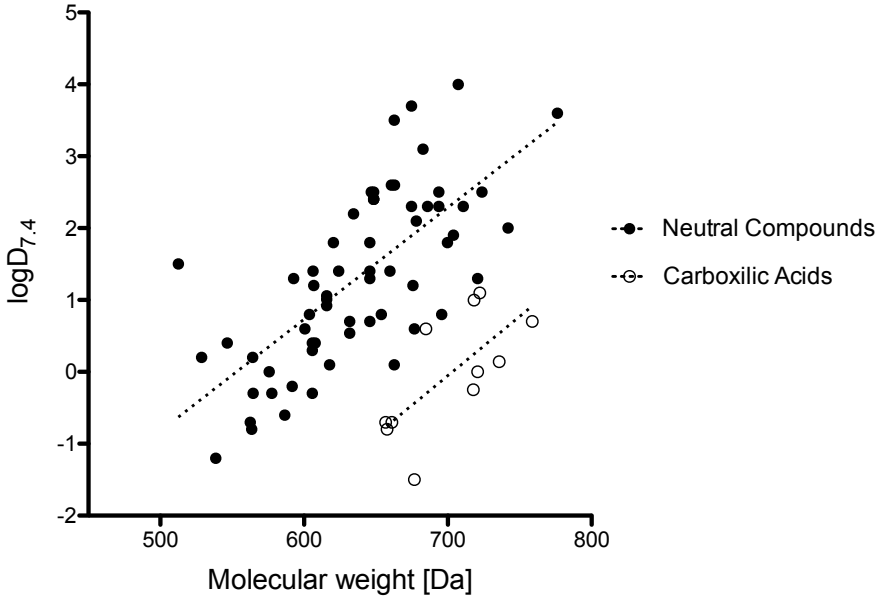


Figure 3.7 Plot of MW [Da] versus log$D_{7.4}$. Data represents the mean only with a maximal standard deviation of 0.2. Log$D_{7.4}$ values were determined at pH 7.4 by a miniaturized shake flask procedure in sextuplicate, compounds with an experimental log$D_{7.4}$ value below -1.5 are excluded from the graph due to the detection limits of the method. Uncharged (filled dots) and carboxylic acids (hollow dots) are divided in two subgroups due to the charge present at pH 7.4. The dashed lines represent the linear regression between MW and log$D_{7.4}$. Both slopes of the regression lines are significantly non-zero with $p < 0.0001$ for neutral compounds and $p = 0.02$ for carboxylic acids.

The lipophilicity of a molecule is the driving force for integration into the phospholipid bilayer of a membrane as the first step to permeate the membrane (37). Especially for carbohydrate mimetics, sufficient lipophilic moieties have to be added to ensure passive permeation by the transcellular route. In Figure 3.8, the strong correlation of the passive permeability measured in a parallel artificial membrane permeation assay (PAMPA) and the lipophilicity of E-selectin antagonists is shown. All molecules showing a moderate or high absorption potential in PAMPA have a log$D_{7.4}$ of 2.0 or above.

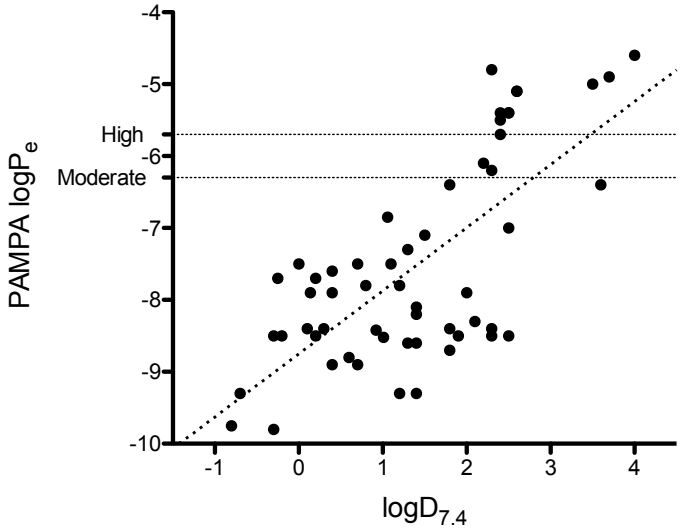


Figure 3.8 Plot of log$D_{7.4}$ versus effective permeation logP_e. Data represents the mean only with a maximal standard deviation of 0.2 for the log$D_{7.4}$ and 0.2 for logP_e for molecules within the range of the thresholds. log$D_{7.4}$ values were determined at pH 7.4 by a miniaturized shake flask procedure in sextuplicate; compounds with an experimental log$D_{7.4}$ below -1.5 are excluded from the graph due to the detection limits of the method. log*P*e (effective permeability) was determined by diffusion through an artificial membrane in PAMPA in quadruplicate at pH 7.4 and compounds with a measured logP_e of -10 or below are excluded due to the detection limits of the assay. The thresholds for high and moderate absorption potential are indicated according to Avdeef *et al.* (35). The dashed slope represents the linear regression between log$D_{7.4}$ and logP_e and is significantly non-zero with $p < 0.0001$.

The permeability of promising compounds was further measured in a Caco-2 cell based assay to evaluate, if active transport is involved during the absorption process. The permeation of the compounds in the absorptive (AB, uptake, filled dots) and secretory (BA, efflux, hollow dots) direction are plotted against the lipophilicity of the molecule in Figure 3.9. Increasing the lipophilicity of E-selectin antagonists leads predominantly to an increase of the efflux, with an efflux ratio (efflux divided by uptake) up to 50 for certain molecules. The main reason for this observation is probably due to the increased incorporation into the membrane of lipophilic compounds, leading to higher activity of P-gp, a major efflux transporter recognizing predominantly xenobiotics in the membrane (38). Besides the majority of E-selectin antagonists showing active efflux, some amidic antagonists showed active uptake. These results are further discussed in manuscript 2, and are pointing out the importance of a cell-based assay implemented in the PADMET-platform.

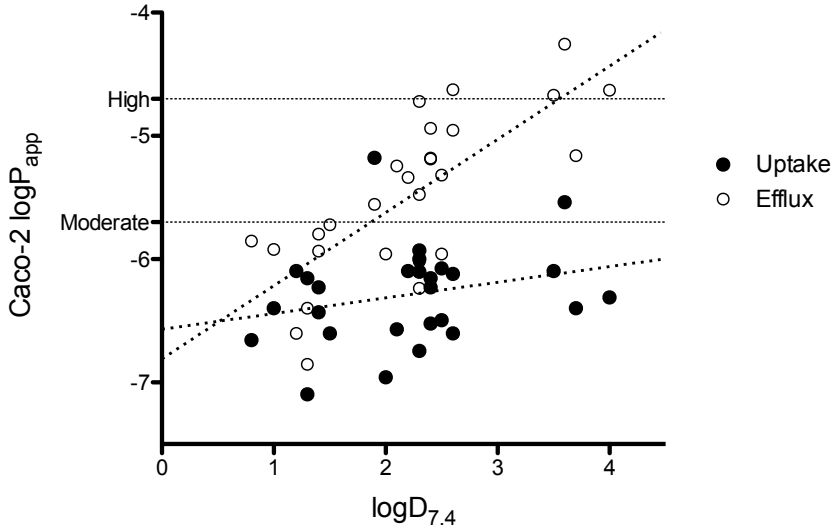


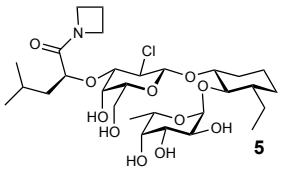
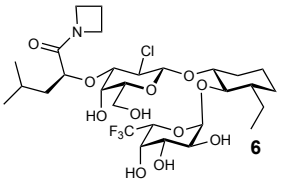
Figure 3.9 Plot of log$D_{7.4}$ versus apparent permeation logP_{app}. Data represents the mean only with a maximal standard deviation of 0.2 for the log$D_{7.4}$ and 0.2 for logP_{app} for molecules within the range of the thresholds. Log$D_{7.4}$ values were determined at pH 7.4 by a miniaturized shake flask procedure in sextuplicate. logP_{app} (apparent permeability) was assessed by permeation through a Caco-2 cell monolayer in the absorptive (apical to basolateral, uptake) and secretory (basolateral to apical, efflux) directions in triplicate. The thresholds for high and moderate absorption potentials for the absorptive direction are indicated according to Hou *et al.* (39). The dashed slope represents the linear regression between log$D_{7.4}$ and logP_{app} for both directions, i.e. uptake and efflux. The linear regression is not significantly non-zero for the uptake with $p = 0.19$ and for efflux it is significantly non-zero with $p < 0.0001$.

3.5.2 Structural characteristics of E-selectin antagonists

The effects of fluorinated fucose

Due to the hydrophilic character of their core structure, E-selectin antagonists generally show good solubility, i.e. > 10 mM in aqueous buffer. However, for analogues with a fluorinated fucose moiety the pharmacokinetic properties, especially the solubility, are largely influenced as illustrated with the antagonists **5** and **6** (Table 3.2). The fluorinated antagonist **6** is more lipophilic compared to the unfluorinated analogue **5**, but the increased lipophilicity does influence passive permeation measured in PAMPA. Furthermore, the unfluorinated **5** is at least thirty times more soluble than **6**. This could be due to the core stabilization and therefore allowing easier stacking of the fluorinated compound in crystalline form. Further structural consequences triggered by fluorination are discussed in manuscript 3.

Table 3.2 The effect of fluorinated fucose moiety on physicochemical properties of E-selectin antagonists. Log$D_{7.4}$ values were determined at pH 7.4 by a miniaturized shake flask procedure in sextuplicate. The thermodynamic solubility was determined in HAB1 buffer at pH 7.4 in triplicate. Effective permeability (P_e), i.e. the diffusion through an artificial membrane was determined by PAMPA in quadruplicate at pH 7.4. Errors are given as the standard deviation.

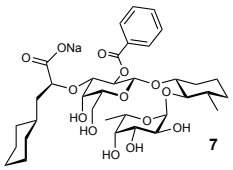
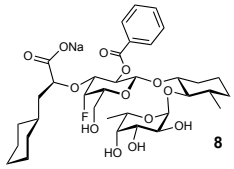
Compound	log$D_{7.4}$	Solubility [µM]	PAMPA logP_e [log cm/s]
 **5**	1.4 ± 0.1	> 10 mM	-8.6 ± 0.1
 **6**	2.1 ± 0.1	312 ± 24 µM	-8.3 ± 0.1

Intramolecular hydrogen bonds

The reduction of the PSA and therefore the hydrophilicity of the core structure of E-selectin antagonists should be improving physicochemical properties. That the removal of hydroxyl groups is not always beneficial is shown in the example of compound **7** and **8** (

Table 3.3). By replacing the C4-hydroxyl of galactose with fluorine in compound **8**, the lipophilicity surprisingly drops by 1 order of magnitude compared to **7**. Since the PSA is reduced, other structural properties must be responsible for the observed loss in lipophilicity. In compound **7**, the carboxylic acid is able to form an intramolecular hydrogen bond with C4-hydroxyl of galactose and therefore induces a reduction of the apparent hydrophilic surface. Preventing this intramolecular hydrogen bond leads to an increase in hydrophilicity. Further structural influences of hydroxyl groups in carbohydrates and their effects on physicochemical properties in E-selectin antagonists are discussed in manuscript 4.

Table 3.3 The influence of replacing C4-hydroxyl by fluorine on physicochemical properties. Log$D_{7.4}$ values were determined at pH 7.4 by a miniaturized shake flask procedure in sextuplicate. *P*e (effective permeability) was determined in the PAMPA (parallel artificial membrane permeability assay), measuring diffusion through an artificial membrane in quadruplicate at pH 7.4. PSAs were calculated using MarvinSketch version 16.10.10.0 using a fragment approach according to Ertl *et al.* (34).

Compound	MW [g/mol]	log$D_{7.4}$	PAMPA logP_e [log cm/s]	PSA [Å^2]
 **7**	718.23	1.0 ± 0.1	<-10	206.9
 **8**	720.76	0.0 ± 0.1	<-10	190.3

3.5.3 Substructures of E-selectin antagonists (Master thesis of Roman Koch)

For a further understanding of the absorption potential of modified E-selectin antagonists, partial substructures of compound **7** were synthesized (**9**-**15**, Figure 3.10). **7** is not permeating over an artificial membrane in PAMPA, even though it has a reasonable lipophilicity with a log$D_{7.4}$ of 1.0. Furthermore, **7** violates the Lipinski's rule of five by the number of hydrogen donors, acceptors and the molecular weight.

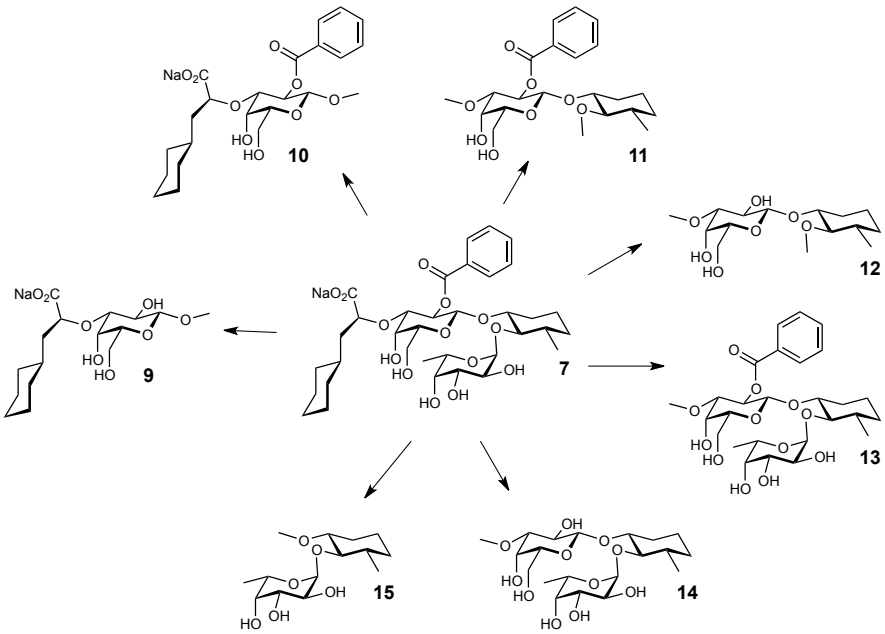


Figure 3.10 Partial structures **9**-**15** of antagonist **7** screened for pharmacokinetic properties.

Table 3.4 summarizes the physicochemical properties of the partial structures **9** - **15**. All tested molecules are highly soluble and compounds **9** **10**, **11**, **12** and **15** fulfill all criteria of the rule of five, all other violate two criteria. However, according the PAMPA classification, only the lipophilic compound **11** can be expected to be highly permeable. The permeation properties of **15** are still good and compound **12** is just below the cut-off of -6.3 with a logP_e of -6.5. All the other substances did not show sufficient permeability at pH 7.4.

The Caco-2 cell permeation assay showed similar results, ranking compound **11** as the most permeable of all tested compounds, followed by compound **15**. All the other compounds had a P_{app} below the cut-off of $2*10^{-6}$. Similar to the PAMPA results, compound **12** did not reach the cut-off for medium permeation.

However, compound **13** performed surprisingly well in the Caco-2 cell assay compared with the PAMPA results with an apparent permeability P_{app} of around $1*10^{-6}$ cm/s, which is only a factor of two below the cut-off. In the PAMPA assay, **13** had a logP_e around -9, which means almost no permeation at all. In the Caco-2 monolayer assay, the apparent permeability of **13** is approximately 3.5 times higher than the rate of excretion, suggesting active uptake. **13** consist of a 2-*O*-benzoylated galactoside moiety linked with a C2 linker to a fucose. The corresponding non-benzoylated derivatives did not show active intake. This suggests, that the benzoyl group in the galactose moiety is involved in this observed active transport. Nevertheless, this set of tested compounds shows the difficulty to get oral bioaviable E-selectin antagonists containing two carbohydrate moieties.

Table 3.4 Physicochemical and permeability properties of fragments **9**-**15** and E-selectin antagonist **7**. Solubility was determined in a kinetic approach in triplicate. Log$D_{7.4}$ were determined at pH 7.4 by a miniaturized shake flask procedure in sextuplicate. P_e = effective permeability: The diffusion through an artificial membrane was determined by PAMPA in quadruplicate at pH 7.4 and thresholds for moderate and high absorption potential are -6.3 and -5.7, respectively (35). Permeation through a cell monolayer was assessed by a Caco-2 cell assay in the absorptive (apical to basolateral, uptake AB) and secretory (basolateral to apical, efflux BA) directions in triplicates and the efflux ratio was determined by dividing efflux by uptake permeability. Thresholds for moderate and high absorption potential for the absorptive direction are 2 and 20 $*10^{-6}$ cm/s, respectively (39).

Cpd.	Solubility [µg/mL]	log$D_{7.4}$	PAMPA logP_e [log cm/s]	Caco2-cell P_{app} [10^{-6} cm/s]		
				uptake AB	efflux BA	efflux ratio
7	>1000	1.0 ± 0.1	< -10	0.4 ± 0.5	1.2 ± 0.6	
9	>1000	< -1.5	-9.8 ± 0.5	0.04 ± 0.01	0.023 ±0.01	0.59
10	>1000	-1,3 ± 0,2	-7.2 ± 0.5	0.46 ± 0.14	0.49 ± 0.05	1.06
11	>1000	2.5 ± 0.1	-4.3 ± 0.0	22.2 ± 0.92	45.6 ± 1.7	2.06
12	>1000	-0.4 ± 0.2	-6.5 ± 0.0	0.95 ± 0.12	1.53 ± 0.08	1.61
13	>1000	0.4 ± 0.1	-8.8 ± 0.3	1.2 ± 0.4	0.35 ± 0.03	0.29
14	798 ± 59	< -1.5	< -10	0.17 ± 0.03	0.08 ± 0.01	2.21
15	798 ± 136	-0.2 ± 0.1	-5.7 ± 0.0	10.5 ± 1.5	13.9 ± 0.5	1.33

3.6 References

1. Springer TA. Traffic signals for lymphocyte recirculation and leukocyte emigration: the multistep paradigm. *Cell*. **1994**;*76(2)*:301-14.
2. Seong SY, Matzinger P. Hydrophobicity: an ancient damage-associated molecular pattern that initiates innate immune responses. *Nat Rev Immunol*. **2004**;*4(6)*:469-78.
3. Takeuchi O, Akira S. Pattern recognition receptors and inflammation. *Cell*. **2010**;*140(6)*:805-20.
4. Strieter RM, Kunkel SL, Bone RC. Role of tumor necrosis factor-alpha in disease states and inflammation. *Crit Care Med*. **1993**;*21(10 Suppl)*:S447-63.
5. Bevilacqua M, Butcher E, Furie B, Furie B, Gallatin M, Gimbrone M, et al. Selectins - a Family of Adhesion Receptors. *Cell*. **1991**;*67(2)*:233-.
6. Lasky LA. Selectins - Interpreters of Cell-Specific Carbohydrate Information during Inflammation. *Science*. **1992**;*258(5084)*:964-9.
7. Carlos TM, Harlan JM. Leukocyte-endothelial adhesion molecules. *Blood*. **1994**;*84(7)*:2068-101.
8. Kansas GS, Saunders KB, Ley K, Zakrzewicz A, Gibson RM, Furie BC, et al. A role for the epidermal growth factor-like domain of P-selectin in ligand recognition and cell adhesion. *J Cell Biol*. **1994**;*124(4)*:609-18.
9. Ley K. The role of selectins in inflammation and disease. *Trends in Molecular Medicine*. **2003**;*9(6)*:263-8.
10. Lorenzon P, Vecile E, Nardon E, Ferrero E, Harlan JM, Tedesco F, et al. Endothelial cell E- and P-selectin and vascular cell adhesion molecule-1 function as signaling receptors. *J Cell Biol*. **1998**;*142(5)*:1381-91.
11. Yoshida M, Westlin WF, Wang N, Ingber DE, Rosenzweig A, Resnick N, et al. Leukocyte adhesion to vascular endothelium induces E-selectin linkage to the actin cytoskeleton. *J Cell Biol*. **1996**;*133(2)*:445-55.
12. Rondaij MG, Bierings R, Kragt A, van Mourik JA, Voorberg J. Dynamics and plasticity of Weibel-Palade bodies in endothelial cells. *Arterioscler Thromb Vasc Biol*. **2006**;*26(5)*:1002-7.
13. Wyble CW, Hynes KL, Kuchibhotla J, Marcus BC, Hallahan D, Gewertz BL. TNF-alpha and IL-1 upregulate membrane-bound and soluble E-selectin through a common pathway. *J Surg Res*. **1997**;*73(2)*:107-12.
14. Labow MA, Norton CR, Rumberger JM, Lombard-Gillooly KM, Shuster DJ, Hubbard J, et al. Characterization of E-selectin-deficient mice: demonstration of overlapping function of the endothelial selectins. *Immunity*. **1994**;*1(8)*:709-20.
15. Rosen SD. Ligands for L-selectin: homing, inflammation, and beyond. *Annu Rev Immunol*. **2004**;*22*:129-56.
16. Angiari S. Selectin-mediated leukocyte trafficking during the development of autoimmune disease. *Autoimmun Rev*. **2015**;*14(11)*:984-95.
17. Koch AE, Burrows JC, Haines GK, Carlos TM, Harlan JM, Leibovich SJ. Immunolocalization of endothelial and leukocyte adhesion molecules in human rheumatoid and osteoarthritic synovial tissues. *Lab Invest*. **1991**;*64(3)*:313-20.
18. Davies MJ, Gordon JL, Gearing AJ, Pigott R, Woolf N, Katz D, et al. The expression of the adhesion molecules ICAM-1, VCAM-1, PECAM, and E-selectin in human atherosclerosis. *J Pathol*. **1993**;*171(3)*:223-9.
19. Takada M, Nadeau KC, Shaw GD, Marquette KA, Tilney NL. The cytokine-adhesion molecule cascade in ischemia/reperfusion injury of the rat kidney. Inhibition by a soluble P-selectin ligand. *J Clin Invest*. **1997**;*99(11)*:2682-90.
20. Blann AD, Tse W, Maxwell SJ, Waite MA. Increased levels of the soluble adhesion molecule E-selectin in essential hypertension. *J Hypertens*. **1994**;*12(8)*:925-8.

21. Dutta P, Hoyer FF, Sun Y, Iwamoto Y, Tricot B, Weissleder R, et al. E-Selectin Inhibition Mitigates Splenic HSC Activation and Myelopoiesis in Hypercholesterolemic Mice With Myocardial Infarction. *Arterioscler Thromb Vasc Biol*. **2016**;*36(9)*:1802-8.
22. Laubli H, Borsig L. Selectins promote tumor metastasis. *Semin Cancer Biol*. **2010**;*20(3)*:169-77.
23. Winkler IG, Barbier V, Nowlan B, Jacobsen RN, Forristal CE, Patton JT, et al. Vascular niche E-selectin regulates hematopoietic stem cell dormancy, self renewal and chemoresistance. *Nat Med*. **2012**;*18(11)*:1651-7.
24. Price TT, Burness ML, Sivan A, Warner MJ, Cheng R, Lee CH, et al. Dormant breast cancer micrometastases reside in specific bone marrow niches that regulate their transit to and from bone. *Sci Transl Med*. **2016**;*8(340)*:340ra73.
25. Natarajan M, Udden MM, McIntire LV. Adhesion of sickle red blood cells and damage to interleukin-1 beta stimulated endothelial cells under flow in vitro. *Blood*. **1996**;*87(11)*:4845-52.
26. Erbe DV, Wolitzky BA, Presta LG, Norton CR, Ramos RJ, Burns DK, et al. Identification of an E-Selectin Region Critical for Carbohydrate Recognition and Cell-Adhesion. *J Cell Biol*. **1992**;*119(1)*:215-27.
27. Binder FP, Lemme K, Preston RC, Ernst B. Sialyl Lewis(x): a "pre-organized water oligomer"? *Angew Chem Int Ed Engl*. **2012**;*51(29)*:7327-31.
28. Schwizer D, Patton JT, Cutting B, Smiesko M, Wagner B, Kato A, et al. Pre-organization of the core structure of E-selectin antagonists. *Chemistry*. **2012**;*18(5)*:1342-51.
29. Veber DF, Johnson SR, Cheng HY, Smith BR, Ward KW, Kopple KD. Molecular properties that influence the oral bioavailability of drug candidates. *J Med Chem*. **2002**;*45(12)*:2615-23.
30. Lipinski CA, Lombardo F, Dominy BW, Feeney PJ. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv Drug Deliv Rev*. **2001**;*46(1-3)*:3-26.
31. Chang J, Patton JT, Sarkar A, Ernst B, Magnani JL, Frenette PS. GMI-1070, a novel pan-selectin antagonist, reverses acute vascular occlusions in sickle cell mice. *Blood*. **2010**;*116(10)*:1779-86.
32. Telen MJ, Wun T, McCavit TL, De Castro LM, Krishnamurti L, Lanzkron S, et al. Randomized phase 2 study of GMI-1070 in SCD: reduction in time to resolution of vaso-occlusive events and decreased opioid use. *Blood*. **2015**;*125(17)*:2656-64.
33. Norman KE, Anderson GP, Kolb HC, Ley K, Ernst B. Sialyl Lewis(x) (sLe(x)) and an sLe(x) mimetic, CGP69669A, disrupt E-selectin-dependent leukocyte rolling in vivo. *Blood*. **1998**;*91(2)*:475-83.
34. Ertl P, Rohde B, Selzer P. Fast calculation of molecular polar surface area as a sum of fragment-based contributions and its application to the prediction of drug transport properties. *J Med Chem*. **2000**;*43(20)*:3714-7.
35. Avdeef A, Bendels S, Di L, Faller B, Kansy M, Sugano K, et al. PAMPA--critical factors for better predictions of absorption. *J Pharm Sci*. **2007**;*96(11)*:2893-909.
36. Bickerton GR, Paolini GV, Besnard J, Muresan S, Hopkins AL. Quantifying the chemical beauty of drugs. *Nat Chem*. **2012**;*4(2)*:90-8.
37. Guimaraes CR, Mathiowetz AM, Shalaeva M, Goetz G, Liras S. Use of 3D properties to characterize beyond rule-of-5 property space for passive permeation. *J Chem Inf Model*. **2012**;*52(4)*:882-90.
38. Aller SG, Yu J, Ward A, Weng Y, Chittaboina S, Zhuo R, et al. Structure of P-glycoprotein reveals a molecular basis for poly-specific drug binding. *Science*. **2009**;*323(5922)*:1718-22.
39. Hou T, Wang J, Zhang W, Xu X. ADME evaluation in drug discovery. 7. Prediction of oral absorption by correlation and classification. *J Chem Inf Model*. **2007**;*47(1)*:208-18.

3.7 Manuscript 1: E-selectin antagonist ester prodrug approach

In this manuscript, the physicochemical and pharmacokinetic properties of ester prodrugs was evaluated *in vitro*. The focus of the work was on the stability of the molecule during the gastrointestinal passage, the permeability and the metabolic pathway in liver microsomes. *In vitro* results were further evaluated by oral bioavailability studies in mice.

Contribution to the Project:

Philipp Dätwyler designed and performed the assays of the pharmacokinetic evaluation *in vitro* and *in vivo* of the ester prodrugs. He further supervised the work of Katja Hostettler during her master thesis and wrote the manuscript.

Philipp Dätwyler, Xiaohua Jiang, Norbert Varga, Beatrice Wagner, Tobias Mühlethaler, Katja Hostettler, Beat Ernst*

Institute of Molecular Pharmacy, Department of Pharmaceutical Sciences, University of Basel, Klingelbergstrasse 50, 4056 Basel, Switzerland

*Corresponding author

E-selectin antagonist prodrugs with enhanced pharmacokinetic profile

Philipp Dätwyler, Xiaohua Jiang, Norbert Varga, Beatrice Wagner, Tobias Mühlethaler, Katja Hostettler, Beat Ernst*

Institute of Molecular Pharmacy, Department of Pharmaceutical Sciences, University of Basel, Klingelbergstrasse 50, 4056 Basel, Switzerland

*Corresponding author

Abbreviations: cell adhesion molecules (CAM), intravenous (IV), human carboxylesterase (CES) sialyl Lewis^x (sLe^x), vaso-occlusive crisis (VOC), microscale thermophoresis (MST), parallel artificial membrane permeation assay (PAMPA), rat or human liver microsomes (RLM or HLM), cytochrome P450, (CYP450), bis(4-nitrophenyl)phosphate (BNPP), nicotinamide adenine dinucleotide phosphate (NADPH).

Keywords: sialyl Lewis^x, E-selectin, pharmacokinetic properties, oral bioavailability, ester prodrugs.

Introduction

Upon an inflammatory stimulus, a family of cell adhesion molecules (CAMs), among others E-selectin, are upregulated on endothelial cells (1, 2). By interacting with their glycoprotein ligands located on leukocytes, E-selectin can mediate the rolling of these leukocytes on the endothelial surface. This first step of the inflammatory cascade is followed by firm adhesion and finally transmigration to the site of inflammation (2, 3). There, the inflammatory stimulus is abolished with oxidative agents, proteases and cytokines. However, when this vital defense mechanism in the event of injuries or infections is excessively used, it can turn deleterious in numerous diseases with an inflammatory component as e.g. stroke, psoriasis or asthma (4, 5). Therefore, blocking E-selectin with a selectin antagonist and as a consequence interrupt the whole inflammatory cascade has been recognized as a promising therapeutic approach for the treatment of inflammatory diseases (6).

The main epitope recognized by E-selectin is the tetrasaccharide sialyl Lewis^x (sLe^x) as part of glycan structures of selectin ligands present on leukocytes (2). Because sLe^x has only a moderate affinitiy (K_D 800 µM), and far from drug-like pharmacokinetic properties, glycomimetics have been developed (7). A successful example is the selectin antagonist Rivipansel currently in clinical phase 3 evaluation for the treatment of vaso-occlusive crisis (VOC) in cickle cell anemia patients (8, 9). Since Rivipansel is not orally bioavailable, an intravenous administration is required. However, for chronic inflammatory diseases with a strong correlation to E-selectin overexpression, such as arthrosclerosis, rheumatoid arthritis and psoriasis, an orally bioavailable treatment would be preferable (10-14).

With carbohydrates as lead structures, inherent problems regarding oral bioavailability are inevitable, mainly because their large polar surface area and their hydrophilicity (15). Although the replacement or removal of polar groups, which are not part of the pharmacophores, lead to a substantial reduction of hydrophilicity (→**1**) (16), the corresponding glycomimetics still violate parameters usually predictive for oral availability [rule of five by Lipinski (17) and Veber's rules (18)]. Specifically, the carboxylic acid function crucial for the interaction with E-selectin prevents oral bioavailability. In summary, since a further reduction of the polarity or size of the glycomimetic **1** is accompanied by a loss of affinity, other approaches to reach sufficient lipophilicity and permeability are required (19, 20).

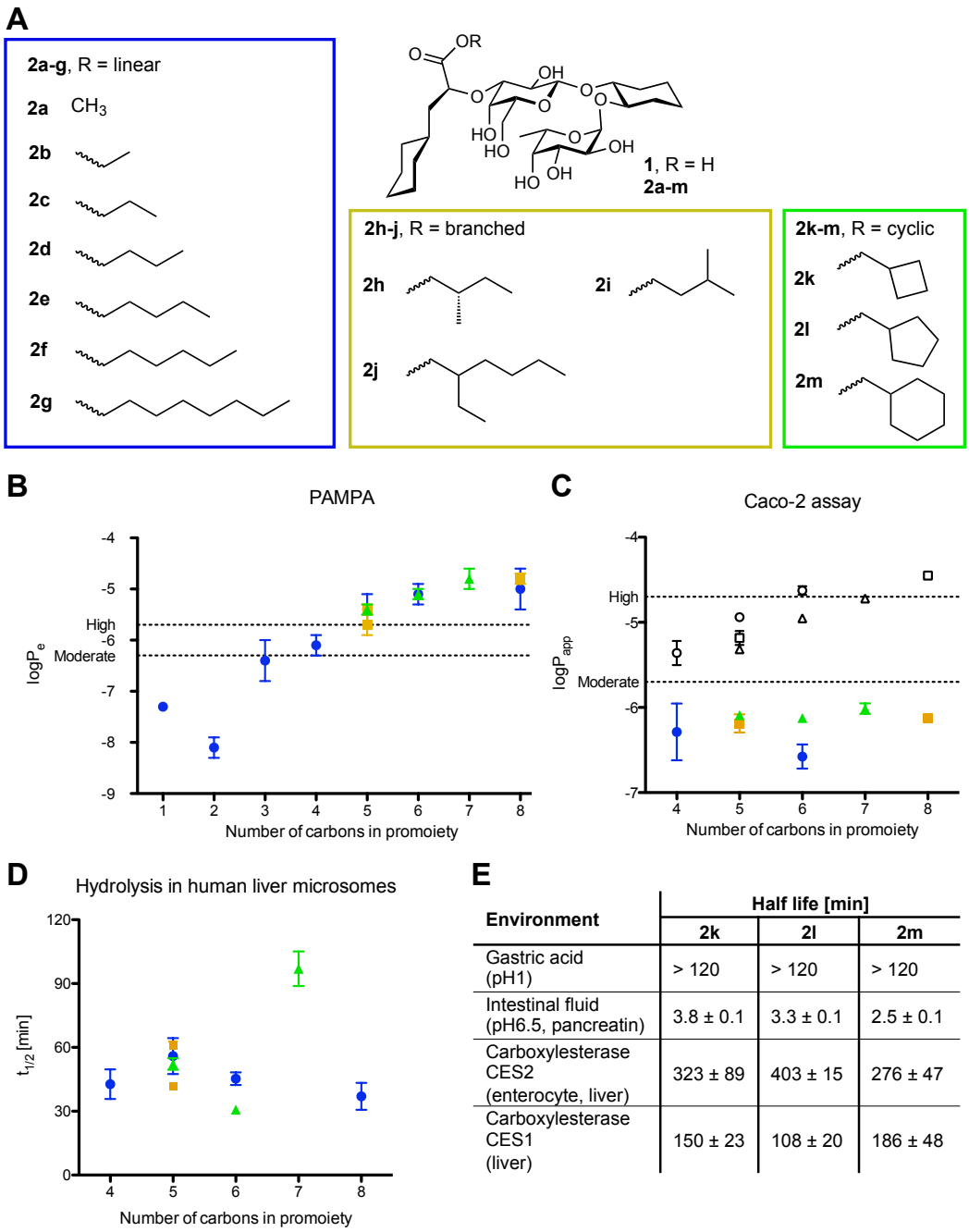
A common approach to overcome insufficient lipophilicity of carboxylic acids are ester prodrugs which are hydrolyzed by carboxylesterases upon absorption to release the active principle (21). To reach their target on endothelial cells adjacent to an inflammatory stimulus, ester prodrugs of E-selectin antagonists should be stable in the gastrointestinal environment and the enterocytes of the small intestine. Once the prodrug is either actively or passively absorbed, the active principle has to be released by hydrolysis (22, 23). In the present study, ester prodrugs based on E-selectin antagonist **1** using aliphatic promoieties were synthesized and evaluated their potential for oral bioavailability *in vitro* and *in vivo.* In a second step, the best prodrug solutions were applied to recently identified E-selectin antagonists with improved affinities but still insufficient pharmacokinetic properties (Manuscript 3,4 and 5).

Results

Different linear, branched and cyclic aliphatic ester prodrugs of the carboxylic acid **1** were synthesized (→ **2a–m**) and their absorption potential and metabolic stability were screened (Figure 1). Prodrugs with more than three carbons in the promoiety exhibit logP_e values in a parallel artificial permeation assay (PAMPA) indicative for a moderate (with four carbons in the promoiety) to a high absorption potential (with 5 or more carbon atoms in the promoiety) (Figure 1B) (24, 25). Furthermore, the passive permeation is depending only on the number of carbon atoms and not of the constitution of the promoiety. Compounds showing moderate to high permeability in PAMPA were further assessed in a cell-based permeation model using colon adenocarcinoma (Caco-2) cells to determine active transport mechanism, *i.e.* uptake (apical to basolateral; a-b) and efflux (b-a) permeability (Figure 1C) (26, 27). Unfortunately, all compounds exhibit active efflux with an efflux ratio (b-a/a-b) over 2.0, leading to an absorption below the threshold for a moderate absorption potential. Overall, the efflux ratio and the absolute uptake were better for prodrugs containing cyclic and branched promoieties compared to linear promoieties.

The metabolic stability of the ester prodrugs **2a**-**m** was assessed with human liver microsomes (28) (Figure 1D). Similar to PAMPA, the shape of the promoieties has a minor effect on the ester hydrolysis, except for **2m**, when the bulky cyclohexylester leads to slower metabolic release. Slow hydrolysis of the ester prodrug is favorable for carbohydrate mimetics, due to the likely fast renal excretion of the active principle (29). Over all, ester prodrug with cyclic promieties **2k**-**m** performed best and were chosen for further **hydrolysis studies (Figure 1E).** These esters

are stable in gastric acid, but are rapidly hydrolyzed in simulated intestinal fluid containing pancreatic enzymes. Furthermore, the molecules are faster hydrolyzed by the carboxylesterase CES1, which is highly expressed in hepatocytes, compared to CES2. CES2 is expressed in enterocytes and could lead to hydrolysis during the absorption, limiting the oral bioavailability of ester prodrugs (30).



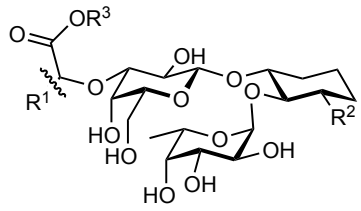
Environment	Half life [min]		
	2k	2l	2m
Gastric acid (pH1)	> 120	> 120	> 120
Intestinal fluid (pH6.5, pancreatin)	3.8 ± 0.1	3.3 ± 0.1	2.5 ± 0.1
Carboxylesterase CES2 (enterocyte, liver)	323 ± 89	403 ± 15	276 ± 47
Carboxylesterase CES1 (liver)	150 ± 23	108 ± 20	186 ± 48

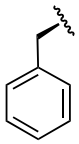
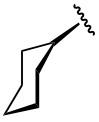
Figure 1 **A** Aliphatic ester prodrugs of **2a-m** screened in **B**, **C** and **D**. Prodrugs with linear aliphatic promoieties (**2a-g**) are indicated as circle, branched (**2h-i**) as rectangle and cyclic (**2k-m**) as triangle. **B** The parallel artificial membrane permeation assay (PAMPA) was performed at pH 7.4 in quadruplicate. Thresholds were used according to Avdeef *et al.*(24) **C** Caco-2 cell permeation assay was performed 21 days after seeding of the cells in absorptive (a-b, full dots) and efflux (b-a, hollow dots) direction in triplicate. Thresholds for absorptive direction were used

according to Hou *et al.*(27) **D** The hydrolysis of ester prodrugs was assessed in human liver microsomes, 0.125 mg/ml at 37°C in triplicate. **E** The hydrolysis of cyclic ester prodrugs **2k**, **2l** and **2m** were measured in different environments in triplicate. (All error bars and ranges are given as the standard deviation. Synthesis and assay procedures are described in the suppl. data.)

A problem inherent to all produgs **2a**-**m** is their instability in intestinal fluid, hydrolyzing the molecules before absorption. Independent of the promoiety, they showed fast and complete hydrolysis with a half-life below 10 minutes in intestinal fluid containing pancreatic enzymes (suppl. data). It can therefore be assumed that the structure of the active principle and not the promoiety is responsible for the recognition by pancreatic enzymes. Therefore, ethyl ester prodrugs of a series of carbohydrate mimetics with different lactic acid derivatives in the 3-position of the galactose moiety were synthesized and tested (Table 1). Apparently, already minor changes of the cyclohexyl lactic acid substituent such as replacing the cyclohexyl moiety by a phenyl ($\rightarrow$ **3**) or isopropyl ($\rightarrow$ **5**) group improves pancreatic stability. When cyclohexyl lactic acid ($\rightarrow$ 2**b**) is replaced by 1-hydroxy-cyclohexaneacetic acid ($\rightarrow$ **7**), stability is further enhanced. Finally, with the glycolic acid ($\rightarrow$ **9**), almost complete pancreatic stability could be achieved, however, for the price of a substantial reduction of affinity of the active principle. On the contrary, an additional methyl group at the cycloheyl moiety replacing GlcNAc ($\rightarrow$ **5**, **7** & **9**) improves not only the lipophilicity of the prodrug but also the affinity of the corresponding active principle (19). Thus, considering both affinity and stability, the prodrugs **7** perfomed best.

Table 1 Prodrug; The intestinal fluid stability of ethyl ester prodrugs was determined in 2.5 mg/ml pancreatin at pH 6.5 in triplicate. After 30 minutes, samples were quenched and analyzed by LC-MS and the concentration of prodrug and active principle was calculated. The error indicates the standard deviation. **Active principle**; The affinity towards E-selectin of the active principle was assessed by micro scale thermophoresis measurements. Synthesis and assay procedures are described in the supplementary information.



		Prodrug		Active principle	
		Intestinal fluid stability		Affinity	
R^1	R^2	Promoiety $R^3 = CH_2CH_3$	Pancreatic stability [%]	R^3 = H	K_D [µM]
	H	**2b**	9.0 ± 5.9	**1**	60.7
	H	**3**	48.0 ± 4.2	**4**	329
	CH_3	**5**	38.0 ± 3.5	**6**	14.9
	CH_3	**7**	74.8 ± 7.9	**8**	33.2
H	CH_3	**9**	93.7 ± 5.7	**10**	133.4

Next, the preferred promoiety and lactic acid moiety were combined and the metabolic pathway in liver microsomes of **11** was explored in liver microsomes (Figure 2A). To answer the question regarding the influence of the lipophilicity of the promoiety on the metabolism, the

phase I metabolism determined in rat liver microsomes (RLM) by adding the co-factor NADPH to activate cytochrome P450s and the unspecific carboxylesterases inhibitor bis(4-nitrophenyl)phosphate (BNPP) to inhibit hydrolysis (31, 32). For the ester prodrug **11**, oxidation in RLM with a half-life of around 10 minutes was observed (Table 2). Two major metabolites showing an increase of the molecular weight by 16 compared to the parent ester were detected by mass spectrometry, indicating one oxidation of the compound. When this RLM reaction mixture containing the metabolites was treated with sodium hydroxide to hydrolyze the ester moiety under basic conditions, unchanged active principle **8** was regained, indicating that the oxidation took place on the promoiety. Furthermore, incubating the mixture of oxidized metabolites in fresh RLM the metabolites can no longer be hydrolyzed by carboxylesterases to yield the active principle **8**. Finally, with RLM/NADPH containing active carboxylesterases, i.e. without BNPP addition, both metabolic steps occur simultaneously, although with different kinetics, leading to 20% of the deprotected active principle **8** and to 80% to the metabolite **12**. Therefore, only 20% of the prodrug is transformed to the active principle in RLM and 80% is lost due to oxidation by CYP450s. Furthermore the prodrug as well as the active principle were tested for their susceptibility for the most frequently occurring phase II metabolism, namely glucoronidation by UDP-glucuronyltransferases (UGTs) and sulfation by sulfotransferases (STs), in RLM and rat liver cytosol (33, 34).Both, prodrug **11** and active principle **8**, did not show any conjugation by UGTs and STs. Furthermore, the active principle is not oxidized by CYP450.

The 20/80 oxidation/hydrolysis ratio of **11** was improved with prodrug **13** or **2j**, yielding a 50/50 and 65/35 ratio in favor for the hydrolysis into the active principles **6** and **1**, respectively (Figure 2B). The complete pharmacokinetic analysis of the ester prodrugs **11**, **13** and **2j** and the affinity of their respectively active principle are listed in Table 2. All three compounds showed high absorption potential in PAMPA, but not sufficient permeability in Caco-2 cell based assay and active efflux, similar to the primary test series **2a-m**.

Next, the ester prodrug **14** was synthesized to have an improved hydrolysis/oxidation ratio and affinity of the active principle. Therefore, the cyclobutyl methyl promoiety of **2j** and the 1-hydroxy-cyclohexaneacetic acid of **13** were combined. Furthermore, other modifications improving the affinity of the active principle are additionally implemented (Manuscript 3, 4 and 5). The C-6 of fucose is fluorinated, C-2 hydroxyl group of the galactose is replaced by chlorine and the side chain of the Glc*N*Ac is elongated. A comparison of its physicochemical and *in vitro* PK/PD properties with the prodrugs discussed above is summarized in Table 2. Those changes

led to an enhanced lipophilicity of the molecule, sufficient stability against pancreatic enzymes and high absorption potential in PAMPA. Unfortunately, the observed efflux ratio was further enhanced. But compared to the initial lead prodrug **11**, the hydrolysis/oxidation ratio of **14** improved to 40/60 and the affinity of the active principle improved by factor 10.

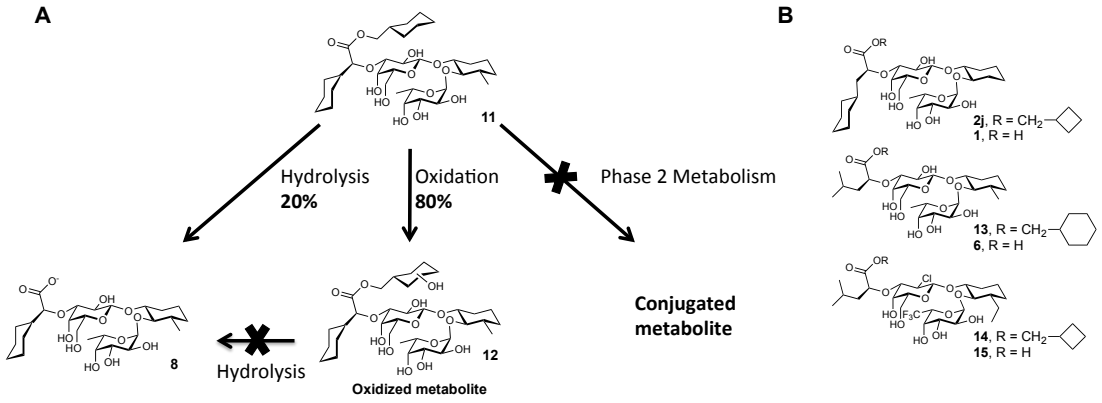


Figure 2 A The metabolic pathway of prodrug **11** in rat liver microsomes. Cytochrome P450s oxidize the promoiety, leading to a metabolite no longer susceptible to hydrolysis by carboxylesterases. **B** Structures of prodrugs **13, 2j** and **14**, and their active principles **6, 1** and **15**, respectively. A synthesis of the prodrugs is described in the Supplementary Information.

Table 2 Summary of the physicochemical and PK/PD data of ester prodrugs **11**, **13**, **2j** and **14**. Errors are indicated as the standard deviation.

Cpd	$logD_{7.4}$	Stability $t_{1/2}$ [min] intestinal Fluid	PAMPA $logP_e$	Caco- cell permeation assay $logP_{app}$			Stability $t_{1/2}$ [min] RLM Oxidation	Stability $t_{1/2}$ [min] RLM Hydrolysis	Affinity K_D [µM] Active principle
				uptake a-b	efflux b-a	efflux ratio b-a/a-b			
11	3.1 ±0.1	> 120	-4.9±0.1	-6.4 ±0.1	-5.2 ±0.1	19	10.7 ±0.3 80 ±2%	23.5 ±1.1 20 ±2%	33.2
13	2.4±0.1	25 ± 1	-5.5 ±0.1	-6.2 ±0.1	-5.2 ±0.1	11	27 ±2 48 ±2%	30 ±2 52 ±2%	14.9
2j	2.5 ±0.1	< 10	-5.4 ±0.1	-6.1±0.1	-5.3 ±0.1	5.8	50 ±20 35 ±5%	8.1 ±0.2 65 ±5%	60.7
14	4.0 ±0.2	77 ± 4	-4.6 ±0.2	-6.3 ±0.1	-4.6 ±0.1	48	62 ±3 58 ±4%	72 ±6 42 ±4%	3.3

Animal Studies

For a further evaluation of the pharmacokinetic properties, *in vivo* bioavailability studies in mice were conducted with the prodrugs **11** and **14**. To reach their target on the endothelial surface of blood vessels adjacent to an inflammatory stimulus, the prodrugs have to fulfill three prerequisites. They have to pass the intestinal membrane to reach the blood circulation, the bioconversion to the active principle has to be complete and slowly over time to elongate the therapeutic window and finally the active principle has to be available at the therapeutic concentrations at the target. Unfortunately, both prodrugs **11** and **14** did not show any oral bioavailability, with a concentration in blood plasma below the detection limit of the prodrug and the corresponding active principle **8** and **15**, respectively (Figure 3C and D). Therefore, the passive permeation observed in PAMPA was not sufficient to reach oral bioavailability. In the case of prodrug **11**, the formation of the prodrug led to an elongation of the blood plasma half-life of the active principle **8** by factor 5 (Figure 3A). After direct injection of the active principle a blood plasma half-life of 6.2 min was observed, whereas after the prodrug injection, the active principle had a half-life of 34.4 min. Furthermore, the bioconversion of the prodrug to the active principle was higher than predicted in liver microsomes. The total exposure (Area under the curve, AUC) of the active principle after injection of the prodrug was adjusted to the molecular weight (MW) only 20% less compared to the direct injection of the active principle. The final prodrug **14** did not show this improvement of the pharmacokinetic properties *in vivo* (Figure 3B. The half-life of the active principle was for both IV administration as prodrug ($t_{1/2}$ = 6.8 min) and active principle ($t_{1/2}$ = 4.2 min) similar and the AUC of the active principle was strongly reduced by formation of a prodrug (33% compared to the direct injection of the active principle).

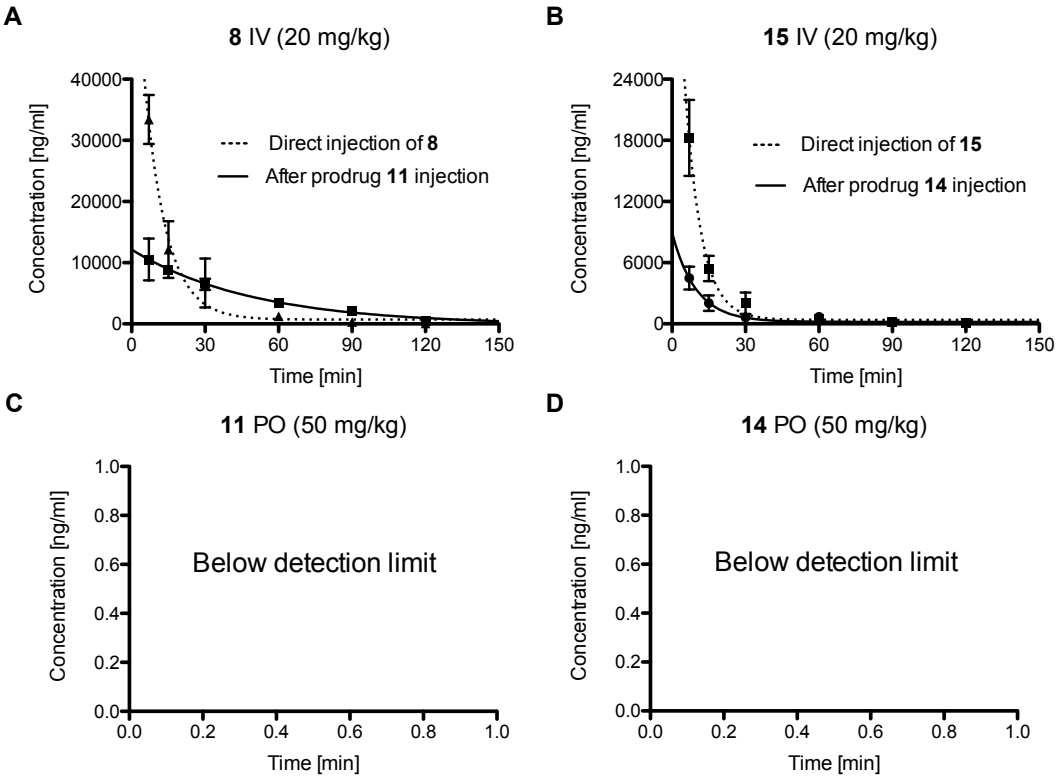


Figure 3 **A** and **B**: Plasma concentration of the active principle **8** (**A**) and **15** (**B**) after intravenous injection 20 mg/kg of the prodrug (continuous line) or active principle (dotted line). Each point represents the mean and standard error of the mean (SEM) of three individual experiments. The half-life and area under the curve was calculated using PKSolver (35). **A**: The half-life of the active principle **8** in plasma was after direct injection 6.2 min (4.5-10.1 min 95% confidence interval; AUC 0-t = 835213 min*ng*ml^{-1}) and prodrug **11** injection 34.4 min (16.6 – 52.4 min 95% confidence interval; AUC 0-t = 573121 min*ng*ml^{-1}, 79% of direct injection adjusted to the molecular weight (MW)). **B**: The half-life of the active principle **15** in plasma was after direct injection 4.6 min (3.1-9.1 min 95% confidence interval; AUC 0-t = 448121 min*ng*ml^{-1}) and prodrug **14** injection 6.8 min (7.2 – 15.5 min 95% confidence interval; AUC 0-t = 134773 min*ng*ml^{-1}, 33% of direct injection adjusted to the MW). **C** and **D**: After oral gavage of 50 mg/kg, the blood plasma concentration of the active principle and prodrug was followed for a period of 180 min in three individual mice. There was no active principle or prodrug detectable over this period of time.

Conclusion and outlook

The *in vitro* pharmacokinetic properties of E-selectin antagonists, especially the passive permeation measured in PAMPA, were improved by forming ester prodrugs. But the observed permeability was not translatable to an *in vivo* situation in mice. Apparently, the efflux ratio observed in Caco-2 cell permeation assay had a strong impact on the oral bioavailability of the prodrugs. The major efflux pump P-glycoprotein (P-gp) preferably recognizes lipophilic xenobiotics bound to the membrane (36).The large lipophilic promoieties of the ester prodrugs are therefore a potent target for P-gp. Over all, a simple addition of lipophilic moieties to improve passive permeation is not sufficient to get oral bioavailable E-selectin antagonists.

Furthermore, the metabolic pathway of ester prodrugs was studied in liver microsomes. Carboxylesterases prefer a lipophilic environment around the ester bond, making ester prodrugs with aliphatic promoieties susceptible for hydrolysis (22). As observed in compound **11**, an oxidation of the promoiety by CYP450 leads to a more hydrophilic promoiety and therefore reduces drastically the hydrolysis rate of the ester bond by carboxylesterases. The reduced bioconversion was however to a lesser extend observable *in vivo* compared to rat liver microsomes. Ester prodrug **11** showed sufficient bioconversion *in vivo* with a prolonged half-life of the active principle. Including modification to improve the hydrolysis/oxidation ratio and affinity of the active principle led to compound **14**. Unfortunately, the *in vitro* predictions were not translatable to *in vivo* studies in mice, leading to a fast elimination of the active principle. Therefore, other approaches to develop an orally bioavailable E-selectin antagonist with high affinity have to be followed.

References

1. Lasky LA. Selectins - Interpreters of Cell-Specific Carbohydrate Information during Inflammation. *Science*. 1992;*258(5084)*:964-9.
2. Erbe DV, Wolitzky BA, Presta LG, Norton CR, Ramos RJ, Burns DK, et al. Identification of an E-Selectin Region Critical for Carbohydrate Recognition and Cell-Adhesion. *J Cell Biol*. 1992;*119(1)*:215-27.
3. Ley K, Kansas GS. Selectins in T-cell recruitment to non-lymphoid tissues and sites of inflammation. *Nat Rev Immunol*. 2004;*4(5)*:325-35.
4. Ley K. The role of selectins in inflammation and disease. *Trends in Molecular Medicine*. 2003;*9(6)*:263-8.
5. Angiari S. Selectin-mediated leukocyte trafficking during the development of autoimmune disease. *Autoimmun Rev*. 2015;*14(11)*:984-95.
6. Rossi B, Constantin G. Anti-selectin therapy for the treatment of inflammatory diseases. *Inflamm Allergy Drug Targets*. 2008;*7(2)*:85-93.
7. Binder FP, Lemme K, Preston RC, Ernst B. Sialyl Lewis(x): a "pre-organized water oligomer"? *Angew Chem Int Ed Engl*. 2012;*51(29)*:7327-31.
8. Chang J, Patton JT, Sarkar A, Ernst B, Magnani JL, Frenette PS. GMI-1070, a novel pan-selectin antagonist, reverses acute vascular occlusions in sickle cell mice. *Blood*. 2010;*116(10)*:1779-86.
9. Telen MJ, Wun T, McCavit TL, De Castro LM, Krishnamurti L, Lanzkron S, et al. Randomized phase 2 study of GMI-1070 in SCD: reduction in time to resolution of vaso-occlusive events and decreased opioid use. *Blood*. 2015;*125(17)*:2656-64.
10. Cybulsky MI, Gimbrone MA, Jr. Endothelial expression of a mononuclear leukocyte adhesion molecule during atherogenesis. *Science*. 1991;*251(4995)*:788-91.
11. Davies MJ, Gordon JL, Gearing AJ, Pigott R, Woolf N, Katz D, et al. The expression of the adhesion molecules ICAM-1, VCAM-1, PECAM, and E-selectin in human atherosclerosis. *J Pathol*. 1993;*171(3)*:223-9.
12. Koch AE, Burrows JC, Haines GK, Carlos TM, Harlan JM, Leibovich SJ. Immunolocalization of endothelial and leukocyte adhesion molecules in human rheumatoid and osteoarthritic synovial tissues. *Lab Invest*. 1991;*64(3)*:313-20.
13. Petzelbauer P, Pober JS, Keh A, Braverman IM. Inducibility and expression of microvascular endothelial adhesion molecules in lesional, perilesional, and uninvolved skin of psoriatic patients. *J Invest Dermatol*. 1994;*103(3)*:300-5.
14. Dutta P, Hoyer FF, Sun Y, Iwamoto Y, Tricot B, Weissleder R, et al. E-Selectin Inhibition Mitigates Splenic HSC Activation and Myelopoiesis in Hypercholesterolemic Mice With Myocardial Infarction. *Arterioscler Thromb Vasc Biol*. 2016;*36(9)*:1802-8.
15. Ernst B, Magnani JL. From carbohydrate leads to glycomimetic drugs. *Nat Rev Drug Discov*. 2009;*8(8)*:661-77.
16. Norman KE, Anderson GP, Kolb HC, Ley K, Ernst B. Sialyl Lewis(x) (sLe(x)) and an sLe(x) mimetic, CGP69669A, disrupt E-selectin-dependent leukocyte rolling in vivo. *Blood*. 1998;*91(2)*:475-83.
17. Lipinski CA, Lombardo F, Dominy BW, Feeney PJ. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv Drug Deliv Rev*. 2001;*46(1-3)*:3-26.
18. Veber DF, Johnson SR, Cheng HY, Smith BR, Ward KW, Kopple KD. Molecular properties that influence the oral bioavailability of drug candidates. *J Med Chem*. 2002;*45(12)*:2615-23.
19. Schwizer D, Patton JT, Cutting B, Smiesko M, Wagner B, Kato A, et al. Pre-organization of the core structure of E-selectin antagonists. *Chemistry*. 2012;*18(5)*:1342-51.
20. Preston RC, Jakob RP, Binder FP, Sager CP, Ernst B, Maier T. E-selectin ligand complexes adopt an extended high-affinity conformation. *J Mol Cell Biol*. 2016;*8(1)*:62-72.
21. Rautio J, Kumpulainen H, Heimbach T, Oliyai R, Oh D, Jarvinen T, et al. Prodrugs: design and clinical applications. *Nat Rev Drug Discov*. 2008;*7(3)*:255-70.
22. Imai T. Human carboxylesterase isozymes: catalytic properties and rational drug design. *Drug Metab Pharmacokinet*. 2006;*21(3)*:173-85.

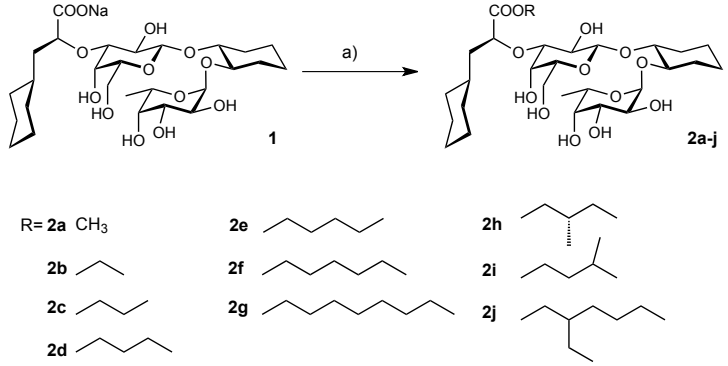
23. Liederer BM, Borchardt RT. Enzymes involved in the bioconversion of ester-based prodrugs. *J Pharm Sci*. 2006;*95(6)*:1177-95.
24. Avdeef A, Bendels S, Di L, Faller B, Kansy M, Sugano K, et al. PAMPA--critical factors for better predictions of absorption. *J Pharm Sci*. 2007;*96(11)*:2893-909.
25. Kansy M, Senner F, Gubernator K. Physicochemical high throughput screening: parallel artificial membrane permeation assay in the description of passive absorption processes. *J Med Chem*. 1998;*41(7)*:1007-10.
26. Hubatsch I, Ragnarsson EG, Artursson P. Determination of drug permeability and prediction of drug absorption in Caco-2 monolayers. *Nat Protoc*. 2007;*2(9)*:2111-9.
27. Hou T, Wang J, Zhang W, Xu X. ADME evaluation in drug discovery. 7. Prediction of oral absorption by correlation and classification. *J Chem Inf Model*. 2007;*47(1)*:208-18.
28. Kleeb S, Pang L, Mayer K, Eris D, Sigl A, Preston RC, et al. FimH antagonists: bioisosteres to improve the in vitro and in vivo PK/PD profile. *J Med Chem*. 2015;*58(5)*:2221-39.
29. Cipolla L, La Ferla B, Airoldi C, Zona C, Orsato A, Shaikh N, et al. Carbohydrate mimetics and scaffolds: sweet spots in medicinal chemistry. *Future Med Chem*. 2010;*2(4)*:587-99.
30. Satoh T, Taylor P, Bosron WF, Sanghani SP, Hosokawa M, La Du BN. Current progress on esterases: from molecular structure to function. *Drug Metab Dispos*. 2002;*30(5)*:488-93.
31. Obach RS, Baxter JG, Liston TE, Silber BM, Jones BC, MacIntyre F, et al. The prediction of human pharmacokinetic parameters from preclinical and in vitro metabolism data. *J Pharmacol Exp Ther*. 1997;*283(1)*:46-58.
32. Taketani M, Shii M, Ohura K, Ninomiya S, Imai T. Carboxylesterase in the liver and small intestine of experimental animals and human. *Life Sci*. 2007;*81(11)*:924-32.
33. Aprile S, Del Grosso E, Grosa G. In vitro and in vivo phase II metabolism of combretastatin A-4: evidence for the formation of a sulphate conjugate metabolite. *Xenobiotica*. 2009;*39(2)*:148-61.
34. Chapman E, Best MD, Hanson SR, Wong CH. Sulfotransferases: structure, mechanism, biological activity, inhibition, and synthetic utility. *Angew Chem Int Ed Engl*. 2004;*43(27)*:3526-48.
35. Zhang Y, Huo M, Zhou J, Xie S. PKSolver: An add-in program for pharmacokinetic and pharmacodynamic data analysis in Microsoft Excel. *Comput Methods Programs Biomed*. 2010;*99(3)*:306-14.
36. Aller SG, Yu J, Ward A, Weng Y, Chittaboina S, Zhuo R, et al. Structure of P-glycoprotein reveals a molecular basis for poly-specific drug binding. *Science*. 2009;*323(5922)*:1718-22.

Supplementary Data

Table S1 Physicochemical and in vitro pharmacokinetic properties of ester prodrugs tested. Measurements were conducted in sexduplicate ($logD_{7.4}$), quadruplicate ($\log P_e$) or triplicate.

Compound	Molecular weight [Da]	$\log D_{7.4}$	PAMPA $\log P_e$	Caco-2 cell permeation assay P_{app} [10^{-6} cm/s]			Stability Intestinal fluid Pancreatin [min]	Stability HLM Hydrolysis [min]	Stability RLM Oxidation [min]	Stability RLM Hydrolysis [min]
				Absorption a-b	Efflux b-a	Efflux ratio b-a/a-b				
1	600.63	< -1.5	< -10							
2a	592.67	1.3 ± 0.1	-7.3 ± 0.0				< 10	153		
2b	606.33	1.4 ± 0.1	-8.1 ± 1.4				< 10	162		
2c	620.34	1.8 ± 0.1	-6.4 ± 0.4					51.5		
2d	634.36	2.2 ± 0.1	-6.1 ± 0.2	0.8 ± 0.1	4.6 ± 1.5	5.7	< 10	42.3		
2e	648.37	2.4 ± 0.1	-5.3 ± 0.3	0.3 ± 0.4	11.5 ± 0.5	36	< 10	52.6		
2f	662.86	2.6 ± 0.1	-5.1 ± 0.2	0.25 ± 0.1	23.7 ± 2.9	94	< 10	45.2		
2g	690.86	> 2.5	-5.0 ± 0.4	0.89 ± 0.51	6.7 ± 2.3	7.5	< 10	36.3		
2h	648.37	2.4 ± 0.1	-5.7 ± 0.2	0.7 ± 0.2	6.6 ± 1.3	9	< 10	60.9		
2i	648.37	2.5 ± 0.1	-5.7 ± 0.1				< 10	41.9		
2j	690.42	> 2.5	-4.8 ± 0.1	0.8 ± 0.7	35.7 ± 1.3	44	< 10			
2k	646.77	2.5 ± 0.1	-5.4 ± 0.1	0.84 ± 0.02	4.8 ± 0.4	5.8	< 10	52.1	50 ± 20	8.1 ± 0.1
2l	660.80	2.6 ± 0.1	-5.1 ± 0.1	0.76 ± 0.03	11.1 ± 0.3	15	< 10	30.7		
2m	674.83	2.3 ± 0.1	-4.8 ± 0.2	1.0 ± 0.2	19.0 ± 0.5	19	< 10	96.3		
3	600.66	0.6 ± 0.1	-8.8 ± 0.2				29			
7	606.71	1.2 ± 0.1	-7.8 ± 0.0				> 120			
11	674.82	3.7 ± 0.1	-4.9 ± 0.1	0.4 ± 0.1	6.9 ± 0.1	19	119		10.7 ± 0.3	23.5 ± 1.1
5	538.60	-1.2 ± 0.1	< -10				> 120			
13	648.37	2.4 ± 0.1	-5.5 ± 0.1	0.59 ± 0.05	6.5 ± 0.7	11	25		27 ± 2	30 ± 2
14	707.17	4.0 ± 0.2	-4.6 ± 0.2	0.49 ± 0.06	23.5 ± 1.3	48	77		62 ± 3	72 ± 6

Synthesis



Scheme S1 Reactions and reaction conditions: a) $KHCO_3$, DMF, 4°C – 15°C, 4 h - 6 d, 15 – 80%. The Synthesis of **1** was described by Kolb *et al.* (1).

General procedure:

2a) To a mixture of **1** (8.0 mg, 0.0133 mmol) and $KHCO_3$ (2.7 mg, 0.0266 mmol in dry DMF (150 µL) at 4°C, Methyliodide (1.24 µL, 0.020 mmol) was added. After 4 h at this temperature the reaction mixture was purified by column chromatography on silica gel (DCM, MeOH 6:1). The obtained product was diluted in dioxane, filtered through an Acrodisc filter (0.45 µm) and lyophilized to afford **2a** (4.0 mg, 50%); $[\alpha]_D^{20}$ -84.74 (*c* 0.40, MeOH).

^{1}H NMR (500 MHz, Methanol-d_4) δ 1.19 (d, *J* = 6.6 Hz, 3H, H-6^F), 0.80 – 1.03, 1.21 – 1.44, 1.46 – 1.79, 1.83 – 1.97, 1.98 – 2.13 (m, 21H, $C_7H_{13}{}^L$, H-3^{cyc}, H-4^{cyc}, H-5^{cyc}, H-6^{cyc}), 3.25 (dd, *J* = 9.5, 3.3 Hz, 1H, H-3^G), 3.41 (ddd, *J* = 6.6, 5.3, 1.3 Hz, 1H, H-5^G), 3.54 (ddd, *J* = 9.9, 8.1, 4.4 Hz, 1H, H-2^{cyc}), 3.59 (dd, *J* = 9.5, 7.7 Hz, 1H, H-2^G), 3.64 – 3.78 (m, 5H, H-1^{cyc}, H-2^F, H-4^F, H-6^G), 3.73 (s, 3 H, OMe), 3.86 (dd, *J* = 10.1, 3.4 Hz, 1H, H-3^F), 3.91 (dd, *J* = 3.4, 1.1 Hz, 1H, H-4^G), 4.27 (d, *J* = 7.7 Hz, 1H, H-1^G), 4.44 (dd, *J* = 9.1, 4.0 Hz, 1H, H-2^L), 4.52 – 4.65 (m, 1H, H-5^F), 4.80 – 4.91 (m, 1H, H-1^F); ^{13}C NMR (126 MHz, MeOD) δ 16.60 (C-6^F), 24.44, 27.19, 27.44, 27.67, 30.09, 30.76, 30.88, 33.54, 34.67, 35.09, 42.33 ($C_7H_{13}{}^L$, C-3^{cyc}, C-4^{cyc}, C-5^{cyc}, C-6^{cyc}), 52.55 (OMe), 62.75 (C-6^G), 67.44 (C-5^F), 68.65 (C-4^G), 70.04 (C-2^F), 71.57 (C-3^F), 72.56 (C-2^G), 73.82 (C-4^F), 76.07 (C-5^G), 77.48 (C-2^{cyc}), 78.56 (C-2^L), 79.24 (C-1^{cyc}), 83.90 (C-3^G), 97.38 (C-1^F), 102.63 (C-1^G), 177.20 (C=O);

ESI-MS *m/z* calcd for $C_{28}H_{48}NaO_{13}$ $[M+Na]^+$: 615.30; found: 615.37.

2b) According to general procedure, **1** (9.8 mg, 0.0163 mmol), $KHCO_3$ (3.2 mg, 0.0326 mmol) and iodoethane (2 µL, 0.0245 mmol), 4°C, 4 d, gave **2b** (7.5 mg, 75%). $[\alpha]_D^{20}$ -94.48 (*c* 0.75, MeOH).

^{1}H NMR (500 MHz, Methanol-d_4) δ 0.83 – 1.03, 1.20 – 1.44, 1.51 – 1.59, 1.62 – 1.75, 1.84 – 1.96, 1.96 – 2.13 (m, 24H, $C_7H_{13}{}^L$, H-3^{cyc}, H-4^{cyc}, H-5^{cyc}, H-6^{cyc}, CH_3), 1.19 (d, *J* = 6.6 Hz, 3H, H-6^F), 3.25 (dd, *J* = 9.5, 3.3 Hz, 1H, H-3^G), 3.41 (ddd, *J* = 6.7, 5.2, 1.2 Hz, 1H, H-5^G), 3.54 (ddd, *J* = 9.9, 8.2, 4.4 Hz, 1H, H-2^{cyc}), 3.59 (dd, *J* = 9.5, 7.7 Hz, 1H, H-2^G), 3.64 – 3.78 (m, 5H, H-1^{cyc}, H-2^F, H-4^F, H-6^G), 3.86 (dd, *J* = 10.1, 3.4 Hz, 1H, H-3^F), 3.91 (dd, *J* = 3.3, 1.2 Hz, 1H, H-4^G), 4.15 – 4.25 (m, 2H, OCH_2), 4.27 (d, *J* = 7.7 Hz, 1H, H-1^G), 4.41 (dd, *J* = 9.1, 4.1 Hz, 1H, H-2^L), 4.58 – 4.63 (m, 1H, H-5^F), 4.76 – 5.02 (m, 1H, H-1^F);

^{13}C NMR (126 MHz, MeOD) δ 14.50 (CH_3), 16.61 (C-6^F), 24.45, 27.22, 27.45, 27.67, 30.10, 30.89, 33.58, 34.70, 35.09, 42.36 (11C, $C_7H_{13}{}^L$, C-3^{cyc}, C-4^{cyc}, C-5^{cyc}, C-6^{cyc}), 62.26 (OCH_2), 62.73 (C-6^G), 67.44 (C-5^F), 68.65 (C-4^G), 70.03 (C-2^F), 71.56 (C-3^F), 72.54 (C-2^G), 73.82 (C-4^F), 76.04 (C-5^G), 77.50 (C-2^{cyc}), 78.66 (C-2^L), 79.26 (C-1^{cyc}), 83.96 (C-3^G), 97.38 (C-1^F), 102.62 (C-1^G), 176.80 (C=O);

ESI-MS *m/z* calcd for $C_{29}H_{50}NaO_{13}$ $[M+Na]^+$: 629.31; found: 629.39.

2c) According to general procedure, **1** (10.1 mg, 0.0168 mmol), $KHCO_3$ (3.4 mg, 0.0336 mmol) and iodopropane (2.5 µL, 0.0252 mmol), 4°C, 3 d, gave **2c** (7.0 mg, 70%). $[\alpha]_D^{20}$ -87.19 (*c* 0.68, MeOH).

^{1}H NMR (500 MHz, Methanol-d_4) δ 0.82 – 0.94, 1.21 – 1.43, 1.49 – 1.76, 1.87 – 1.96, 1.99 – 2.10 (m, 23H, $C_7H_{13}{}^L$, CH_2, H-3^{cyc}, H-4^{cyc}, H-5^{cyc}, H-6^{cyc}), 0.96 (t, *J* = 7.5 Hz, 3H, CH_3), 1.19 (d, *J* = 6.7 Hz, 3H, H-6^F),

3.25 (dd, *J* = 9.5, 3.3 Hz, 1H, H-3^{G}), 3.39 – 3.43 (m, 1H, H-5^{G}), 3.51 – 3.57 (m, 1H, H-2^{cyc}), 3.59 (dd, *J* = 9.5, 7.7 Hz, 1H, H-2^{G}), 3.64 – 3.77 (m, 5H, H-1^{cyc}, H-2^{F}, H-4^{F}, H-6^{G}), 3.86 (dd, *J* = 10.1, 3.4 Hz, 1H, H-3^{F}), 3.91 (d, *J* = 3.2 Hz, 1H, H-4^{G}), 4.05 – 4.17 (m, 2H, OCH_2), 4.27 (d, *J* = 7.7 Hz, 1H, H-1^{G}), 4.43 (dd, *J* = 9.0, 4.2 Hz, 1H, H-2^{L}), 4.60 (q, *J* = 6.6 Hz, 1H, H-5^{F}), 4.78 – 4.90 (m, 1H, H-1^{F});

^{13}C NMR (126 MHz, MeOD) δ 10.74 (CH_3), 16.60 (C-6^{F}), 23.04, 24.44, 27.22, 27.44, 27.67, 30.09, 30.89, 33.63, 34.72, 35.08, 42.45 (12C, C_7H_{13}^{L}, CH_2, C-3^{cyc}, C-4^{cyc}, C-5^{cyc}, C-6^{cyc}), 62.73 (C-6^{G}), 67.44 (C-5^{F}), 67.82 (OCH_2), 68.66 (C-4^{G}), 70.04 (C-2^{F}), 71.57 (C-3^{F}), 72.55 (C-2^{G}), 73.83 (C-4^{F}), 76.06 (C-5^{G}), 77.49 (C-2^{cyc}), 78.67 (C-2^{L}), 79.25 (C-1^{cyc}), 83.94 (C-3^{G}), 97.38 (C-1^{F}), 102.64 (C-1^{G}), 176.87 (C=O); ESI-MS *m/z* calcd for $C_{30}H_{52}NaO_{13}$ $[M+Na]^+$: 643.33; found: 643.47.

2d) According to general procedure, **1** (9.9 mg, 0.0165 mmol), $KHCO_3$ (3.3 mg, 0.0330 mmol) and iodobutane (2.8 µL, 0.0248 mmol), 4°C, 3 d, gave **2d** (6.0 mg, 60%). $[\alpha]_D^{20}$ -95.80 (*c* 0.61, MeOH).

^{1}H NMR (500 MHz, Methanol-*d*$_4$) δ 0.82 – 0.94, 1.21 - 1.48, 1.50 – 1.77, 1.87 – 1.96, 1.98 – 2.13 (m, 25H, C_7H_{13}^{L}, H-3^{cyc}, H-4^{cyc}, H-5^{cyc}, H-6^{cyc}, 2 x CH_2), 0.96 (t, *J* = 7.4 Hz, 3H, CH_3), 1.19 (d, *J* = 6.6 Hz, 3H, H-6^{F}), 3.25 (dd, *J* = 9.5, 3.3 Hz, 1H, H-3^{G}), 3.41 (ddd, *J* = 6.6, 5.1, 1.2 Hz, 1H, H-5^{G}), 3.54 (ddd, *J* = 9.9, 8.2, 4.5 Hz, 1H, H-2^{cyc}), 3.59 (dd, *J* = 9.5, 7.7 Hz, 1H, H-2^{G}), 3.64 – 3.77 (m, 5H, H-1^{cyc}, H-2^{F}, H-4^{F}, H-6^{G}), 3.86 (dd, *J* = 10.1, 3.4 Hz, 1H, H-3^{F}), 3.91 (dd, *J* = 3.4, 1.1 Hz, 1H, H-4^{G}), 4.10 – 4.22 (m, 2H, OCH_2), 4.27 (d, *J* = 7.7 Hz, 1H, H-1^{G}), 4.42 (dd, *J* = 8.9, 4.2 Hz, 1H, H-2^{L}), 4.56 – 4.64 (m, 1H, H-5^{F}), 4.82 – 4.88 (m, 1H, H-1^{F});

^{13}C NMR (126 MHz, MeOD) δ 14.00 (CH_3), 16.60 (C-6^{F}), 20.21, 24.44, 27.23, 27.44, 27.67, 30.09, 30.89, 31.79, 33.65, 34.73, 35.06, 42.42 (13C, C_7H_{13}^{L}, C-3^{cyc}, C-4^{cyc}, C-5^{cyc}, C-6^{cyc}, 2 x CH_2), 62.74 (C-6^{G}), 66.02 (OCH_2), 67.43 (C-5^{F}), 68.66 (C-4^{G}), 70.03 (C-2^{F}), 71.56 (C-3^{F}), 72.55 (C-2^{G}), 73.82 (C-4^{F}), 76.06 (C-5^{G}), 77.49 (C-2^{cyc}), 78.68 (C-2^{L}), 79.25 (C-1^{cyc}), 83.93 (C-3^{G}), 97.38 (C-1^{F}), 102.64 (C-1^{G}), 176.85 (C=O);

ESI-MS *m/z* calcd for $C_{31}H_{54}NaO_{13}$ $[M+Na]^+$: 657.35; found: 657.43.

2e) According to general procedure, **1** (9.0 mg, 0.0150 mmol), $KHCO_3$ (3.0 mg, 0.0300 mmol) and iodopentane (2.9 µL, 0.0225 mmol), 4°C, 3 d, gave **2e** (7.9 mg, 80%). $[\alpha]_D^{20}$ -86.93 (*c* 0.79, MeOH).

^{1}H NMR (500 MHz, Methanol-*d*$_4$) δ 0.81 – 1.03, 1.21 – 1.45, 1.47 – 1.78, 1.85 – 1.97, 1.99 – 2.11 (m, 30H, C_7H_{13}^{L}, H-3^{cyc}, H-4^{cyc}, H-5^{cyc}, H-6^{cyc}, 3 x CH_2, CH_3), 1.19 (d, *J* = 6.7 Hz, 3H, H-6^{F}), 3.25 (dd, *J* = 9.5, 3.3 Hz, 1H, H-3^{G}), 3.41 (ddd, *J* = 6.8, 5.3, 1.2 Hz, 1H, H-5^{G}), 3.54 (ddd, *J* = 9.8, 8.1, 4.4 Hz, 1H, H-2^{cyc}), 3.59 (dd, *J* = 9.5, 7.7 Hz, 1H, H-2^{G}), 3.62 – 3.78 (m, 5H, H-1^{cyc}, H-2^{F}, H-4^{F}, H-6^{G}), 3.86 (dd, *J* = 10.1, 3.4 Hz, 1H, H-3^{F}), 3.91 (dd, *J* = 3.1, 1.1 Hz, 1H, H-4^{G}), 4.09 – 4.22 (m, 2H, OCH_3), 4.27 (d, *J* = 7.7 Hz, 1H, H-1^{G}), 4.42 (dd, *J* = 8.7, 4.3 Hz, 1H, H-2^{L}), 4.60 (q, *J* = 6.6 Hz, 1H, H-5^{F}), 4.79 – 4.92 (m, 1H, H-1^{F});

^{13}C NMR (126 MHz, MeOD) δ 14.33 (CH_3), 16.61 (C-6^{F}), 23.35, 24.44, 24.47, 27.23, 27.44, 27.67, 29.27, 29.42, 30.09, 30.89, 33.67, 34.74, 35.05, 42.42 (C_7H_{13}^{L},C-3^{cyc}, C-4^{cyc}, C-5^{cyc}, C-6^{cyc}, 3 x CH_2), 62.76 (C-6^{G}), 66.30 (OCH_2), 67.43 (C-5^{F}), 68.68 (C-4^{G}), 70.03 (C-2^{F}), 71.56 (C-3^{F}), 72.54 (C-2^{G}), 73.82 (C-4^{F}), 76.07 (C-5^{G}), 77.48 (C-2^{cyc}), 78.69 (C-2^{L}), 79.25 (C-1^{cyc}), 83.92 (C-3^{G}), 97.37 (C-1$^{F)}$, 102.64 (C-1^{G}), 176.84 (C=O);

ESI-MS *m/z* calcd for $C_{32}H_{56}NaO_{13}$ $[M+Na]^+$: 671.36; found: 671.42.

2f) According to general procedure, **1** (9.9 mg, 0.0165 mmol), $KHCO_3$ (3.3 mg, 0.0330 mmol) and iodohexane (3.6 µL, 0.0248 mmol), 4°C, 4 d, gave **2f** (5.0 mg, 48%). $[\alpha]_D^{20}$ -73.51 (*c* 0.80, MeOH).

^{1}H NMR (500 MHz, Methanol-*d*$_4$) δ 0.83 – 1.02, 1.21 – 1.42, 1.50 – 1.77, 1.88 – 1.95, 1.99 – 2.10 (m, 32H, C_7H_{13}^{L}, H-3^{cyc}, H-4^{cyc}, H-5^{cyc}, H-6^{cyc}, 4 x CH_2, CH_3), 1.19 (d, *J* = 6.6 Hz, 3H, H-6^{F}), 3.25 (dd, *J* = 9.5, 3.3 Hz, 1H, H-3^{G}), 3.41 (ddd, *J* = 6.7, 5.2, 1.2 Hz, 1H, H-5^{G}), 3.54 (ddd, *J* = 9.8, 8.2, 4.4 Hz, 1H, H-2^{cyc}), 3.59 (dd, *J* = 9.5, 7.7 Hz, 1H, H-2^{G}), 3.63 – 3.78 (m, 5H, H-1^{cyc}, H-2^{F}, H-4^{F}, H-6^{G}), 3.86 (dd, *J* = 10.1, 3.4 Hz, 1H, H-3^{F}), 3.91 (dd, *J* = 3.4, 1.1 Hz, 1H, H-4^{G}), 4.08 – 4.22 (m, 2H, OCH_2), 4.27 (d, *J* = 7.7 Hz, 1H, H-1^{G}), 4.42 (dd, *J* = 8.8, 4.3 Hz, 1H, H-2^{L}), 4.60 (q, *J* = 6.6 Hz, 1H, H-5^{F}), 4.81 – 4.89 (m, 1H, H-1^{F});

^{13}C NMR (126 MHz, MeOD) δ 14.36 (CH_3), 16.60 (C-6^{F}), 23.64, 24.44, 26.76, 27.22, 27.44, 27.67, 29.68, 30.09, 30.89, 32.56, 33.67, 34.73, 35.05, 42.43 (15C, C_7H_{13}^{L}, C-3^{cyc}, C-4^{cyc}, C-5^{cyc}, C-6^{cyc}, 4 x CH_2), 62.77 (C-6^{G}), 66.30 (OCH_2), 67.43 (C-5^{F}), 68.68 (C-4^{G}), 70.03 (C-2^{F}), 71.56 (C-3^{F}), 72.54 (C-2^{G}), 73.82 (C-4^{F}),

76.08 (C-5^{G}), 77.48 (C-2^{cyc}), 78.68 (C-2^{L}), 79.25 (C-1^{cyc}), 83.92 (C-3^{G}), 97.38 (C-1^{F}), 102.64 (C-1^{G}), 176.84 (C=O);

ESI-MS *m/z* calcd for $C_{33}H_{58}NaO_{13}$ $[M+Na]^+$: 685.38; found: 685.45.

2g) According to general procedure, **1** (9.8 mg, 0.0163 mmol), $KHCO_3$ (3.2 mg, 0.0326 mmol) and iodooctane (4.4 µL, 0.0245 mmol), 4 - 15°C, 4 h, gave **2g** (8.0 mg, 70%). %). $[\alpha]_D^{20}$ -136.14 (*c* 0.53, MeOH).

^{1}H NMR (500 MHz, Methanol-d_4) δ 0.85 – 1.03, 1.21 – 1.42, 1.50 – 1.76, 1.86 – 1.95, 1.97 – 2.10 (m, 36H, C_7H_{13}^{L}, H-3^{cyc}, H-4^{cyc}, H-5^{cyc}, H-6^{cyc}, 6 x CH_2, CH_3), 1.19 (d, *J* = 6.7 Hz, 3H, H-6^{F}), 3.25 (dd, *J* = 9.5, 3.3 Hz, 1H, H-3$^{G)}$, 3.41 (ddd, *J* = 6.7, 5.2, 1.3 Hz, 1H, H-5^{G}), 3.54 (ddd, *J* = 9.8, 8.1, 4.4 Hz, 1H, H-2^{cyc}), 3.59 (dd, *J* = 9.5, 7.7 Hz, 1H, H-2^{G}), 3.64 – 3.78 (m, 5H, H-1^{cyc}, H-2^{F}, H-4^{F}, H-6^{G}), 3.86 (dd, *J* = 10.1, 3.4 Hz, 1H, H-3^{F}), 3.91 (dd, *J* = 3.3, 1.1 Hz, 1H, H-4^{G}), 4.08 – 4.22 (m, 2H, OCH_2), 4.27 (d, *J* = 7.7 Hz, 1H, H-1^{G}), 4.43 (dd, *J* = 8.7, 4.3 Hz, 1H, H-2^{L}), 4.57 – 4.63 (m, 1H, H-5^{F}), 4.77 – 4.96 (m, 1H, H-1^{F});

^{13}C NMR (126 MHz, MeOD) δ 14.45 (CH_3), 16.60 (H-6^{F}), 23.72, 24.44, 27.09, 27.22, 27.44, 27.67, 29.71, 30.09, 30.29, 30.36, 30.90, 32.96, 33.68, 34.74, 35.05, 42.44 (17C, C_7H_{13}^{L}, C-3^{cyc}, C-4^{cyc}, C-5^{cyc}, C-6^{cyc}, 6 x CH_2), 62.77 (C-6^{G}), 66.30 (OCH_2), 67.44 (C-5^{F}), 68.69 (C-4^{G}), 70.03 (C-2^{F}), 71.56 (C-3^{F}), 72.54 (C-2^{G}), 73.82 (C-4^{F}), 76.08 (C-5^{G}), 77.48 (C-2^{cyc}), 78.69 (C-2^{L}), 79.26 (C-1^{cyc}), 83.91 (C-3^{G}), 97.37 (C-1^{F}), 102.64 (C-1^{G}), 176.84 (C=O);

ESI-MS *m/z* calcd for $C_{35}H_{62}NaO_{13}$ $[M+Na]^+$: 713.41; found: 713.45.

2h) According to general procedure, **1** (9.8 mg, 0.0163 mmol), $KHCO_3$ (3.3 mg, 0.0326 mmol) and (S)-(+)-1-Iodo-2-methylbutane (3.2 µL, 0.0245 mmol), 4°C, 6 d, gave **2h** (1.5 mg, 15%). $[\alpha]_D^{20}$ -45.07 (*c* 0.16, MeOH);

^{1}H NMR (500 MHz, Methanol-d_4) δ 0.77 – 1.04, 1.19, 1.20 – 1.51, 1.51 – 1.78, 1.87 – 1.96, 1.98 – 2.11 (m, 30H, C_7H_{13}^{L}, H-3^{cyc}, H-4^{cyc}, H-5^{cyc}, H-6^{cyc}, CH, CH_2, 2 x CH_3), (d, *J* = 6.6 Hz, 3H, H-6^{F}), 3.26 (dd, *J* = 9.5, 3.3 Hz, 1H, H-3^{G}), 3.40 (ddd, *J* = 6.5, 5.1, 1.1 Hz, 1H, H-5^{G}), 3.50 – 3.56 (m, 1H, H-2^{cyc}), 3.59 (dd, *J* = 9.5, 7.7 Hz, 1H, H-2^{G}), 3.63 – 3.77 (m, 5H, H-1^{cyc}, H-2^{F}, H-4^{F}, H-6G), 3.86 (dd, *J* = 10.1, 3.4 Hz, 1H, H-3^{F}), 3.90 – 3.92 (m, 1H, H-4^{G}), 3.97 – 4.05 (m, 2H, OCH_2), 4.27 (d, *J* = 7.7 Hz, 1H, H-1^{G}), 4.44 (dd, *J* = 8.7, 4.3 Hz, 1H, H-2^{L}), 4.58 – 4.63 (m, 1H, H-5^{F}), 4.75 – 5.02 (m, 1H, H-1^{F});

^{13}C NMR (126 MHz, MeOD) δ 11.56, 16.59 (CH_3), 16.78 (C-6^{F}), 24.43, 27.06, 27.22, 27.42, 27.65, 30.08, 30.67, 33.69, 34.74, 35.04, 35.53, 38.05, 42.47 (C_7H_{13}^{L}, C-3^{cyc}, C-4^{cyc}, C-5^{cyc}, C-6^{cyc}, CH, CH_2), 62.74 (C-6^{G}), 67.43 (C-5^{F}), 68.67 (C-4^{G}), 70.03 (C-2^{F}), 70.67 (OCH_2), 71.56 (C-3^{F}), 72.56 (C-2^{G}), 73.82 (C-4^{F}), 76.07 (C-5^{G}), 77.50 (C-2^{cyc}), 78.67 (C-2^{L}), 79.26 (C-1^{cyc}), 83.88 (C-3^{G}), 97.37 (C-1^{F}), 102.63 (C-1^{G});

ESI-MS *m/z* calcd for $C_{32}H_{56}NaO_{13}$ $[M+Na]^+$: 671.36; found: 671.47.

2i) According to general procedure, **1** (11.0 mg, 0.0183 mmol), $KHCO_3$ (3.7 mg, 0.0366 mmol) and 1-iodo-3-methylbutane (3.63 µL, 0.0275 mmol), 4°C, 6 d, gave **2i** (8.2 mg, 69%). $[\alpha]_D^{20}$ -78.88 (*c* 0.82, MeOH).

^{1}H NMR (500 MHz, Methanol-d_4) δ 0.84 – 1.02, 1.21 – 1.42, 1.51 – 1.59, 1.61 – 1.77, 1.87 – 1.95, 1.99 – 2.12 (m, 24H, C_7H_{13}^{L}, H-3^{cyc}, H-4^{cyc}, H-5^{cyc}, H-6^{cyc}), 0.96 (d, *J* = 6.6 Hz, 3H, CH_3), 0.97 (d, *J* = 6.6 Hz, 3H, CH_3), 1.19 (d, *J* = 6.6 Hz, 3H, H-6^{F}), 3.25 (dd, *J* = 9.5, 3.3 Hz, 1H, H-3^{G}), 3.41 (ddd, *J* = 6.6, 5.2, 1.3 Hz, 1H, H-5^{G}), 3.54 (ddd, *J* = 9.8, 8.1, 4.4 Hz, 1H, H-2^{cyc}), 3.59 (dd, *J* = 9.5, 7.7 Hz, 1H, H-2^{G}), 3.62 – 3.78 (m, 5H, H-1^{cyc}, H-2^{F}, H-4^{F}, H-6^{G}), 3.86 (dd, *J* = 10.1, 3.4 Hz, 1H, H-3^{F}), 3.91 (dd, *J* = 3.4, 1.1 Hz, 1H, H-4^{G}), 4.12 – 4.25 (m, 2H, OCH_2), 4.27 (d, *J* = 7.7 Hz, 1H, H-1^{G}), 4.42 (dd, *J* = 8.8, 4.3 Hz, 1H, H-2^{L}), 4.60 (q, *J* = 6.4 Hz, 1H, H-5^{F}), 4.78 – 4.89 (m, 1H, H-1^{F});

^{13}C NMR (126 MHz, MeOD) δ 16.61(H-6^{F}), 22.73, 22.82 (CH_3), 24.44, 24.47, 26.23, 27.23, 27.44, 27.66, 30.09, 30.90, 33.67, 34.74, 35.05, 38.49, 42.40 (C_7H_{13}^{L}, C-3^{cyc}, C-4^{cyc}, C-5^{cyc}, C-6^{cyc}, CH, CH_2), 62.75 (C-6^{G}), 64.76 (OCH_2), 67.44 (C-5^{F}), 68.67 (C-4^{G}), 70.03 (C-2^{F}), 71.57 (C-3^{F}), 72.55 (C-2^{G}), 73.82 (C-4^{F}), 76.07 (C-5^{G}), 77.49 (C-2^{cyc}), 78.70 (C-2^{L}), 79.26 (C-1^{cyc}), 83.92 (C-3^{G}), 97.38 (C-1^{F}), 102.63 (C-1^{G}), 176.82 (C=O);

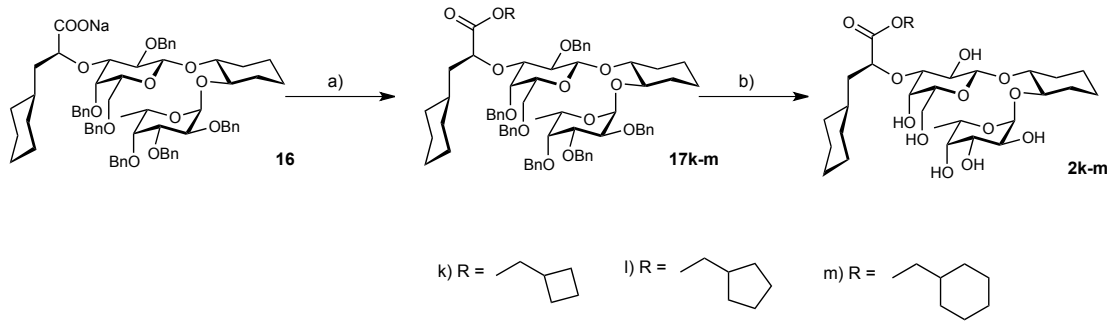
ESI-MS *m/z* calcd for $C_{32}H_{56}NaO_{13}$ $[M+Na]^+$: 671.36; found: 671.44.

2j) According to general procedure, **1** (9.9 mg, 0.0165 mmol), $KHCO_3$ (3.3 mg, 0.0330 mmol) and 2-ethyl hexyl iodide (4.4 µL, 0.0247 mmol), 4°C, 6 d, gave **2j** (2.5 mg, 21%). $[\alpha]_D^{20}$ -40.76 (*c* 0.32, MeOH);

^{1}H NMR (500 MHz, Methanol-d_4) δ 0.82 – 1.00, 1.20 – 1.44, 1.52 – 1.77, 1.87 – 1.95, 1.99 – 2.10, 2.31 – 2.38 (m, 36H, $C_7H_{13}^{L}$, H-3^{cyc}, H-4^{cyc}, H-5^{cyc}, H-6^{cyc}, CH, 4 x CH_2, 2 x CH_3), 1.19 (d, *J* = 6.7 Hz, 3H, H-6^{F}), 3.25 (dd, *J* = 9.5, 3.3 Hz, 1H, H-3^{G}), 3.38 – 3.43 (m, 1H, H-5^{G}), 3.51 – 3.57 (m, 1H, H-2^{cyc}), 3.59 (dd, *J* = 9.5, 7.7 Hz, 1H, H-2^{G}), 3.62 – 3.78 (m, 5H, H-1^{cyc}, H-2^{F}, H-4^{F}, H-6^{G}), 3.86 (dd, *J* = 10.1, 3.4 Hz, 1H, H-3^{F}), 3.91 (dd, *J* = 3.4, 1.1 Hz, 1H, H-4^{G}), 4.06 (ddd, *J* = 11.0, 5.5, 2.3 Hz, 1H, OCH_2), 4.13 (ddd, *J* = 10.1, 5.6, 4.3 Hz, 1H, OCH_2), 4.27 (d, *J* = 7.7 Hz, 1H, H-1^{G}), 4.44 (dd, *J* = 8.5, 4.4 Hz, 1H, H-2^{L}), 4.58 – 4.63 (m, 1H, H-5^{F}), 4.80 – 4.91 (m, 1H, H-1^{F});

^{13}C NMR (126 MHz, MeOD) δ 11.37, 14.41 (CH_3), 16.60 (C-6^{F}), 24.05, 24.45, 24.93, 27.20, 27.40, 27.66, 30.09, 30.77, 31.62, 33.74, 34.76, 35.03, 40.28, 42.48, 43.88 (16C, $C_7H_{13}^{L}$, C-3^{cyc}, C-4^{cyc}, C-5^{cyc}, C-6^{cyc}, CH, 4 x CH_2), 62.79 (C-6^{G}), 67.44 (C-5^{F}), 68.25 (OCH_2), 68.71 (C-4^{G}), 70.03 (C-2^{F}), 71.55 (C-3^{F}), 72.56 (C-2^{G}), 73.83 (C-4^{F}), 76.11 (C-5^{G}), 77.49 (C-2^{cyc}), 78.70 (C-2^{L}), 79.28 (C-1^{cyc}), 83.89 (C-3^{G}), 97.38 (C-1^{F}), 102.65 (C-1^{G}), 176.90 (C=O);

ESI-MS *m/z* calcd for $C_{35}H_{62}NaO_{13}$ $[M+Na]^+$: 713.41; found: 713.53.



Scheme S2 Reactions and reaction conditions: a) DCM, 1-chloro-N,N-2-trimethylpropenylamin, rt, 2 h, 46 -73%; b) $Pd(OH)_2$, dioxane, water, H_2, rt, 1 -2 h, 62-93%. The synthesis of **16** was described by Norman *et. al* (2).

17k) To a mixture of **16** (21.3 mg, 0.019 mmol) in dry DCM (1 mL) under an argon atmosphere at rt, 1-chloro-N,N-2-trimethylpropenylamin (3.2 µL, 0.022 mmol) was added. After stirring for 2 h triethylamine (13 µL, 0.094 mmol) and cyclobutanemethanol (17.6 µL, 0.187 mmol) were added and the mixture was stirred for another hour. Concentration in *vacuo* and column chromatography on silica gel (petroleum ether/ethyl acetate 6:1) of the remains gave **17k** (14.0 mg, 63%). $[\alpha]_D^{20}$ -80.61 (*c* 0.70, DCM);

^{1}H NMR (500 MHz, Chloroform-*d*) δ 0.77 – 0.98, , 1.12 – 1.40, 1.51 – 1.79, 1.79 – 1.91, 1.91 – 2.10 (m, 27H, $C_7H_{13}^{L}$, H-3^{cyc}, H-4^{cyc}, H-5^{cyc}, H-6^{cyc}, 3 x CH_2 (C_5H_9)), 1.10 (d, *J* = 6.5 Hz, 3H, H-6^{F}), 2.59 (hept, *J* = 7.5 Hz, 1H, CH (C_5H_9)), 3.34 (m, 1H, H-4^{F}), 3.40 (dd, *J* = 9.6, 3.1 Hz, 1H, H-3^{G}), 3.47 – 3.57 (m, 2H, H-5^{G}, H-2^{cyc}), 3.59 (dd, *J* = 8.9, 5.2 Hz, 1H, H-6a^{G}), 3.62 – 3.73 (m, 3H, H-6b^{G}, H-1^{cyc}, H-2^{G}), 3.91 – 3.94 (m, 2H, H-2^{F}, H-3^{F}), 4.02 (dd, *J* = 10.9, 6.9 Hz, 1H, OCH_2 (C_5H_9)), 4.09 – 4.16 (m, 2H, H-4^{G}, OCH_2 (C_5H_9)), 4.19 (d, *J* = 11.5 Hz, 1H, A of AB, CH_2Ph), 4.34 (d, *J* = 11.9 Hz, 1H, A of AB, CH_2Ph), 4.40 (d, *J* = 7.7 Hz, 1H, H-1^{G}), 4.43 (d, *J* = 11.9 Hz, 1H, B of AB, CH_2Ph), 4.49 (t, *J* = 6.4 Hz, 1H, H-2^{L}), 4.53 – 4.75 (m, 8H, H-5^{F}, CH_2Ph), 4.84 (m, 1H, H-1^{F}), 5.04 (d, *J* = 10.9 Hz, 1H, B of AB, CH_2Ph), 5.08 (d, *J* = 11.0 Hz, 1H, B of AB, CH_2Ph), 7.10 – 7.48 (m, 30H, 6 x C_6H_5);

^{13}C NMR (126 MHz, $CDCl_3$) δ 16.58 (C-6^{F}), 18.31, 23.43, 24.71, 24.80, 26.21, 26.24, 26.34, 29.41, 29.69, 30.05, 33.49, 33.60, 33.91, 34.20, 41.25 ($C_7H_{13}^{L}$, C-3^{cyc}, C-4^{cyc}, C-5^{cyc}, C-6^{cyc}, 4C C_5H_9), 65.94 (C-5^{F}), 68.58, 68.63 (C-6^{G}, OCH_2 C_5H_9), 72.55, 72.89, 72.93, 73.46, 74.61, 75.01, 75.19 (C-5^{G}, CH_2Ph), 75.71, 75.87, 76.10 (C-4^{G}, C-2^{F}, C-2^{cyc}), 77.69 (C-1^{cyc}), 78.78 (C-4^{F}), 79.22 (C-2^{L}), 79.84 (C-3^{F}), 80.72 (C-2^{G}), 81.49 (C-3^{G}), 94.73 (C-1^{F}), 101.16 (C-1^{G}), 127.15, 127.20, 127.25, 127.30, 127.35, 127.59, 127.83, 127.88, 128.01, 128.15, 128.18, 128.25, 128.36, 128.45, 138.10, 138.80, 138.98, 139.29, 139.36, 139.61 (36C, 6 x C_6H_5), 174.22 (C=O);

ESI-MS *m/z* calcd for $C_{74}H_{90}NaO_{13}$ $[M+Na]^+$: 1209.63; found: 1209.86.

17l) To a mixture of **16** (11.5 mg, 0.010 mmol) in dry DCM (500 µL) under a argon atmosphere at rt, 1-chloro-N,N-2-trimethylpropenylamin (1.7 µL, 0.012 mmol) was added. After stirring for 2 h triethylamine (7 µL, 0.051 mmol) and cyclopentanemethanol (11 µL, 0.101 mmol) were added and the mixture was stirred for another hour. Concentration in *vacuo* and column chromatography on silica gel (petroleum ether/ethyl acetate 6:1) of the remains gave **17l** (9.0 mg, 75%). $[\alpha]_D^{20}$ -69.18 (*c* 0.90, DCM);

^{1}H NMR (500 MHz, Chloroform-*d*) δ 0.77 – 0.95, 1.12 – 1.40, 1.40 – 1.82, 1.92 – 2.05 (m, 29H, $C_7H_{13}^L$, H-3^{cyc}, H-4^{cyc}, H-5^{cyc}, H-6^{cyc}, 4 x CH_2 C_6H_{11}), 1.10 (d, *J* = 6.5 Hz, 3H, H-6^F), 2.18 (hept, *J* = 7.5 Hz, 1H, CH C_6H_{11}), 3.34 (m, 1H, H-4^F), 3.41 (dd, *J* = 9.6, 3.1 Hz, 1H, H-3^G), 3.47 – 3.57 (m, 2H, H-5^G, H-2^{cyc}), 3.59 (dd, *J* = 8.9, 5.2 Hz, 1H, H-$6a^G$), 3.61 – 3.73 (m, 3H, H-$6b^G$, H-2^G, H-1^{cyc}), 3.91 – 3.96 (m, 3H, H-2^F, H-3^F, OCH_2 C_6H_{11}), 4.02 (dd, *J* = 10.7, 7.3 Hz, 1H, OCH_2 C_6H_{11}), 4.13 (d, *J* = 3.0 Hz, 1H, H-4^G), 4.19 (d, *J* = 11.4 Hz, 1H, A of AB, CH_2Ph), 4.33 (d, *J* = 12.0 Hz, 1H, A of AB, CH_2Ph), 4.38 – 4.45 (m, 2H, H-1^G, CH_2Ph), 4.49 (t, *J* = 6.4 Hz, 1H, H-2^L), 4.52 – 4.75 (m, 8H, H-5^F, CH_2Ph), 4.84 (m, 1H, H-1^F), 5.05 (d, *J* = 10.9 Hz, 1H, B of AB, CH_2Ph), 5.08 (d, *J* = 11.1 Hz, 1H, B of AB, CH_2Ph), 7.11 – 7.54 (m, 30H, 6 x C_6H_5);

^{13}C NMR (126 MHz, $CDCl_3$) δ 16.59 (C-6^F), 23.43, 25.31, 25.44, 26.21, 26.24, 26.35, 29.06, 29.37, 29.44, 29.69, 30.05, 33.52, 33.59, 34.21, 38.42, 42.16 ($C_7H_{13}^L$, C-3^{cyc}, C-4^{cyc}, C-5^{cyc}, C-6^{cyc}, 5C C_6H_{11}), 65.95 (C-5^F), 68.63, 68.74 (C-6^G, OCH_2 C_6H_{11}), 72.55, 72.89, 72.94, 73.45, 74.60, 75.01, 75.20 (C-5^G, CH_2Ph), 75.72 , 75.90, 76.11 (C-2^F, C4^G, C-2^{cyc}), 77.68 (C-1^{cyc}), 78.77 (C-4^F), 79.16 (C-2^L), 79.84 (C-3^F), 80.73 (C-2^G), 81.46 (C-3^G), 94.74 (C-1^F), 101.16 (C-1^G), 127.15, 127.21, 127.25, 127.29, 127.35, 127.58, 127.84, 127.88, 128.02, 128.15, 128.18, 128.25, 128.36, 128.45, 138.10, 138.81, 138.98, 139.29, 139.37, 139.62 (36C, 6 x C_6H_5), 174.21 (C=O);

ESI-MS *m/z* calcd for $C_{75}H_{92}NaO_{13}$ $[M+Na]^+$: 1223.64; found: 1223.84.

17m) To a mixture of **16** (26.0 mg, 0.023 mmol) in dry DCM (1 mL) under a argon atmosphere at rt, 1-chloro-N,N-2-trimethylpropenylamin (3.9 µL, 0.027 mmol) was added. After stirring for 2 h triethylamine (16 µL, 0.114 mmol) and cyclohexanemethanol (28 µL, 0.228 mmol) were added and the mixture was stirred for another 3 hours. Concentration in *vacuo* and column chromatography on silica gel (petroleum ether/ethyl acetate 6:1) of the remains gave **17m** (12.7 mg, 46%).$[\alpha]_D^{20}$ -77.28 (*c* 0.60, DCM);

^{1}H NMR (500 MHz, Chloroform-*d*) δ 0.79 – 1.01, 1.12 – 1.39, 1.39 – 1.51, 1.52 – 1.77, 1.91 – 2.06 (m, 32H, $C_7H_{13}^L$, H-3^{cyc}, H-4^{cyc}, H-5^{cyc}, H-6^{cyc}, 5 x CH_2, CH C_7H_{13}), 1.10 (d, *J* = 6.4 Hz, 3H, H-6F), 3.35 (m, 1H, H-4F), 3.41 (dd, *J* = 9.6, 3.0 Hz, 1H, H-3^G), 3.49 (dd, *J* = 7.6, 5.4 Hz, 1H, H-5^G), 3.51 – 3.56 (m, 1H, H-2^{cyc}), 3.58 (dd, *J* = 9.0, 5.2 Hz, 1H, H-$6a^G$), 3.61 – 3.72 (m, 3H, H-$6b^G$, H-2^G, H-1^{cyc}), 3.87 (dd, *J* = 10.7, 6.6 Hz, 1H, OCH_2 C_7H_{13}), 3.91 – 3.96 (m, 3H, H-2F, H-3F, OCH_2 C_7H_{13}), 4.12 (d, *J* = 3.0 Hz, 1H, H-4^G), 4.20 (d, *J* = 11.4 Hz, 1H, A of AB, CH_2Ph), 4.34 (d, *J* = 11.9 Hz, 1H, B of AB, CH_2Ph), 4.40 (d, *J* = 7.8 Hz, 1H, H-1^G), 4.42 (d, *J* = 12.0 Hz, 1H, A of AB, CH_2Ph), 4.49 (t, *J* = 6.4 Hz, 1H, H-2^L), 4.52 – 4.75 (m, 8H, H-5F, CH_2Ph), 4.84 (m, 1H, H-1F), 5.05 (d, *J* = 10.9 Hz, 1H, A of AB, CH_2Ph), 5.08 (d, *J* = 11.0 Hz, 1H, A of AB, CH_2Ph), 7.15 – 7.43 (m, 30H, 6 x C_6H_5);

^{13}C NMR (126 MHz, $CDCl_3$) δ 16.59 (C-6^F), 23.43, 25.56, 25.57, 26.22, 26.24, 26.29, 26.35, 29.68, 33.52, 33.61, 34.19, 37.03, 41.22 (17C, $C_7H_{13}^L$, C-3^{cyc}, C-4^{cyc}, C-5^{cyc}, C-6^{cyc}, 6C C_7H_{13}), 65.94 (C-5^F), 68.65 (C-6^G), 69.84 (OCH_2 C_7H_{13}), 72.54, 72.88, 72.97, 73.45, 74.61, 75.01, 75.18 (C-5^G, CH_2Ph), 75.71, 75.90, 76.10 (C-4^G, C-2^F, C-2^{cyc}), 77.66 (C-1^{cyc}), 78.76 (C-4^F), 79.12 (C-2^L), 79.83 (C-3^F), 80.72 (C-2^G), 81.44 (C-3^G), 94.74 (H-1^F), 101.13 (H-1^G), 127.14, 127.22, 127.25, 127.31, 127.34, 127.36, 127.58, 127.59, 127.84, 127.87, 128.01, 128.14, 128.18, 128.25, 128.36, 128.44, 138.09, 138.80, 138.98, 139.27, 139.36, 139.60 (36C, 6 x C_6H_5), 174.15 (C=O);

ESI-MS *m/z* calcd for $C_{76}H_{94}NaO_{13}$ $[M+Na]^+$: 1237.66; found: 1237.89.

2k) A mixture of **17k** (14.0 mg, 0.012 mmol) and $Pd(OH)_2$ (12.9 mg) in dioxane/water (1.0 mL, 4:1) was stirred under a hydrogen atmosphere. After 1 h the mixture was filtered through a pad of celite and

evaporated to dryness. Column chromatography on silica gel (DCM/Methanol 6:1) gave **2k** (7.0 mg, 92%). $[\alpha]_D^{20}$ -93.80 (*c* 0.70, MeOH);

^{1}H NMR (500 MHz, Methanol-d_4) δ 0.83 – 1.03, 1.21 – 1.43, 1.51 – 1.60, 1.61 – 1.76, 1.76 – 1.85, 1.86 – 2.00, 2.00 – 2.13 (m, 27H, $C_7H_{13}^L$, H-3^{cyc}, H-4^{cyc}, H-5^{cyc}, H-6^{cyc}, 3 x CH_2 C_5H_9), 1.19 (d, *J* = 6.6 Hz, 3H, H-6^F), 2.66 (hept, *J* = 7.4 Hz, 1H, CH C_5H_9), 3.25 (dd, *J* = 9.5, 3.3 Hz, 1H, H-3^G), 3.40 (ddd, *J* = 6.6, 5.1, 1.1 Hz, 1H, H-5^G), 3.54 (ddd, *J* = 9.8, 8.1, 4.4 Hz, 1H, H-2^{cyc}), 3.59 (dd, *J* = 9.5, 7.7 Hz, 1H, H-2^G), 3.63 – 3.78 (m, 5H, H-2^F, H-4^F, H-6^G, H-1^{cyc}), 3.86 (dd, *J* = 10.1, 3.4 Hz, 1H, H-3^F), 3.91 (dd, *J* = 3.3, 1.1 Hz, 1H, H-4^G), 4.08 (dd, *J* = 10.9, 6.6 Hz, 1H, OCH_2 C_5H_9), 4.17 (dd, *J* = 10.9, 6.8 Hz, 1H, OCH_2 C_5H_9), 4.27 (d, *J* = 7.7 Hz, 1H, H-1^G), 4.43 (dd, *J* = 8.7, 4.4 Hz, 1H, H-2^L), 4.60 (q, *J* = 6.7 Hz, 1H, H-5^F), 4.87 (m, 1H, H-1^F);

^{13}C NMR (126 MHz, MeOD) δ 16.59 (C-6^F), 19.16, 24.45, 25.62, 25.68, 27.21, 27.42, 27.65, 30.07, 30.88, 33.68, 34.74, 35.02, 35.47, 42.49 (15C, $C_7H_{13}^L$, C-3^{cyc}, C-4^{cyc}, C-5^{cyc}, C-6^{cyc}, 4C C_5H_9), 62.72 (C-6^G), 67.42 (C-5^F), 68.65 (C-4^G), 69.89 (OCH_2 C_5H_9), 70.02 (C-2^F), 71.55 (C-3^F), 72.55 (C-2^G), 73.81 (C-4^F), 76.06 (C-5^G), 77.47 (C-2^{cyc}), 78.71 (C-2^L), 79.24 (C-1^{cyc}), 83.88 (C-3^G), 97.36 (C-1^F), 102.63 (C-1^G), 176.87 (C=O);

ESI-MS *m/z* calcd for $C_{32}H_{54}NaO_{13}$ $[M+Na]^+$: 669.35; found: 669.41.

2l) A mixture of **17l** (9.0 mg, 0.008 mmol) and $Pd(OH)_2$ (9.0 mg) in dioxane/water (1.0 mL, 4:1) was stirred under a hydrogen atmosphere. After 2 h the mixture was filtered through a pad of celite and evaporated to dryness. Column chromatography on silica gel (DCM/Methanol 6:1) gave **2l** (3.2 mg, 64%). $[\alpha]_D^{20}$ -126.74 (*c* 0.32, MeOH);

^{1}H NMR (500 MHz, Methanol-d_4) δ 0.82 – 1.04, 1.20 – 1.43, 1.49 – 1.84, 1.86 – 1.96, 1.99 – 2.11 (m, 29H, $C_7H_{13}^L$, H-3^{cyc}, H-4^{cyc}, H-5^{cyc}, H-6^{cyc}, 4 x CH_2 C_6H_{11})), 1.19 (d, *J* = 6.6 Hz, 3H, H-6^F), 2.24 (hept, *J* = 7.5 Hz, 1H, CH C_6H_{11}), 3.25 (dd, *J* = 9.5, 3.3 Hz, 1H, H-3^G), 3.41 (ddd, *J* = 6.6, 5.1, 1.1 Hz, 1H, H-5^G), 3.54 (ddd, *J* = 9.8, 8.1, 4.4 Hz, 1H, H-2^{cyc}), 3.59 (dd, *J* = 9.5, 7.7 Hz, 1H, H-2^G), 3.63 – 3.78 (m, 5H, H-6^G, H-2^F, H-4^F, H-1^{cyc}), 3.86 (dd, *J* = 10.1, 3.3 Hz, 1H, H-3^F), 3.91 (dd, *J* = 3.3, 1.1 Hz, 1H, H-4^G), 4.03 (dd, *J* = 10.5, 7.1 Hz, 1H, OCH_2 C_6H_{11}), 4.08 (dd, *J* = 10.7, 7.1 Hz, 1H, OCH_2 C_6H_{11}), 4.27 (d, *J* = 7.7 Hz, 1H, H-1^G), 4.43 (dd, *J* = 8.7, 4.3 Hz, 1H, H-2^L), 4.60 (q, *J* = 8.5, 7.5 Hz, 1H, H-5^F), 4.85 (m, 1H, H-1^F);

^{13}C NMR (126 MHz, MeOD) δ 16.59 (C-6^F), 24.43, 26.32, 27.22, 27.43, 27.66, 30.08, 30.29, 30.37, 33.68, 34.74, 35.04, 39.89, 42.46 (16C, $C_7H_{13}^L$, C-3^{cyc}, C-4^{cyc}, C-5^{cyc}, C-6^{cyc}, 5C C_6H_{11}), 62.72 (C-6^G), 67.42 (C-5^F), 68.65 (C-4^G), 70.02, 70.06 (C-2^F, OCH_2 C_6H_{11}), 71.55 (C-3^F), 72.55 (C-2^G), 73.81 (C-4^F), 76.06 (C-5^G), 77.48 (C-2^{cyc}), 78.69 (C-2^L), 79.25 (C-1^{cyc}), 83.90 (C-3^G), 97.37 (C-1^F), 102.63 (C-1^G), 176.85 (C=O);

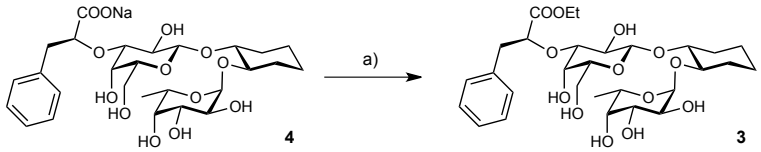
ESI-MS *m/z* calcd for $C_{33}H_{56}NaO_{13}$ $[M+Na]^+$: 683.36; found: 683.46.

2m) A mixture of **17m** (12.0 mg, 0.010 mmol) and $Pd(OH)_2$ (10.0 mg) in dioxane/water (1.0 mL, 4:1) was stirred under a hydrogen atmosphere. After 2 h the mixture was filtered through a pad of celite and evaporated to dryness. Column chromatography on silica gel (DCM/Methanol 6:1) gave **2m** (6.2 mg, 92%). $[\alpha]_D^{20}$ -81.79 (*c* 0.36, MeOH);

^{1}H NMR (500 MHz, Methanol-d_4) δ 0.82 – 1.09, 1.20 – 1.43, 1.52 – 1.61, 1.60 – 1.80, 1.88 – 1.95, 1.99 – 2.10 (32H, $C_7H_{13}^L$, H-3^{cyc}, H-4^{cyc}, H-5^{cyc}, H-6^{cyc}, 5 x CH_2, CH C_7H_{13}), 1.19 (d, *J* = 6.7 Hz, 3H, H-6^F), 3.25 (dd, *J* = 9.5, 3.3 Hz, 1H, H-3^G), 3.40 (ddd, *J* = 6.6, 5.2, 1.1 Hz, 1H, H-5^G), 3.54 (ddd, *J* = 9.8, 8.1, 4.4 Hz, 1H, H-2^{cyc}), 3.59 (dd, *J* = 9.5, 7.7 Hz, 1H, H-2^G), 3.63 – 3.77 (m, 5H, H-6^G, H-2^F, H-4^F, H-1^{cyc}), 3.86 (dd, *J* = 10.1, 3.4 Hz, 1H, H-3^F), 3.91 (dd, *J* = 3.3, 1.1 Hz, 1H, H-4^G), 3.94 (dd, *J* = 10.7, 6.4 Hz, 1H, OCH_2 C_7H_{13}), 3.99 (dd, *J* = 10.7, 6.5 Hz, 1H, OCH_2 C_7H_{13}), 4.27 (d, *J* = 7.7 Hz, 1H, H-1^G), 4.43 (dd, *J* = 8.6, 4.4 Hz, 1H, H-2^L), 4.60 (m, 1H, H-5^F), 4.85 (m, 1H, H-1^F);

^{13}C NMR (126 MHz, MeOD) δ 16.59 (C-6^F), 24.43, 24.45, 26.77, 26.78, 27.23, 27.43, 27.65, 30.10, 30.74, 30.76, 30.89, 33.70, 34.74, 35.03, 38.53, 42.47 (17C, $C_7H_{13}^L$, C-3^{cyc}, C-4^{cyc}, C-5^{cyc}, C-6^{cyc}, 6C C_7H_{13}), 62.73 (C-6^G), 67.42 (C-5^F), 68.65 (C-4^G), 70.02 (C-2^F), 71.25 (OCH_2 C_7H_{13}), 71.55 (C-3^F), 72.54 (C-2^G), 73.81 (C-4^F), 76.06 (C-5^G), 77.47 (C-2^{cyc}), 78.67 (C-2^L), 79.24 (C-1^{cyc}), 83.88 (C-3^G), 97.36 (C-1^F), 102.62 (C-1^G), 176.81 (C=O);

ESI-MS *m/z* calcd for $C_{34}H_{58}NaO_{13}$ $[M+Na]^+$: 697.38; found: 697.55.



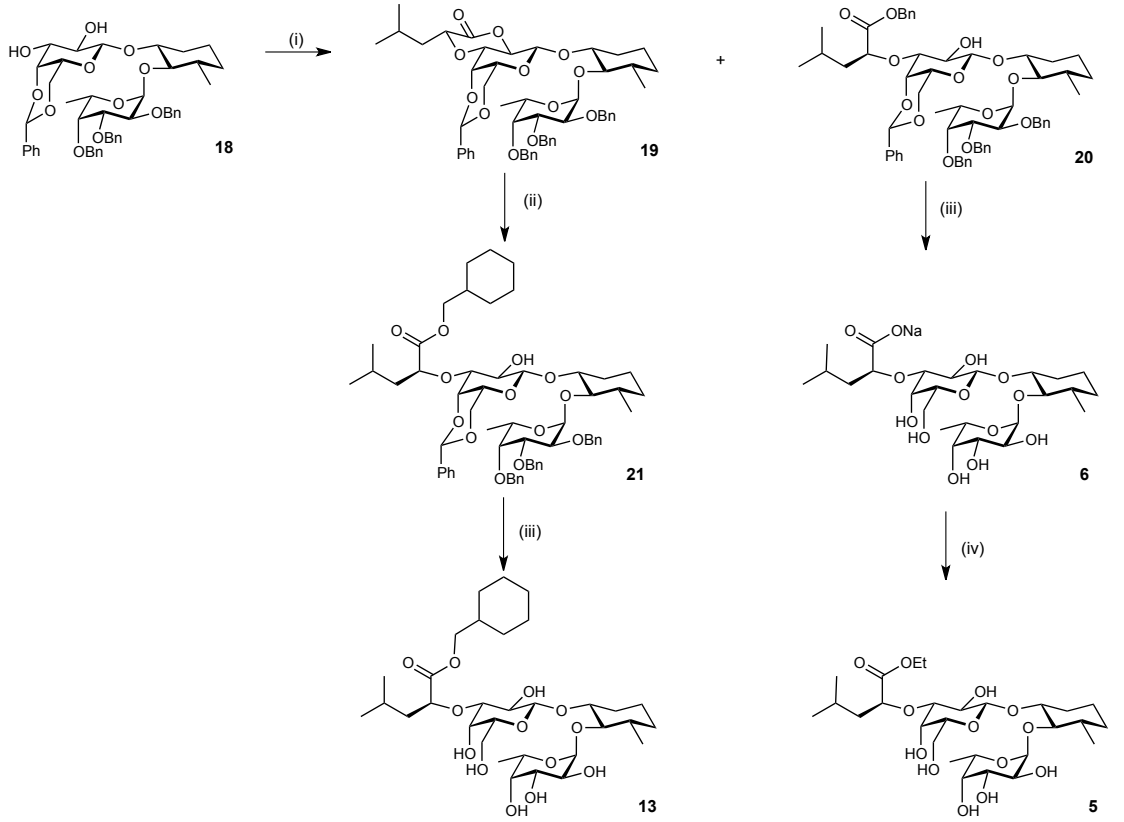
Scheme S3 Reactions and reaction conditions: a) $KHCO_3$, DMF, 4°C – 15°C, 4 h - 6 d, 15 – 80%. The synthesis of **4** was described by Kolb *et al.* (3).

3) According to general procedure (described for compound **2a**), **4** (14.1 mg, 0.0237 mmol), $KHCO_3$ (4.7 mg, 0.0474 mmol) and iodoethane (2.85 µL, 0.0356 mmol), 4°C, 1 d, gave **3** (10.9mg mg, 77%). $[\alpha]_D^{20}$ -60.70 (*c* 1.09, MeOH);

^{1}H NMR (500 MHz, Methanol-d_4) δ 0.84 – 0.95, 1.20 – 1.43, 1.67 – 1.76, 2.00 – 2.08 (m, 8H, H-3^{cyc}, H-4^{cyc}, H-5^{cyc}, H-6^{cyc}), 1.15 – 1.19 (m, 6H, H-6^F, CH_3), 3.12 (d, *J* = 6.3 Hz, 2H, CH_2Ph), 3.34 – 3.37 (m, 1H, H-3^G), 3.40 (dd, *J* = 6.8, 5.4 Hz, 1H, H-5^G), 3.49 – 3.60 (m, 2H, H-2^G, H-2^{cyc}), 3.62 – 3.77 (m, 5H, H-6^G, H-2^F, H-4^F, H-1^{cyc}), 3.85 (dd, *J* = 10.1, 3.4 Hz, 1H, H-3^F), 3.90 (d, *J* = 3.2 Hz, 1H, H-4^G), 4.11 (q, *J* = 7.1 Hz, 2H, CH_2), 4.26 (d, *J* = 7.6 Hz, 1H, H-1^G), 4.53 – 4.64 (m, 2H, H-5^F, H-2^L), 4.82 – 4.87 (m, 1H, H-1^F), 7.15 – 7.32 (m, 5H, C_6H_5);

^{13}C NMR (126 MHz, MeOD) δ 14.36 (CH_3), 16.60 (C-6^F), 24.44, 30.06, 30.92, 40.34 (C-3^{cyc}, C-4^{cyc}, C-5^{cyc}, C-6^{cyc}), 62.25 (OCH_3), 62.68 (C-6^G), 67.41 (C-5^F), 68.48 (C-4^G), 70.03 (C-2^F), 71.56 (C-3^F), 72.36 (C-2^G), 73.82 (C-4^F), 76.05 (C-5^G), 77.40 (C-2^{cyc}), 79.28 (C-^{1cyc}), 81.33 (C-2^L), 83.97 (C-3^G), 97.32 (C-1^F), 102.54 (C-1^G), 127.65, 129.24, 130.76, 138.17 (6C, C_6H_5), 175.06 (C=O);

ESI-MS *m/z* calcd for $C_{29}H_{44}NaO_{13}$ $[M+Na]^+$: 623.27; found: 623.37.



Scheme S4 Reagents and reaction conditions: (i) i. nBu2SnO, MeOH, reflux, 3 h, ii. CsF, DME, rt, 18 h, 25%; (ii) DMAP/cyclohexylmethanol, 82°C, overnight (79%); (iii) Pd(OH)2/C, THF, H2 balloon, JXH-VIII-018 (91%), JXH-VIII-005 (87%); (iv)) KHCO3/EtI/DMF, 0-4°C, 24 h, (54%). The synthesis of **18** was described by Norman *et al.* (2).

19 and **20**) A suspension of **18** (95.6 mg, 0.12 mmol) and Bu_2SnO (34.35 mg, 0.138 mmol) in dry MeOH (5 mL) was refluxed for 3 h. The solvent was removed under reduced pressure and the resulting foam

dried *in vacuo* 4 h. CsF (42 mg, 0.24 mmol) was dried *in vacuo* at 100°C for 30 min and flushed with argon. Then benzyl (*R*)-4-Methyl-2-(((trifluoromethyl)sulfonyl)oxy)pentanoate (85 mg, 0.24 mmol), CsF (36 mg) and DME (3.0 mL) were added to the tin acetal at rt. The reaction mixture was stirred at rt overnight. The solvent was removed under reduced pressure. The residue was purified by silica gel chromatography (petroleum ether/ethyl acetate, 6:4 - 3:1) to afford **20** (85 mg, 70%) as white foam and **19** (20 mg) as white foam. $[\alpha]_D^{20}$ -108.3 (c 0.77, DCM);

^{1}H NMR (500 MHz, $CDCl_3$) δ 7.57 (d, *J* = 7.2 Hz, 2H), 7.40 – 7.12 (m, 23H), 5.48 (s, 1H, OC*H*Ph), 5.16 (d, *J* = 12.2 Hz, 1H, OC*H*$_2$Ph), 5.08 (d, *J* = 12.2 Hz, 1H, OC*H*$_2$Ph), 4.97 (d, *J* = 3.0 Hz, 1H, Fuc-H1), 4.85 (d, *J* = 6.6 Hz, 1H, Fuc-H5), 4.82 (d, *J* = 11.7 Hz, 1H, OC*H*$_2$Ph), 4.71 (d, *J* = 11.7 Hz, 1H, OC*H*$_2$Ph), 4.66 – 4.55 (m, 3H, Lactic-H2, OC*H*$_2$Ph), 4.32 – 4.24 (m, 2H, Gal-H6a, OC*H*$_2$Ph, Gal-H1), 4.20 (d, *J* = 3.3 Hz, 1H, Gal-H4), 4.01 – 3.88 (m, 4H, Gal-H6b, Fuc-H2, Fuc-H3, Gal-H2), 3.65 – 3.60 (m, 1H, MeCy-H1), 3.61 (d, *J* = 11.2 Hz, 1H, OC*H*$_2$Ph), 3.47 (dd, *J* = 9.7, 3.4 Hz, 1H, Gal-H3), 3.25 (s, 1H, Gal-H5), 3.22 (s, 1H, Fuc-H4), 3.19 (t, *J* = 9.5 Hz, 1H, MeCy-H2), 2.38 (d, *J* = 1.5 Hz, 1H, 2-OH), 2.10 – 2.03 (m, 1H), 1.96 (dt, *J* = 13.5, 6.7 Hz, 1H), 1.80 (ddd, *J* = 14.5, 9.6, 5.2 Hz, 1H), 1.69 – 1.53 (m, 4H), 1.38 – 1.16 (m, 2H), 1.09 (d, *J* = 6.4 Hz, 3H, MeCy), 1.04 (d, *J* = 6.5 Hz, 3H, Fuc-H6), 1.02 – 0.98 (m, 1H), 1.00 (d, *J* = 1.8 Hz, 3H, CH_3), 0.99 (d, *J* = 1.9 Hz3H, CH_3);

^{13}C NMR (126 MHz, $CDCl_3$) δ 173.71 (C=O), 139.62, 139.43, 138.61, 138.32, 135.63, 128.59, 128.54, 128.51, 128.35, 128.33, 128.11, 128.03, 127.92, 127.82, 127.42, 127.34, 127.03, 126.81, 125.90, 101.14 (Gal-C1), 99.52 (*C*HPh), 98.49 (Fuc-C1), 82.26 (MeCy-C2), 79.89 (MeCy-C1), 79.56 (Fuc-C2), 79.02 (Gal-C3), 78.61 (Fuc-C4), 78.02 (Lac-C2), 75.45 (Fuc-C3), 75.05 (Gal-C4), 74.72 (O*C*H$_2$Ph), 74.42 (O*C*H$_2$Ph), 71.25 (O*C*H$_2$Ph), 71.08 (Gal-C2), 69.27 (Gal-C6), 66.56 (Gal-C5), 66.46 (O*C*H$_2$Ph), 66.09 (Fuc-C5), 41.92, 39.44, 33.64, 31.41, 24.65, 23.28, 23.17, 21.80 (8C), 18.74 (*Me*Cy), 16.46 (Fuc-C6);

MS (ESI) *m*/*z*: calcd for $C_{60}H_{72}NaO_{13}$ $[M+Na]^+$: 1023.49; found: 1023.56.

21) A mixture of **19** (19.0 mg, 0.019 mmol) and DMAP (catalytic amount) in Cyclohexylmethanol (0.5 mL) was stirred at 82°C overnight. The solvent was removed under high vacuum with heating and the residue was purified by silica gel chromatography (petroleum ether/ethyl acetate, 6:4 - 3:1) to afford **21** (16.7 mg, 79%) as white foam. $[\alpha]_D^{20}$ -96.7 (c 0.7, DCM);

^{1}H NMR (500 MHz, $CDCl_3$) δ 7.58 (d, *J* = 7.3 Hz, 2H), 7.39 – 7.11 (m, 18H), 5.62 (s, 1H, OC*H*Ph), 4.97 (d, *J* = 3.4 Hz, 1H, Fuc-H1), 4.87 (q, *J* = 6.4 Hz, 1H, Fuc-H5), 4.82 (d, *J* = 11.6 Hz, 1H, OC*H*$_2$Ph), 4.71 (d, *J* = 11.7 Hz, 1H, OC*H*$_2$Ph), 4.65 – 4.58 (m, 2H, OC*H*$_2$Ph), 4.57 (dd, *J* = 9.5, 4.0 Hz, 1H, Lactic-H2), 4.38 (d, *J* = 3.3 Hz, 1H, Gal-H4), 4.32 – 4.29 (m, 1H, Gal-H6a), 4.28 (d, *J* = 7.7 Hz, 1H, Gal-H1), 4.27 (d, *J* = 11.3 Hz, 1H, OC*H*$_2$Ph), 4.08 (dd, *J* = 11.9, 1.1 Hz, 1H, Gal-H6b), 3.98 – 3.92 (m, 3H, Fuc-H3, Fuc-H2, Gal-H2), 3.91 – 3.82 (m, 2H, OCH$_2$Cyclohexyl), 3.66 – 3.61 (m, 1H, MeCy-H1), 3.61 (d, *J* = 11.3 Hz, 1H, OC*H*$_2$Ph), 3.51 (dd, *J* = 9.7, 3.4 Hz, 1H, Gal-H3), 3.36 (s, 1H, Gal-H5), 3.23 (s, 1H, Fuc-H4), 3.19 (t, *J* = 9.5 Hz, 1H, MeCy-H2), 2.41 (d, *J* = 1.3 Hz, 1H, 2-OH), 2.06 (m, 1H), 2.02 – 1.91 (m, 1H), 1.81 – 1.56 (m, 11H), 1.39 – 1.13 (m, 6H), 1.09 (d, *J* = 6.4 Hz, 3H, Fuc-H6), 1.05 (d, *J* = 6.5 Hz, 3H, MeCy), 1.01 (d, *J* = 2.5 Hz, 3H), 1.00 (d, *J* = 2.7 Hz, 3H), 1.11 – 0.85 (m, 2H);

^{13}C NMR (126 MHz, MeOD) δ ^{13}C NMR (126 MHz, $CDCl_3$) δ 174.14 (C=O), 139.65, 139.45, 138.64, 138.35, 128.62, 128.55, 128.13, 128.05, 127.93, 127.84, 127.44, 127.36, 127.04, 126.84, 125.95 (Ar-C), 101.18 (Gal-C1), 99.62 (*C*HPh), 98.53 (Fuc-C1), 82.29 (MeCy-C2), 79.92 (MeCy-C1), 79.59 (Fuc-C3), 78.99 (Gal-C3), 78.64 (Fuc-C4), 77.97 (Lac-C2), 75.47 (Fuc-C2), 75.17 (Gal-C4), 74.75 (O*C*H$_2$Ph), 74.45 (O*C*H$_2$Ph), 71.26 (O*C*H$_2$Ph), 71.12 (Gal-C2), 69.96 (CO$_2$*C*H$_2$Cy), 69.37 (Gal-C6), 66.65 (Gal-C5), 66.12 (Fuc-C5), 42.07, 39.46, 36.99, 33.66, 31.43, 29.64, 29.59, 26.29, 25.58, 24.72, 23.30, 23.22, 21.87 (14C), 18.77 (*Me*Cy), 16.50 (Fuc-C6);

MS (ESI) *m*/*z*: calcd for $C_{60}H_{78}NaO_{13}$ $[M+Na]^+$: 1029.53; found: 1029.77.

13) A suspension of **21** (16.0 mg, 0.0145 mmol) and $Pd(OH)_2$/C (5.0 mg, 10% Pd) in THF (4.0 mL) was hydrogenated with H_2 balloon at rt for 2hr. The reaction mixture was filtered through celite, the solvent was removed under reduced pressure. The residue was purified by silica gel chromatography (DCM/MeOH, 19:1 – 9:1) to afford JXH-VIII-018 as white solid. The solid was redissolved in H_2O/CH_3CN

(4:1), filtered through Acrodisc® syringe filter with 0.45 µm PTFE member and the filtrate was lyophilized to provide **13** (8.6 mg, 91%) as white solid.

^{1}H NMR (500 MHz, MeOD) δ 4.99 (d, J = 4.0 Hz, 1H, Fuc-H1), 4.89 (q, J = 6.3 Hz, 1H, Fuc-H5), 4.42 (dd, J = 9.4, 4.0 Hz, 1H, Lactic-H2), 4.25 (d, J = 7.8 Hz, 1H, Gal-H1), 4.03 – 3.93 (m, 2H, OCO2CH2Cy), 3.91 (d, J = 2.6 Hz, 1H, Gal-H4), 3.83 (dd, J = 10.2, 3.3 Hz, 1H, Fuc-H3), 3.77 – 3.70 (m, 2H, Fuc-H2, Gal-H6a), 3.69 (d, J = 2.6 Hz, 1H, Fuc-H4), 3.67 – 3.61 (m, 3H, Gal-H6b, Gal-H2, MeCy-H1), 3.39 (t, J = 6.3 Hz, 1H, Gal-H5), 3.25 (dd, J = 9.5, 3.2 Hz, 1H, Gal-H3), 3.19 (t, J = 9.3 Hz, 1H, MeCy-H2), 2.11 (m, 1H), 2.04 – 1.93 (m, 1H), 1.81 – 1.58 (m, 10H), 1.56 – 1.47 (m, 1H), 1.35 – 1.21 (m, 6H), 1.18 (d, J = 6.6 Hz, 3H, Fuc-H6), 1.13 (d, J = 6.3 Hz, 3H, MeCy), 1.09 – 1.00 (m, 2H), 0.98 (d, J = 6.6 Hz, 3H), 0.95 (d, J = 6.7 Hz, 3H).

^{13}C NMR (126 MHz, MeOD) δ 176.69 (C=O), 102.43 (Gal-C1), 100.48 (Fuc-C1), 84.68 (MeCy-C2), 83.80 (Gal-C3), 79.96 (MeCy-C1), 79.21 (Lac-C2), 75.95 (Gal-C5), 73.83 (Fuc-C4), 72.20 (Gal-C2), 71.38 (Fuc-C3), 71.26 (CO_2CH_2Cy), 70.34 (Fuc-C2), 68.67 (Gal-C4), 67.55 (Fuc-C5), 62.84 (Gal-C6), 43.91, 40.42, 38.52, 34.95, 31.92, 30.75, 30.73, 27.43, 26.77, 25.42, 24.23, 23.67, 22.25, 19.57, 16.73. MS (ESI) *m/z*: calcd for $C_{32}H_{56}NaO_{13}$ $[M+Na]^+$: 671.36; found: 671.45.

6) A suspension of **20** (68 mg, 0.0679 mmol) and $Pd(OH)_2$/C (17 mg, 10% Pd) in THF (5.0 mL) was hydrogenated with H_2 balloon at rt for 2.5h. The reaction mixture was filtered through celite, the solvent was removed under reduced pressure. The residue was purified by silica gel chromatography (DCM/(MeOH:H_2O, 10:1), 7:3 - 6:4), ion-exchanged with Dowex® Marathon™ C sodium form, then purified with reversed-phase column chromatography (RP-**18**, MeOH/H_2O) to afford **6** (34 mg, 87%) as white solid.

^{1}H NMR (500 MHz, D_2O) δ 5.11 (d, J = 4.0 Hz, 1H, Fuc-H1), 4.87 (q, J = 6.6 Hz, 1H, Fuc-H5), 4.50 (d, J = 8.0 Hz, 1H, Gal-H1), 3.96 (dd, J = 9.8, 3.7 Hz, 1H, Lactic-H2), 3.95 – 3.87 (m, 2H, Gal-H4, Fuc-H3), 3.87 – 3.79 (m, 2H, Fuc-H4, Fuc-H2), 3.79 – 3.70 (m, 3H, Gal-H6a, Gal-H6b, MeCy-H1), 3.65 – 3.61 (m, 2H, Gal-H2, Gal-H5), 3.43 (dd, J = 9.6, 3.1 Hz, 1H, Gal-H3), 3.24 (t, J = 9.6 Hz, 1H, MeCy-H2), 2.15 (m, 1H), 1.88 – 1.84 (m, 1H), 1.75 – 1.56 (m, 4H), 1.55 – 1.43 (m, 1H), 1.37 – 1.25 (m, 2H), 1.21 (d, J = 6.6 Hz, 3H), 1.10 (t, J = 6.4 Hz, 3H), 1.15 – 1.06 (m, 1H), 0.95 (d, J = 2.8 Hz, 3H), 0.94 (d, J = 2.9 Hz, 3H);

^{13}C NMR (126 MHz, D_2O) δ 182.62 (C=O), 99.77 (Gal-C1), 98.92 (Fuc-C1), 84.24 (MeCy-C2), 82.75 (Gal-C3), 79.75 (Lac-C2), 78.68 (MeCy-C1), 74.17 (Gal-C5), 71.99 (Fuc-C4), 69.82 (Gal-C2), 69.27 (Fuc-C3), 68.20 (Fuc-C2), 66.50 (Fuc-C5), 66.34 (Gal-C4), 61.73 (Gal-C6), 42.74, 38.75, 33.17, 30.31, 23.87, 22.67, 22.58, 20.92 (8xC), 18.12 (MeCy), 15.55 (Fuc-C6);

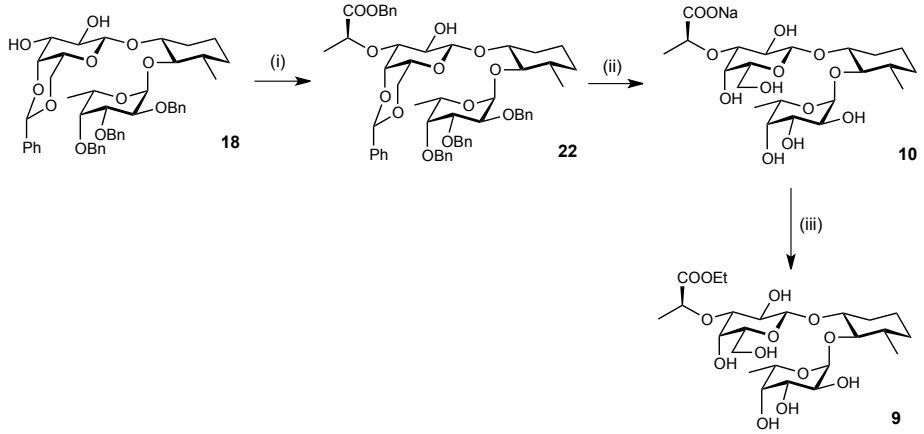
MS (ESI) m/z: calcd for $C_{25}H_{44}NaO_{13}$ [M+H]+: 575.27; found: 575.28

5) Compound **6** (9.6 mg, 0.016 mmol) in dry DMF (100 µL) was added $KHCO_3$ (3.36 mg, 0.032 mmol) and EtI (2.68 µL) at 0°C. The reaction mixture was stirred at 4 °C for 24 hrs. The solvent was removed under high vacuum and the residue was purified by silica gel chromatography (DCM/MeOH, 9:1 - 6:1), the desired product was dissolved in acetonitrile/H_2O and then filtered through 0.45 um filter, the filtrate was lyophilized to afford **5** (5.2 mg, 54%) as white foam.

^{1}H NMR (500 MHz, MeOD) δ 4.99 (d, J = 4.0 Hz, 1H, Fuc-H1), 4.88 (q, J = 6.5 Hz, 1H, Fuc-H5), 4.39 (dd, J = 9.6, 3.9 Hz, 1H, Lactic-H2), 4.25 (d, J = 7.8 Hz, 1H, Gal-H1), 4.24 – 4.14 (m, 2H, $CO_2CH_2CH_3$), 3.92 (d, J = 2.6 Hz, 1H, Gal-H4), 3.83 (dd, J = 10.3, 3.3 Hz, 1H, Fuc-H3), 3.75 – 3.61 (m, 6H, Fuc-H2, Gal-H6a, Fuc-H4, Gal-H6b, Gal-H2, MeCy-H1), 3.39 (t, J = 6.0 Hz, 1H, Gal-H5), 3.25 (dd, J = 9.5, 3.2 Hz, 1H, Gal-H3), 3.19 (t, J = 9.3 Hz, 1H, MeCy-H2), 2.12 – 2.09 (m, 1H, CHCH3CH3), 2.04 – 1.95 (m, 1H), 1.74 – 1.58 (m, 4H), 1.54 – 1.46 (m, 1H), 1.34 – 1.21 (m, 3H), 1.27 (t, J = 7.1 Hz, 3H, OCH_2CH_3), 1.18 (d, J = 6.6 Hz, 3H, Fuc-H6), 1.13 (d, J = 6.3 Hz, 3H, MeCy), 1.38 – 1.23 (m, 1H), 0.97 (d, J = 6.6 Hz, 3H, $CHCH_3CH_3$), 0.95 (d, J = 6.7 Hz, 3H, CHCH3CH3);

^{13}C NMR (126 MHz, MeOD) δ 176.64 (C=O), 102.43 (Gal-C1), 100.48 (Fuc-C1), 84.67 (MeCy-C2), 83.86 (Gal-C3), 79.96 (MeCy-C1), 79.20 (Lac-C2), 75.92 (Gal-C5), 73.83 (Fuc-C4), 72.17 (Gal-C2), 71.38 (Fuc-C3), 70.34 (Fuc-C2), 68.65 (Gal-C4), 67.55 (Fuc-C5), 62.82 (Gal-C6), 62.24 (OCH_2CH_3), 43.79 ($CH_3CH_3CHCH_2$), 40.42 (Cy-C3), 34.94 (Cy-C4), 31.92 (Cy-C6), 25.37 (CH3CH3CHCH2), 24.22 (Cy-C5), 23.68, 22.18 ($CH_3CH_3CHCH_2$), 19.57 (CH_3Cy), 16.72 (Fuc-C6), 14.47 (OCH_2CH_3);

MS (ESI) m/z: calcd for $C_{27}H_{48}NaO_{13}$ [M+Na]+: 603.30; found: 603.32.



Scheme S5 Reactions and reaction conditions: (i) i. Bu_2SnO, MeOH, reflux, 3 h, ii. CsF, DME, r.t., 1 h, (90%); (ii) $Pd(OH)_2$/MeOH, H_2 balloon, rt, (87%); (iii) $KHCO_3$/EtI/DMF, 0°C, 24 h, (70%). The synthesis of **18** was described by Norman *et al.* (2).

22) A suspension of **18** (258 mg, 0.323 mmol) and Bu_2SnO (96.73 mg, 0.388 mmol) in dry MeOH (5 mL) was refluxed for 3 h. The solvent was removed under reduced pressure and the resulting foam dried *in vacuo* 4 h. CsF (98 mg, 0.647 mmol) was dried *in vacuo* at 100°C for 30 min and flushed with argon. Then benzyl (*R*)-2-(((trifluoromethyl)sulfonyl)oxy)propanoate (152 mg, 0.485 mmol), CsF (98 mg) and DME (4.0 mL) were added to the tin acetal at rt. The reaction mixture was stirred at rt for 1 h. The solvent was removed under reduced pressure. The residue was purified by silica gel chromatography (petroleum ether/ethyl acetate, 6:4 - 1:1) to afford **22** (289 mg, 90%) as white foam.

^{1}H NMR (500 MHz, $CDCl_3$) δ 7.57 (d, *J* = 7.4 Hz, 2H), 7.44 – 7.10 (m, 24H), 5.49 (s, 1H, PhCH), 5.19 (d, *J* = 12.2 Hz, 1H, OCH_2Ph), 5.10 (d, *J* = 12.2 Hz, 1H, OCH_2Ph), 4.97 (d, *J* = 2.8 Hz, 1H, Fuc-H1), 4.85 (d, *J* = 6.5 Hz, 1H, Fuc-H5), 4.82 (d, *J* = 11.7 Hz, 1H, OCH_2Ph), 4.71 (d, *J* = 11.7 Hz, 1H, OCH_2Ph), 4.66 – 4.55 (m, 3H, Lactic-H2, OCH_2Ph), 4.29 (d, *J* = 11.3 Hz, 1H, OCH_2Ph), 4.27 – 4.24 (m, 3H, Gal-H6a, Gal-H1, Gal-H4), 3.97 – 3.90 (m, 4H, Gal-H6b, Fuc-H3, Fuc-H2, Gal-H2), 3.65 (d, *J* = 11.2 Hz, 1H, OCH_2Ph), 3.62 (m, 1H, Cy-H1), 3.46 (dd, *J* = 9.8, 3.3 Hz, 1H, Gal-H3), 3.25 (m, 2H, Fuc-H4, Gal-H5), 3.20 (t, *J* = 9.5 Hz, 1H, Cy-H2), 2.44 (s, 1H, 2-OH), 2.06 (m, 1H, Cy-H6a), 1.71 – 1.56 (m, 3H, Cy-H5a, Cy-H4a, Cy-H3), 1.53 (d, *J* = 6.9 Hz, 3H, Me-Lactic), 1.35 – 1.17 (m, 2H, Cy-H6b, Cy-H5b), 1.09 (d, *J* = 6.4 Hz, 3H, Me-Cy), 1.05 (d, *J* = 6.5 Hz, 3H, Fuc-H6), 1.07 – 0.98 (m, 1H, Cy-H4b);

^{13}C NMR (126 MHz, $CDCl_3$) δ 173.62 (C=O), 139.62, 139.43, 138.64, 138.27, 135.60, 128.60, 128.42, 128.32, 128.14, 128.06, 127.95, 127.86, 127.47, 127.38, 127.06, 126.86, 125.98 (Ar-C), 101.41 (Gal-C1), 99.67 (Ph*C*H), 98.46 (Fuc-C1), 82.20 (Cy-C2), 80.16 (Cy-C1), 79.60 (Fuc-C3, Gal-C3), 78.64 (Fuc-C4), 75.91 (Lactic-C2), 75.53 (Fuc-C2), 75.20 (Gal-C4), 74.77 (O*C*H_2Ph), 74.44 (O*C*H_2Ph), 71.30 (O*C*H_2Ph), 70.88 (Gal-C2), 69.31 (Gal-C6), 66.60 (Gal-C5), 66.59 (O*C*H_2Ph), 66.13 (Fuc-C5), 39.47 (Cy-C3), 33.67 (Cy-C4), 31.47 (Cy-C6), 23.32 (Cy-C5), 18.78, 18.66 (Lactic-Me, Cy-*C*H_3), 16.61 (Fuc-C6);

MS (ESI) *m/z*: calcd for $C_{57}H_{66}NaO_{13}$ $[M+Na]^+$: 981.44; found: 981.61.

10) A suspension of **22** (28.8 mg, 0.03 mmol) and $Pd(OH)_2$/C (10 mg, 10% Pd) in MeOH (2.0 mL) was hydrogenated with H_2 balloon at rt overnight. The reaction mixture was filtered through celite, the solvent was removed under reduced pressure. The residue was purified by silica gel chromatography (DCM/(MeOH:H_2O, 10:1), 7:3 - 6:4), ion-exchanged with Dowex® Marathon™ C sodium form, then purified with reversed-phase column chromatography (RP-18, MeOH/H_2O) to afford 10 (14 mg, 87%) as white solid.

^{1}H NMR (500 MHz, MeOD) δ 5.00 (d, *J* = 4.0 Hz, 1H, Fuc-H1), 4.86 (m, 1H, Fuc-H5), 4.31 (d, *J* = 7.8 Hz, 1H, Gal-H1), 4.11 (q, *J* = 6.8 Hz, 1H, Lactic-H2), 3.93 (d, *J* = 2.6 Hz, 1H, Gal-H4), 3.85 (dd, *J* = 10.3, 3.3 Hz, 1H, Fuc-H3), 3.78 – 3.58 (m, 6H, Gal-H6a, Fuc-H2, Fuc-H4, Gal-H6b, Cy-H1, Gal-H2), 3.43 (t, *J* = 5.8 Hz, 1H, Gal-H5), 3.26 – 3.12 (m, 1H, Gal-H3), 3.21 (t, *J* = 9.3 Hz, 1H, Cy-H2), 2.12 (m, 1H, Cy-H6a), 1.74 – 1.56 (m, 3H, Cy-H5a, Cy-H4a, Cy-H3), 1.43 (d, *J* = 6.9 Hz, 3H Me-Lactic), 1.40 – 1.22 (m, 2H, Cy-H6b, Cy-H5b), 1.19 (d, *J* = 6.6 Hz, 3H, Fuc-H6), 1.13 (d, *J* = 6.3 Hz, 3H, Me-Cy), 1.12 – 1.04 (m, 1H, Cy-H4b);

^{13}C NMR (126 MHz, MeOD) δ 181.17 (C=O), 102.25 (Gal-C1), 100.25 (Fuc-C1), 84.47 (Cy-C2), 83.97 (Gal-C3), 79.93 (Cy-C1), 77.56 (Lactic-C2), 75.82 (Gal-C5), 73.82 (Fuc-C4), 71.38 (Fuc-C3), 71.23 (Gal-C2), 70.32 (Fuc-C2), 67.78 (Gal-C4), 67.52 (Fuc-C5), 63.04 (Gal-C6), 40.34 (Cy-C3), 34.93 (Cy-C4), 31.87 (Cy-C6), 24.20 (Cy-C5), 19.59 (CH$_3$-Lactic, Cy-CH$_3$), 16.76 (Fuc-C6);

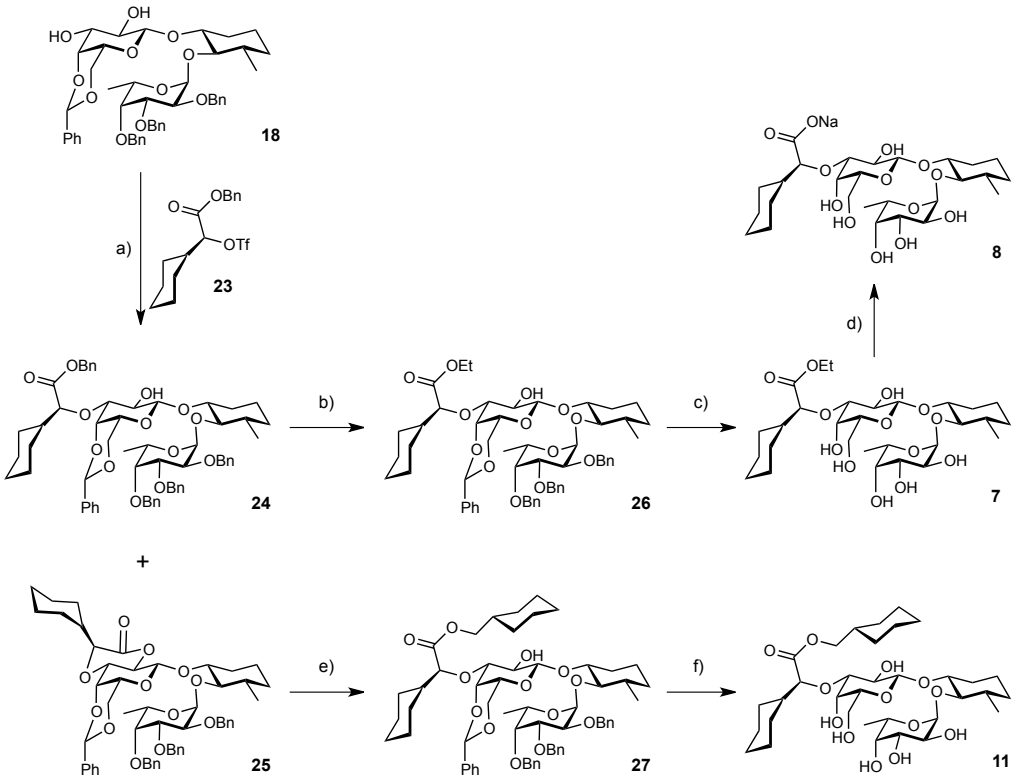
MS (ESI) *m*/*z*: calcd for $C_{22}H_{38}NaO_{13}$ [M+H]$^+$: 533.22; found: 533.22.

9) Compound **10** (12 mg, 0.022 mmol) in dry DMF (225 µL) was added $KHCO_3$ (3.6 mg, 0.044 mmol) and EtI (6.6 µL) at 0°C. The reaction mixture was stirred at 4 °C for 24 hrs. The solvent was removed under high vacuum and the residue was purified by silica gel chromatography (DCM/MeOH, 9:1 - 6:1) to afford **9** (8.5 mg, 70%) as white foam.

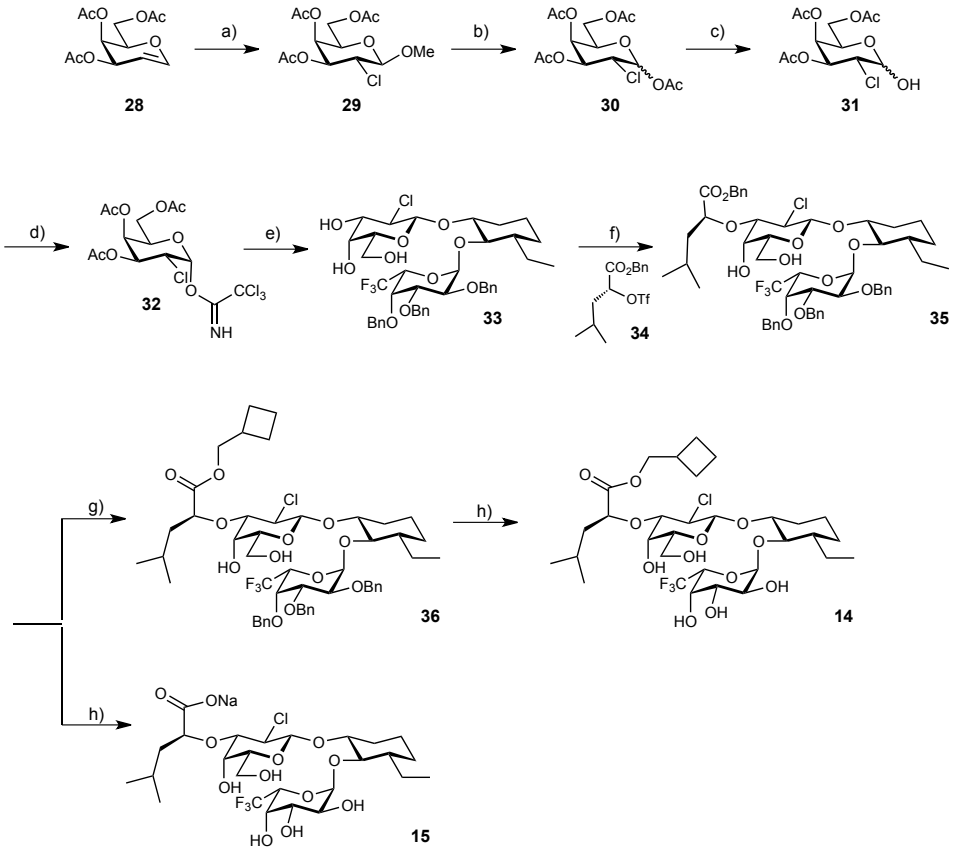
^{1}H NMR (500 MHz, MeOD) δ 4.99 (d, *J* = 4.0 Hz, 1H, Fuc-H1), 4.88 (q, *J* = 6.5 Hz, 1H, Fuc-H5), 4.39 (q, *J* = 6.9 Hz, 1H, Lactic-H2), 4.27 (d, *J* = 7.8 Hz, 1H, Gal-H1), 4.26 – 4.14 (m, 2H, CO_2C*H*$_2$CH$_3$), 3.94 (d, *J* = 2.9 Hz, 1H, Gal-H4), 3.84 (dd, *J* = 10.2, 3.3 Hz, 1H, Fuc-H3), 3.80 – 3.52 (m, 6H, Gal-H6a, Fuc-H2, Fuc-H4, Gal-H6b, MeCy-H1, Gal-H2), 3.40 (t, *J* = 6.0 Hz, 1H, Gal-H5), 3.31 – 3.27 (m, 1H, Gal-H3), 3.19 (t, *J* = 9.3 Hz, 1H, MeCy-H2), 2.16 – 2.06 (m, 1H), 1.70 – 1.57 (m, 3H), 1.43 (d, *J* = 6.9 Hz, 3H, CH_3CH), 1.38 – 1.23 (m, 3H), 1.28 (t, *J* = 7.1 Hz, 3H, OCH$_2C$*H*$_3$), 1.18 (d, *J* = 6.6 Hz, 3H, Fuc-H6), 1.13 (d, *J* = 6.3 Hz, 3H, MeCy), 1.07 – 1.03 (m, 1H);

^{13}C NMR (126 MHz, MeOD) δ 176.34 (C=O), 102.37 (Gal-C1), 100.43 (Fuc-C1), 84.56 (MeCy-C2), 83.95 (Gal-C3), 80.02 (MeCy-C1), 76.65 (Lac-C2), 75.97 (Gal-C5), 73.83 (Fuc-C4), 71.87 (Gal-C2), 71.39 (Fuc-C3), 70.34 (Fuc-C2), 68.44 (Gal-C4), 67.51 (Fuc-C5), 62.82 (Gal-C6), 62.31 (OCH$_2$CH$_3$), 40.41 (Cy-C3), 34.93 (Cy-C4), 31.94 (Cy-C6), 24.22 (Cy-C5), 19.57 (CH$_3$Cy), 19.13 (CHCH$_3$), 16.74 (Fuc-C6), 14.44 (OCH$_2C$H$_3$);

MS (ESI) *m*/*z*: calcd for $C_{24}H_{42}NaO_{13}$ [M+Na]$^+$: 561.25; found: 561.45



Scheme S6 Reactions and reaction conditions: a) i. Bu_2SnO, MeOH, reflux, 4 h; ii. **13,** CsF, DME, rt, 18 h, benzylester **14** = 58%, lactone **15** = 18%; b) NaOEt, MeOH, DCM, 0°C, 1h, 94%; after 2 steps; c) H_2 (1 bar), $Pd(OH)_2/C$, THF, rt., 2 h, 69%; d) NaOH, H_2O, rt, 2h, 85%. e) cyclohexylmethanol, DMAP, 80°C, 18h, 62%; f) H_2 (1 bar), $Pd(OH)_2/C$, THF, rt., 2 h, 78%; The synthesis of **18** was described by Norman *et al.* (2).



Scheme S7 Reagent and conditions: a) i. Cl_2, DCM, 0 °C, 0.5 h; ii. Ag_2CO_3, MeOH, rt, 2.5 h, 80% (2 steps); b) Ac_2O, H_2SO_4, rt, 18 h, 81-90%; c) NH_4OAc, DMF, rt, 18 h, 64% or 7N NH_3 in MeOH, 0 °C 5 h, 60%; d) DBU, Cl_3CCN, DCM, rt, 1.5 h, 81%; e) i. 4Å MS, TMSOTf, DCM, -18 °C to rt, 4 h; ii. CH_3ONa, CH_3OH, rt, 1 h, 34% (2 steps); f) i. nBu_2SnO, MeOH, reflux, 3 h, ii. CsF, DME, rt, 18 h, 25% (2 steps); g) cyclobutylmethanol/DMAP, 85 °C, 48 h, 70%; h) $Pd(OH)_2/C$, THF, H_2 balloon **14** (87%), **15** (92%).

36) Compound **35** (16.0 mg, 0.016 mmol) and DMAP (2.1 mg) was dissolved in cyclobutylmethanol (0.2 mL). The reaction mixture was stirred at 85°C for 48 hrs. The solvent was removed under high vaccum and the residue (11.6 mg) was used for next step without further purification.

14) A suspension of **36** (11.0 mg, 0.011 mmol) and $Pd(OH)_2/C$ (3.2 mg, 10% Pd) in THF (3.0 mL) was hydrogenated with H_2 balloon at rt for 2h. The reaction mixture was filtered through celite, the solvent was removed under reduced pressure. The residue was purified by silica gel chromatography (DCM/MeOH 19:1 - 9:1), the desired product was dissolved in acetonitrile/H_2O and then filtered through 0.45 um filter, the filtrate was lyophilized to provide **14**(6.9 mg, 87%) as white fluffy solid. $[\alpha]_D^{20}$ -32.0 (c 0.18, MeOH);

^{1}H NMR (500 MHz, CD_3OD) δ 5.24 (q, *J* = 6.5 Hz, 1H, Fuc-H5), 5.14 (d, *J* = 3.0 Hz, 1H, Fuc-H1), 4.47 (d, *J* = 8.4 Hz, 1H, Gal-H1), 4.29 (dd, *J* = 9.1, 3.4 Hz, 1H, Lactic-H2), 4.21 – 4.14 (m, 2H, OCH_2^cBu, Fuc-H4), 4.11 (dd, *J* = 10.7, 6.8 Hz, 1H, OCH_2^CBu), 3.90 (m, 2H, Gal-H4, Fuc-H3), 3.86 – 3.79 (m, 2H, Fuc-H2, Gal-H2), 3.76 (dd, *J* = 11.3, 6.9 Hz, 1H, Gal-H6a), 3.68 (dd, *J* = 11.3, 5.0 Hz, 1H, Gal-H6b), 3.66 – 3.57 (m, 1H, Cy-H1), 3.48 (m, 2H, Gal-H5, Gal-H3), 3.39 (t, *J* = 9.3 Hz, 1H, Cy-H2), 2.66 (dt, *J* = 14.8, 7.4 Hz, 1H), 2.06 (m, 4H), 2.00 – 1.86 (m, 3H), 1.78 (m, 5H), 1.53 (m, 2H), 1.42 – 1.17 (m, 3H), 1.06 – 0.95 (m, 7H), 0.91 (t, *J* = 7.3 Hz, 3H); ^{13}C NMR (126 MHz, CD_3OD) δ 176.59 (C=O), 102.39 (Gal-C1), 100.06 (Fuc-C1), 85.52 (Gal-C3), 81.64 (Cy-C1), 80.75 (Cy-C2), 79.50 (Lactic-C2), 76.00 (Gal-C5), 70.19, 69.95, 69.77, 69.25 (Fuc-C2, Fuc-C4, Fuc-C3, Fuc-C5), 68.37 (O OCH_2^cBu), 62.58 (Gal-C6), 61.25 (Gal-C2), 46.45, 44.03, 35.41, 32.01, 29.81, 25.67, 25.62, 25.23, 24.77, 24.01, 23.66, 22.21, 19.16, 10.86 (14xC);

MS (ESI) *m/z*: calcd for $C_{31}H_{50}ClF_3NaO_{12}$ $[M+Na]^+$: 729.28; found: 729.42.

15) A suspension of **35** (2.7 mg, 0.0027 mmol) and $Pd(OH)_2/C$ (2.0 mg, 10% Pd) in THF (2.0 mL) was hydrogenated with H_2 balloon at rt for 3 hrs. The reaction mixture was filtered through celite, the solvent was removed under reduced pressure. To the residue was added a drop of 1N NaOH aqueous solution and then purified with reversed-phase column chromatography (RP-18, MeOH/H_2O) to afford **15** (1.65 mg, 92%) as white solid.

^1H NMR (500 MHz, D_2O) δ 5.29 (q, *J* = 6.4 Hz, 1H, Fuc-H5), 5.27 (d, *J* = 3.3 Hz, 1H, Fuc-H1), 4.69 (d, *J* = 8.6 Hz, 1H, Gal-H1), 4.29 (d, *J* = 1.3 Hz, 1H), 4.00 (dd, *J* = 10.4, 3.4 Hz, 1H), 3.96 – 3.89 (m, 3H), 3.83 (dd, *J* = 10.0, 8.7 Hz, 1H), 3.79 – 3.71 (m, 3H), 3.71 – 3.68 (m, 1H), 3.66 (dd, *J* = 10.1, 3.1 Hz, 1H), 3.45 (t, *J* = 9.6 Hz, 1H), 2.24 – 2.16 (m, 1H), 2.05 – 1.95 (m, 1H), 1.85 (m, 1H), 1.81 – 1.66 (m, 3H), 1.60 – 1.44 (m, 2H), 1.42 – 1.30 (m, 2H), 1.24 (q, *J* = 13.1 Hz, 2H), 1.07 – 1.00 (m, 1H), 0.95 (d, *J* = 6.6 Hz, 3H), 0.94 (d, *J* = 6.7 Hz, 3H), 0.90 (t, *J* = 7.4 Hz, 3H);

MS (ESI) *m/z*: calcd for $C_{26}H_{41}ClF_3NaO_{12}$ $[M-Na]^-$: 637.22; found: 637.44.

Methods

Log$D_{7.4}$ determination (shake-flask method)

Equal amounts of TRIS-HCl buffer (0.1 M, pH 7.4) and 1-octanol were mixed and vigorously shaken for 5 minutes to saturate the phases. The mixture was left until complete separation of the phases occurred and the buffer was retrieved. The stock solutions of test compounds (DMSO, 10 mM) were diluted to 10 µM in saturated buffer (Final DMSO concentration; 0.1 %). 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 to water ratio, respectively. Each ratio was measured in triplicate and simultaneous measurements were conducted with propranolol as control. The plate was sealed with aluminium 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 with Equation 1:

$$logD_{7.4} = log\left(\frac{10\ \mu M - c_b}{c_b} \times \frac{1}{o:b}\right) \quad \text{(eq. 1)}$$

Parallel artificial membrane permeation assay (PAMPA):

Effective permeability (logP_e) was determined in a 96-well format with PAMPA. For each compound, measurements were performed at pH7.4 in quadruplicate. Four wells of a deep well plate were filled with 500 µL PRISMA HT buffer (pH 7.4, pION P/N 110 151). Then, analyte dissolved in DMSO (10 mM) was added to the buffer to yield 50 µM solutions (Containing 0.5 % DMSO). To exclude precipitation, the optical density (OD) was measured at 650 nm, and solutions exceeding OD 0.01 were filtrated. Afterwards, 200 µL was transferred to each well of the donor plate of the PAMPA sandwich (pIon Billerica, USA, P/N 110 163). The filter membranes at the bottom of the acceptor plate were infused with 5 µL of GIT-0 Lipid Solution (pIon, P/N 110 669) and 200 µL of Acceptor buffer (pION, P/N 110 139) was filled into each acceptor well. The sandwich was assembled, placed in the GutBoxTM (pION), and left undisturbed for 16 h. Then, it was disassembled and the compound concentration of donor, acceptor and reference compartment was measured by LC-MS (see below). Effective permeability (logP_e) was calculated from the compound flux deduced from the concentration in the donor and acceptor compartment as well as the retention in the filter area with the aid of the PAMPA Explorer Software (pIon, version 3.5) (4, 5).

Caco-2 cell permeation assay

Caco-2 cells were cultivated in tissue culture flasks (BD Biosciences, Franklin Lakes, NJ, USA) with DMEM high glucose medium, containing L-glutamine (2 mM), nonessential amino acids (0.1 mM), Penicillin (100 U/ml), Streptomycin (100 µg/ml), and fetal bovine serum (10%). The cells were kept at 37°C in humidified air containing 5% CO_2, and the medium was changed every second day. When approximately 90% confluence was reached, the cells were split in a 1:10 ratio and distributed to new tissue culture flasks. At passage numbers between 60 and 65, they were seeded at a density of 5.3×10^5 cells per well to Transwell 6-well plates (Corning Inc., Corning, NY, USA) with 2.5 ml of culture medium in the basolateral and 1.5 ml in the apical compartment. The medium was renewed on alternate days. Permeation experiments were performed between days 19 and 21 post seeding. Previous to the experiment, the integrity of the Caco-2 monolayers was evaluated by measuring the transepithelial electrical resistance (TEER) with an Endohm tissue resistance instrument (World Precision Instruments Inc., Sarasota, FL, USA). Experiments were performed in the apical-to-basolateral (absorptive) and basolateral-to-apical (secretory) directions in triplicates. Transport medium (DMEM without sodium pyruvate and phenol red) was withdrawn from the donor compartments of three wells and replaced by the same volume of compound stock solutions to reach an initial sample concentration of 62.5 µM. The transwell plate was then shaken (250 rpm) in the incubator. Samples (40 µl) were withdrawn after 15, 30, and 60 min from the donor and acceptor compartments and the concentrations were determined

by LC-MS. Apparent permeability coefficients (P_{app}) were calculated according to Equation 2, where dQ/dt is the permeability rate, A the surface area of the monolayer, and c_0 the initial compound concentration in the donor compartment.

$$P_{app} = \frac{dQ}{dt} * \frac{1}{A*c_0} \quad \text{(eq. 2)}$$

At the end of the experiment, TEER values were assessed again for each well and results from wells with values below 300 Ωcm^2 were discarded (6).

Simulated intestinal fluid stability assay

Incubations were performed in triplicate in a 96-well format on an Eppendorf Thermomixer Comfort. The reaction mixture (270 µl) consisting of porcine pancreatin in TRIS-HCl buffer (0.1 M, pH 6.5) was preheated (37°C, 500 rpm, 10 min), and the incubation was initiated by adding 30 µl of compound solution (200 µM) in TRIS-HCl buffer. The final concentration of the compound was 20 µM, and the pancreatin concentration was 2.5 mg/ml. At the beginning of the experiment (t = 0 min) and after an incubation time of 5, 10, 20, 30, and 60 min, samples (40 µl) were transferred to 120 µl of ice-cooled MeOH and centrifuged (3600 rpm, 4 °C, 10 min). The supernatant was transferred to a 96-well plate for LC-MS analysis. The metabolic degradation was assessed as percentage remaining compound versus incubation time. The concentration of the free acid was measured by LC-MS as well. Negative control experiments were performed in parallel by incubating the compound in assay buffer. Positive control experiments were conducted simultaneously with **2b**.

Human liver microsomal stability assay

Incubations were performed in triplicate in a 96-well format on an Eppendorf Thermomixer Comfort. The reaction mixture (270 µl) consisting of human liver microsomes, TRIS-HCl buffer (0.1 M, pH 7.4) and $MgCl_2$ (2 mM) was preheated (37°C, 500 rpm, 10 min), and the incubation was initiated by adding 30 µl of compound solution (20 µM) in TRIS-HCl buffer. The final concentration of the compound was 2 µM, and the microsomal concentration was 0.125 mg/ml. At the beginning of the experiment (t = 0 min) and after an incubation time of 5, 10, 20, 40, and 60 min, samples (40 µl) were transferred to 120 µl of ice-cooled MeOH and centrifuged (3600 rpm, 4 °C, 10 min). The supernatant was transferred to a 96-well plate for LC-MS analysis. The metabolic degradation was assessed as percentage remaining compound versus incubation time. The concentration of the free acid was measured by LC-MS as well. Negative control experiments were performed in parallel by preincubating the microsomes with the specific carboxylesterase inhibitor Bis(4-nitrophenyl)phosphate (1 mM) for 5 min before addition of the compound solution. Positive control experiments were conducted simultaneously with FimH ester prodrugs (7).

Recombinant carboxylesterases hydrolysis study

Incubations were performed in a 96-deepwell on an eppendorf Thermomixer comfort. The reaction mixture (180 µl) consisting of human carboxylesterase 1b or human carboxylesterase 2 (Sigma Aldrich), TRIS-HCl buffer (0.1 M, pH 7.4), test compound and reference compounds was preheated (40°C, 500 rpm, 10 min). Concurrently, controls experiments were performed with buffer. The incubation was initiated by adding 30 µl of the compound solution (200 µM) to TRIS buffer and human carboxylesterase 1 or 2. The final concentration of the compound was 20 µM and the concentration of the carboxylesterase was 100 µg/ml. At time points 0, 10, 30, 60, 90 and 120 min 25 µl of samples was transferred to 75 µl of icecooled MeOH. After the last time point, the quenching plate was put in the freezer (10 min) and later centrifuged (4 °C, 3600 rpm, 10 min). On an agilent 96-well plate the supernatant was transferred for LC-MS analysis. The metabolic release was calculated as percentage remaining compound versus incubation time.

Oxidation in rat liver microsomes assay

Incubations were performed in triplicate in a 96-well format on an Eppendorf Thermomixer Comfort. The reaction mixture (240 µl) consisting of rat liver microsomes, TRIS-HCl buffer (0.1 M, pH 7.4), $MgCl_2$ (2 mM), Compound and Bis(4-nitrophenyl)phosphate (0.1 mM) was preheated (37°C, 500 rpm, 10 min), and the incubation was initiated by first adding 30 µl of compound solution (200 µM) in TRIS-HCl buffer and then adding 30 µl of nicotinamide adenine dinucleotide phosphate (NADPH, 10 mM). The final concentration of the compound was 20 µM, 1 mM NADPH and the microsomal concentration was 0.5 mg/ml. At the beginning of the experiment (t = 0 min) and after an incubation time of 5, 10, 20, 40, and 60 min, samples (40 µl) were transferred to 120 µl of ice-cooled MeOH and centrifuged (3600 rpm, 4 °C, 10 min). The supernatant was transferred to a 96-well plate for LC-MS analysis. The metabolic degradation was assessed as percentage remaining compound versus incubation time. Control experiments were performed in parallel by preincubating the microsomes by not adding NADPH. Positive control experiments were conducted simultaneously with Midazolam.

Hydrolysis in rat liver microsomes assay

Assays were conducted according to human liver microsomal stability assay with 0.5 mg/ml rat liver microsomes and 20 µM compound concentration in the reaction mixture.

Simultaneous oxidation and hydrolysis in rat liver microsomes assay

Assays were conducted according to oxidation in rat liver microsomes assay in triplicate without adding BNPP to the reaction mixture. After 60 minutes the percentage of hydrolyzed molecule compared to the initial incubated prodrug was calculated.

Glucoronidation in rat liver microsomes assay

Incubations and sample taking were performed according to oxidations in rat liver microsomes. 15 minutes prior measurements, Alamethicin and Saccharolactone was added. The incubation was initiated by first adding 30 µl of compound solution (200 µM) in TRIS-HCl buffer and then adding 30 µl of UDP-α-D-glucuronic acid (UDPGA, 50 mM). The final concentrations were of the compound 20 µM, 5 mM UDPGA, 0.25 mg/ml Alamethicin, 1 mM Saccharolactone and the microsomal concentration was 0.5 mg/ml. Control experiments were performed in parallel by preincubating the microsomes by not adding UDPGA. Positive control experiments were conducted simultaneously with Testosterone.

Sulfation in rat liver microsomes assay

Incubations and sample taking were performed according to oxidations , butin rat liver cytosol. The incubation was initiated by first adding 30 µl of compound solution (200 µM) in TRIS-HCl buffer and then adding 30 µl of 3'-phosphoadenosine-5'-phosphosulfate (PAPS, 2 mM). The final concentrations were of the compound 20 µM, 0.2 mM PAPS. Control experiments were performed in parallel by preincubating the microsomes by not adding PAPS. Positive control experiments were conducted simultaneously with Hespericine.

Metabolite screening and hydrolysis susceptibility

A primary assay was conducted according to the oxidation in rat liver microsomes assay with 40 µM prodrug concentration and 2 mM NADPH. After 60 minutes, the complete reaction mixture was quenched in cold methanol in a 3 to 1 methanol to reaction mixture ratio. The mixture was concentrated *in vacuo* and resuspended in TRIS-HCl, to obtain a theoretical concentration of 100 µM of metabolite, assuming 100% oxidation occurred. Samples of this mixture were taken for metabolite analysis by LC-MS. Additional samples were added to 0.5 M sodium hydroxide in a 1 to 1 ratio, and after 30 minutes, the reaction was neutralized to verify the place of oxidation by measuring the concentration of active principle. Finally, the susceptibility of the metabolite was measured by incubating the mixture according to hydrolysis in rat liver microsomes assay with 0.5 mg/ml rat liver microsomes and a theoretical concentration of 10 µM metabolite. Over a time of 60 minutes, samples were taken analyzed for the active principle by LC-MS.

Micro scale thermophoresis

Reagents, devices and compounds. Monolith™ NT.115, Protein Labeling Kit BLUE-NHS (Amine Reactive), and Standard Treated Capillaries were purchased from Nanotemper Technologies GmbH (Munich, Germany). HEPES, NaOH, NaCl, $CaCl_2 \cdot 2H_2O$, and Tween 20 were purchased from Sigma-Aldrich Chemie GmbH (Steinheim, Germany). Slide-A-Lyzer cassettes (10 kDa MWCO) and bovine serum albumin (BSA) standard ampules were obtained from Thermo Fisher Scientific (Rockford, IL, USA):

E-selectin production. Cloning, transfection, expression and purification were previously described for the E-selectin$_{LEC2}$ construct by Preston *et al.* (8).

Labeling. E-selectin$_{LEC2}$ was labeled using the amine reactive protein labeling kit BLUE-NHS. Buffer exchange and labeling were performed according to the manufacturer's protocol. To protect the 1ysines in the binding site from being labeled, the protein was saturated with 600 µM of compound **1**. The labeled protein was dialyzed over night against assay buffer (10mM HEPES pH 7.4, 150 mM NaCl, 1 mM $CaCl_2$) using Slide-A-Lyzer dialysis cassettes (10 kDa MWCO). Protein concentration was determined by HPLC-UV against a BSA standard.(9, 10)

Microscale Thermophoresis. MST experiments were carried out at 25° C with 50% LED power, 30% MST power, a laser on time of 30 sec, and a laser off time of 5 sec using standard treated capillaries. Ligands were dissolved in assay buffer containing 0.05% v/v Tween 20. A 16-point dilution series of ligand starting at 0.5 mM for compound **1** and **3b** at 2.5 mM, **4b** at 5 mM, **5b** at 20 mM and **6b** at 0.25 mM was prepared and subsequently mixed 1:1 with a solution of 0.1 µM E-Selectin$_{LEC2}$ and incubated for 10 mins at room temperature. Datapoints were normalized using the bound and unbound borders achieved by NanoTemper Analysis 1.5.41 software (NanoTemper Technologies GmbH, Munich, Germany) and analyzed/illustrated with GraphPad Prism 5.0 (GraphPad Software, La Jolla, CA, USA). The measurements were globally fitted using Equation 3 for single site binding (11).

$$[PL] = \frac{(C_P + C_L + K_D) - \sqrt{(C_P + C_L + K_D) - 4C_P C_L}}{2C_P} \qquad \text{(eq. 3)}$$

Where [*PL*] is the protein-ligand complex concentration and K_D is the dissociation constant. C_P represents the total concentration of protein and C_L the total concentration of ligand.

In vivo pharmacokinetic studies

Male and female Balb/c mice weighting between 19 and 25 g were obtained from Janvier (Labs Le Genest-Saint-Isle, France) and were housed 3 per cage. The mice were kept under specific pathogen-free conditions in the Animal House of the Biozentrum, University of Basel, and animal experimentation guidelines according to the regulations of the Swiss veterinary law were followed. After 7 days of acclimatization, 8- to 14-week-old mice were used for the studies. Animals had free access to chow and water at any time except 2 hours prior dosage and were kept in a 12 h/12 h light/dark cycle. For administration volumes and sampling the good practice guidelines were followed.

The pharmacokinetic studies were performed by intravenous (IV) administration of the active principle or ester prodrugs at a dosage of 20 mg/kg or oral dosage of 50 mg/kg of the ester prodrugs by oral gavage. The compounds were diluted in 0.9 % NaCl Solution and injected into the tail vein. Tween 80 and DMSO was added to maximal 2%, respectively 5% if solubility problems occurred. Blood samples were obtained from the tail veins 7, 15, 30, 60, 90, min and 2 h after injection. Blood samples (approximately 20 µL) were transformed into tubes containing 1 µL 1 M EDTA solution and centrifuged at 2000 rpm for 10 minutes to obtain blood plasma. 5 µL of blood plasma was diluted in 50 µL cold MeOH containing 1 µM internal standarts to precipitate proteins. The supernatants were analyzed by LC-MS (See below).

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 C18 column (2.1 x 50 mm) with a 3 µm-particle size (Waters Corp., Milford, MA, USA). The mobile phase consisted of eluent A: 10 mM ammonium acetate, pH 5.0 in 95:5, H2O: ACN; and eluent B: ACN containing 0.1% formic acid. The flow rate was maintained at 0.6 mL/min. The gradient was ramped from 95% A/5% B to 5% A/95% B over 1 min, and then hold at 5% A/95% B for 0.1 min. The system was then brought back to 95% A/5% B, resulting in a total duration of 4 min. Fragmentor voltage and collision energy were optimized for the analysis of compounds in multiple reaction monitoring mode in positive mode for ester prodrugs or negative mode for active principles.

References

1. Kolb HC, Ernst B. Development of Tools for the Design of Selectin Antagonists. *Chemistry - A European Journal*. 1997;*3(10)*:1571-8.
2. Norman KE, Anderson GP, Kolb HC, Ley K, Ernst B. Sialyl Lewis(x) (sLe(x)) and an sLe(x) mimetic, CGP69669A, disrupt E-selectin-dependent leukocyte rolling in vivo. *Blood*. 1998;*91(2)*:475-83.
3. Kolb HC, Ernst B. Development of Tools for the Design of Selectin Antagonists. *Chemistry*. 1997;*3(10)*:7.
4. Avdeef A, Bendels S, Di L, Faller B, Kansy M, Sugano K, et al. PAMPA--critical factors for better predictions of absorption. *J Pharm Sci*. 2007;*96(11)*:2893-909.
5. Kansy M, Senner F, Gubernator K. Physicochemical high throughput screening: parallel artificial membrane permeation assay in the description of passive absorption processes. *J Med Chem*. 1998;*41(7)*:1007-10.
6. Hubatsch I, Ragnarsson EG, Artursson P. Determination of drug permeability and prediction of drug absorption in Caco-2 monolayers. *Nat Protoc*. 2007;*2(9)*:2111-9.
7. Klein T, Abgottspon D, Wittwer M, Rabbani S, Herold J, Jiang X, et al. FimH antagonists for the oral treatment of urinary tract infections: from design and synthesis to in vitro and in vivo evaluation. *J Med Chem*. 2010;*53(24)*:8627-41.
8. Preston RC, Jakob RP, Binder FP, Sager CP, Ernst B, Maier T. E-selectin ligand complexes adopt an extended high-affinity conformation. *J Mol Cell Biol*. 2016;*8(1)*:62-72.
9. Mesch S, Lemme K, Koliwer-Brandl H, Strasser DS, Schwardt O, Kelm S, et al. Kinetic and thermodynamic properties of MAG antagonists. *Carbohydr Res*. 2010;*345(10)*:1348-59.
10. Bitsch F, Aichholz R, Kallen J, Geisse S, Fournier B, Schlaeppi J-M. Identification of natural ligands of retinoic acid receptor-related orphan receptor α ligand-binding domain expressed in Sf9 cells—a mass spectrometry approach. *Analytical Biochemistry*. 2003;*323(1)*:139-49.
11. Cooper A. Biophysical Chemistry. 2004.

3.8 Manuscript 2: Amidic E-selectin antagonists

The bioisosteric replacement of the carboxylic acid of previously designed E-selectin antagonists is a promising approach. Not only the affinity, but also the pharmacokinetic properties of the compounds is enhanced. In this manuscript, the physicochemical and pharmacokinetic properties of amidic E-selectin antagonists is evaluated *in vitro* and *in vivo* with a focus on the observed active transport of the molecule.

Contribution to the Project:

Philipp Dätwyler designed and performed the assays of the pharmacokinetic evaluation *in vitro* and *in vivo* of the E-selectin antagonists. He further performed some affinity measurements by microscale thermophoresis and wrote the manuscript.

Philipp Dätwyler, Beatrice Wagner, Xiaohua Jiang, Tobias Mühlethaler, Beat Ernst*

Institute of Molecular Pharmacy, Department of Pharmaceutical Sciences, University of Basel, Klingelbergstrasse 50, 4056 Basel, Switzerland

*Corresponding author

The potential for oral bioavailability of amidic E-selectin antagonists

Philipp Dätwyler, Beatrice Wagner, Xiaohua Jiang, Tobias Mühlethaler, Beat Ernst*

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

* Corresponding author

Abbreviations: sialyl Lewis^x (sLe^x), vaso-occlusive crisis (VOC), intravenous (IV) multiple myeloma (MM), microscale thermophoresis (MST), parallel artificial membrane permeation assay (PAMPA).

Keywords: sialyl Lewis^x, E-selectin, pharmacokinetic properties, oral bioavailability, bioisosteres.

Abstract

E-selectin is a C-type lectin expressed on endothelial cells of blood vessels adjacent to inflammatory stimuli. It interacts with structures containing the tetrasaccharide sialyl Lewis^x (sLe^x) expressed on leukocytes, initiating the inflammatory adhesion cascade and the extravasation of the leukocytes to the side of inflammation. This process is highly regulated under physiological conditions, but alterations of the inflammatory adhesion cascade can lead to chronic inflammatory diseases. Previous developments of E-selectin antagonists starting from the natural ligand sLe^x led to non-orally bioavailable antagonists due to their hydrophilic core structure and a carboxylic acid. Herein, we present a lead development study with the goal to improve the pharmacokinetic properties of E-selectin antagonist by replacing the carboxylic acid with more lipophilic amides. Due to observed active uptake transport in enterocytes, the first orally bioavailable E-selectin antagonist with high affinity was developed.

Introduction

Cell-cell adhesion mediated by C-type lectins plays a crucial role in the inflammation cascade (1). In prolonged inflammation E-selectin is the main mediator between the site of inflammation and leucocytes (2). E-selectin is expressed on the vascular endothelium at sites of inflammation and binds cell surface glycoproteins of leukocytes, inducing their rolling which leads to the arrest and extravasation on the side of inflammation (3). This recognition is necessary for the initiation of the inflammatory cascade and therefore for host defense. However, excessive amplification of this process is associated with various diseases with an inflammatory component (4, 5). Therefore, blocking E-selectin and hence, inhibiting the initial step of leucocyte extravasation is a potential treatment of various acute and chronic inflammatory diseases (6).

The main motive recognized by E-selectin is the tetrasaccharide sialyl Lewis^x (sLe^x), present on glycoproteins located on leukocytes (7). Several E-selectin antagonists with improved affinities and pharmacokinetic properties have been developed. Recent studies showed a positive effect of the intravenously applied pan-selectin antagonist Rivipansel for the treatment of vaso-occlusive crisis (VOC) in sickle cell anemia patients in a phase II study and GMI1271 in resistance multiple myeloma (MM) in mice (8, 9). VOC and MM are severe and acute conditions requiring intensive care therefore therefore intravenous (IV) administrations are tolerated. However, for chronic inflammatory diseases, an orally bioavailable treatment would be more convenient. Many different chronic diseases with a strong correlation to E-selectin overexpression, such as arthrosclerosis, rheumatoid arthritis and psoriasis, still constitute an urgent unmet medical need (10-13).

Carbohydrates as lead structures for oral bioavailable drug candidates have the inherent problems of a large polar surface area leading to high hydrophilicity (14). A consequent reduction of the polar motives not required for binding in sLe^x and replacing them by lipophilic residues led to the antagonist **1** (15). Compound **1** is still a highly hydrophilic and large molecule, containing two carbohydrate moieties. Additionally, the carboxylic acid leads to a charged molecule, which is not fulfilling the criteria for oral bioavailability. Since further reduction of the polarity or size of the molecule induces a loss in affinity, alternative approaches to reach sufficient lipophilicity and permeability are needed (16, 17). Replacing

the carboxylic acid by an uncharged bioisostere offers the opportunity to improve the pharmacokinetic profile of E-selectin antagonists.

Methods

A detailed description of methods and procedures, as well as the synthesis of compounds **2a-p** and **3** is presented in the supplementary data.

Microscale thermophoresis: Cloning, transfection, expression and purification were previously described for the E-selectin$_{LEC2}$ construct by Preston *et al.* (17). Microscale thermophoresis (MST) experiments were carried out at 25° C with 50% LED power, 30% MST power, an initial delay of 5 sec, a laser on time of 30 sec, and a laser off time of 5 sec using standard treated capillaries and a 16 point dilution series.

Log$D_{7.4}$: The lipophilicity of molecules was assessed in sexduplicate at two different octanol-buffer ratios at pH 7.4. The octanol-buffer ratio and reference compound were chosen individually to minimize the inherent error of the assay.

Parallel artificial membrane permeation assay (PAMPA): Passive permeability of the molecules was measured at room temperature and pH 7.4 in quadruplicate. The starting incubation concentration was 25 µM and the incubation time was 16 h (unstirred). The evaluation was done according to Avdeef *et al.* (18).

Caco-2 cell based permeation assay: Experiments were run 21 days after seeding the Caco-2 cells on transwell plates in triplicate for the absorptive (Uptake) and secretory (Efflux) direction. The initial concentration of the incubation was 62.5 µM and samples were taken after 30 min incubations. The evaluation was done according to Hubatsch *et al.* (19) and Hou *et al.* (20).

Solubility: The thermodynamic solubility of the molecules was determined in aqueous Tris-HCl buffer (0.1 M pH 7.4) in triplicate. After adding the solution to solids, the mixture was put on ultrasonification for 30 minutes and after 24 h, the mixture was filtered.

Metabolic stability: The metabolic stability of compound **2l** was assessed in rat liver microsomes at a concentration of 0.5 mg/ml and compound concentration of 20 µM in triplicate. The reaction was initiated by adding the co-factor β-nicotinamide adenine dinucleotide phosphate (NADPH) to a final concentration of 1 mM. After 60 minutes, the incubation was stopped by adding ice-cold methanol.

Animal oral bioavailability studies: The oral bioavailability of compounds was assessed in male and female 8 to 12 week Balb/C mice. The formulation of the compound was done according to Thackaberry *et al.* to minimize acute toxic effects (21). The formulated compound was given either by an oral gavage and or by intravenous injection in triplicate and concentrations in the blood plasma was measured over a period of 3 h. The evaluation was done with PKSolver, using a one compartment model (22). The studies were permitted by the Swiss government (Animal study Nr. A 2865) and guidelines according to the regulations of the Swiss veterinary law were followed.

Results and discussion

By a consequent reduction of the polar surface area of sLe^x, compound **1** was obtained by Norman *et al.* (15). Compound **1** is still a highly hydrophilic and large molecule, containing two carbohydrate moieties and a carboxylic acid. Accordingly, the measured log$D_{7.4}$ is below -1.5 leading to insufficient passive permeation. Forming amidic bioisosteres allows the replacement of the charged carboxylic acid with a neutral moiety and furthermore to introduce larger lipophilic residues. Figure 1 summarizes different amidic derivatives **2a-p**, which were screened for their affinity towards E-selectin and permeability properties (Table 1).

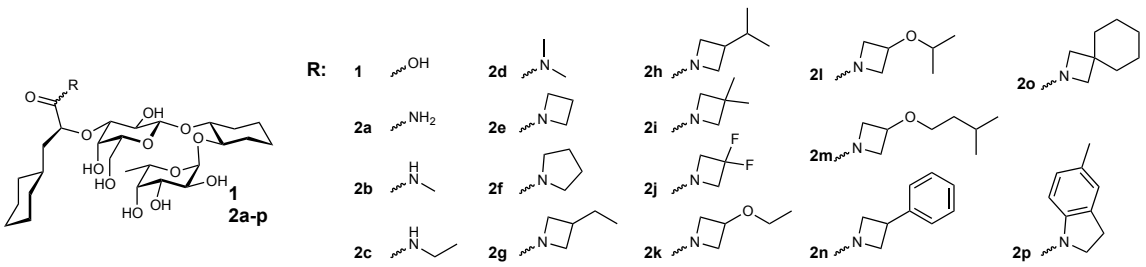


Figure 1 E-selectin antagonist **1** and different amidic bioisosteres thereof (**2a-p**).

Replacement of carboxylate in **1** by an amide (→**2a**) induces a 3-fold decrease of affinity towards E-selectin, but an increase of log$D_{7.4}$ close to zero. A further aliphatic elongation does not only improve the lipophilicity of the compounds, but also the affinity. Especially the cyclobutylamine **2e** and further lipophilic derivatives thereof at position 2' of the ring improve the affinity to E-selectin to the range of 15 to 25 µM. Only the flexible elongations in compound **2m** led to a reduced affinity. Furthermore, these lipophilic residues increased the log$D_{7.4}$ to measured values above 2.0. A log$D_{7.4}$ value between 2 and 3 is considered to be optimal for membrane permeability of small molecules (23). Therefore, the passive

permeation properties of the molecules were tested in a PAMPA (24). Although the lipophilicity increases with the size of the amidic residue, none of the molecules were sufficiently permeating the artificial membrane to predict a moderate to high absorption potential and no correlation between the lipophilicity and permeability was observed (Figure 2A, **A**) (18). Selected molecules showing good affinity or high lipophilicity were furthermore tested in a Caco-2 cell based permeation assay in the absorptive (uptake) or secretory (efflux) direction in order to assess their absorption potential and active transport. Surprisingly, **2h**, **2k** and **2m** showed sufficient uptake indicating a moderate absorption potential. Furthermore, compound **2h** and **2l** are actively taken up with an uptake ratio above 2.0. Both molecules consist of a terminal isopropyl moiety, similar to the terminal position of the amino acid L-valine. Di- and tripeptides including L-valine, as well as L-valin ester prodrugs are known to be actively taken up by enterocytes over the Pept1-transporter (25, 26). To verify the active transport, the Caco-2 transport assay was performed with **2h** and co-incubation of 2 mM enalapril, a known inhibitor of Pept1 (27). This led to a reduction of the uptake ratio to 1.9 ± 0.5, which is below the threshold for a predicted active transportation. Therefore, we assume that Pept1 is major contributor for the observed active transport of **2h** and **2l** in the Caco-2 cell based transport assay.

Table 1 Pharmacodynamic and pharmacokinetic properties of E-selectin antagonists. The affinities of the E-selectin antagonists were measured by microscale thermopheresis. Octanol-water distribution coefficients (log$D_{7.4}$) were determined at pH 7.4 by a miniaturized shake flask procedure in sextuplicate. The effective permeability (logP_e) through an artificial membrane was determined by the PAMPA in quadruplicate at pH 7.4.The thresholds for moderate and high absorption potential are -6.3 and -5.7, respectively (18). The apparent permeation (P_{app}) through a cell monolayer was assessed by a Caco-2 cell assay in the absorptive (apical to basolateral, Uptake) and secretory (basolateral to apical, Efflux) directions in triplicates and the uptake ratio was determined by dividing uptake by efflux. Thresholds for moderate and high absorption potential for the absorptive direction are 2 and 20 $*10^{-6}$ cm/s (20). Uptake ratio was calculated by dividing uptake AB by efflux BA Compound showing an uptake ratio of 2.0 or more are considered to show active uptake (19). The given ranges are standard deviation.

Cpd.	Affinity K_D MST [µM] (95% CI)	log$D_{7.4}$	PAMPA logP_e	Caco-2, P_{app} [10^{-6} cm/s]		
				uptake AB	efflux BA	uptake ratio AB/BA
1	60.7 (52.2-70.6)	< -1.5	< -10			
2a	166 (145-189)	-0.3 ± 0.1	-8.5 ± 0.3			
2b	112 (98.6-128)	-0.2 ± 0.1	-8.5 ± 0.3			
2c	107 (93.5 -127)	0.4 ± 0.0	-7.6 ± 0.1			
2d	32.6 (30.2-39.5)	0.3 ± 0.0	-8.4 ± 0.3			
2e	18.2 (14.9-19.3)	0.1 ± 0.1	-7.5 ± 0.3			
2f	20.5 (17.7-25.8)	0.7 ± 0.1	-7.5 ± 0.1			
2g	21.5 (16.7-27.7)	1.3 ± 0.1	-8.6 ± 0.4	0.08 ± 0.01	0.14 ± 0.03	0.6 ± 0.2
2h	23.4 (15.4-35.6)	1.4 ± 0.1	-8.2 ± 0.2	5.6 ± 0.7	1.15 ± 0.33	4.8 ± 1.9
2i	38.2 (32.4-45.7)	1.4 ± 0.1	-9.3 ± 0.5			
2j	15.3 (13.1-17.9)	0.8 ± 0.0	< -10.0			
2k	23.6 (16.7-33.5)	1.1 ± 0.1	-8.5 ± 0.2	8.9 ± 1.6	4.8 ± 4.4	1.9 ± 2.1
2l	15.0 (13.1-17.4)	1.2 ± 0.1	-9.3 ± 0.5	0.8 ± 0.2	0.25 ± 0.05	3.3 ± 0.7
2m	45.9 (34.13-61.8)	1.9 ± 0.1	-8.5 ± 0.2	6.6 ± 1.4	7.3 ± 7.8	0.9 ± 1.2
2n	16.2 (12.9-20.3)	2.3 ± 0.1	-8.5 ± 0.3	1.0 ± 0.3	1.18 ± 0.08	0.8 ± 0.4
2o	22.4 (17.2-29.0)	2.3 ± 0.1	-8.4 ± 0.3	0.09 ± 0.03	0.50 ± 0.07	0.2 ± 0.1
2p		2.5 ± 0.1	-7.0 ± 0.1	0.31 ± 0.02	1.10 ± 0.10	0.3 ± 0.1

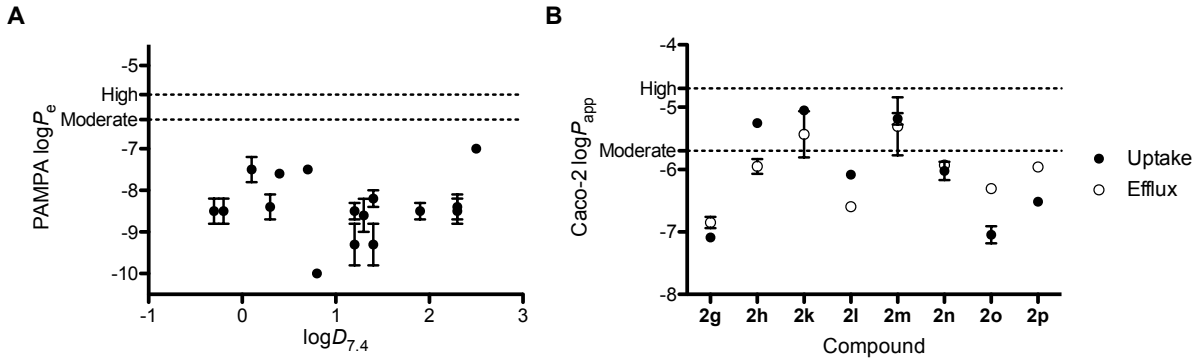


Figure 2 A The log$D_{7.4}$ of amidic E-selectin antagonists **2a-p** versus the effective permeation logP_e measured in PAMPA. Thresholds for moderate and high absorption potential are indicated as dotted lines and the range is given as standard deviation**. B** The apparent permeation logP_{app} is measured in a Caco-2 cell based permeation assay. Thresholds for moderate and high absorption potential are indicated as dotted lines and the range is given as standard deviation**.**

To verify the observed permeability predicting either an active uptake or a moderate passive absorption potential, the oral bioavailability of E-selectin antagonists was tested in a mouse model (Table 2). Compound **2n** was used as reference, since it was not sufficiently permeable even though its high lipophilicity. Besides the reference compound **2n**, the oral bioavailability of **2k** was below the detection limit. Overall, none of the compounds predicted to have moderate absorption potential in the Caco-2 cell based assay showed oral bioavailability above 3%. Only compound **2l** has an oral bioavailability of 7.7 %, although it did not show sufficient permeability in both permeation assays.

Table 2 Oral bioavailability and plasma half life ($t_{1/2}$, after IV injection) of amidic E-selectin antagonists in mice. It was assessed by comparing the area under curve (AUC) of the blood plasma concentration in mice after intravenous injection (IV) and oral gavage (PO) in triplicate. Errors given are the standard deviation.

Cpd.	$t_{1/2}$ [min]	Oral bioavailability (*F*) [%]
2h	7.5 ± 2.1	1.6 ± 0.8
2k	8.2 ± 0.4	< 0.5
2l	10.3 ± 2.0	7.7 ± 0.3
2m	5.0 ± 1.4	2.0 ± 1.5
2n	8.4 ± 1.7	< 0.5

Based on these data, the terminal cyclobutylamine-2-isoproxy moiety was chosen as preferred amidic replacement of the carboxylic acid. Furthermore, the stability against oxidation by cytochrome P450 metabolic enzymes in rat liver microsomes as well as the solubility in aqueous solution of **2l** was assessed. With a solubility of 1.5 ± 0.1 mg/ml and a stability of 88 ± 3 % of starting concentration remaining after 60 minutes incubating in liver microsomes, compound **2l** met the required pharmacokinetic properties for further investigations. Furthermore, other structural modifications known to enhance the affinity of E-selectin antagonists were implemented to form compound **3** (Figure 3A). The hydroxyl group at 2-position of galactose was replaced by chloride (manuscript 5), the cyclohexyl linker between the galactose and the fucose moiety was stabilized by adding an ethyl moiety (16), and the fucose was trifluorinated at the 6-position to improve the internal no-conventional hydrogen bond (manuscript 3). All three structural changes are enhancing the lipophilic character of the molecule, leading to a measured log$D_{7.4}$ of 3.6 and improved passive permeation in PAMPA (Table 3). Unfortunatelly, those changes lead also to a dramatically reduced solubility and an active efflux in the Caco-2 cells assay. However, despite the active efflux, an oral bioavailability of 4.3 ± 0.3 % was reached.

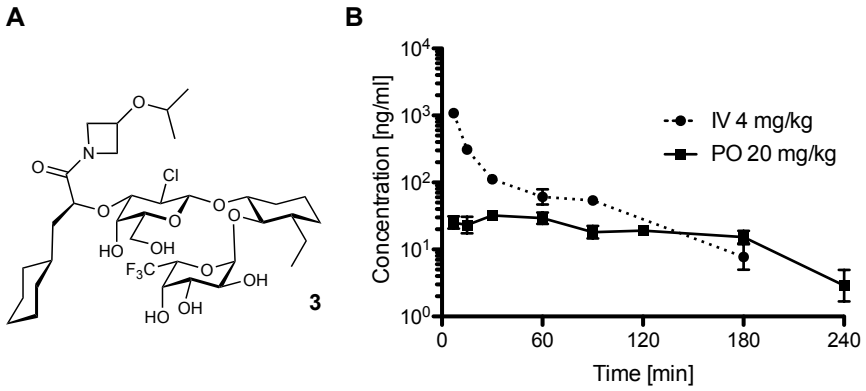


Figure 3 A Compound **3** with implemented modification to improve affinity. **B** Oral bioavailability of amidic E-selectin antagonist **3** in mice. Intravenous injection (IV) and oral gavage (PO) were given in triplicate. Error bars are given as standard deviation.

Table 3 Pharmacodynamic and pharmacokinetic properties of E-selectin antagonist **3**. The E-selectin affinity was measured by isothermal titration calorimetry (ITC). Octanol-water distribution coefficients (log$D_{7.4}$) were determined at pH 7.4 by a miniaturized shake flask procedure in sextuplicate. The effective permeability (P_e) through an artificial membrane was determined by the PAMPA in quadruplicate at pH 7.4. The apparent permeation (P_{app}) through a cell monolayer was assessed by a Caco-2 cell assay in the absorptive (apical to basolateral, Uptake) and secretory (basolateral to apical, Efflux) directions in triplicates and the uptake ratio was determined by dividing Uptake by Efflux. The oral bioavailability was assessed by comparing the AUC of the concentration in blood plasma in mice after intravenous injection (IV) and oral gavage (PO) in triplicate. All ranges given are as standard deviation.

Cpd	Affinity ITC	log$D_{7.4}$	PAMPA logP_e	Solubility [µg/mL]	Caco-2 P_{app} [10^{-6} cm/s]			*In vivo*	
					uptake AB	efflux BA	uptake ratio	$t_{1/2}$ [min]	*F*
3	To be measured	3.6 ± 0.1	-6.4 ± 0.2	23 ± 2	2.9 ± 0.4	55.5 ± 2.7	0.05 ± 0.01	4.8 ± 0.6	4.3 ± 0.3

Conclusions

By replacing the carboxylic acid moiety of E-selectin antagonists by more lipophilic amides, not only the pharmacodynamic, but also the pharmacokinetic properties were enhanced. A simple addition of lipophilic moieties was not sufficient to improve the permeability of the antagonists, since the core structure fails the established the rules of Lipinski *et al.* (28) and Veber *et al.* (29) to predict oral bioavailability. Even though the antagonists showed insufficient passive permeability, specific amidic replacements lead according to the Caco-2 cell assay to active uptake. In the case of the cyclobutylamine-2-isoproxy moiety ($\rightarrow$ **2l**), the active uptake resulted in a 7 % oral bioavailability in mice. When additional structural modifications were implemented to improve the affinity (**2l** $\rightarrow$ **3**), oral bioavailability although to a slightly reduced proportion was observed. To our knowledge, this is the first reported highly affine and oral bioavailable E-selectin antagonist.

References

1. Lasky LA. Selectins - Interpreters of Cell-Specific Carbohydrate Information during Inflammation. *Science*. **1992**;*258(5084)*:964-9.
2. Erbe DV, Wolitzky BA, Presta LG, Norton CR, Ramos RJ, Burns DK, et al. Identification of an E-Selectin Region Critical for Carbohydrate Recognition and Cell-Adhesion. *J Cell Biol*. **1992**;*119(1)*:215-27.
3. Ley K, Kansas GS. Selectins in T-cell recruitment to non-lymphoid tissues and sites of inflammation. *Nat Rev Immunol*. **2004**;*4(5)*:325-35.
4. Angiari S. Selectin-mediated leukocyte trafficking during the development of autoimmune disease. *Autoimmun Rev*. **2015**;*14(11)*:984-95.
5. Ley K. The role of selectins in inflammation and disease. *Trends in Molecular Medicine*. **2003**;*9(6)*:263-8.
6. Rossi B, Constantin G. Anti-selectin therapy for the treatment of inflammatory diseases. *Inflamm Allergy Drug Targets*. **2008**;*7(2)*:85-93.
7. Ley K, Gaehtgens P, Fennie C, Singer MS, Lasky LA, Rosen SD. Lectin-like cell adhesion molecule 1 mediates leukocyte rolling in mesenteric venules in vivo. *Blood*. **1991**;*77(12)*:2553-5.
8. Telen MJ, Wun T, McCavit TL, De Castro LM, Krishnamurti L, Lanzkron S, et al. Randomized phase 2 study of GMI-1070 in SCD: reduction in time to resolution of vaso-occlusive events and decreased opioid use. *Blood*. **2015**;*125(17)*:2656-64.
9. Natoni A, Smith TAG, Keane N, McEllistrim C, Connolly C, Jha A, et al. E-selectin ligands recognised by HECA452 induce drug resistance in myeloma, which is overcome by the E-selectin antagonist, GMI-1271. *Leukemia*. **2017**.
10. Cybulsky MI, Gimbrone MA, Jr. Endothelial expression of a mononuclear leukocyte adhesion molecule during atherogenesis. *Science*. **1991**;*251(4995)*:788-91.
11. Davies MJ, Gordon JL, Gearing AJ, Pigott R, Woolf N, Katz D, et al. The expression of the adhesion molecules ICAM-1, VCAM-1, PECAM, and E-selectin in human atherosclerosis. *J Pathol*. **1993**;*171(3)*:223-9.
12. Koch AE, Burrows JC, Haines GK, Carlos TM, Harlan JM, Leibovich SJ. Immunolocalization of endothelial and leukocyte adhesion molecules in human rheumatoid and osteoarthritic synovial tissues. *Lab Invest*. **1991**;*64(3)*:313-20.
13. Petzelbauer P, Pober JS, Keh A, Braverman IM. Inducibility and expression of microvascular endothelial adhesion molecules in lesional, perilesional, and uninvolved skin of psoriatic patients. *J Invest Dermatol*. **1994**;*103(3)*:300-5.
14. Ernst B, Magnani JL. From carbohydrate leads to glycomimetic drugs. *Nat Rev Drug Discov*. **2009**;*8(8)*:661-77.
15. Norman KE, Anderson GP, Kolb HC, Ley K, Ernst B. Sialyl Lewis(x) (sLe(x)) and an sLe(x) mimetic, CGP69669A, disrupt E-selectin-dependent leukocyte rolling in vivo. *Blood*. **1998**;*91(2)*:475-83.
16. Schwizer D, Patton JT, Cutting B, Smiesko M, Wagner B, Kato A, et al. Pre-organization of the core structure of E-selectin antagonists. *Chemistry*. **2012**;*18(5)*:1342-51.
17. Preston RC, Jakob RP, Binder FP, Sager CP, Ernst B, Maier T. E-selectin ligand complexes adopt an extended high-affinity conformation. *J Mol Cell Biol*. **2016**;*8(1)*:62-72.
18. Avdeef A, Bendels S, Di L, Faller B, Kansy M, Sugano K, et al. PAMPA--critical factors for better predictions of absorption. *J Pharm Sci*. **2007**;*96(11)*:2893-909.
19. Hubatsch I, Ragnarsson EG, Artursson P. Determination of drug permeability and prediction of drug absorption in Caco-2 monolayers. *Nat Protoc*. **2007**;*2(9)*:2111-9.
20. Hou T, Wang J, Zhang W, Xu X. ADME evaluation in drug discovery. 7. Prediction of oral absorption by correlation and classification. *J Chem Inf Model*. **2007**;*47(1)*:208-18.
21. Thackaberry EA, Wang X, Schweiger M, Messick K, Valle N, Dean B, et al. Solvent-based formulations for intravenous mouse pharmacokinetic studies: tolerability and recommended solvent dose limits. *Xenobiotica*. **2014**;*44(3)*:235-41.

22. Zhang Y, Huo M, Zhou J, Xie S. PKSolver: An add-in program for pharmacokinetic and pharmacodynamic data analysis in Microsoft Excel. *Comput Methods Programs Biomed*. **2010**;*99(3)*:306-14.
23. Guimaraes CR, Mathiowetz AM, Shalaeva M, Goetz G, Liras S. Use of 3D properties to characterize beyond rule-of-5 property space for passive permeation. *J Chem Inf Model*. **2012**;*52(4)*:882-90.
24. Kansy M, Senner F, Gubernator K. Physicochemical high throughput screening: parallel artificial membrane permeation assay in the description of passive absorption processes. *J Med Chem*. **1998**;*41(7)*:1007-10.
25. Adibi SA. The oligopeptide transporter (Pept-1) in human intestine: biology and function. *Gastroenterology*. **1997**;*113(1)*:332-40.
26. Terada T, Inui K. Peptide transporters: structure, function, regulation and application for drug delivery. *Curr Drug Metab*. **2004**;*5(1)*:85-94.
27. Knutter I, Wollesky C, Kottra G, Hahn MG, Fischer W, Zebisch K, et al. Transport of angiotensin-converting enzyme inhibitors by H+/peptide transporters revisited. *J Pharmacol Exp Ther*. **2008**;*327(2)*:432-41.
28. Lipinski CA. Drug-like properties and the causes of poor solubility and poor permeability. *J Pharmacol Toxicol Methods*. **2000**;*44(1)*:235-49.
29. Veber DF, Johnson SR, Cheng HY, Smith BR, Ward KW, Kopple KD. Molecular properties that influence the oral bioavailability of drug candidates. *J Med Chem*. **2002**;*45(12)*:2615-23.

3.9 **Manuscript 3:** Non-conventional hydrogen bond

The core of sialyl Lewis^x is stabilized by an intramolecular non-conventional hydrogen bond formed between H-C5 of the L-fucose moiety and O5 of D-galactose. By replacing the methyl group in the 5-position of L-Fuc by the electron-withdrawing CF_3-group, the polarization of the neighboring H-C is increased and thus leading to an increased quality of the non-conventional hydrogen bond. This is expected to have an impact on the affinity and physicochemical properties of newly synthesized E-selectin antagonists.

Contribution to the Project:

Philipp Dätwyler measured the stability of E-selectin antagonists in acidic conditions and contributed to the writing of the manuscript.

Norbert Varga,[1)] Xiaohua Jiang,[1)] Tobias Mühletaler,[1)] Beatrice Wagner,[1)] Martin Smieško,[1)] Roman Jakob,[2)] **Philipp Dätwyler,**[1)] Pascal Zihlmann,[1)] Timm Maier,[2)] Oliver Schwardt,[1)] Beat Ernst[1)]*

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

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

*Corresponding author

Getting in Shape for Binding – Rational Design of a Non-Conventional Hydrogen Bond in Sialyl Lewis^x-Mimetics

Norbert Varga,[1)] Xiaohua Jiang,[1)] Tobias Mühletaler,[1)] Beatrice Wagner,[1)] Martin Smieško,[1)] Roman Jakob,[2)] Philipp Dätwyler,[1)] Pascal Zihlmann,[1)] Timm Maier,[2)] Oliver Schwardt,[1)] Beat Ernst[1)]*

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

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

* Corresponding author

Introduction

Hydrogen bonding is important in many chemical processes. It is responsible for water's unique solvent capabilities, for holding complementary strands of DNA together, and for determining the three-dimensional structure of folded proteins including enzymes and antibodies (1-3). Principally, a hydrogen bond is formed between a proton donor AH containing an electronegative donor atom A and an acceptor B exhibiting at least one lone pair of electrons. Although the electronegativity of carbon and hydrogen are almost identical, there were some early suggestions that C-H groups can form weak, so-called non-classical or non-conventional hydrogen bonds when sufficiently activated by adjacent electronegative substituents (4-6).

Some years ago, we discovery by NMR experiments that the trisaccharide core of Lewis^x (**1**, Le^x, Figure 1A) is stabilized by an intramolecular non-conventional hydrogen bond formed between H-C5 of the L-fucose moiety and O5 of D-galactose. Quantum mechanical calculations exhibited that the C–H···O hydrogen bond represents approx. 40% of the total energy stabilizing the Le^x core (7). The remaining stabilizing energy is contributed by the exo-anomeric effects reducing the conformational flexibility of the glycosidic bonds, a repulsion exerted by the *N*-acetate group of D-Glc*N*Ac pushing the L-Fuc moiety underneath D-Gal and finally a strong stacking interaction between L-Fuc and the D-Gal (Figure 1A) (8, 9). These findings were later confirmed by Battistel *et al.* based on long range HSQC NMR experiments (10). Recently, we reported that non-conventional hydrogen bonds are not limited to the Lewis core but stabilize a wide range of fucosylated glycoepitopes and thereby constitute a general secondary structural element stabilizing carbohydrate conformations (11).

Intrigued by this observation, we planned to further enhance the strength of the non-conventional hydrogen bonds donated by H-C5 by appropriate substitution of L-fucose. We selected sialyl Lewis^x (**2**, sLe^x, Figure 1B) as model compound because of its important role as ligands of E- and P-selectin in inflammatory diseases, e.g. ischemia-reperfusion injuries, asthma, or psoriasis (12-14). Blocking the selectin-leukocyte interaction with selectin antagonists was shown to have a significant therapeutic impact (15-17).

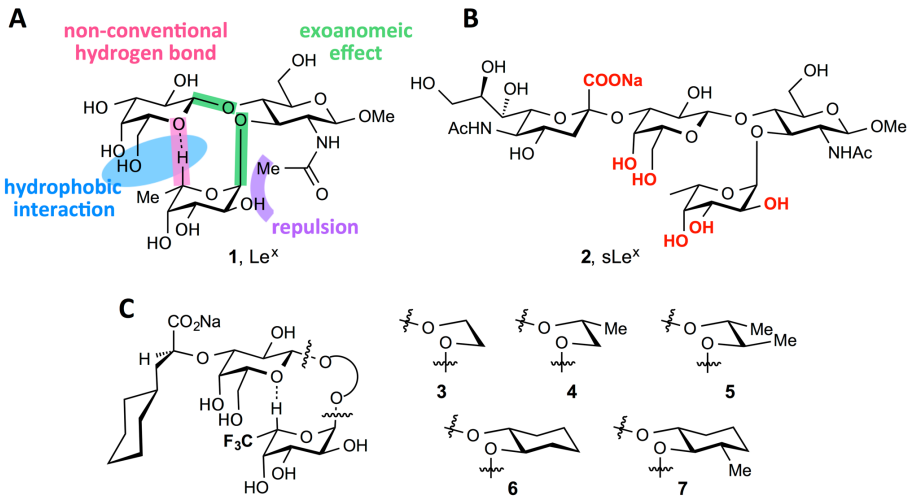


Figure 1 A Trisaccharide Lewis^x (**1**) stabilized by exoanomeric effects, repulsion between the *N*Ac-group of D-Glc*N*Ac and L-fucose, a strong stacking interaction between the L-Fuc and the D-Gal moiety and the non-conventional H-bond $C5^{Fuc}$-H···$O5^{Gal}$, (7-9). **B** Sialyl Lewis^x (**2**), the tetrasaccharide epitope recognized by selectins; pharmacophores are highlighted in red. **C** E-selectin antagonists with an electron-withdrawing CF_3-group in the 5-position of L-Fuc moiety (**3** - **7**) exhibiting improved non-conventional H-bonds between H-$C5^{Fuc}$ and O-5^{Gal}.

However, due to the shallow binding site of E- and P-selectin and the considerable distances between the pharmacophores (Figure 1B, highlighted in red), their pre-organization in the bioactive form turned out to be a prerequisite for high binding affinity (9, 18-21). In order to further stabilize the Le^x core and therewith the degree of pre-organization, we replaced in the current study the methyl group in the 5-position of L-Fuc by the electron-withdrawing CF_3-group. This is expected to increase the polarization of the neighboring H–C(5) bond, and thus the quality of the non-conventional hydrogen bond (Figure 1C).

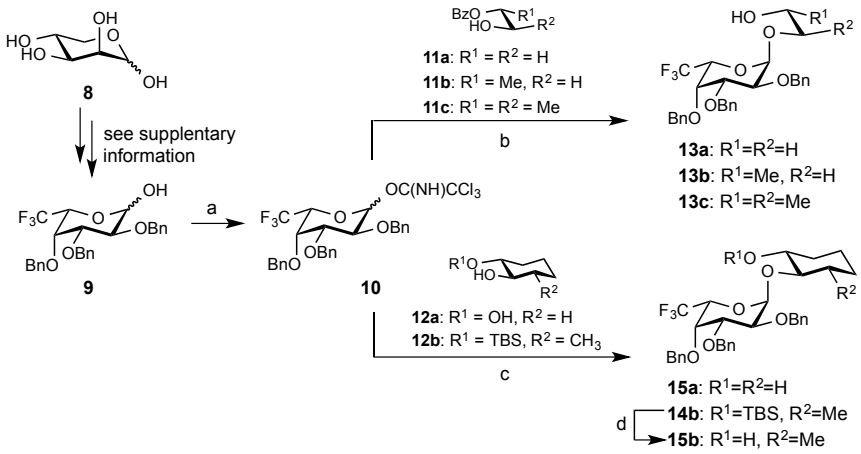
Results

For the chemical syntheses of the sLe^x mimetics **3** - **7** we primarily used a 2+2 strategy, i.e. Fucα1-3Glc*N*Ac mimetics were glycosylated with a D-galactosyl donor carrying (1*S*)-carboxy-2-cyclohexylethyl in its 3-position (22). Only for antagonist **5** a linear approach was applied, i.e. starting from a Glc*N*Ac mimic the subunits were introduced one by one. For comparison purposes, analogs equipped L-fucose instead of 6,6,6-trifluoro-L-fucose were prepared as well (23).

Synthesis of 6,6,6-trifluoro-L-fucose. A previous reported preparation of 6,6,6-trifluoro-L-fucose afforded a mixture of furanose and pyranose derivatives (24). To prevent the undesired formation of the furanose form, a strategy used for the synthesis of the trifluoromethyl analog of L-rhamnopyranose was applied, yielding 2,3,4-tribenzyl-6,6,6-trifluoro-L-fucose (**9**) starting

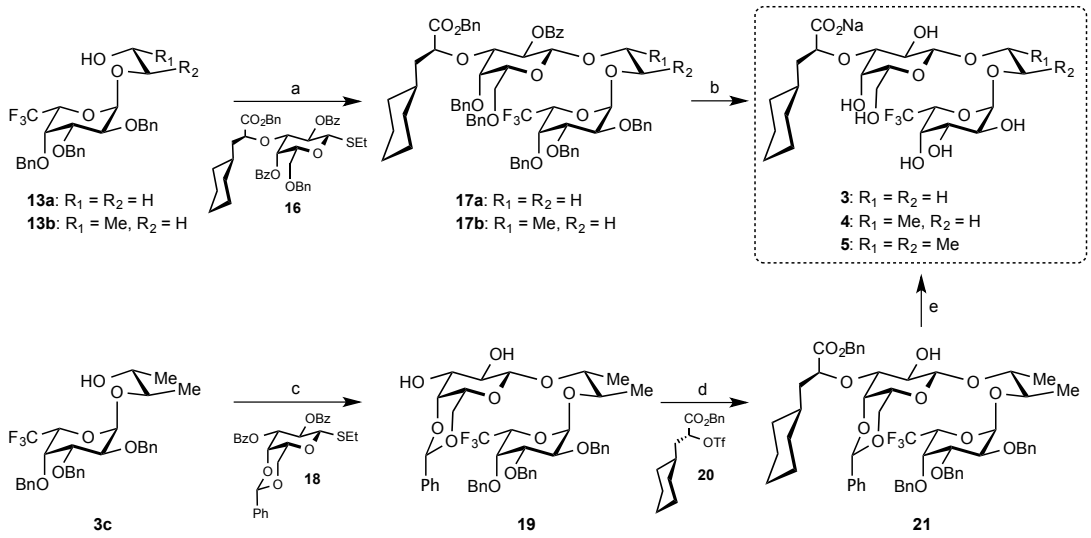
from D-lyxose (**8**) in an overall yield of 12% (Scheme 1, see also Supplementary Information) (25).

Synthesis of trifluorinated sLe^x mimetics. For the synthesis of disaccharide mimics **13a-c** and **15a,b** (Scheme 1), 2,3,4-tribenzyl-6,6,6-trifluoro-L-fucose (**9**) was transformed into an α/β-mixture of the glycosyl donor **10** using Cl_3CCN in the presence of K_2CO_3 at 0 °C. Fucosylation of the non-cyclic acceptors **11a-c** and cyclic acceptors **12a,b** promoted with TBSOTf in DCM/1,4-dioxane yielded mainly the α-fucosides **13a-c**, **15a** and, via the TBS-protected **14b**, the pseudodisaccharide **15b**.



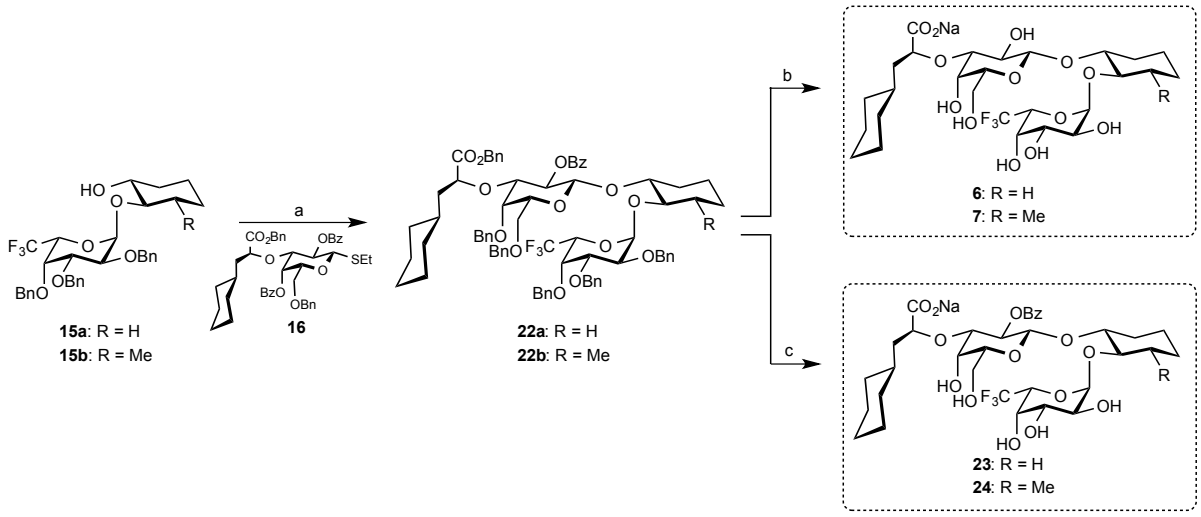
Scheme 1 a) Cl_3CCN, K_2CO_3, DCM, 0 °C, 18 h, (90%, α/β = 1:4); b) i. **11a-c**, TBSOTf, DCM/1,4-dioxane (1:2), 4 Å mol. sieves, 0 °C to rt, 18 h; ii. NaOMe, MeOH/DCM (1:1), rt, 3 h, (**13a**: 42%, **13b**: 50%, **13c**: 41%, 2 steps); c) **12a** (excess), TBSOTf, DCM, 4 Å mol. sieves, 0 °C to rt, 18 h, (**15a**: 56%) or **12b**, TBSOTf, DCM/1,4-dioxane (1:2), 4 Å mol. sieves, 0 °C to rt, 18 h (**14b**), d. HF-py, MeCN, 0 °C, 48 h, (**15b**: 50%, 2 steps).

For the synthesis of the test compounds **3** - **5** (Scheme 2), two different strategies were applied. In the first strategy, the pseudodisaccharides **13a,b** were galactosylated with donor **16** using DMTST as promotor (**17a,b**) (26). Deprotection yielded the test compounds **3** and **4**. In a second approach, pseudodisaccharide **13c** was first galactosylated with donor **18** followed by alkylation of the 3'-position with triflate **20** (27, 28). Final deprotection by catalytic hydrogenation yielded the test compound **5**.



Scheme 2 a) DMTST, DCM, 4 Å mol. sieves, rt, 24 h, (**17a**: 75%, **17b**: 87%); b) i. H_2 (1 bar), $Pd(OH)_2$/C, THF, rt, 2 h, ii. LiOH/MeOH (1:1), H_2O, 60 °C microwave, 10 h, (**3**: 45%, **4**: 55%,); c) i. DMTST, DCM, 4 Å mol. sieves, rt, 18 h; ii. NaOMe, MeOH/DCM (1:1), rt, 18 h, (82%, 2 steps); d) Bu_2SnO, MeOH, reflux, 4 h; **20**, CsF**,** DME, rt, 18 h (45%); e) H_2 (1 bar), $Pd(OH)_2$/C, THF, rt, 2 h, (**5**: 33%).

Finally, the synthesis of sLe^x mimetics containing cyclohexane-1,2-diol or 1-methyl-cyclohexane-2,3-diol instead of Glc*N*Ac is described in scheme 3.

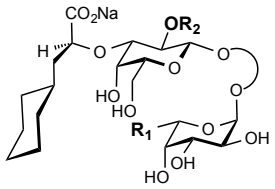


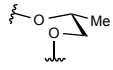
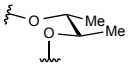
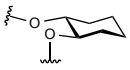
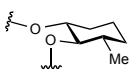
Scheme 3 a) DMTST, DCM, 4 Å mol. sieves, rt, 24 h, (**22a**: 68%, **22b**: 46%); b) i. H_2 (1 bar), $Pd(OH)_2$/C, THF, rt, 2 h, ii. LiOH/MeOH (1:1), H_2O, 60 °C microwave, 10 h, (**6**: 69%, **7**: 40%); c) i. H_2 (1 bar), $Pd(OH)_2$/C, THF, rt, 2 h, ii. NaOMe, MeOH, 3 h, rt, (**23**: 40%, **24**: 34%).

Affinity determination by microscale thermophoresis. With the microscale thermophoresis assay (MST) dissociation constant (K_D) can be determined based on the thermophoretic behavior of E-selectin in the presence of different amounts of antagonist (29). For better comparison purposes, the analogues **25** - **30** containing natural L-fucose were tested alongside with the new antagonists **3** - **7** (Table 1). In both series, **3** - **7** (containing 6,6,6-trifluoro-L-fucose) and **25** - **30** (containing L-fucose), the apparent affinity gains clearly demonstrate the beneficial

contribution of the pre-organization of the Le^x core mimics to affinity (9, 26). However, MST measurements revealed that selectin antagonists with a CF_3-group performed better compared to their counterparts without the modification on the L-fucose moiety. Finally, a benzoyl group in 2'-position led to a 3-fold improvement of affinity (see e.g. Table 1, entry 7 vs. entry 9), most probably because of an improved pre-organization of the acid pharmacophore (see Supplementary Information, Figure SQ). The beneficial effects from the individual SAR observations, i.e. a constrained linker, an improved non-conventional hydrogen bond and an improved pre-organization of the acid moiety proved to be additive and surmount to a gain of affinity by a factor of 2000 when realized in one molecule (see **25** vs. **24**).

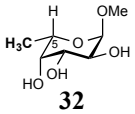
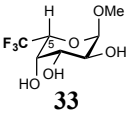
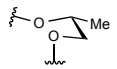
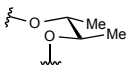
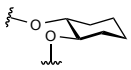
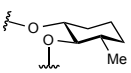
Table 1 Affinities for E-selectin (Lec-EGF-SCR2) were determined by a microscale thermophoresis (MST) assay (30); **25** & **27-31** have been published earlier; [a)]for the synthesis of **26** see supplementary information.



Entry	Ligand	Spacer	R_1	R_2	K_D [µM]	Improvement
1	**3**		CF_3	H	879	5.6-fold
2	**25**[31]		CH_3	H	4922	
3	**4**		CF_3	H	161	3.1-fold
4	**26**[a)]		CH_3	H	499	
5	**5**		CF_3	H	55.6	2.7-fold
6	**27**[26]		CH_3	H	150	
7	**6**		CF_3	H	21.5	2.8-fold
8	**28**[20]		CH_3	H	60.7	
9	**23**		CF_3	Bz	7.2	2.6-fold
10	**29**[31]		CH_3	Bz	19.1	
11	**7**		CF_3	H	5.2	2.6-fold
12	**30**[9]		CH_3	H	13.7	
13	**24**		CF_3	Bz	2.4	1.8-fold
14	**31**[9]		CH_3	Bz	4.3	

Chemical shift as an indicator for proximity (Table 2). The ^{1}H chemical shift of H5 of the fucose moiety containing either CH_3 or CF_3 at the 5-position was monitored as a reporter for the strength of the non-conventional H-bond between H-C5^{Fuc} and O5^{Gal} as this values is highly sensitive to the distance of the investigated H-bond interaction and represents a population-weighted average of all conformations present in solution (9, 31).

Table 2 Comparison of the H5 chemical shifts of the 6,6,6-trifluoro-L-fucosyl moiety (**3**-**7**) and the fucosyl moiety (**25**-**28**, **30**). For the full chemical shift assignments see Supplementary Information and Ref 41 (for methyl α-L-fucoside **32**).

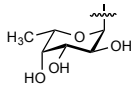
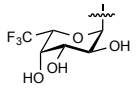
Entry	Linker	Compound	δ [ppm] of H5^{Fuc}	Δδ [ppm] rel. to 32	Compound	δ [ppm] of H5^{trifluorFuc}	Δδ [ppm] rel. to 33
1	-	 **32**	3.93	-	 **33**	4.19	-
2		**25**	4.12[31]	0.19	**3**	4.57	0.38
3		**26**	4.16	0.23	**4**	4.58	0.39
4		**27**	4.35[26]	0.42	**5**	4.87	0.68
5		**28**	4.50[26]	0.57	**6**	5.10	0.91
6		**30**	4.84[9]	0.91	**7**	5.32-5.38	1.13-1.19

In the L-Fuc series, the chemical shift of H5^{Fuc} for both antagonists with the non-cyclic linkers, i.e. with 1,2-ethandiol (**25**) and (2*R*)-1,2-propandiol (**26**) (entry 2 and 3), exhibits values almost identical to methyl L-fucoside (**32**) (entry 1), indicating that the fraction of conformations exhibiting a close proximity between H-C5^{Fuc} and O5^{Gal} is low. In contrast, in the trifluoro-L-Fuc series, a powerful downfield shift of almost 0.5 ppm compared to 6,6,6-trifluoro-α-L-fucoside (**33**) indicates that the fraction of conformations stabilized by the non-conventional H-bond is substantially increased (**3** & **4**), or in other words the average H-bond distance substantially decreased. An additional restriction of the conformational flexibility by introducing the (2*R*,3*R*)-2,3-butanediol linker (entry 4) and to an even larger extent, by cyclic D-Glc*N*Ac replacements (entry 5 and 6) induced substantial downfield shifts of the H5^{Fuc} signal in both series. In these

cases, the spatial proximity of $H5^{Fuc}$ and $O5^{Gal}$ is further pronounced, indicating an increased pre-organization of the antagonists in the bioactive conformation. Thus, the polarization of the H-C5 bond induced by the electron withdrawing potential of the CF_3-group likely leads to a strengthening of the non-classical H-bond. In summary, by improving the pre-organization of the core of sLe^x-derived E-selectin mimetics, we achieved an almost 1'000-fold improvement of the dissociation constants K_D when the flexible L-fucose derivative **25** (Table 1, entry 2) is compared with the rigid, CF_3-analogue **7** (entry 11).

Isothermal Titration Calorimetry (ITC, Table 3). The substitution of the methyl group (C6) by an electron withdrawing trifluoromethyl group (**30** → **7**) polarizes the adjacent hydrogen at 5-position of L-Fuc and thereby improves the core pre-organization. In ITC experiments, this effect is observed as a decreased entropic cost (**30** → **7**; $-T\Delta\Delta S°$ = -1.6 kJ mol^{-1}) upon binding as well as an enthalpic gain (**30** → **7**; $\Delta\Delta H°$ = -1.0 kJ mol^{-1}), most likely due to improved protein–ligand interactions. Furthermore, larger size of the trifluormethyl group might also better shield the Ca^{2+} cavity, lowering the local dielectric constant and therefore strengthening the existing electrostatic interactions.

Table 3 Thermodynamic parameters of E-selectin binding to sLe^x (**2**) and glycomimetic antagonists (**7** & **30**) determined by isothermal titration calorimetry (ITC). All measurements were carried out at 298.15 K. Affinities measured with microscale thermophoresis (MST) are shown for comparison.

Cpd.		MST: K_D [µM]	ITC: K_D [µM]	$\Delta G°$ [kJ mol^{-1}]	$\Delta H°$ [kJ mol^{-1}]	$-T\Delta S°$ [kJ mol^{-1}]
2	sLe^x	739	778	-17.7	4.9	-22.6
30		13.7	17.8	-27.1	-5.9	-21.2
7		5.2	6.39	-29.6	-6.9	-22.8

C-H···O non-conventional hydrogen bond confirmed by molecular dynamics (MD) simulations and *ab initio* calculations. The pairwise analysis of preferred conformations of derivatives **3** & **25**, **6** & **28**, and **7** & **30** obtained from 0.48 µs MD simulations resulted in a consistent shortening of the average FucH5-GalO5 distance when the methyl group of L-Fuc was replaced by a CF_3 group from 5.84 Å in **25** to 5.72 Å in **3**, from 2.98 Å in **28** to 2.84 Å in **6**, and from 2.84 Å in **30**

to 2.69 Å in **7**. This correlates well with the trend in measured chemical shifts of FucH5 (increased deshielding with shorter FucH5-GalO5 distance).

The quantification of the intramolecular stabilization using our established high-level *ab initio* approach (MP2/aug-cc-pVTZ counterpoise / B3LYP/6-31G(d,p) (7) showed a strong increase of the C-H···O hydrogen bond interaction from 1.87 kcal/mol in **28** (*N.B.* in Le^x (**1**) it is 1.76 kcal/mol) to 3.55 kcal/mol in the CF_3-counterpart **6**. The total stacking interaction between the Fuc and the Gal moiety is also improved from 4.35 kcal/mol in **28** (*N.B.* in Le^x (**1**) it is 4.52 kcal/mol) to 6.65 kcal/mol in **6**, clearly demonstrating the key effect of the CF_3 group on the stabilization of the bioactive core conformation.

Co-crystallization of E-selectin (Lec-EGF-SCR2) with the antagonists 6 and 7 (30).

When the CH_3-group of L-Fuc was replaced by aromatic and heteroaromatic substituents, only a minor influence of affinity was observed (32). This is in accordance with the substituent in the 5-position of the fucose moiety pointing to the surrounding water. In addition, all substituents can still exert a hydrophobic contact with the β-face of the above Gal moiety, essential for the pre-organization of the core. Thus, the replacement of CH_3 group by the electron-withdrawing CF_3 group leads mainly to an improved polarization of the C5-H bond leading to a further stabilization of the core by an improvement of the non-conventional hydrogen bond formed with the ring oxygen of the adjacent D-Gal moiety. When the co-crystal structures of the CF_3-derivatives **6** and **7** were compared with E-selectin co-crystalized with the Fuc derivatives **30** and **28,** a shortening of the $C5^{Fuc}$-$O5^{Gal}$ distance by 0.10 to 0.13 Å (from 3.42/3.44 Å to 3.32/3.31 Å) was detected (30). Compared to 3.69 Å for the C-H bond (1.09 Å) and the van der Waals radii of $H5^{Fuc}$ and $O5^{Gal}$ on side and an optimal H-bond of 2.8 Å on the other side, this is a substantial effect (33).

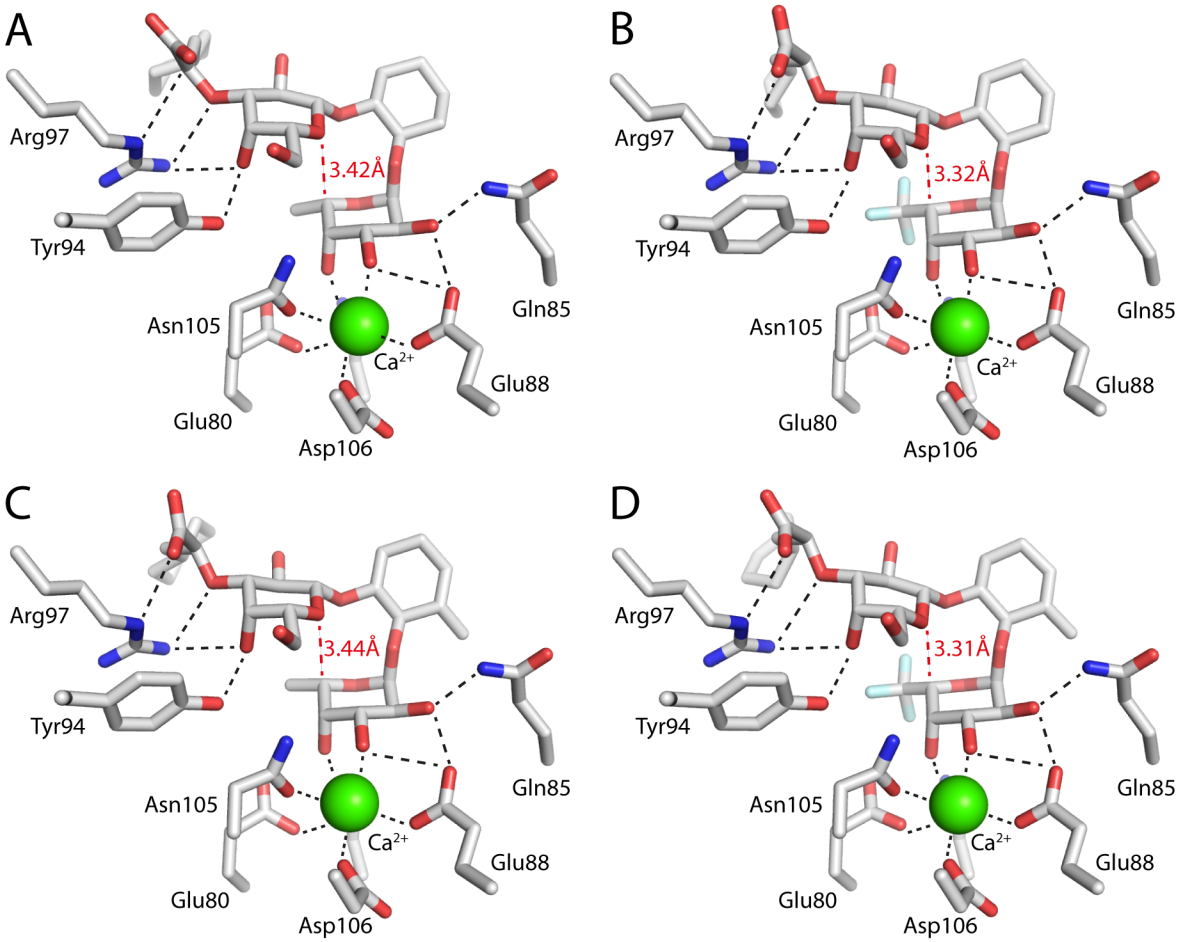


Figure 3 Close-up view of the ligand interaction in co-crystal structures of E-selectin with **28** (A), **6** (B), **30** (C), and **7** (D). All ligands establish similar interaction pattern with E-selectin (consisting of lectin domain, EGF-like domain and two consensus repeats) (30). Ligand interacting amino acids are labeled, the Ca^{2+} is shown as green sphere, oxygen and nitrogen atoms are colored red and blue, respectively. The distance between C5 of L-fucose and the O5 of D-galactose are indicated in red.

Acid stability assay. Natural polysaccharides containing acidic carbohydrates have been reported to be unstable in acidic conditions (34). To examine the stability of glycomimetics and their potential to surpass the gastric passage without degradation for a potential oral dosage form, stability assays in acidic conditions have been conducted (Table 4). The measured stability of all compounds was much higher than required for the gastric passage, since the experiments had to be conducted over a duration of 10 days to obtain visual degradation (35). Furthermore, the introduction of the electron-withdrawing CF_3 group leads to higher acidic stability of compound **3** and **6** compared to their corresponding nonfluorinated counterparts. The incorporation of the CF_3 group therefore not only increases the affinity of the glycomimetics, but additionally enhances the stability of the compounds due to the core stabilization.

Table 4. Stability of glycomimetics counterparts **3**, & **25** and **6** & **28** in 1 M HCl aqueous solution (pH approximately 0) and at physiological pH 7.4 (TRIS-HCl buffer 1.0 M) was measured in triplicate over a period of 10 days at room temperature (25 °C). Standard deviation is indicated and stabilities reported over 95% were considered as stable.

Entry	Ligand	R_1	Stability 1 M HCl	Stability pH 7.4
1	**3**	CF_3	> 95 %	> 95 %
2	**25**	CH_3	79.7 ± 0.1 %	> 95 %
3	**6**	CF_3	83.8 ± 1.4 %	> 95 %
4	**28**	CH_3	54.7 ± 1.9 %	> 95 %

Conclusions

Increasingly, non-conventional hydrogen bonds are considered to play a significant role in molecular recognition involved in the binding of small molecules to large ones. In addition, a number of investigations were focusing on the question whether these weak interactions are also conformationally influential (36).

References

1. Pauling L. The nature of the chemical bond and the structure of molecules and crystals; an introduction to modern structural chemistry: Cornell University Press; 1960 1960.
2. Stahl N, Jencks WP. Hydrogen bonding between solutes in aqueous solution. *Journal of the American Chemical Society*. 1986;*108(14)*:4196-205.
3. Takahashi O, Kohno Y, Nishio M. Relevance of weak hydrogen bonds in the conformation of organic compounds and bioconjugates: evidence from recent experimental data and high-level ab initio MO calculations. *Chem Rev*. 2010;*110(10)*:6049-76.
4. Bailey WF, Lambert KM, Stempel ZD, Wiberg KB, Mercado BQ. Controlling the Conformational Energy of a Phenyl Group by Tuning the Strength of a Nonclassical CH...O Hydrogen Bond: The Case of 5-Phenyl-1,3-dioxane. *J Org Chem*. 2016;*81(24)*:12116-27.
5. Rovira C, Novoa JJ. Strength and directionality of the C(sp3)H···S(sp3) interaction. An ab initio study using the H2S···CH4 model complex. *Chemical Physics Letters*. 1997;*279(3-4)*:140-50.
6. Desiraju GR, Gavezzotti A. From molecular to crystal structure; polynuclear aromatic hydrocarbons. *Journal of the Chemical Society, Chemical Communications*. 1989*(10)*.
7. Zierke M, Smiesko M, Rabbani S, Aeschbacher T, Cutting B, Allain FH, et al. Stabilization of branched oligosaccharides: Lewis(x) benefits from a nonconventional C-H...O hydrogen bond. *J Am Chem Soc*. 2013;*135(36)*:13464-72.
8. Lemieux RU, Koto S, Voisin D. The Exo-Anomeric Effect. Anomeric Effect. ACS Symposium Series 1979;17-29.
9. Schwizer D, Patton JT, Cutting B, Smiesko M, Wagner B, Kato A, et al. Pre-organization of the core structure of E-selectin antagonists. *Chemistry*. 2012;*18(5)*:1342-51.
10. Battistel MD, Azurmendi HF, Frank M, Freedberg DI. Uncovering Nonconventional and Conventional Hydrogen Bonds in Oligosaccharides through NMR Experiments and Molecular Modeling: Application to Sialyl Lewis-X. *J Am Chem Soc*. 2015;*137(42)*:13444-7.
11. Aeschbacher T, Zierke M, Smiesko M, Collot M, Mallet JM, Ernst B, et al. A Secondary Structural Element in a Wide Range of Fucosylated Glycoepitopes. *Chemistry*. 2017;*23(48)*:11598-610.
12. Ley K. The role of selectins in inflammation and disease. *Trends in Molecular Medicine*. 2003;*9(6)*:263-8.
13. Takada M, Nadeau KC, Shaw GD, Marquette KA, Tilney NL. The cytokine-adhesion molecule cascade in ischemia/reperfusion injury of the rat kidney. Inhibition by a soluble P-selectin ligand. *J Clin Invest*. 1997;*99(11)*:2682-90.
14. Koch AE, Burrows JC, Haines GK, Carlos TM, Harlan JM, Leibovich SJ. Immunolocalization of endothelial and leukocyte adhesion molecules in human rheumatoid and osteoarthritic synovial tissues. *Lab Invest*. 1991;*64(3)*:313-20.
15. Ernst B, Magnani JL. From carbohydrate leads to glycomimetic drugs. *Nat Rev Drug Discov*. 2009;*8(8)*:661-77.
16. Chang J, Patton JT, Sarkar A, Ernst B, Magnani JL, Frenette PS. GMI-1070, a novel pan-selectin antagonist, reverses acute vascular occlusions in sickle cell mice. *Blood*. 2010;*116(10)*:1779-86.
17. Bevilacqua MP, Nelson RM, Mannori G, Cecconi O. Endothelial-leukocyte adhesion molecules in human disease. *Annu Rev Med*. 1994;*45*:361-78.
18. Binder FP, Lemme K, Preston RC, Ernst B. Sialyl Lewis(x): a "pre-organized water oligomer"? *Angew Chem Int Ed Engl*. 2012;*51(29)*:7327-31.
19. Somers WS, Tang J, Shaw GD, Camphausen RT. Insights into the molecular basis of leukocyte tethering and rolling revealed by structures of P- and E-selectin bound to SLe(X) and PSGL-1. *Cell*. 2000;*103(3)*:467-79.
20. Kolb HC, Ernst B. Development of Tools for the Design of Selectin Antagonists. *Chemistry - A European Journal*. 1997;*3(10)*:1571-8.
21. Kolb HC, Ernst B. Recent progresses in the glycodrug area. *Pure and Applied Chemistry*. 1997;*69(9)*.
22. Nicolaou KC, Hummel CW, Iwabuchi Y. Total synthesis of sialyl dimeric Lex. *Journal of the American Chemical Society*. 1992;*114(8)*:3126-8.

23. Nicolaou KC, Hummel CW, Bockovich NJ, Wong CH. Stereocontrolled synthesis of sialyl Lex, the oligosaccharide binding ligand to ELAM-1 (sialyl =N-acetylneuramin). *Journal of the Chemical Society, Chemical Communications*. 1991*(13)*.
24. Bansal RC, Dean B, Hakomori S-i, Toyokuni T. Synthesis of trifluoromethyl analogue of L-fucose and 6-deoxy-D-altrose. *Journal of the Chemical Society, Chemical Communications*. 1991*(12)*.
25. Crich D, Vinogradova O. Synthesis and glycosylation of a series of 6-mono-, di-, and trifluoro S-phenyl 2,3,4-tri-O-benzyl-thiorhamnopyranosides. Effect of the fluorine substituents on glycosylation stereoselectivity. *J Am Chem Soc*. 2007;*129(38)*:11756-65.
26. Ernst B, Wagner B, Baisch G, Katopodis A, Winkler T, Öhrlein R. Substrate specificity of fucosyl transferase III: An efficient synthesis of sialyl Lewisx-, sialyl Lewisa-derivatives and mimetics thereof. *Canadian Journal of Chemistry*. 2000;*78(6)*:892-904.
27. Osborn HM, Brome VA, Harwood LM, Suthers WG. Regioselective C-3-O-acylation and O-methylation of 4,6-O-benzylidene-beta-D-gluco- and galactopyranosides displaying a range of anomeric substituents. *Carbohydr Res*. 2001;*332(2)*:157-66.
28. Thoma G, Kinzy W, Bruns C, Patton JT, Magnani JL, Banteli R. Synthesis and biological evaluation of a potent E-selectin antagonist. *J Med Chem*. 1999;*42(23)*:4909-13.
29. Jerabek-Willemsen M, Wienken CJ, Braun D, Baaske P, Duhr S. Molecular interaction studies using microscale thermophoresis. *Assay Drug Dev Technol*. 2011;*9(4)*:342-53.
30. Preston RC, Jakob RP, Binder FP, Sager CP, Ernst B, Maier T. E-selectin ligand complexes adopt an extended high-affinity conformation. *J Mol Cell Biol*. 2016;*8(1)*:62-72.
31. Thoma G, Banteli R, Jahnke W, Magnani JL, Patton JT. A Readily Available, Highly Potent E-Selectin Antagonist. *Angew Chem Int Ed Engl*. 2001;*40(19)*:3644-7.
32. Titz A, Patton J, Smiesko M, Radic Z, Schwardt O, Magnani JL, et al. Probing the carbohydrate recognition domain of E-selectin: the importance of the acid orientation in sLex mimetics. *Bioorg Med Chem*. 2010;*18(1)*:19-27.
33. Batsanov SS. *Inorganic Materials*. 2001;*37(9)*:871-85.
34. Smidsrød O, Haug A, Larsen B, von Hofsten B, Williams DH, Bunnenberg E, et al. The Influence of pH on the Rate of Hydrolysis of Acidic Polysaccharides. *Acta Chemica Scandinavica*. 1966;*20*:1026-34.
35. Hopkins A. The pattern of gastric emptying: a new view of old results. *J Physiol*. 1966;*182(1)*:144-9.
36. Zehavi U, Sharon N. Synthesis of methyl 2,4-diacetamido-2,4,6-trideoxy hexopyranosides. *The Journal of Organic Chemistry*. 1972;*37(13)*:2141-5.

3.10 Manuscript 4: Bridging oligosaccharides and glycomimetics with protein receptors

In this manuscript, the development and evaluation of an *in silico* tool to categorize the quality of hydrogen bonds in carbohydrate lectin interaction is described. This tool guides lead compound developments by reducing the polarity of carbohydrate mimetics without loosing affinity. The poor pharmacokinetic properties of carbohydrates are primarily due to their polar surface consisting of hydroxyl groups. The categorization by the *in silico* tool was evaluated *in vitro* in three different projects.

Contribution to the Project:

Philipp Dätwyler measured the physicochemical properties of E-selectin antagonists and contributed to the writing of the manuscript.

Christoph P. Sager, Pascal Zihlmann, Philipp Dätwyler, Tobias Mühlethaler, Xiaohua Jiang, Jonathan Cramer, Martin Smieško *, Beat Ernst *

Institute of Molecular Pharmacy, Department of Pharmaceutical Sciences,
University of Basel, Klingelbergstrasse 50, 4056 Basel, Switzerland

* Corresponding author

Bridging oligosaccharides and glycomimetics with protein receptors — Acca-Bruca, an *in-silico* tool to categorize carbohydrate–lectin hydrogen bond interactions

Christoph P. Sager, Pascal Zihlmann, Philipp Dätwyler,

Tobias Mühlethaler, Xiaohua Jiang, Jonathan Cramer, Martin Smieško *, Beat Ernst *

Institute of Molecular Pharmacy, Department of Pharmaceutical Sciences,

University of Basel, Klingelbergstrasse 50, 4056 Basel, Switzerland

Keywords: carbohydrate-lectin interactions, hydrogen bond, glycomimetics, drug development, pharmacokinetics

Abstract

The past failure of carbohydrate-based drug candidates are mostly attributed with low affinity but also poor pharmacokinetics. One reason for the low affinity of carbohydrates are the high desolvation costs, linked to the numerous hydroxyl groups of carbohydrates. Furthermore, common pharmacokinetic parameters including the polar surface area (PSA) as well as the number of hydrogen bond acceptors and donors, that are usually above drug-like values. To reduce the likeliness of failure, hydroxyl groups making no or only a minimal contribution to binding should be replaced. Therefore, we developed the software Acca-Bruca and correlated thermodynamic data for 37 monodeoxygenated ligands with the crystal structures of the parent compound. The data resulted in a linear-regression model with a good correlation ($R2 = 0.85$) that was further simplified by a categorization system with three classes. Finally, the model and the categorization system were integrated in Acca-Bruca. To assess the predicting capabilities of Acca-Bruca, we performed two prospective studies on carbohydrate–lectin interactions and considered one example retrospectively. In addition, we report on preliminary results on non-carbohydrate ligands, analyzed with Acca-Bruca.

Introduction

Carbohydrates play an important role as energy source, as structural element in the form of cellulose or chitin, and as key elements in diverse physiological processes such as cell recognition and cell adhesion. For these processes carbohydrates bind to lectins, exhibiting high selectivity because of the three-dimensional structure of the oligosaccharides and their combinatorial complexity (1-6).

Lectins involved in cell adhesion, such as selectins, siglecs, and galectins, emerged as drug targets in the early 1990's, only to be abandoned shortly afterwards and deemed undruggable (7-13). Past and present failures of carbohydrate-based drug candidates are mostly attributed to issues like poor pharmacokinetics (PK) and low affinity (14). Carbohydrates contain numerous hydroxyls and charged groups, leading to their inherent polarity leading to insufficient permeability and oral bioavailability (7, 15). Furthermore, desolvation costs of the polar surface of carbohydrates and their binding sites are high and have to be compensated by the energy gained from the carbohydrate–lectin interactions (16). Cabani *et al.* calculated the cost of desolvating an isolated hydroxyl group to be 25 kJ/mol (17). The overall cost is scaled down for vicinal groups, leading to a total desolvation cost of 34 kJ/mol for two adjacent hydroxyl groups. This downscale effect applies to monosaccharides with 3 to 4 hydroxyl groups in proximity to each other, though the costs will add up with every additional unit of an oligosaccharide. The list of orally bioavailable glycomimetic drugs is therefore short and exclusively derived from monosaccharides: Oseltamivir (Tamiflu®) for the prevention of influenza infections, miglitol (Glyset®) and sodium–glucose cotransporter 2 (SGLT2) inhibitors used in diabetes, miglustat (Zavesca®) to treat Gaucher's disease, topiramate (Topamax®) for the treatment of epilepsy, and nucleosidic reverse transcriptase inhibitors (NRTIs) to treat HIV patients (18-23).

To harness the therapeutic potential of carbohydrate-lectin interactions, it will be necessary to address the problems originating from the polar nature of carbohydrates. However, hydrogen bonds (H-bonds) formed by hydroxyl groups substantially contribute to specificity and affinity of carbohydrate-lectin interactions and basically cannot be removed without careful inspection (24). In order for a hydroxyl group to exert a beneficial influence on binding, its desolvation cost has to be lower than the energy gained by newly formed protein-ligand interactions. This is only achieved by (i) optimal H-bonding distance and geometry, (ii) multiple H-bonds formed

to/from a single hydroxyl group, and (iii) their buriedness since the dielectric constant ε is lower in buried cavities of the binding site (ε ≈ 5-10) than on protein surface (ε ≈ 20) or in bulk water (ε ≈ 80) (25). Therefore, a H-bond is energetically more valuable in buried protein cavities (26). Additionally, occluded H-bonds show slower exchange rates when shielded from bulk water due to a penalized transition state (27).

Early pharmacokinetic profiling in lead optimization became an integral part of modern drug design reducing preclinical research costs and improving time-to-market efficiency (28).To support carbohydrate-based drug design we developed Acca-Bruca, an open source program for analyzing crystal structures and subsequently categorizing H-bond interactions to assist medicinal chemists in their decision-making process leading from oligosaccharides to glycomimetics.

Results

Acca-Bruca descriptors. In order to build our H-bonds categorizing tool on a solid basis, we analyzed six carbohydrate–lectin interactions with published co-crystal structures (Table 1). Thermodynamic data for 37 monodeoxygenated ligands bidning to those proteins were available from published literature. Each data point corresponds to the contribution of a single hydroxyl group.

Acca-Bruca scoring algorithm relies on six descriptors available from co-crystal structures. The descriptors are defined as follows:

Atom buriedness as the number of protein heavy atoms within a sphere of a 8 Å radius around the ligand's heavy atoms.

Average ligand buriedness is defined as the sum of buriedness of all ligand heavy atoms divided by the number of ligand heavy atoms.

Geometry-corrected H-bond energy is calculated with the directional H-bond potential function from the Yeti force-field (25).

Number of interactions is counted from the highest possible sum of H-bond energies.

Charge type refers to the type of interaction, e.g. a normal H-bond is assigned value 0, a charge-assisted H-bond is assigned value 1, and a salt-bridge formed between two charged functional groups (ligand-protein) is denoted with value 2.

BFactDiff is calculated as the difference of B-factors of two interacting atoms. Atoms without a B-factor, for example in NMR solution structures, are treated as if there was no difference in the B-factor.

Table 1 Carbohydrate–lectin co-crystals with corresponding thermodynamic data for monodeoxygenated ligands.

Target	Data Points	PDB code	Resolution [Å]	References
FimH	3	4XO8	1.7	(29)
Arabinose Binding Protein	4	5ABP	1.8	(30, 31)
Lectin IV	7	1LED	2.0	(32, 33)
Concanavalin A	10	1CVN	2.3	(34, 35)
Artocarpin	3	1VBO	2.4	(36, 37)
DGL	10	1DGL	2.4	(38, 39)

Linear-regression model. A linear-regression model was built using Canvas 2.6 (40, 41) to relate the experimental differential thermodynamic fingerprint $\Delta\Delta G_{Exp}$ of the parent molecule and its monodeoxygenated-derivatives with the six descriptors (see above; atom buriedness, average ligand buriedness, H-bond geometry, number of interactions, charge type, and B-factor difference) calculated by Acca-Bruca. It was not possible to describe carbohydrates satisfactory by using only the atom buriedness, as they often bind on the lectin surfaces. Therefore, atoms that are deeply buried upon binding but are part of a rather solvent exposed ligand had to be penalized as described by Equation 1. Furthermore, the B-factor differences were applied as a penalty for the geometric H-bond energies, as dissimilar B-factors would have the most pronounced effect on the interaction geometry. The remaining three descriptors *Number of interactions*, *ChargeType*, and *Geometry-corrected H-bond energy* were used to build the model described by Equation 2. However, to avoid overestimation of an individual descriptor, three models relying on a combination of two descriptors (i. *Interactions* and *ChargeType*; ii. *Interactions* and *Geometry*; iii. *ChargeType* and *Geometry*) were built. The average of the three models has led to the following parameters for Equation 2: a = 2.43, b = 4.02, c = 1.45, and d = -0.44.

$$Buriedness = \frac{\sqrt{AtomBuriedness * AverageLigandBuriedness}}{100} \quad \text{Eq. 1}$$

$$Score = Buriedness * (a * Interactions + b * ChargeType + c * Geometry * BFactDiff) + d \quad \text{Eq. 2}$$

To facilitate the interpretation of the predicted ΔΔ*G* values from the linear-regression model, three classes of hydroxyl groups were defined. All hydroxyl groups that were substantial for binding and their removal would lead to an affinity loss by a factor of 40 or more, belong to *class I*. These hydroxyl groups should not be removed from a compound, unless they are related to toxicity or metabolic instability.

Class II includes the hydroxyl groups that were scored between a factor of 4 and 40 (3.4 kJ/mol and 9.1 kJ/mol, respectively) in terms of loss in affinity. Removing a class II hydroxyl group can improve PK properties of a carbohydrate mimetic, however the loss in affinity must be considered and weighed against the associated PK benefits.

Class III includes all groups that were considered worthwhile removing, as they could improve PK properties with only a small loss in affinity by a factor of 4 or less, or even a gain through reduced desolvation costs.

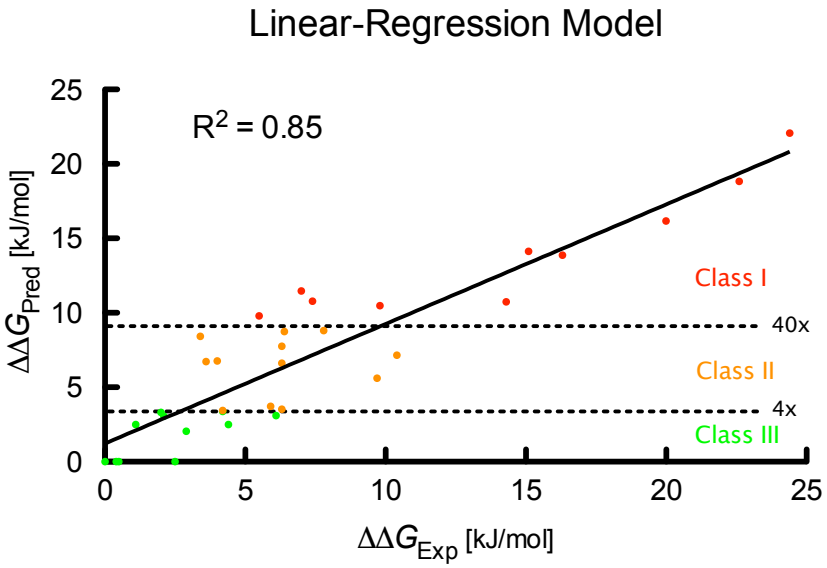


Figure 1 Linear-regression model. Data points in red are considered substantial and therefore belong to class I. Data points in orange belong to the intermediate class II, which can be removed to improve PK properties but will lead to a minor affinity loss. Data points in green belong to class III, which are dispensable. Dashed lines show a factor of 4 and 40 in change of affinity (3.4 kJ/mol and 9.1 kJ/mol, respectively). Goodness of fit, R^2 = 0.85.

Although the model was built with carbohydrate hydroxyl groups in mind, support for other interactions, e.g. salt bridges between arginine and carboxyl groups frequently occurring in sialylated carbohydrates, were implemented as well. The robustness of the resulting model was validated by splitting the whole data set ten times randomly into a training- and a test-set (3:1). The ten test-sets were predicted with an average Q^2 of 0.83, while the models average R^2 was 0.85 (Table S1). Furthermore, $\Delta\Delta G_{Exp}$ values were scrambled to show the model's sensitivity towards the experimental data. The scramble test did not yield a predictive model (R^2 = 0.04, Figure S1).

Naïve Bayes classification. To cross-validate our approach and anticipate potential over fitting of the linear-regression model we also built a conditional probability model. This statistical approach is not predicated on training the model beforehand. However, interpretation of data which has not been encountered before is difficult.

The probability model included all previous 37 data points and three Bayes classifiers, which span the experimental $\Delta\Delta G$ values and that describe the probability for a data point belonging to the classes I-III. Class I was defined as > 9.1 kJ/mol with 9 data points, class II as between 3.4 kJ/mol and 9.1 kJ/mol with 16 data points, and class III as < 3.4 kJ/mol including 12 data points. Four descriptors were included in this model; number of H-bonds, charge type, H-bond geometry, and buriedness calculated by the Equation 1. The model predicted 32 out of 37 data points correctly with two data points in class I, two data points in class II, and one data point in class III incorrectly classified either one class too high or too low. This was caused by the fact that the incorrectly predicted data points were untypical representatives for their class, for example slightly more buried than others. Therefore, the probability for this descriptor combination was lower for the correct class and higher for the incorrect class.

Acca-Bruca output. The classification generated by Acca-Bruca can be visualized in a molecule viewer able to color atoms by b-factor. The b-factor column in the pdb output is used to color the assigned atoms in red, yellow and green for the classes I-III, respectively. Unclassified atoms, such as carbon atoms are colored in blue (Figure 2). Furthermore, the classification is available as a tabular output within the program, referenced by the pdb atom and residue name from the used input structure.

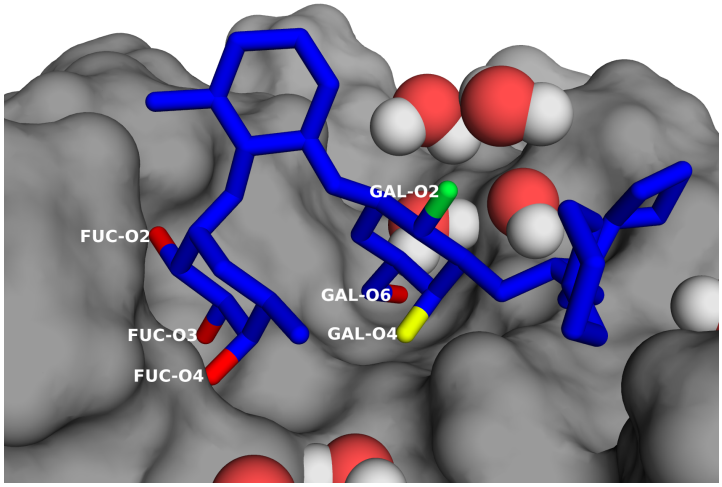


Figure 2. Exemplary output of compound **1A** produced by Acca-Bruca, visualized in PyMol. The ligand is colored by b-factor, whereas red, yellow and green atoms stand for the classes I-III, respectively. Unclassified atoms, such as carbon atoms are colored in blue. The protein is shown as gray surface, water molecules as spheres.

Prospective PK study – E-selectin antagonists. After successful phase II clinical trials of the pan-selectin antagonist Rivipansel for the treatment of vaso-occlusive crises in patients with sickle-cell disease, the interest to develop an orally bioavailable selectin antagonist arose (42). Several projects to improve pharmacokinetic properties were conducted, but so far no orally bioavailable selectin antagonist has been identified (43-46). E-, P-, and L-selectin are a family of Ca^{2+}-dependent C-type lectins mediating cell-cell adhesion that play a role in inflammation (47). E-selectin is expressed on the vascular endothelium. The minimal binding epitope is the tetrasaccharide sialyl Lewis^x (48), which was also the starting point for mimetic lead structures such as **1A** (Table 2, Figure 2).

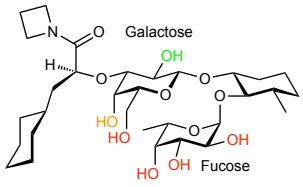
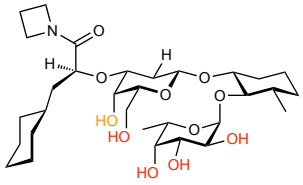
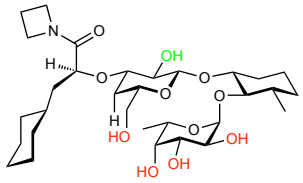
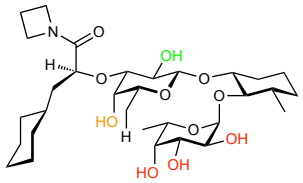
The large and hydrophilic core structure of **1A** demands a consequent reduction in size and polar surface area (PSA) to increase the drug-likeness of the antagonists. In addition, a recent study on E-selectin co-crystallized with sLe^x revealed a conformational change of E-selectin to a high-affinity state upon ligand binding (49). This is important, because the hydroxyl group classification directly depends on the input crystal structure, from which Acca-Bruca cannot predict such conformational changes. Since we assume that this high-affinity conformation is generally adopted upon crystallization with E-selectin antagonists, we docked lead compound **1A** to the high-affinity state co-crystal structure (1.9 Å, PDB ID: 4C16). The three hydroxyls of fucose interacting with the calcium ion and the negatively charged surrounding (Glu80, Glu88, Glu107) and the 6-hydroxyl group of galactose forming a hydrogen bond with Glu92 were

identified as important for the lectin-antagonist interaction and therefore assigned to class I. However, the 4-hydroxyl group of galactose forming a single interaction with Tyr48 was assigned to class II and the 2-hydroxyl group not involved in any direct interaction with the lectin to class III.

To test our prediction and quantify the influence of a single hydroxyl group on the octanol-water distribution-coefficient (log$D_{7.4}$), the polar surface area (PSA), and the passive permeability (logP_e) measured in a parallel artificial membrane permeation assay (PAMPA) (50), we synthesized a series of galactose mono deoxy-derivatives (**1B-1D**, Table 2, see also Schemes 1-3, supporting information). The experimental affinities shown in Table 2 confirm the hydroxyl group classification of Acca-Bruca. Compared to compound **1A (**2.8 µM**)**, compound **1B** (Gal-2H) exhibits a 3-fold (7.8 µM), compound **1C** (Gal-4H) a 20-fold (53.8 µM) and compound **1D** (Gal-6H) complete loss of affinity (1700 µM). The removal of one hydroxyl increases the log$D_{7.4}$ values by 0.4 (**1C**) to 0.5 (**1B** and **1D**). But more importantly, the PSA is reduced below the threshold value of 140 Å^2 recommended by Veber (51). These improvements of the physicochemical properties resulted in better passive permeation, even though all compounds are still below the accepted threshold for sufficient permeability (logP_e > -6.3) (52). Interestingly, the strongest effects regarding the observed permeability are shown by compound **1D** (Gal-6H), which has a reduced rotatable bond count compared to compounds **1B** and **1C** (Gal-2H and Gal-4H) and the smallest PSA among the tested compounds. However, the gain in permeability does not justify the loss of affinity that accompanies the removal of the hydroxyl group of compound **1D** (Gal-6H).

Compound **1B** (Gal-2H) exhibits a lower molecular weight and improved lipophilicity, PSA and permeability (logP_e) without a pronounced loss in affinity and is therefore a good starting point for further development of orally bioavailable E-selectin antagonists.

Table 2 E-selectin antagonist lead structure (**1A**) and its galactose mono deoxy-derivatives (**1B** - **1D**). Class I hydroxyl groups are shown in red, class II in orange, and class III in green.

Compound	Structure	Affinity [µM]	log$D_{7.4}$	PSA [Å^2]	PAMPA (logP_e)
1A		2.8	0.54 ±0.03	149	Below detection limit
1B Gal-2H		7.8	1.01 ±0.03	135	-8.52 ±0.04
1C Gal-4H		53.8	0.92 ±0.02	128	-8.42 ±0.02
1D Gal-6H		1700	1.06 ±0.02	122	-6.85 ±0.02

Prospective lead optimization – Siglec-8 ligands. Kroezen *et al.* will publish results of their lead optimization program regarding pulmonary deliverable Siglec-8 ligands for the treatment of asthma (Manuscript ready to publish). Siglec-8 is a CD33-related, human immune-inhibitory receptor that down-regulates immune response under inflammatory conditions. Already slightly hydrophobic molecules with a log*P* around 1 are absorbed in the lung within minutes and have a higher lung bioavailability compared to oral delivery (53).

Because the lead structure 6'sulfo-sLe^x (**2A,** Neu5Acα2–3[6S]Galβ1–4[Fucα1–3]Glc*N*Ac) is very polar (clog*P* = -9.7) we tried to improve PK properties early on in the lead optimization process without sacrificing affinity. Thus, 6'sulfo-sLe^x (**2A**) bound to human Siglec-8 (54), was analyzed by Acca-Bruca. Because the hydroxyl groups of fucose and *N*-acetyl-glucosamine of compound **2A** were all assigned to class III, we focused on the partial structure, disaccharide Neu5Ac-α2-3(6-O-sulfo)GalOMe (**2B**) and prepared deoxy-derivatives thereof. In a competitive binding assay IC_{50} values were determined and subsequently ligand efficiencies (LE) were calculated. LE

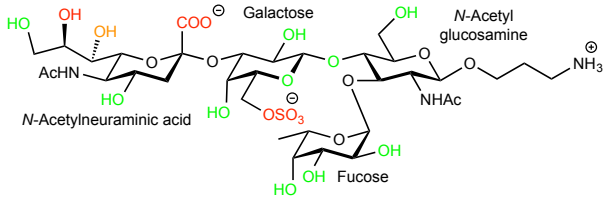
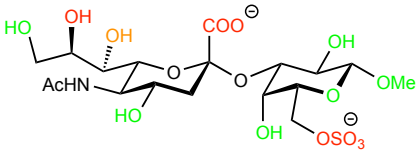
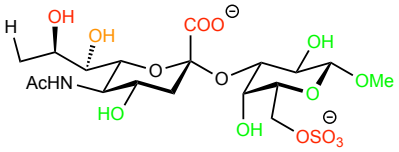
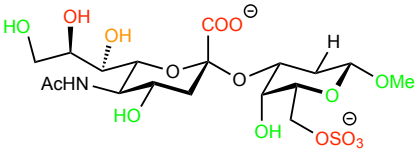
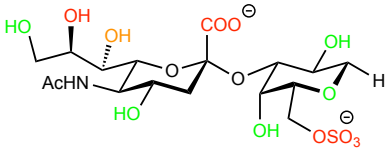
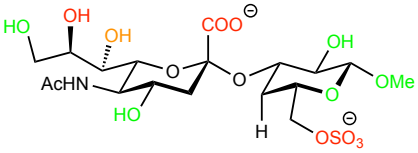
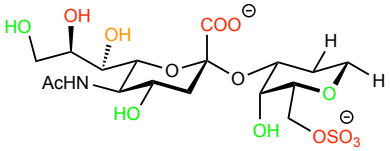
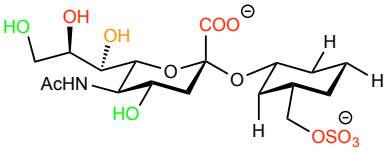
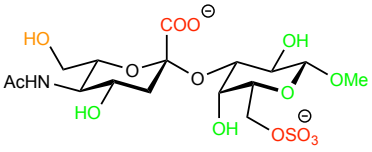
is defined as the binding affinity per ligand heavy atom and is routinely used in early stages of a lead optimization process to evaluate the effectiveness of structural modifications (55). To illustrate change in PK properties we also report the calculated octanol-water partition coefficient (clog*P*) and PSA in Table 3.

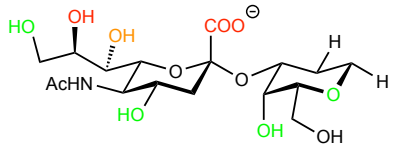
The IC_{50} of compound **2B** decreased by a factor of 2 compared to compound **2A**. However, compound **2B** shows a more than four order of magnitude increase of clog*P* accompanied by a decrease of the PSA by 70 $Å^2$, and an increase of LE by 0.04, overall improvements of PK parameters essential for lung bioavailability.

The hydroxyl group in position 9 of the sialic acid is exposed to solvent. Hence its removal ($\rightarrow$ **2C**, Sia-9H) leads to the largest change in clog*P* (+1.1) and PSA (-22 $Å^2$), while the affinity is preserved. Galactose modifications show less distinct changes in PSA because they are relatively small compared to the contribution of the sulfate group. The small improvement in affinity is most likely due to reduced desolvation costs of compounds **2E** and **2F** (Gal-1H and Gal-4H, both LE = 0.14). In contrast, compound **2D** (Gal-2H, LE = 0.12) does not show an affinity improvement, likely because the reduced desolvation costs are lost and a stabilizing intramolecular H-bond to the sialic acid ring oxygen is no longer possible, leading to higher desolvation costs. Additional improvements of LE, clog*P*, and PSA were achieved by combined galactose modifications. While the LE of the dideoxy-derivative **2G** (Gal-1,2H) is increased to 0.15, the LE of trideoxy **2H** (Gal-1,2,4H) additionally lacking the ring oxygen, is even further increased to 0.17, with a distinct increase in clog*P* to -2.5 and a decrease of PSA to 187 $Å^2$. Removing the hydroxyl group at C8 of sialic acid in compound **2I** as well as removing the sulfate group in compound **2J**, both classed as class I, renders the ligand inactive as predicted.

Drug-like, rule of five compliant molecules with affinities in the range of 10 nM, and an average number of 38 heavy atoms correspond to a LE of 0.29 (56). The observed improvement in LE of 0.04 from the tetrasaccharide **2A** (LE = 0.08) to the disaccharide **2B** (LE = 0.12) and additional improvement of 0.05 obtained for compound **2H** (Gal-1,2,4H, LE = 0.17) can therefore be regarded as a significant step forward in lead optimization. The improvements also emphasizes the importance of proper group analysis, as the removal of two hydroxyl groups, a methoxy group, and the ring oxygen, all assigned to class III, govern this change.

Table 3 Siglec-8 ligands analyzed by Acca-Bruca. Class I groups are shown in red, class II in orange, and class III in green.

Name	Structure	IC_{50} [µM]	LE	clog*P*	PSA
2A 6'sulfo-sLe^x		303	0.08	-9.7	308
2B		681	0.12	-5.2	238
2C Sia-9H		572	0.13	-4.1	216
2D Gal-2H		626	0.12	-4.3	228
2E Gal-1H		354	0.14	-5.4	240
2F Gal-4H		251	0.14	-4.5	234
2G Gal-1,2H		190	0.15	-4.6	216
2H Gal-1,2,4H		109	0.17	-2.5	187
2I Sia-8,9H		> 2000	-	-3.9	208

2J Gal-1,2H-6OH		> 2000	-	-4.5	176

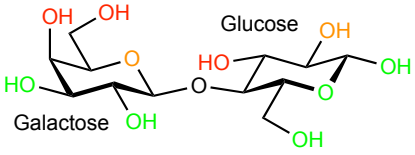
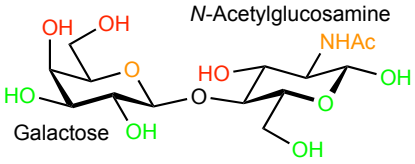
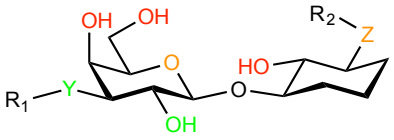
Retrospective lead optimization – Galectin-3 antagonists. Overexpression of human galectin-3 (hGal-3) is related to cancer drug resistance and plays an important role in various stages of cancer progression, but is also implicated in inflammatory processes and is therefore regarded as a valuable therapeutic target for the treatment of cancer and pharmacological modulation of inflammation (10). Human galectin-3 naturally binds to D-galactose and galactose-based disaccharides like lactose (**3A**, Lac) and *N*-acetyllactosamine (**3B**, Lac*N*Ac) in the milli- to micromolar range. A recent patent review by Blanchard *et al.* summarized the current strategies in industry and academia towards a selective hGal-3 inhibitor .[62] While most modifications were made on C3 of galactose to facilitate an arginine–arene interaction, Lund University patented a scaffold that includes a substituted cyclohexanol linked to the galactose as shown in compound **3C** (57).

Intrigued by this finding we classified the lactose (**3A**) hydroxyl groups using a high-resolution crystal structure (3ZSJ, 0.86 Å) of lactose binding to hGal-3 and compared them with the patented scaffold. Analyzing the crystal structure (1KJL, 1.4 Å) of Lac*N*Ac (**3B**) bound to hGal-3 yielded the same classification (58).

Interestingly, the removed groups on the cyclohexanol ring are in a good agreement with our predicted classification. The 1- and 6-hydroxyls of glucose as well as the ring oxygen were assigned to class III (**3A** and **3B**), which is reflected in the absence of these groups in the patented molecules.

To further increase the binding affinity of galectin-3 antagonists, Gal-3-hydroxyl classified as class III interaction was modified to bear an arene moiety (R_1 in **3C**) that can form a cation-π stacking interaction with Arg144 of hGal-3. Glc-2-hydroxyl, assigned to class II, was either modified to an amide, like in Lac*N*Ac (**3B**) or a triazole. This allowed to introduce an arene substituent (R_2 in **3C**) that can interact with Arg186, also by cation-π stacking. In this example, the two groups were not removed to improve PK properties, but modified to extend the binding mode towards other parts of the protein that offered new interactions and therefore higher affinity.

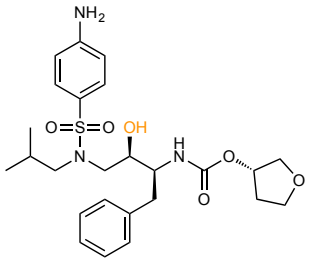
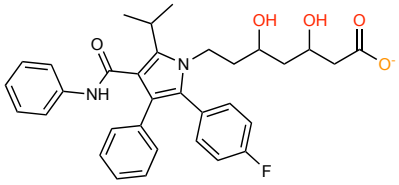
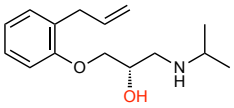
Table 4 Galectin-3 Antagonists analyzed by Acca-Bruca; class I hydroxyl groups are shown in red, class II in orange, and class III in green.

Name	Structure
3A Lactose	
3B *N*-Acteyllactosamine	
3C Forskarpatent I Syd Ab	

Acca-Bruca – scope. Although the linear-regression model was built for carbohydrates, the application of Acca-Bruca may not be limited to those only. Instead, a preliminary analysis on marketed drugs containing hydroxyl groups suggests that Acca-Bruca can also predict H-bonds which are not carbohydrate related. Orally bioavailable molecules therapeutically used, had to undergone rigorous cycles of optimization and their hydroxyl groups must be regarded as valuable, otherwise they would have been removed. Based on this premise, we analyzed molecules from three different drug classes containing at least one hydroxyl group: amprenavir (**4A**), atorvastatin (**4B**), and alprenolol (**4C**) (Table 5).

The analysis with Acca-Bruca categorized the hydroxyl group of amprenavir (**4A**) only as class II. Indeed, this particular hydroxyl group forms only one H-bond with Asp25. However, since it mimics the transition state recognized by asparte proteases, it is essential for the HIV protease inhibition. The hydroxyl groups of atorvastatin (**4B**) and alprenolol (**4C**), however, were assigned to class I. Both hydroxyl groups of atorvastatin (**4B**) participate in three good H-bonds with the charged amino acids (O3 with Arg590 and O5 with Asp690, Glu559, Lys691 and Asn755). The hydroxyl group of alprenolol (**4C**) forms two H-bonds, donating one to the charged Asp113 and accepting one from Asn312. The hydroxyl groups of **4B** and **4C** are in good agreement with our premise and shows that Acca-Bruca is sensible not only to carbohydrate hydroxyl groups.

Table 5. Marketed drugs from three different classes analyzed by Acca-Bruca. Class I hydroxyl groups are shown in red, class II in orange, and class III in green.

Name	Structure	PDB Code
4A Amprenavir		3S45
4B Atorvastatin		1HWK
4C Alprenolol		3NYA

Conclusions

Based on thermodynamic data from 37 monodeoxygenated ligands and six descriptors calculated from their respective co-crystal structures, a linear-regression model for H-bond scoring with a correlation of R^2 = 0.85 was built. Randomized training- and test-sets, a scramble-test confirming the sensitivity towards the selected descriptors, and a conditional probability model excluding the risk of potential over-fitting verified the validity of the linear-regression model. The resulting score for a specific hydroxyl group was assigned to class I (substantial), class II (intermediate) and class III (dispensable) to facilitate the read out. Finally, the linear-regression model and categorization criteria were implemented into Acca-Bruca for convenience.

Three carbohydrate-lectin interactions were analyzed regarding the potential for improving pharmacokinetic properties and for lead optimization.

With Acca-Bruca, the change in affinity associated with the removal of specific hydroxyl groups of E-selectin antagonist **1A** could be correctly assessed and correlates well with experimentally determined affinities. In addition, when hydroxyl groups were assigned to class III, their removal allows pharmacokinetic improvements, e.g. the removal of a single hydroxyl group leads to increased log$D_{7.4}$ and improved passive permeation, compared to the parent compound. Thus, antagonist **1B** exhibits improved PK properties with only a slight loss in affinity

and is therefore a good starting point for further development of orally bioavailable E-selectin antagonists.

Based on prediction by Acca-Bruca, the siglec-8 lead compound 6'sulfo-sLe^x (**2A**), a highly polar tetrasaccharide (clog*P* = -9.7), was optimized in regards to ligand efficiency (LE) and pharmacokinetic properties. The overall LE improvement of 0.09 is not only based on the reduction from the tetrasaccharide to the disaccharide (**2A** to **2B**), but also the removal of four polar atoms (**2B** → **2H**), most likely due to a reduced cost of desolvation.

Additionally, current strategies in human galectin-3 drug design could be rationalized using Acca-Bruca. A patented scaffold replacing the glucose unit with a cyclohexanol ring overlaps with our predicted assignment of hydroxyl groups.

Furthermore, Acca-Bruca was applied to three different non carbohydrate-based drugs containing at least one hydroxyl group. The hydroxyl groups of the two drugs atorvastatin (**4B**) and alprenolol (**4C**) were predicted as class I hydroxyls, whereas the hydroxyl group of amprenavir (**4A**) was predicted as class II. Although only classified as class II, the hydroxyl group of amprenavir (**4A**) is essential for the inhibition of HIV proteases, since it mimics the transition state of the enzymatic hydrolysis. Albeit intriguing, further evaluation is needed before using Acca-Bruca for non-carbohydrates.

The development of Acca-Bruca and the implementation of a classification system for hydroxyl groups of carbohydrates allows medicinal chemists to readily assess carbohydrate-lectin interactions. Additionally, the open source nature of the program allows for individual adaptations.

The potential to reduce the time spent on early lead optimization cycles is essential for lowering investment and risk of failure, both aspects relevant for academia and pharmaceutical industry alike. Finally, Acca-Bruca supports the challenge to develop oligosaccharides into oral carbohydrate-based drugs.

Methods

Crystal structure preparation and model building. Crystal structures with resolution above 2.0 Å were preprocessed by the Protein Preparation Wizard. The system was solvated with the TIP3P solvent model within an orthorhombic box and minimized using Desmond 4.4 with 2500 iterations and a convergence threshold of 0.05 kcal/mol/Å (59-61). Otherwise, default

parameters were used. After minimization, waters with less than 2 non-water interactions were discarded. Crystal structures with resolution below 2.0 Å were directly analyzed by Acca-Bruca.

NMR solution structure preparation. The structure of Siglec-8 (PDB accession code 2N7B) was preprocessed and minimized as described above. Waters with less than 2 non-water interactions were again discarded after minimization. Hydroxyl groups were classified by Acca-Bruca based on the described non-linear regression model.

Polar surface area calculation. Even simple oligosaccharides can have a rather complex 3D structure resulting in variable exposure of the polar functional groups to the surrounding environment. For their proper characterization in terms of the surface area available for interaction with the solvent molecules (SASA) and particularly the polar fraction of it (PSA), a specialized 3D-based algorithm was developed.

A 3D grid of points (density 0.25 Å) centered at the molecular centroid is created around the ligand. Molecular solvent accessible surface is constructed by analyzing overlaps of a solvent probe (radius of 1.4 Å) with the molecule (solute) at each of the grid points. The Van der Waals radii for atoms recommended by the Cambridge Crystal Data Bank are used (62). Next, the accessible surface is adjusted by subtracting the probe radius to obtain the effective Van der Waals surface. Each point of the final surface is assigned either to the polar (all O and N, polar H atoms, S atom in -SH or $-SO_2-$ groups) or the non-polar (remaining atoms) surface depending on the nearest atom.

In order to minimize the bias associated with grid placement, the final result is averaged from four distinct calculations using altered spatial orientation (original orientation plus three additional orientations obtained by x, y, z rotations). Using the steps and settings described above, the algorithm produces reliable PSA values (± 0.5 $Å^2$), which for simple conformational (3D) independent molecular species compare well with their standardized contributions (63).

Author Contributions

The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

Funding Sources

This work was supported by the Swiss National Science Foundation (grant number 146202)

Notes

The name Acca-Bruca originates from a swiss-italian pharmaceutical sciences student that made a statement about the importance of hydrogen bonds during a lecture. In German, H-bonds are often called H-Brücke and she phrased her sentence using the Italian phonetic sound for H /'akka/ and the word Bruca, which doesn't have a meaning in neither Italian, German nor English, but sounds very similar to Brücke. Everybody could follow her statement, although with a smile.

The Acca-Bruca source code is available under an open source license on GitHub.

Acknowledgements

‡ Dedicated to the memory of Prof. Dr. Angelo Vedani.

We thank Dr. Deniz Eris for comments on the manuscript and editorial assistance.

Abbreviations

ADMET, absorption distribution metabolism excretion and toxicity; LE, ligand efficiency; NRTI, non nucleosidic reverse transcriptase inhibitor; PAMPA, parallel artificial membrane permeation assay; PK, pharmacokinetic; PSA, polar surface area; SGLT2, sodium glucose transporter L2.

References

1. Sharon N, Lis H. Lectins as cell recognition molecules. *Science*. **1989**;*246(4927)*:227-34.
2. Sharon N, Lis H. Carbohydrates in cell recognition. *Sci Am*. **1993**;*268(1)*:82-9.
3. Cummings RD, Smith DF. The selectin family of carbohydrate-binding proteins: structure and importance of carbohydrate ligands for cell adhesion. *Bioessays*. **1992**;*14(12)*:849-56.
4. Varki A. Sialic acids as ligands in recognition phenomena. *FASEB J*. **1997**;*11(4)*:248-55.
5. Vestweber D. Ligand-specificity of the selectins. *J Cell Biochem*. **1996**;*61(4)*:585-91.
6. Williams SJ, Davies GJ. Protein--carbohydrate interactions: learning lessons from nature. *Trends Biotechnol*. **2001**;*19(9)*:356-62.
7. Ernst B, Magnani JL. From carbohydrate leads to glycomimetic drugs. *Nat Rev Drug Discov*. **2009**;*8(8)*:661-77.
8. Crocker PR, Paulson JC, Varki A. Siglecs and their roles in the immune system. *Nat Rev Immunol*. **2007**;*7(4)*:255-66.
9. Ingrassia L, Camby I, Lefranc F, Mathieu V, Nshimyumukiza P, Darro F, et al. Anti-galectin compounds as potential anti-cancer drugs. *Curr Med Chem*. **2006**;*13(29)*:3513-27.
10. Oberg CT, Leffler H, Nilsson UJ. Inhibition of galectins with small molecules. *Chimia (Aarau)*. **2011**;*65(1-2)*:18-23.
11. Mousa SA, Cheresh DA. Recent advances in cell adhesion molecules and extracellular matrix proteins: potential clinical implications. *Drug Discovery Today*. **1997**;*2(5)*:187-99.
12. Aretz J, Wamhoff EC, Hanske J, Heymann D, Rademacher C. Computational and experimental prediction of human C-type lectin receptor druggability. *Front Immunol*. **2014**;*5*:323.
13. Hopkins AL, Groom CR. The druggable genome. *Nat Rev Drug Discov*. **2002**;*1(9)*:727-30.
14. Meanwell NA. Improving drug candidates by design: a focus on physicochemical properties as a means of improving compound disposition and safety. *Chem Res Toxicol*. **2011**;*24(9)*:1420-56.
15. Lipinski CA. Drug-like properties and the causes of poor solubility and poor permeability. *J Pharmacol Toxicol Methods*. **2000**;*44(1)*:235-49.
16. Bryce RA, Hillier IH, Naismith JH. Carbohydrate-protein recognition: molecular dynamics simulations and free energy analysis of oligosaccharide binding to concanavalin A. *Biophys J*. **2001**;*81(3)*:1373-88.
17. Cabani S, Gianni P, Mollica V, Lepori L. Group contributions to the thermodynamic properties of non-ionic organic solutes in dilute aqueous solution. *Journal of Solution Chemistry*. **1981**;*10(8)*:563-95.
18. Kim CU, Lew W, Williams MA, Wu H, Zhang L, Chen X, et al. Structure-activity relationship studies of novel carbocyclic influenza neuraminidase inhibitors. *J Med Chem*. **1998**;*41(14)*:2451-60.
19. Campbell LK, Baker DE, Campbell RK. Miglitol: assessment of its role in the treatment of patients with diabetes mellitus. *Ann Pharmacother*. **2000**;*34(11)*:1291-301.
20. Tahrani AA, Barnett AH, Bailey CJ. SGLT inhibitors in management of diabetes. *The Lancet Diabetes & Endocrinology*. **2013**;*1(2)*:140-51.
21. Weinreb NJ, Barranger JA, Charrow J, Grabowski GA, Mankin HJ, Mistry P. Guidance on the use of miglustat for treating patients with type 1 Gaucher disease. *Am J Hematol*. **2005**;*80(3)*:223-9.
22. Maryanoff BE, Nortey SO, Gardocki JF, Shank RP, Dodgson SP. Anticonvulsant O-alkyl sulfamates. 2,3:4,5-Bis-O-(1-methylethylidene)-beta-D-fructopyranose sulfamate and related compounds. *J Med Chem*. **1987**;*30(5)*:880-7.
23. Cihlar T, Ray AS. Nucleoside and nucleotide HIV reverse transcriptase inhibitors: 25 years after zidovudine. *Antiviral Res*. **2010**;*85(1)*:39-58.
24. Steiner T. The hydrogen bond in the solid state. *Angew Chem Int Ed Engl*. **2002**;*41(1)*:49-76.

25. Vedani A, Dunitz JD. Lone-pair directionality in hydrogen-bond potential functions for molecular mechanics calculations: the inhibition of human carbonic anhydrase II by sulfonamides. *Journal of the American Chemical Society*. **1985**;*107(25)*:7653-8.
26. Fitch CA, Karp DA, Lee KK, Stites WE, Lattman EE, Garcia-Moreno EB. Experimental pK(a) values of buried residues: analysis with continuum methods and role of water penetration. *Biophys J*. **2002**;*82(6)*:3289-304.
27. Schmidtke P, Luque FJ, Murray JB, Barril X. Shielded hydrogen bonds as structural determinants of binding kinetics: application in drug design. *J Am Chem Soc*. **2011**;*133(46)*:18903-10.
28. Schwardt O, Kolb H, Ernst B. Drug discovery today. *Curr Top Med Chem*. **2003**;*3(1)*:1-9.
29. Sauer MM, Jakob RP, Eras J, Baday S, Eris D, Navarra G, et al. Catch-bond mechanism of the bacterial adhesin FimH. *Nat Commun*. **2016**;*7*:10738.
30. Quiocho FA, Wilson DK, Vyas NK. Substrate specificity and affinity of a protein modulated by bound water molecules. *Nature*. **1989**;*340(6232)*:404-7.
31. Daranas AH, Shimizu H, Homans SW. Thermodynamics of binding of D-galactose and deoxy derivatives thereof to the L-arabinose-binding protein. *J Am Chem Soc*. **2004**;*126(38)*:11870-6.
32. Delbaere LT, Vandonselaar M, Prasad L, Quail JW, Wilson KS, Dauter Z. Structures of the lectin IV of Griffonia simplicifolia and its complex with the Lewis b human blood group determinant at 2.0 A resolution. *J Mol Biol*. **1993**;*230(3)*:950-65.
33. Lemieux RU. How Water Provides the Impetus for Molecular Recognition in Aqueous Solution. *Accounts of Chemical Research*. **1996**;*29(8)*:373-80.
34. Naismith JH, Field RA. Structural basis of trimannoside recognition by concanavalin A. *J Biol Chem*. **1996**;*271(2)*:972-6.
35. Gupta D, Dam TK, Oscarson S, Brewer CF. Thermodynamics of lectin-carbohydrate interactions. Binding of the core trimannoside of asparagine-linked carbohydrates and deoxy analogs to concanavalin A. *J Biol Chem*. **1997**;*272(10)*:6388-92.
36. Jeyaprakash AA, Srivastav A, Surolia A, Vijayan M. Structural basis for the carbohydrate specificities of artocarpin: variation in the length of a loop as a strategy for generating ligand specificity. *J Mol Biol*. **2004**;*338(4)*:757-70.
37. Rani PG, Bachhawat K, Reddy GB, Oscarson S, Surolia A. Isothermal titration calorimetric studies on the binding of deoxytrimannoside derivatives with artocarpin: implications for a deep-seated combining site in lectins. *Biochemistry*. **2000**;*39(35)*:10755-60.
38. Rozwarski DA, Swami BM, Brewer CF, Sacchettini JC. Crystal Structure of the Lectin fromDioclea grandifloraComplexed with Core Trimannoside of Asparagine-linked Carbohydrates. *Journal of Biological Chemistry*. **1998**;*273(49)*:32818-25.
39. Dam TK, Oscarson S, Brewer CF. Thermodynamics of binding of the core trimannoside of asparagine-linked carbohydrates and deoxy analogs to Dioclea grandiflora lectin. *J Biol Chem*. **1998**;*273(49)*:32812-7.
40. Sastry M, Lowrie JF, Dixon SL, Sherman W. Large-scale systematic analysis of 2D fingerprint methods and parameters to improve virtual screening enrichments. *J Chem Inf Model*. **2010**;*50(5)*:771-84.
41. Duan J, Dixon SL, Lowrie JF, Sherman W. Analysis and comparison of 2D fingerprints: insights into database screening performance using eight fingerprint methods. *J Mol Graph Model*. **2010**;*29(2)*:157-70.
42. Telen MJ, Wun T, McCavit TL, De Castro LM, Krishnamurti L, Lanzkron S, et al. Randomized phase 2 study of GMI-1070 in SCD: reduction in time to resolution of vaso-occlusive events and decreased opioid use. *Blood*. **2015**;*125(17)*:2656-64.
43. Pedatella S, De Nisco M, Ernst B, Guaragna A, Wagner B, Woods RJ, et al. New sialyl Lewis(x) mimic containing an alpha-substituted beta(3)-amino acid spacer. *Carbohydr Res*. **2008**;*343(1)*:31-8.
44. Titz A, Patton J, Smiesko M, Radic Z, Schwardt O, Magnani JL, et al. Probing the carbohydrate recognition domain of E-selectin: the importance of the acid orientation in sLex mimetics. *Bioorg Med Chem*. **2010**;*18(1)*:19-27.

45. Binder FP, Lemme K, Preston RC, Ernst B. Sialyl Lewis(x): a "pre-organized water oligomer"? *Angew Chem Int Ed Engl*. **2012**;*51(29)*:7327-31.
46. Schwizer D, Patton JT, Cutting B, Smiesko M, Wagner B, Kato A, et al. Pre-organization of the core structure of E-selectin antagonists. *Chemistry*. **2012**;*18(5)*:1342-51.
47. Lasky LA. Selectins - Interpreters of Cell-Specific Carbohydrate Information during Inflammation. *Science*. **1992**;*258(5084)*:964-9.
48. Foxall C, Watson SR, Dowbenko D, Fennie C, Lasky LA, Kiso M, et al. The three members of the selectin receptor family recognize a common carbohydrate epitope, the sialyl Lewis(x) oligosaccharide. *J Cell Biol*. **1992**;*117(4)*:895-902.
49. Preston RC, Jakob RP, Binder FP, Sager CP, Ernst B, Maier T. E-selectin ligand complexes adopt an extended high-affinity conformation. *J Mol Cell Biol*. **2016**;*8(1)*:62-72.
50. Kansy M, Senner F, Gubernator K. Physicochemical high throughput screening: parallel artificial membrane permeation assay in the description of passive absorption processes. *J Med Chem*. **1998**;*41(7)*:1007-10.
51. Veber DF, Johnson SR, Cheng HY, Smith BR, Ward KW, Kopple KD. Molecular properties that influence the oral bioavailability of drug candidates. *J Med Chem*. **2002**;*45(12)*:2615-23.
52. Avdeef A, Bendels S, Di L, Faller B, Kansy M, Sugano K, et al. PAMPA--critical factors for better predictions of absorption. *J Pharm Sci*. **2007**;*96(11)*:2893-909.
53. Patton JS, Byron PR. Inhaling medicines: delivering drugs to the body through the lungs. *Nat Rev Drug Discov*. **2007**;*6(1)*:67-74.
54. Propster JM, Yang F, Rabbani S, Ernst B, Allain FH, Schubert M. Structural basis for sulfation-dependent self-glycan recognition by the human immune-inhibitory receptor Siglec-8. *Proc Natl Acad Sci U S A*. **2016**;*113(29)*:E4170-9.
55. Perola E. An analysis of the binding efficiencies of drugs and their leads in successful drug discovery programs. *J Med Chem*. **2010**;*53(7)*:2986-97.
56. Schultes S, Graaf d, Krämer. Ligand efficiency as a guide in fragment hit selection and optimization. *Drug Discov Today Technol*. **2010**;*7(3)*:e147-202.
57. Leffler HN, U. J.; Wachenfeldt, H. V, Novel galactoside inhibitors of galectins. **2012**.Patent
58. Sorme P, Arnoux P, Kahl-Knutsson B, Leffler H, Rini JM, Nilsson UJ. Structural and thermodynamic studies on cation-Pi interactions in lectin-ligand complexes: high-affinity galectin-3 inhibitors through fine-tuning of an arginine-arene interaction. *J Am Chem Soc*. **2005**;*127(6)*:1737-43.
59. Sastry GM, Adzhigirey M, Day T, Annabhimoju R, Sherman W. Protein and ligand preparation: parameters, protocols, and influence on virtual screening enrichments. *J Comput Aided Mol Des*. **2013**;*27(3)*:221-34.
60. Shivakumar D, Williams J, Wu Y, Damm W, Shelley J, Sherman W. Prediction of Absolute Solvation Free Energies using Molecular Dynamics Free Energy Perturbation and the OPLS Force Field. *J Chem Theory Comput*. **2010**;*6(5)*:1509-19.
61. Guo Z, Mohanty U, Noehre J, Sawyer TK, Sherman W, Krilov G. Probing the alpha-helical structural stability of stapled p53 peptides: molecular dynamics simulations and analysis. *Chem Biol Drug Des*. **2010**;*75(4)*:348-59.
62. Groom CR, Bruno IJ, Lightfoot MP, Ward SC. The Cambridge Structural Database. *Acta Crystallogr B Struct Sci Cryst Eng Mater*. **2016**;*72(Pt 2)*:171-9.
63. Ertl P, Rohde B, Selzer P. Fast calculation of molecular polar surface area as a sum of fragment-based contributions and its application to the prediction of drug transport properties. *J Med Chem*. **2000**;*43(20)*:3714-7.

3.11 Manuscript 5: Towards a nanomolar E-selectin antagonist

In this manuscript, the stepwise approach towards highly affine E-selectin antagonists is described. By replacing structural group from the natural ligand sLe^x, the lead development process improved the affinity by pre-organization, orientation of functional groups, desolvation, and rigidification without having an effect on the entropic properties. Furthermore, lower molecular weight, reduced numbers of hydrogen bond acceptors and donors, and a more hydrophobic character increased the drug-likeness of the final E-selectin antagonist.

Contribution to the project:

Philipp Dätwyler was measuring and evaluating in the physicochemical and pharmacokinetic properties of compound **9**. He further contributed to the writing and arrangement of the publication.

Tobias Mühlethaler,‡ Pascal Zihlmann,‡ Norbert Varga, Roman Peter Jakob, Martin Smieško, **Philipp Dätwyler**, Bea Wagner, Xiaohua Jiang, Timm Maier and Beat Ernst*

‡ Author Contributions*:* These authors contributed equally to this work.

*Corresponding author.

Towards a Nanomolar E-selectin Antagonist: Structures, Thermodynamics and Group Contributions

Tobias Mühlethaler,[#‡] Pascal Zihlmann,[#‡] Norbert Varga,[#] Roman Peter Jakob,[†] Martin Smieško,[#] Philipp Dätwyler,[#] Bea Wagner,[#] Xiaohua Jiang,[#] Timm Maier[†] and Beat Ernst[#*]

‡ Author Contributions*:* These authors contributed equally to this work.

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

[†] University of Basel, Institute of Structural Biology,
Klingelbergstr. 70, 4056 Basel, Switzerland

*Corresponding author.

E-selectin is a cell-adhesion lectin expressed on the surface of vascular endothelial cells involved in the recruitment of leukocytes to the site of inflammation. By interacting with the tetrasaccharide sialyl Lewis^x ([Neu5Ac(α2-3)Galβ(1-4)[Fucα(1-3)]GlcNAc, sLe^x, **1**) in the glycan of glycoproteins on leukocytes, E-selectins mediate the initial contact and enable leukocyte-rolling along the vessel endothelium (1). This process is crucial to defend the body against infections, but excessive extravasation of leukocytes can harm the tissue and is associated with a broad variety of diseases, such as atherosclerosis (2), rheumatoid arthritis (3), and asthma bronchiale (4). Furthermore, selectins are associated with sickle cell disease (5) and carcinoma cell metastasis (6, 7). Hence, blocking this interaction is a promising strategy to suppress an inflammation at the beginning of the cascade (8-10).

To develop a potent E-selectin antagonist we are optimizing ligands based on the sLe^x epitope. This approach culminated in the intravenously applied glycomimetic drug rivipansel (Glycomimetics, Inc.) that is currently in clinical phase III for the treatment of an acute sickle cell crisis (11, 12). In our previous work (13), we discussed the thermodynamics of glycomimetic E-selectin antagonists (compounds **1**-**4**). We showed that binding of sLe^x to E-selectin is entropy driven arising from the release of bound water molecules from the binding interface to bulk water and a well pre-organized ligand conformation in solution.

In sLe^x, the D-Glc*N*Ac moiety interacts only weakly through van der Waals interactions with the protein and acts rather as a spacer between L-Fuc and D-Gal to stabilize the bioactive conformation of sLe^x. Its replacement by the more rigid and hydrophobic (1*R*,2*R*,3*S*)-3-methylcyclohexane-1,2-diol stabilized the core conformation of compounds **3** and **4** and improved the affinity 15-18 fold. Core-preorganization turned out to improve both, enthalpy and entropy (**1**→**3**; $\Delta\Delta H°$ = -4.0 kJ mol^{-1}; $-T\Delta\Delta S°$ = -3.2 kJ mol^{-1}), due to a reduced conformational entropy loss upon ligand binding, a more stable interaction with E-selectin and reduced desolvation costs of the hydrophobic mimic of D-GlcNAc. The replacement of D-Neu5Ac by a smaller, hydrophobic and unsubstituted (*S*)-cyclohexyl lactic acid further improved affinity by a factor of 2-3, as seen in compounds **2** and **4**. The carboxylic acid forming a salt bridge to Arg97 gained more flexibility allowing for an optimal orientation resulting in an enthalpic gain (**1**→**2**; $\Delta\Delta H°$ = -7.2 kJ mol^{-1}), largely compensated by an entropic loss (**1**→**2**; $-T\Delta\Delta S°$ = +4.5 kJ mol^{-1}). Progress in the development of E-selectin antagonists, new analysis methods allowing for global analyses of multiple ITC experiments (14, 15), and co-crystal structures of E-selectin with sLe^x and glycomimetics antagonists (Figure 1A) (16) require the

resumption of this important topic.

Surprisingly, the hydroxy group in 2-position of L-Fuc was not found to be involved in a direct interaction with the E-selectin in a soaked crystal structure with sLe^x (PDB-code: 1G1T) from Somers *et al.* (2000) (17), but it was found to be crucial for binding in structure-activity relationship studies (18, 19). This discrepancy could only recently be explained when Preston *et al.* (2015) published a co-crystal structure of E-selectin in complex with sLe^x (PDB-code: 4CSY) and compound **4** (PDB-code: 4C16) (16), disclosing substantial conformational changes at the binding site compared to the soaked crystal structure (Figure 1B). Loop composed of residues 81-89 in the binding site moves up to 10 Å towards the binding site. Glu88 replaces Asn83, coordinates with Ca^{2+} and forms a hydrogen bond with 2-OH of L-Fuc. Additionally, Gln85 and Glu107 shift into hydrogen bonding distance of 2-OH and 3-OH of L-Fuc, respectively. The changes are not only limited to the binding site, but the relative position between the lectin domain and the EGF-like domain is significantly altered. This induced fit is part of a catch-bond mechanism, which was demonstrated for P- and L-selectin as well (20, 21).

To understand the impact of protein-ligand interactions, we have determined seven E-selectin complex structures (Table S1) ranging from 778 µM to 0.34 µM (Table 1). In general, all ligands bind in a very similar conformation to E-selectin and establish the same protein-ligand interactions (Figure 1A). Thus, the differences in the binding affinity between the ligands does not derive from the formation of new or optimized protein-ligand interactions but rather from a change in pre-organization of the ligand in solution and optimized desolvation. We found that contributions of modifications of E-selectin antagonists with a stabilized core conformation (D-Glc*N*Ac or its rigid mimetic (1*R*,2*R*,3*S*)-3-methylcyclohexane-1,2-diol) are additive instead of cooperative (Figure 2, Table 1). In a previous publication we found modifications of E-selectin antagonists containing a (1*R*,2*R*)-cyclohexane-1,2-diol, a more flexible mimetic of D-Glc*N*Ac, to be non-additive (13). Although the concept of additivity is very popular in medicinal chemistry, the separation of energy into individual group contributions is often not applicable. According to Williamson *et al.* the additive approach is not valid, if non-covalent interactions are either mutually reinforcing (positively cooperative) or mutually weakening (negatively cooperative) (22, 23). Gerhard Klebe and co-workers demonstrated that the local hydration pattern of a binding interaction may result in non-additivity (24, 25). However, the high degree of pre-organization of core-stabilized E-selectin antagonists and their identical binding mode allowed for a deconvolution of the binding affinity into individual contributions of distinct functional

groups. In this publication, we present the most successful modifications of each carbohydrate subunit of sLe^x (compound **1**) and their respective benefits for the binding interaction with E-selectin.

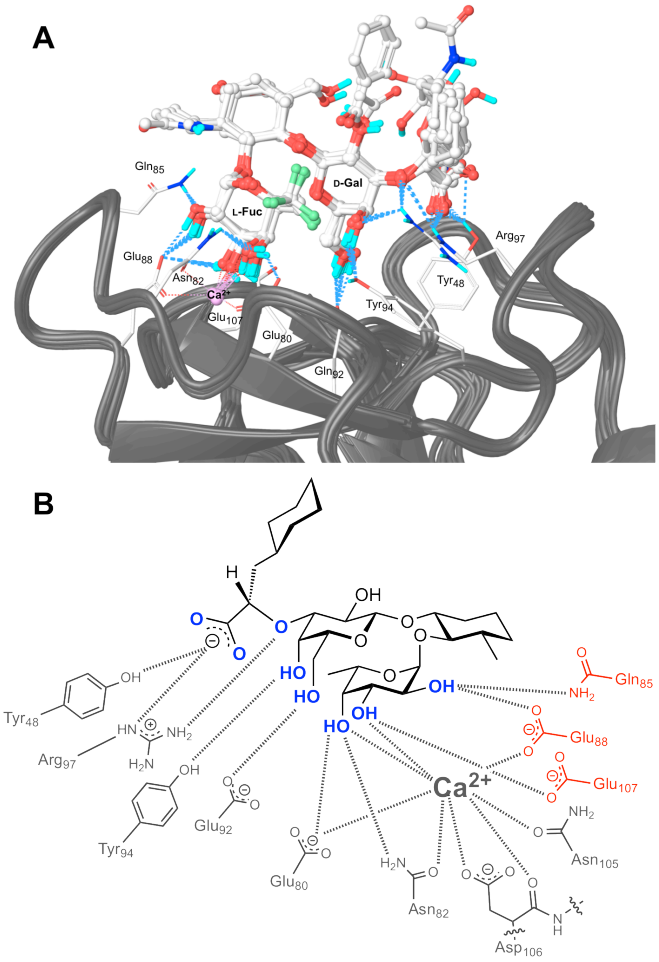


Figure 11 **A** Close-up view of the co-crystal structures of E-selectin with sLe^x (**1**) and E-selectin with its glycomimetics **2**-**9**. All ligands form similar interactions with the protein. **B** Detailed representation of the interaction between glycomimetic compound **4** (PDB: 4C16) and E-selectin as observed in the crystal structure. Amino acids forming an induced fit upon ligand binding are shown in red. Pharmacophores are highlighted in blue.

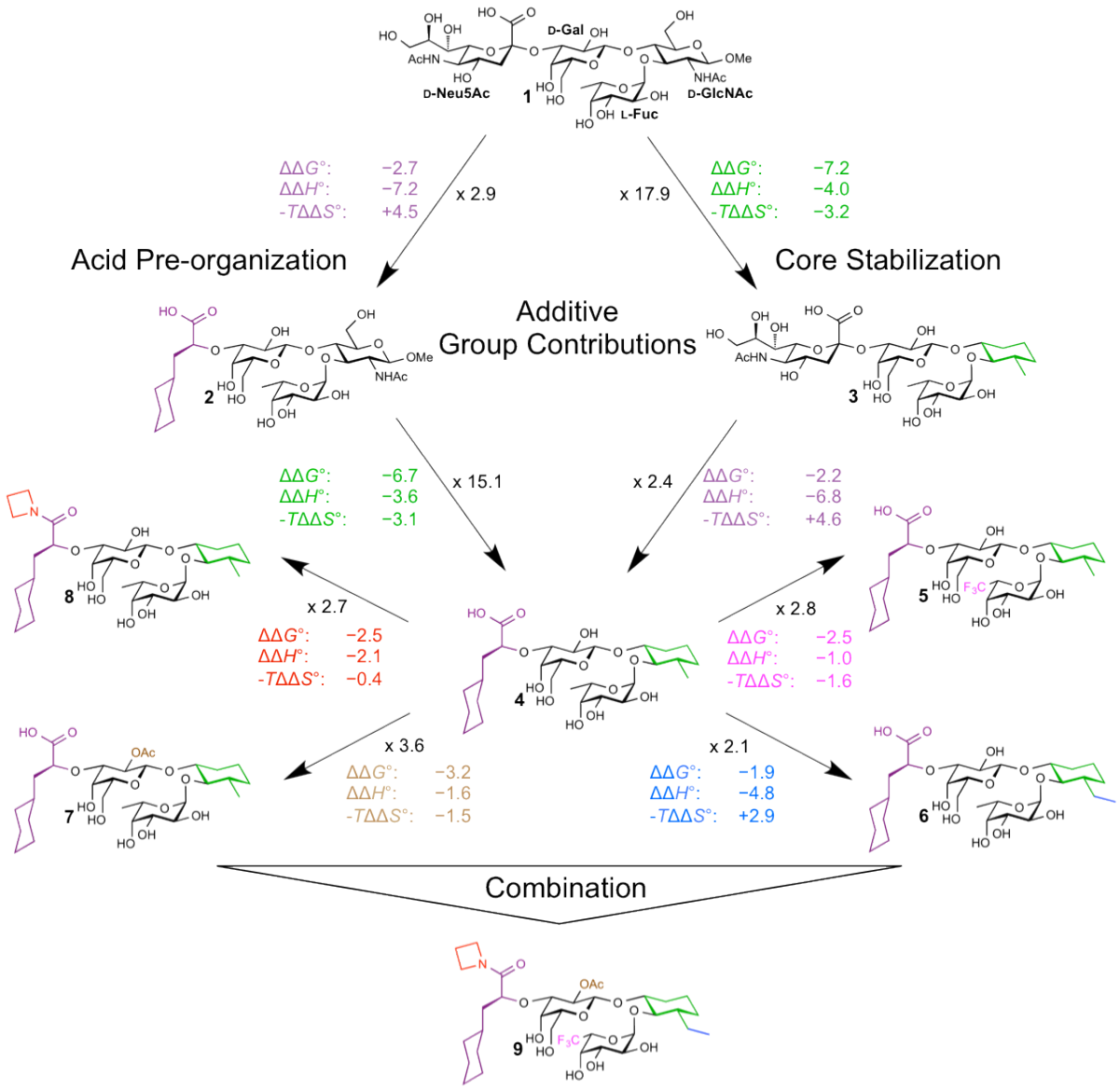


Figure 2 Thermodynamic contributions of individual groups in E-selectin antagonists. Group contribution of (*S*)-cyclohexyl lactic acid (purple); (*R*,*R*,)-3-methylcyclohexane-1,2-diol (green); acetylation in 2'-position of galactose (brown); ethyl substitution in 3'-position of (*R*,*R*,)-3-cyclohexane-1,2-diol (blue); trifluormethyl substitution in 6'-position of fucose (magenta) and an azetidine amide substitution of the carboxylic acid (red). The black factors indicate the affinity improvement of a given ligand modification. The differences in the thermodynamic properties are given as ΔΔ-values in kJ mol^{-1}.

Table 5 Thermodynamic parameters of E-selectin binding to sLe^x (**1**) and glycomimetic antagonists (**2**-**9**) determined by isothermal titration calorimetry (ITC). All measurements were carried out at 298.15 K. Affinities measured with microscale thermophoresis (MST) are shown for comparison. Confidence intervals and stoichiometries (*N*-values) are listed in Table S2.

Cmpd	MST: K_D [µM]	ITC: K_D [µM]	$\Delta G°$ [kJ mol^{-1}]	$\Delta H°$ [kJ mol^{-1}]	$-T\Delta S°$ [kJ mol^{-1}]
1	739	778	-17.7	4.9	-22.6
2	269	269	-20.4	-2.3	-18.1
3	-	43.4	-24.9	0.9	-25.8
4	13.7	17.8	-27.1	-5.9	-21.2
5	5.25	6.39	-29.6	-6.9	-22.8
6	9.53	8.36	-29.0	-10.7	-18.3
7	3.41	4.96	-30.3	-7.5	-22.7
8	2.82	6.56	-29.6	-8.0	-21.6
9	0.339	0.282 (0.310*)	-37.4 -(-37.2*)	-15.1 (-15.4*)	-22.3 (-21.8*)

*The values in brackets of **9** are calculated from the single group contributions of compounds **5**-**8** and demonstrate the additivity of the individual modifications on the rigid sLe^x core structure

*L-Fuc modification for core pre-organization (compound **5**).* While all hydroxyl groups of L-Fuc form essential interactions with the protein and/or the structural Ca^{2+} ion in the E-selectin binding site, a nonconventional intramolecular C-H···O hydrogen bond between H-C(5) of L-Fuc and O(5) of D-Gal helps by locking the core in a bioactive conformation (26). The substitution of the methyl group (C6) by a strongly electron withdrawing trifluormethyl group (**4**→**5**) strongly polarizes the adjacent hydrogen at 5-position of L-Fuc and thereby improves the core pre-organization. This leads to decreased entropic costs (**4**→**5**; $-T\Delta\Delta S°$ = -1.6 kJ mol^{-1}) upon binding as well as an enthalpic gain (**4**→**5**; $\Delta\Delta H°$ = -1.0 kJ mol^{-1}), most likely due to improved protein-ligand interactions (Table 1). Enthalpy is additionally favoured by decreased desolvation costs of a trifluormethyl group compared to the original methyl group.

Furthermore, higher hydrophobicity and size of the trifluormethyl group might also better shield the Ca^{2+} cavity, lowering the local dielectric constant in the cavity and therefore strengthening the existing electrostatic interactions.

*GlcNAc modification to reduce the desolvation penalty and enhance the rigidity (compound **6**).* Replacing the Glc*N*Ac moiety by a cyclohexan-1,2-diol preserves the sLe^x-like conformation while improving the affinity due to a lower desolvation penalty. An additional methyl substitution in 3-position further rigidifies the core structure (13). Elongating the alkyl substituent to an ethyl group (**4→6**) is found to improve affinity by 2-fold (27). The ethyl group does not further rigidify the cyclohexane, but it forms enthalpically beneficial hydrophobic contacts (**4→6**; $\Delta\Delta H°$ = -4.8 kJ mol^{-1}) with Gln85 largely compensated by an entropic loss (**4→6**; $-T\Delta\Delta S°$ = +2.9 kJ mol^{-1}) due to conformational restrictions upon binding. However, the low electron density of this ethyl group (PDB-code: XXXX) suggests some residual flexibility even in the bound state.

*D-Gal modification for acid orientation (compound **7**).* The hydroxyl group in 2-position of D-Gal does not interact with the protein in the bound state. Any hydrophobic substituent in this position, e.g. an acetate (compound **7**) or a benzoate,(27) was found to be favourable for the affinity compared to the hydroxyl group (Table 1). These substituents do not directly interact with the protein but prevent the ligand from forming an unfavourable intramolecular hydrogen bond with the acid in solution. This allows for a more bound-like conformation being present in solution (**4→7**; $-T\Delta\Delta S°$ = -1.5 kJ mol^{-1}) having an optimal lactic acid orientation (**4→7**; $\Delta\Delta H°$ = -1.6 kJ mol^{-1}) for the salt bride formation with the positively charged Arg97.

*Neu5Ac substitution to lower desolvation costs (compound **8**).* The carboxylic acid of compounds **1** - **7** can form an intramolecular hydrogen bond with one of the hydroxy groups in 2 and 4 position of the D-Gal. These hydrogen bonds have to be broken upon binding which results in an enthalpic penalty. This penalty can be circumvent by substituting the carboxylic acid with an azetidine amide, which furthermore lowers the desolvation enthalpy (**4→8**; $\Delta\Delta H°$ = -2.1 kJ mol^{-1}). Likely, the amide most likely has a slightly weaker interaction with Arg97 which leads to an enthalpic loss but an entropic gain due to retained flexibility of both protein and ligand (**4→8**; $-T\Delta\Delta S°$ = -0.4 kJ mol^{-1}).

*Combining all beneficial modifications to derive a drug-like molecule (compound **9**).* Any single modification of compounds **5**-**8** improves the affinity by a factor of only 2.1 - 3.6-fold compared

to compound **4**. But, due to the rigid core structure and the consequent additivity the minor improvements combined in one derivative result in an affinity improvement of a factor of more than 63-fold (**4→9**). The summed up individual thermodynamic contributions of modifications **5**-**8** to calculate the theoretical thermodynamics of compound **9** (Table 1) differ by less than 0.5 kJ mol^{-1} from the experimentally determined thermodynamics demonstrating high additivity effects. Comparing our final compound **9** with the natural binding motif sLe^x (compound **1**) we were able to improve the originally unfavourable enthalpy by -20.0 kJ mol^{-1} (**1→9**), whereas the favorably entropy could be maintained ($-T\Delta\Delta S°$ = +0.3 kJ mol^{-1}) (Figure 3). Furthermore, according to Lipinski's rule of 5 for oral bioavailability (MW ≤ 500 Da; H-bond acceptors ≤ 5; H-bond acceptors ≤ 10; log*P* ≤ 5) the drug-likeness of glycomimetic **9** is greatly improved compared to sLe^x (28). Compound **9** fulfils the requirements for H-bond donation (reduced from 14 to 5 H-bond donors) and lipophilicity (log*P* of -7.4 increased to 2.0). Lipophilicity is increased mainly due to the replacement of the charged carboxylic acid by a neutral amide and the introduction of the trifluormethyl group. Moreover, the number of H-bond acceptors (reduced from 25 to 14) and the molecular weight (reduced from 821 Da to 742 Da) were consequently improved. In addition, compound **9** is stable in rat liver microsomes against hydrolysis of the ester bond by esterases as well as oxidation by cytochrome P450s. However, low solubility (44 ± 8 µM) and low permeation of compound **9** over a colorectal adenocarcinoma cell (Caco-2) (29) membrane still leave room for further optimization.

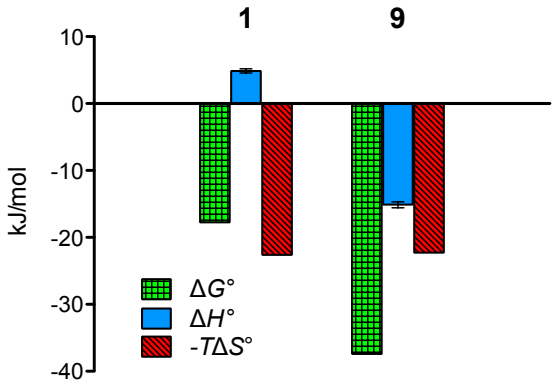


Figure 3. Thermodynamic fingerprints of the natural binding motif sialyl Lewis^x (compound **1**) and the glycomimetic compound **9** derived from ITC measurements. Ligand pre-organization, desolvation, rigidification and pharmacophore orientation improved binding enthalpy by 20 kJ/mol, while the beneficial entropy could be maintained. Without the formation of a new interaction affinity could be improved by almost 3000-fold.

In summary, in 20 years of optimizing E-selectin antagonists we were able to improve the affinity almost 3000-fold compared to the lead structure sLe^x and it is probably the first example in carbohydrate-based drug discovery of an oligosaccharide being rationally improved to a sub-micromolar binder. Biophysical analysis and X-ray crystallography guided the enthalpic optimization process where we did not alter pharmacophores of the glycomimetic ligands, but improved the affinity only by pre-organization, orientation of functional groups, desolvation, and rigidification. Furthermore, lower molecular weight, reduced numbers of hydrogen bond acceptors and donors, and a more hydrophobic character increased the drug-likeness of our E-selectin antagonists. Our approach presents a general approach to improve carbohydrate-based lead structures in drug discovery.

Acknowledgements. The authors thank the staff of the Swiss Light Source, Paul Scherrer Institute, Villigen, Switzerland.

References

1. Ley K, Kansas GS. Selectins in T-cell recruitment to non-lymphoid tissues and sites of inflammation. *Nat Rev Immunol*. **2004**;*4(5)*:325-36.
2. Davies MJ, Gordon JL, Gearing AJH, Pigott R, Woolf N, Katz D, et al. The expression of the adhesion molecules ICAM-1, VCAM-1, PECAM, and E-selectin in human atherosclerosis. *J Pathol*. **1993**;*171(3)*:223-29.
3. Koch AE, Burrows J, Haines G, Carlos T, Harlan J, Leibovich S. Immunolocalization of endothelial and leukocyte adhesion molecules in human rheumatoid and osteoarthritic synovial tissues. *Lab Invest*. **1991**;*64(3)*:313-20.
4. Kobayashi T, Hashimoto S, Imai K, Amemiya E, Yamaguchi M, Yachi A, et al. Elevation of serum soluble intercellular adhesion Molecule-1 (sICAM-1) and sE-Selectin levels in bronchial asthma. *Clin Exp Immunol*. **1994**;*96(1)*:110-5.
5. Natarajan M, Udden M, McIntire L. Adhesion of sickle red blood cells and damage to interleukin-1 beta stimulated endothelial cells under flow in vitro. *Blood*. **1996**;*87(11)*:4845-52.
6. St Hill CA. Interactions between endothelial selectins and cancer cells regulate metastasis. *Front Biosci*. **2010**;*16*:3233-51.
7. Kannagi R, Izawa M, Koike T, Miyazaki K, Kimura N. Carbohydrate-mediated cell adhesion in cancer metastasis and angiogenesis. *Cancer Sci*. **2004**;*95(5)*:377-84.
8. Ernst B, Magnani JL. From carbohydrate leads to glycomimetic drugs. *Nat Rev Drug Discov*. **2009**;*8(8)*:661-77.
9. Kaila N, Thomas BE. Design and synthesis of sialyl Lewis^x mimics as E- and P-selectin inhibitors. *Med Res Rev*. **2002**;*22(6)*:566-601.
10. Simanek EE, McGarvey GJ, Jablonowski JA, Wong CH. Selectin-Carbohydrate Interactions: From Natural Ligands to Designed Mimics. *Chem Rev*. **1998**;*98(2)*:833-62.
11. Chang J, Patton JT, Sarkar A, Ernst B, Magnani JL, Frenette PS. GMI-1070, a novel pan-selectin antagonist, reverses acute vascular occlusions in sickle cell mice. *Blood*. **2010**;*116(10)*:1779-86.
12. Telen MJ, Wun T, McCavit TL, De Castro LM, Krishnamurti L, Lanzkron S, et al. Randomized phase 2 study of GMI-1070 in SCD: reduction in time to resolution of vaso-occlusive events and decreased opioid use. *Blood*. **2015**;*125(17)*:2656-64.
13. Binder FP, Lemme K, Preston RC, Ernst B. Sialyl Lewis^x: A "Pre-Organized Water Oligomer"? *Angew Chem Int Ed*. **2012**;*51(29)*:7327-31.
14. Burnouf D, Ennifar E, Guedich S, Puffer B, Hoffmann G, Bec G, et al. kinITC: a new method for obtaining joint thermodynamic and kinetic data by isothermal titration calorimetry. *J Am Chem Soc*. **2012**;*134(1)*:559-65.
15. Houtman JC, Brown PH, Bowden B, Yamaguchi H, Appella E, Samelson LE, et al. Studying multisite binary and ternary protein interactions by global analysis of isothermal titration calorimetry data in SEDPHAT: application to adaptor protein complexes in cell signaling. *Protein Sci*. **2007**;*16(1)*:30-42.
16. Preston RC, Jakob RP, Binder FPC, Sager CP, Ernst B, Maier T. E-selectin ligand complexes adopt an extended high-affinity conformation. *J Mol Cell Biol*. **2016**;*8(1)*:62-72.
17. Somers WS, Tang J, Shaw GD, Camphausen RT. Insights into the molecular basis of leukocyte tethering and rolling revealed by structures of P-and E-selectin bound to SLe^x and PSGL-1. *Cell*. **2000**;*103(3)*:467-79.
18. Ramphal JY, Zheng Z-L, Perez C, Walker LE, DeFrees SA, Gaeta FC. Structure-activity relationships of sialyl Lewis x-containing oligosaccharides. 1. Effect of modifications of the fucose moiety. *J Med Chem*. **1994**;*37(21)*:3459-63.
19. Brandley BK, Kiso M, Abbas S, Nikrad P, Srivasatava O, Foxall C, et al. Structure—function studies on selectin carbohydrate ligands. Modifications to fucose, sialic acid and sulphate as a sialic acid replacement. *Glycobiology*. **1993**;*3(6)*:633-41.

20. Marshall BT, Long M, Piper JW, Yago T, McEver RP, Zhu C. Direct observation of catch bonds involving cell-adhesion molecules. *Nature*. **2003**;*423(6936)*:190-3.
21. Sarangapani KK, Yago T, Klopocki AG, Lawrence MB, Fieger CB, Rosen SD, et al. Low force decelerates L-selectin dissociation from P-selectin glycoprotein ligand-1 and endoglycan. *J Biol Chem*. **2004**;*279(3)*:2291-8.
22. Williams DH, Stephens E, O'Brien DP, Zhou M. Understanding Noncovalent Interactions: Ligand Binding Energy and Catalytic Efficiency from Ligand-Induced Reductions in Motion within Receptors and Enzymes. *Angew Chem Int Ed*. **2004**;*43(48)*:6596-616.
23. Williams DH, Zhou M, Stephens E. Ligand Binding Energy and Enzyme Efficiency from Reductions in Protein Dynamics. *J Mol Biol*. **2006**;*355(4)*:760-7.
24. Biela A, Betz M, Heine A, Klebe G. Water Makes the Difference: Rearrangement of Water Solvation Layer Triggers Non-additivity of Functional Group Contributions in Protein–Ligand Binding. *ChemMedChem*. **2012**;*7(8)*:1423-34.
25. Baum B, Muley L, Smolinski M, Heine A, Hangauer D, Klebe G. Non-additivity of Functional Group Contributions in Protein–Ligand Binding: A Comprehensive Study by Crystallography and Isothermal Titration Calorimetry. *J Mol Biol*. **2010**;*397(4)*:1042-54.
26. Zierke M, Smieško M, Rabbani S, Aeschbacher T, Cutting B, Allain FHT, et al. Stabilization of Branched Oligosaccharides: Lewis^x Benefits from a Nonconventional C–H···O Hydrogen Bond. *J Am Chem Soc*. **2013**;*135(36)*:13464-72.
27. Schwizer D, Patton JT, Cutting B, Smieško M, Wagner B, Kato A, et al. Pre-organization of the Core Structure of E-Selectin Antagonists. *Chem Eur J*. **2012**;*18(5)*:1342-51.
28. Lipinski CA, Lombardo F, Dominy BW, Feeney PJ. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings1. *Adv Drug Del Rev*. **2001**;*46(1–3)*:3-26.
29. Hubatsch I, Ragnarsson EG, Artursson P. Determination of drug permeability and prediction of drug absorption in Caco-2 monolayers. *Nat Protoc*. **2007**;*2(9)*:2111-9.

4 FimH antagonists

4.1 Urinary tract infection

Urinary tract infection (UTI) is defined as bacterial infection of the otherwise sterile urinary tract. The infection either takes places in the lower urinary tract, i.e. in the bladder and urethra (Cystitis) or in the upper urinary tract, i.e. in the kidneys (Pyelonephritis) (1,2). Clinically, a distinction between uncomplicated and complicated UTIs is made. In the case of an uncomplicated UTI, no structural or neurological abnormalities of the urinary tract arise, whereas complicated UTIs are accompanied by factors compromising the urinary tract or host defense, such as neurological diseases, immunosuppression, renal failure, renal transplantation, pregnancy and the presence of foreign bodies such as catheters (3). The vast majority of complicated UTIs are related to catheter induced UTIs (CAUTI) (4,5).

UTIs are one of the most common infections, affecting approximately 150 million peoples every year (6,7). Bacteria can colonize the urethra in both men and woman, but compared to men, the short urethra of women makes it easier for bacteria to reach the bladder before removal by micturition. Therefore, women are affected by UTI up to 5 times more often compared to men (8,9). The incidence rate of UTI in women for one year is 11% and the risk of suffering from UTI once during lifetime is approximately 50% (10). Other risk factors are sexual activity, vaginal infections, prior UTIs, diabetes and genetic susceptibility (8,9,11-14). The primary symptoms of UTIs are strangury, burning micturition and bacteriuria or pyuria. If untreated, UTI can lead to hematuria, fever and even to a potentially life-threating sepsis (15-17).

4.1.1 Current treatment options

In most cases, an antibiotic treatment is prescribed to erase the symptoms of an acute cystitis and to prevent the infection of the kidneys (18). Due to the inherent problems of bacteria resistance to antibiotics, the recommendations for the treatment changes over time. For a long time, treatment with trimethoprim-sulfomethoxazol or fluoroquinolone-derivatives like ofloxacine and norfloxacine were first-line therapies (10). Latest recommendations advise against the use of these antibiotics. Currently, fosfomycin-trometamole in a single dose treatment or nitrofurantione over a period of three days are considered as first line therapeutics for the treatment of UTI (19, 20). For peoples suffering from recurrent UTIs, preventive antibiotics are indicated either dosed daily or post coitus (21, 22).

Other preventions besides antibiotics are heavily discussed. Whereas probiotics have been shown to be ineffective, the use of cranberry juice showed varying effectivity in meta analyses (23-25). Pure D-mannose was recently marketed as medical product to prevent UTIs and showed effectiveness in one study, but crucial data about pharmacodynamics and –kinetics (PK/PD) are not available and the study suffers from a low number of recruitments (26). Future treatment possibilities might include vaccinations (27, 28). Another approach is presented below.

4.2 Uropathogenic *Escherichia coli* (UPEC)

In approximately 75% of all ambulatory UTI cases, uropathogenic *Escherichia coli* (UPEC) are the causing pathogens. Besides UPEC, *Klebsiella*, *Enterecoccus* and *Staphylococcus* bacterial strains are the major pathogen groups (6, 10, 29). *E. coli* are gram-negative bacteria, which are part of the common intestinal flora of humans (30). Mutated *E. coli* strains can become pathogenic; thus intestinal pathogenic *E. coli* mainly cause diarrhea and extraintestinal pathogenic *E. coli* (ExPEC) are not harmful in the intestines but colonize and harm parts of the body outside the gut. UPEC as shown in Figure 4.1 are part of ExPEC(31).

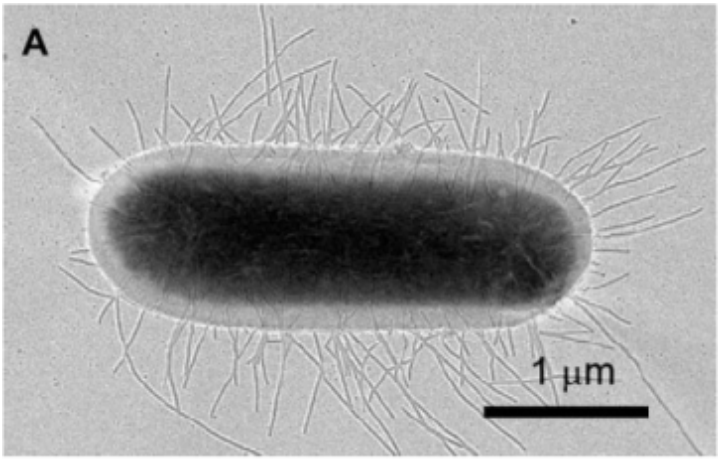


Figure 4.1. Electron micrograph of an uropathogenic E. coli expressing type 1 pili used from Capitani *et al.* (32).

UPEC have developed the ability to colonize bladder and kidney cells. As a first step of the pathogenesis, UPEC attach to the luminal surface of the bladder (33, 34). This binding is mediated by type 1 pili interacting with oligomannosides on the urinary bladder mucosa (35). Once adhered to the cell surface, UPEC invade the host cell and replicate inside the cell to form a biofilm-like intracellular bacterial community (IBC), in which UPEC are in a quiescent mode, resistant to the innate immune response and antibiotic treatment (36, 37). An IBC can persist for months until quiescent bacteria reactivate, exit the cell, and infect new bladder cells. The formation of IBC can be a source for recurrent urinary-tract infections (38).

4.2.1 Type 1 Pili

Type 1 pili are the primary virulence factor of UPEC to promote cystitis by binding to uroplakin Ia on the mucosal surface of the human urinary bladder (39). This initial step of adhesion prevents the rapid clearance of UPEC from the urinary tract by the bulk flow of urine and at the same time enables the invasion of the host cells (35). Type 1 pili are filamentous protein complexes consisting of a helical structure (500-3000 FimA subunits) and on the tip the three subunits FimF, FimG, and FimH (Figure 4.2) (40-42). The pilus rod is assembled through the chaperon-usher pathway (43), i.e. the chaperon protein FimC in the periplasm binds pilus subunits, folds them to FimA, and delivers them to the transmembrane assembly platform FimD, where FimA is released to be incorporated into the pilus (32).

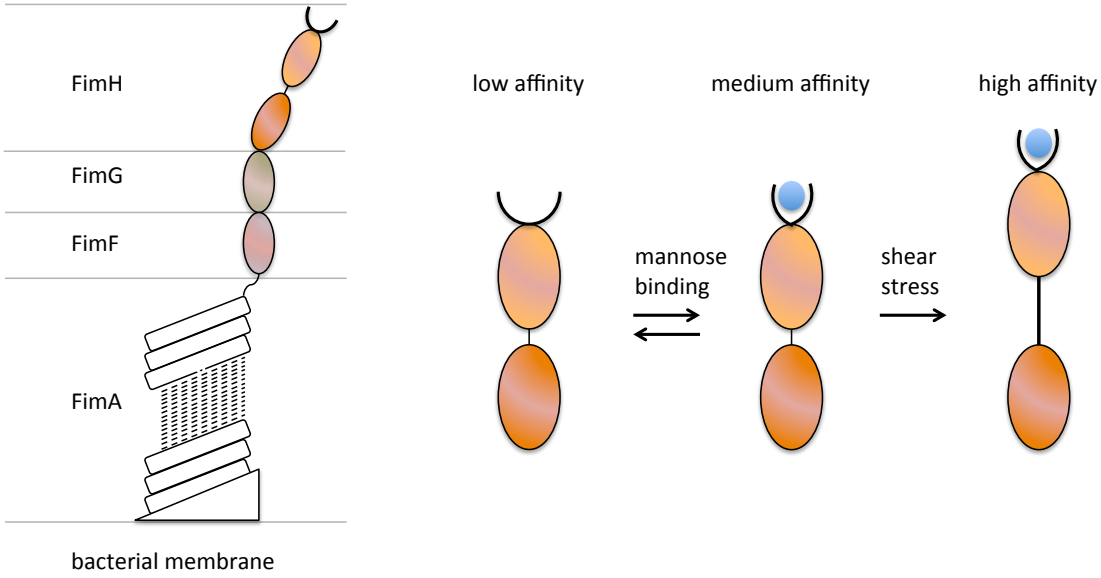


Figure 4.2 The type 1 pilus is consisting of a helical structure formed by FimA subunits and a terminal tip consisting of FimF, FimG and the lectin domain FimH with the terminal carbohydrate recognition domain (CRD). Upon binding of mannosides, FimH is undertaking a conformational change into a medium affinity state. Once shear stress is applied by urinary flow, a high affinity conformation is formed.

4.2.2 Lectin FimH

FimH on the distal tip of the type 1 pili consists of the carbohydrate recognition domain (CRD) which recognizes the mannosides on the luminal surface of the bladder cells. FimH is a polypeptide of 279 amino acids, folding into two domains containing only β-strands. The C-terminal domain (pili domain) anchors FimH to FimG and the N-terminal domain (lectin domain) contains the CRD (44). The lectin domain is an elongated, 11 β-stranded domain with a jelly roll-like topology. On the tip, the mannose-binding pocket is located (45). FimH binds to mannosides in a so-called catch bond mechanism (46). Under shear stress, the FimH domains are separated and the affinity towards mannose residues increase up to 100-fold (47). Due to this effect, UPEC is able to prevent its clearance from the bladder by urinary flow. When the flow ceases, the bacteria can transmigrate again over the urothelial surface (48, 49).

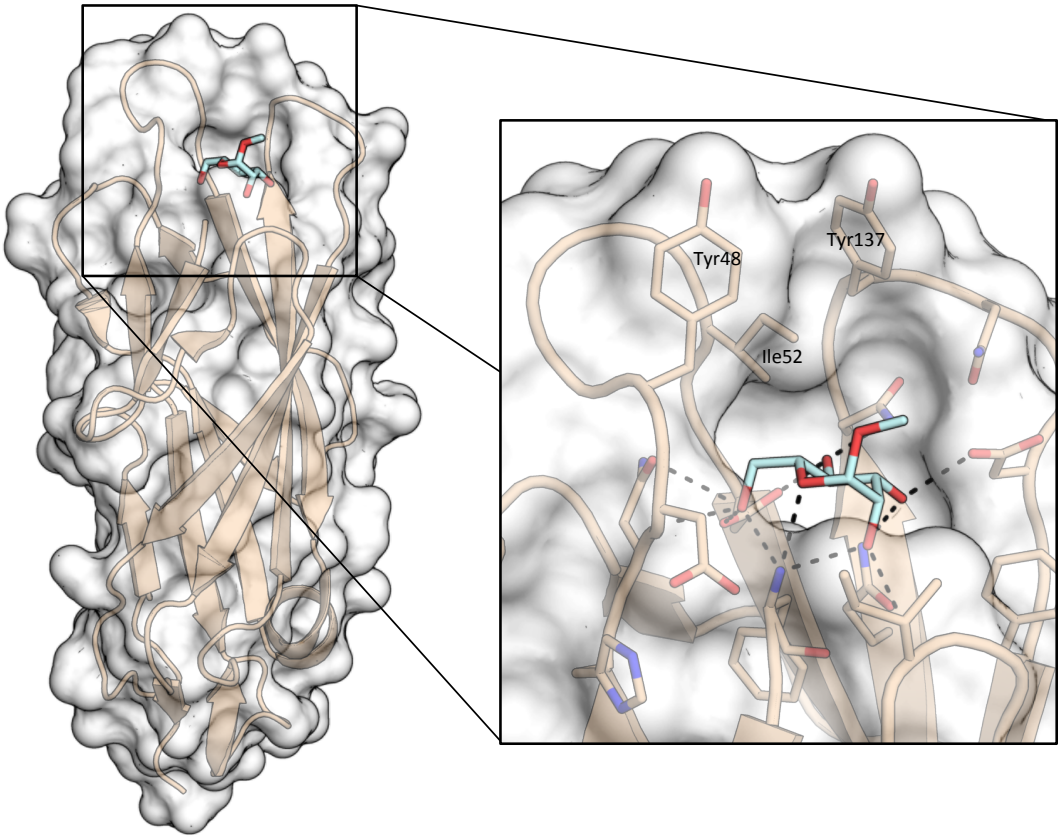


Figure 4.3. Crystal structure of FimH lectin domain binding α-D-methyl-manno-pyranoside (**1**) (44). The lipophilic residues Tyr47, Ile52 and Tyr137 play a crucial role in the development of highly affine FimH antagonists (PDB: 5JCQ).

4.3 FimH Antagonists

Increasing resistance against antibiotics require new approaches for the treatment of UTI (50-53). The bacterial lectin FimH interacting with oligomannosides on the bladder surface is a potential therapeutic target. The adherence of the UPEC to the cell surface followed by the cell invasion can be prevented by α-D-manno-pyranosides, blocking the CRD and thus preventing the interaction with the urothelial surface (54, 55). More than three decades ago, proof of concept studies to treat UTIs with α-D-methyl-manno-pyranoside (**1**)(Figure 4.3) have been conducted in mice (56). Over the last decades, FimH antagonists have been further improved regarding their affinity as well as their pharmacokinetic properties (Figure 4.4.) (57-70). First systematic studies to improve the affinity of FimH antagonists have been conducted with aliphatic aglycones, e.g. to obtain *n*-heptyl-α-D-manno-pyranoside **2** (59). Starting from the *p*-nitrophenyl α-D-manno-pyranoside (**3a**), highly affine FimH antagonists could be obtained by a further aromatic elongation of the aglycone with a squaric acid (→ **4a** & **4b**) (60, 64), phenyl (→**5a** - **5f**) (61, 63, 66) or indolinyl (→**6a** & **6b**) (65). The binding mode of monovalent FimH

antagonists can be rationalized on the basis of the structure of the CRD. In the binding pocket α-D-mannose establishes an extended network of hydrogen bonds or hydrophobic interactions (71). Furthermore, the entrance to the mannose binding pocket is formed by three hydrophobic amino acids (Tyr48, Ile52 and Tyr137; Figure 4.3) and can host the aliphatic and aromatic aglycones. The high affinity of biphenyl-α-D-manno-pyranosides could be achieved due to π-π stacking interactions of the outer aromatic ring with Tyr48 and/or Tyr137. An electron withdrawing substituent on the outer aromatic ring additionally improves these interactions (61,70). Further strategies to achieve high-affinity FimH antagonists are based on combinatorial chemistry (→ **7**) or C-mannosides (→ **8**) (68,69). To overcome permeability or solubility problems, ester prodrugs (→ **9**) or phosphate prodrugs (→ **10**) were synthesized (61, 67).

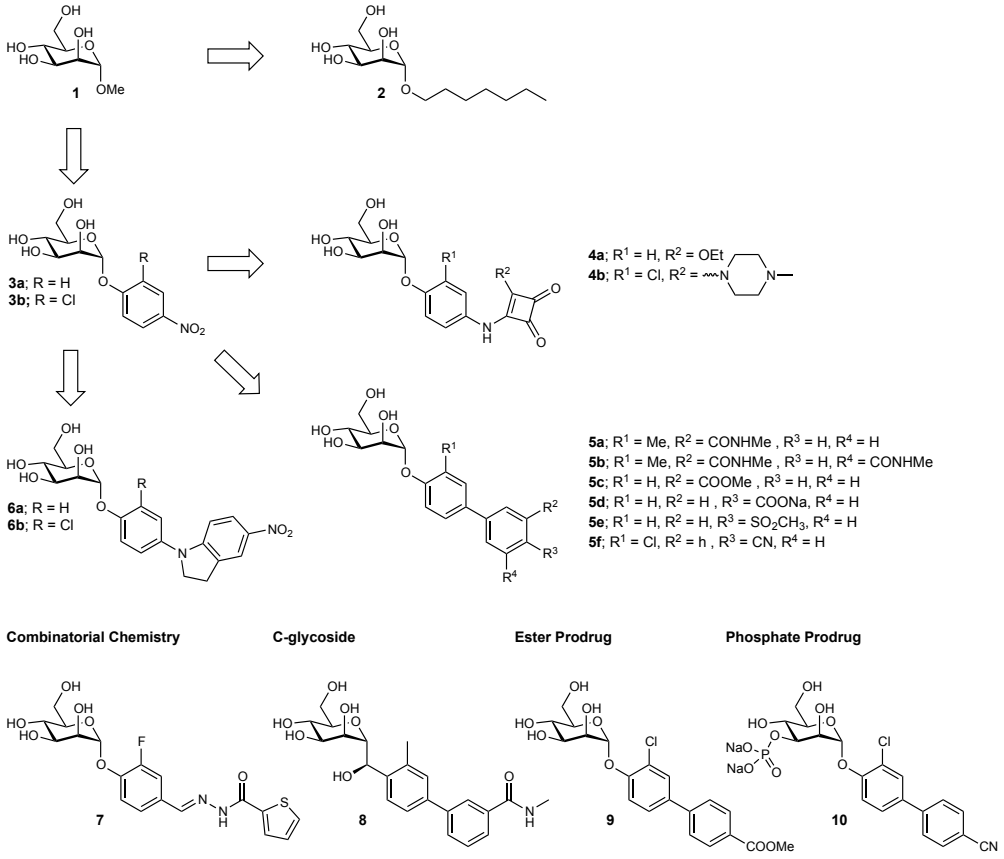


Figure 4.4. Monovalent FimH antagonists showing nanomolar affinities.

4.4 References

1. Flores-Mireles AL, Walker JN, Caparon M, Hultgren SJ. Urinary tract infections: epidemiology, mechanisms of infection and treatment options. *Nat Rev Microbiol*. 2015;*13(5)*:269-84.
2. Barber AE, Norton JP, Spivak AM, Mulvey MA. Urinary tract infections: current and emerging management strategies. *Clin Infect Dis*. 2013;*57(5)*:719-24.
3. Bent S, Nallamothu BK, Simel DL, Fihn SD, Saint S. Does this woman have an acute uncomplicated urinary tract infection? *JAMA*. 2002;*287(20)*:2701-10.
4. Trautner BW, Hull RA, Darouiche RO. Prevention of catheter-associated urinary tract infection. *Curr Opin Infect Dis*. 2005;*18(1)*:37-41.
5. Warren JW. Catheter-associated urinary tract infections. *Infect Dis Clin North Am*. 1997;*11(3)*:609-22.
6. Foxman B. The epidemiology of urinary tract infection. *Nat Rev Urol*. 2010;*7(12)*:653-60.
7. Stamm WE, Norrby SR. Urinary tract infections: disease panorama and challenges. *J Infect Dis*. 2001;*183 Suppl 1*:S1-4.
8. Hooton TM, Scholes D, Hughes JP, Winter C, Roberts PL, Stapleton AE, et al. A prospective study of risk factors for symptomatic urinary tract infection in young women. *N Engl J Med*. 1996;*335(7)*:468-74.
9. Schaeffer AJ, Nicolle LE. Urinary Tract Infections in Older Men. *N Engl J Med*. 2016;*374(22)*:2192.
10. Fihn SD. Clinical practice. Acute uncomplicated urinary tract infection in women. *N Engl J Med*. 2003;*349(3)*:259-66.
11. Scholes D, Hooton TM, Roberts PL, Stapleton AE, Gupta K, Stamm WE. Risk factors for recurrent urinary tract infection in young women. *J Infect Dis*. 2000;*182(4)*:1177-82.
12. Hooton TM. Recurrent urinary tract infection in women. *Int J Antimicrob Agents*. 2001;*17(4)*:259-68.
13. Hooton TM. Pathogenesis of urinary tract infections: an update. *J Antimicrob Chemother*. 2000;*46 Suppl 1*:1-7; discussion 63-5.
14. Ragnarsdottir B, Lutay N, Gronberg-Hernandez J, Koves B, Svanborg C. Genetics of innate immunity and UTI susceptibility. *Nat Rev Urol*. 2011;*8(8)*:449-68.
15. Stamm WE, Hooton TM. Management of urinary tract infections in adults. *N Engl J Med*. 1993;*329(18)*:1328-34.
16. Prentiss KA, Newby PK, Vinci RJ. Adolescent female with urinary symptoms: a diagnostic challenge for the pediatrician. *Pediatr Emerg Care*. 2011;*27(9)*:789-94.
17. Kalra OP, Raizada A. Approach to a patient with urosepsis. *J Glob Infect Dis*. 2009;*1(1)*:57-63.
18. Vogel T, Verreault R, Gourdeau M, Morin M, Grenier-Gosselin L, Rochette L. Optimal duration of antibiotic therapy for uncomplicated urinary tract infection in older women: a double-blind randomized controlled trial. *CMAJ*. 2004;*170(4)*:469-73.
19. Gupta K, Hooton TM, Naber KG, Wullt B, Colgan R, Miller LG, et al. International clinical practice guidelines for the treatment of acute uncomplicated cystitis and pyelonephritis in women: A 2010 update by the Infectious Diseases Society of America and the European Society for Microbiology and Infectious Diseases. *Clin Infect Dis*. 2011;*52(5)*:e103-20.
20. Fosfomycin for urinary tract infections. *Med Lett Drugs Ther*. 1997;*39(1005)*:66-8.
21. Stapleton A, Latham RH, Johnson C, Stamm WE. Postcoital antimicrobial prophylaxis for recurrent urinary tract infection. A randomized, double-blind, placebo-controlled trial. *JAMA*. 1990;*264(6)*:703-6.
22. Albert X, Huertas I, Pereiro, II, Sanfelix J, Gosalbes V, Perrota C. Antibiotics for preventing recurrent urinary tract infection in non-pregnant women. *Cochrane Database Syst Rev*. 2004*(3)*:CD001209.
23. Schwenger EM, Tejani AM, Loewen PS. Probiotics for preventing urinary tract infections in adults and children. *Cochrane Database Syst Rev*. 2015*(12)*:CD008772.
24. Jepson RG, Williams G, Craig JC. Cranberries for preventing urinary tract infections. *Cochrane Database Syst Rev*. 2012;*10*:CD001321.

25. Wang CH, Fang CC, Chen NC, Liu SS, Yu PH, Wu TY, et al. Cranberry-containing products for prevention of urinary tract infections in susceptible populations: a systematic review and meta-analysis of randomized controlled trials. *Arch Intern Med*. 2012;*172(13)*:988-96.
26. Kranjcec B, Papes D, Altarac S. D-mannose powder for prophylaxis of recurrent urinary tract infections in women: a randomized clinical trial. *World J Urol*. 2014;*32(1)*:79-84.
27. Mike LA, Smith SN, Sumner CA, Eaton KA, Mobley HL. Siderophore vaccine conjugates protect against uropathogenic Escherichia coli urinary tract infection. *Proc Natl Acad Sci U S A*. 2016;*113(47)*:13468-73.
28. Huttner A, Hatz C, van den Dobbelsteen G, Abbanat D, Hornacek A, Frolich R, et al. Safety, immunogenicity, and preliminary clinical efficacy of a vaccine against extraintestinal pathogenic Escherichia coli in women with a history of recurrent urinary tract infection: a randomised, single-blind, placebo-controlled phase 1b trial. *Lancet Infect Dis*. 2017;*17(5)*:528-37.
29. Laupland KB, Ross T, Pitout JD, Church DL, Gregson DB. Community-onset urinary tract infections: a population-based assessment. *Infection*. 2007;*35(3)*:150-3.
30. Tenaillon O, Skurnik D, Picard B, Denamur E. The population genetics of commensal Escherichia coli. *Nat Rev Microbiol*. 2010;*8(3)*:207-17.
31. Mulvey MA. Adhesion and entry of uropathogenic Escherichia coli. *Cell Microbiol*. 2002;*4(5)*:257-71.
32. Capitani G, Eidam O, Glockshuber R, Grutter MG. Structural and functional insights into the assembly of type 1 pili from Escherichia coli. *Microbes Infect*. 2006;*8(8)*:2284-90.
33. Berglund J, Knight SD. Structural basis for bacterial adhesion in the urinary tract. *Adv Exp Med Biol*. 2003;*535*:33-52.
34. Dhakal BK, Kulesus RR, Mulvey MA. Mechanisms and consequences of bladder cell invasion by uropathogenic Escherichia coli. *Eur J Clin Invest*. 2008;*38 Suppl 2*:2-11.
35. Martinez JJ, Mulvey MA, Schilling JD, Pinkner JS, Hultgren SJ. Type 1 pilus-mediated bacterial invasion of bladder epithelial cells. *EMBO J*. 2000;*19(12)*:2803-12.
36. Cegelski L, Marshall GR, Eldridge GR, Hultgren SJ. The biology and future prospects of antivirulence therapies. *Nat Rev Microbiol*. 2008;*6(1)*:17-27.
37. Justice SS, Hung C, Theriot JA, Fletcher DA, Anderson GG, Footer MJ, et al. Differentiation and developmental pathways of uropathogenic Escherichia coli in urinary tract pathogenesis. *Proc Natl Acad Sci U S A*. 2004;*101(5)*:1333-8.
38. Mysorekar IU, Hultgren SJ. Mechanisms of uropathogenic Escherichia coli persistence and eradication from the urinary tract. *Proc Natl Acad Sci U S A*. 2006;*103(38)*:14170-5.
39. Johnson JR. Virulence factors in Escherichia coli urinary tract infection. *Clin Microbiol Rev*. 1991;*4(1)*:80-128.
40. Klemm P, Christiansen G. Three fim genes required for the regulation of length and mediation of adhesion of Escherichia coli type 1 fimbriae. *Mol Gen Genet*. 1987;*208(3)*:439-45.
41. Russell PW, Orndorff PE. Lesions in two Escherichia coli type 1 pilus genes alter pilus number and length without affecting receptor binding. *J Bacteriol*. 1992;*174(18)*:5923-35.
42. Hanson MS, Brinton CC, Jr. Identification and characterization of E. coli type-1 pilus tip adhesion protein. *Nature*. 1988;*332(6161)*:265-8.
43. Nuccio SP, Baumler AJ. Evolution of the chaperone/usher assembly pathway: fimbrial classification goes Greek. *Microbiol Mol Biol Rev*. 2007;*71(4)*:551-75.
44. Geibel S, Waksman G. Crystallography and electron microscopy of chaperone/usher pilus systems. *Adv Exp Med Biol*. 2011;*715*:159-74.
45. Choudhury D, Thompson A, Stojanoff V, Langermann S, Pinkner J, Hultgren SJ, et al. X-ray structure of the FimC-FimH chaperone-adhesin complex from uropathogenic Escherichia coli. *Science*. 1999;*285(5430)*:1061-6.
46. Thomas WE, Trintchina E, Forero M, Vogel V, Sokurenko EV. Bacterial adhesion to target cells enhanced by shear force. *Cell*. 2002;*109(7)*:913-23.
47. Sauer MM, Jakob RP, Eras J, Baday S, Eris D, Navarra G, et al. Catch-bond mechanism of the bacterial adhesin FimH. *Nat Commun*. 2016;*7*:10738.

48. Eris D, Preston RC, Scharenberg M, Hulliger F, Abgottspon D, Pang L, et al. The Conformational Variability of FimH: Which Conformation Represents the Therapeutic Target? *Chembiochem*. 2016;*17(11)*:1012-20.
49. Mayer K, Eris D, Schwardt O, Sager CP, Rabbani S, Kleeb S, et al. Urinary Tract Infection: Which Conformation of the Bacterial Lectin FimH Is Therapeutically Relevant? *J Med Chem*. 2017;*60(13)*:5646-62.
50. Nolan LK, Li G, Logue CM. Origin and Dissemination of Antimicrobial Resistance among Uropathogenic Escherichia coli. *Microbiol Spectr*. 2015;*3(5)*.
51. Sanchez GV, Master RN, Karlowsky JA, Bordon JM. In vitro antimicrobial resistance of urinary Escherichia coli isolates among U.S. outpatients from 2000 to 2010. *Antimicrob Agents Chemother*. 2012;*56(4)*:2181-3.
52. anresis.ch – Sentinel Surveillance of Antibiotic Resistance in Switzerland: University of Bern; 2017 [Available from: www.search.ifik.unibe.ch]
53. Ghosh H, Doijad S, Falgenhauer L, Fritzenwanker M, Imirzalioglu C, Chakraborty T. blaCTX-M-27-Encoding Escherichia coli Sequence Type 131 Lineage C1-M27 Clone in Clinical Isolates, Germany. *Emerg Infect Dis*. 2017;*23(10)*:1754-6.
54. Hartmann M, Lindhorst TK. The Bacterial Lectin FimH, a Target for Drug Discovery - Carbohydrate Inhibitors of Type 1 Fimbriae-Mediated Bacterial Adhesion. *European Journal of Organic Chemistry*. 2011;*2011(20-21)*:3583-609.
55. Sharon N. Carbohydrates as future anti-adhesion drugs for infectious diseases. *Biochim Biophys Acta*. 2006;*1760(4)*:527-37.
56. Aronson M, Medalia O, Schori L, Mirelman D, Sharon N, Ofek I. 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*. 1979;*139(3)*:329-32.
57. Ofek I, Mirelman D, Sharon N. Adherence of Escherichia coli to human mucosal cells mediated by mannose receptors. *Nature*. 1977;*265(5595)*:623-5.
58. Firon N, Ashkenazi S, Mirelman D, Ofek I, Sharon N. Aromatic alpha-glycosides of mannose are powerful inhibitors of the adherence of type 1 fimbriated Escherichia coli to yeast and intestinal epithelial cells. *Infect Immun*. 1987;*55(2)*:472-6.
59. Bouckaert J, Berglund J, Schembri M, De Genst E, Cools L, Wuhrer M, et al. Receptor binding studies disclose a novel class of high-affinity inhibitors of the Escherichia coli FimH adhesin. *Mol Microbiol*. 2005;*55(2)*:441-55.
60. Sperling O, Fuchs A, Lindhorst TK. 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)*:3913-22.
61. Klein T, Abgottspon D, Wittwer M, Rabbani S, Herold J, Jiang X, et al. FimH antagonists for the oral treatment of urinary tract infections: from design and synthesis to in vitro and in vivo evaluation. *J Med Chem*. 2010;*53(24)*:8627-41.
62. Han Z, Pinkner JS, Ford B, Obermann R, Nolan W, Wildman SA, et al. Structure-based drug design and optimization of mannoside bacterial FimH antagonists. *J Med Chem*. 2010;*53(12)*:4779-92.
63. Cusumano CK, Pinkner JS, Han Z, Greene SE, Ford BA, Crowley JR, et al. Treatment and prevention of urinary tract infection with orally active FimH inhibitors. *Sci Transl Med*. 2011;*3(109)*:109ra15.
64. Jiang X, Abgottspon D, Kleeb S, Rabbani S, Scharenberg M, Wittwer M, et al. Antiadhesion therapy for urinary tract infections--a balanced PK/PD profile proved to be key for success. *J Med Chem*. 2012;*55(10)*:4700-13.
65. Fiege B, Rabbani S, Preston RC, Jakob RP, Zihlmann P, Schwardt O, et al. The tyrosine gate of the bacterial lectin FimH: a conformational analysis by NMR spectroscopy and X-ray crystallography. *Chembiochem*. 2015;*16(8)*:1235-46.
66. Kleeb S, Pang L, Mayer K, Eris D, Sigl A, Preston RC, et al. FimH antagonists: bioisosteres to improve the in vitro and in vivo PK/PD profile. *J Med Chem*. 2015;*58(5)*:2221-39.
67. Kleeb S, Jiang X, Frei P, Sigl A, Bezencon J, Bamberger K, et al. FimH Antagonists: Phosphate Prodrugs Improve Oral Bioavailability. *J Med Chem*. 2016;*59(7)*:3163-82.

68. Mydock-McGrane L, Cusumano Z, Han Z, Binkley J, Kostakioti M, Hannan T, et al. Antivirulence C-Mannosides as Antibiotic-Sparing, Oral Therapeutics for Urinary Tract Infections. *J Med Chem*. 2016;*59(20)*:9390-408.
69. Frei P, Pang L, Silbermann M, Eris D, Muhlethaler T, Schwardt O, et al. Target-directed Dynamic Combinatorial Chemistry: A Study on Potentials and Pitfalls as Exemplified on a Bacterial Target. *Chemistry*. 2017;*23(48)*:11570-7.
70. Mydock-McGrane LK, Hannan TJ, Janetka JW. Rational design strategies for FimH antagonists: new drugs on the horizon for urinary tract infection and Crohn's disease. *Expert Opin Drug Discov*. 2017;*12(7)*:711-31.
71. Hung CS, Bouckaert J, Hung D, Pinkner J, Widberg C, DeFusco A, et al. Structural basis of tropism of Escherichia coli to the bladder during urinary tract infection. *Mol Microbiol*. 2002;*44(4)*:903-15.

4.5 Manuscript 6: Tyrosin-perfluoroarene interactions of FimH antagonists

Previous work showed the potential for high affinity of biphenyl-α-D-manno-pyranosides , due to π-π stacking interactions with a tyrosine gate. An bioisosteric replacement of the carboxylic acid improved the pharmacokinetic properties as well as affinity. In this Manuscript, a further improvement of the affinity was achieved by forming perfluorinated aglycones. The molecule furthermore showed promising physicochemical properties.

Contribution to the project:

Philipp Dätwyler was measuring and evaluating in the physicochemical properties of perfluorinated FimH antagonists. He further contributed to the writing and arrangement of the publication.

Wojciech Schönemann, Jonathan Cramer, Tobias Mühlethaler, Brigitte Fiege, Marleen Silbermann, Said Rabbani, **Philipp Dätwyler**, Pascal Zihlmann, Roman P. Jakob, Christoph P. Sager, Martin Smieško, Oliver Schwardt, Timm Maier, Beat Ernst*

*Corresponding author

Tyrosine-perfluoroarene interactions lead to picomolar 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)] University of Basel, Institute of Molecular Pharmacy, Pharmacenter of the University of Basel, Klingelbergstrasse 50, 4056, Basel, Switzerland

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

*Corresponding author

Keywords Urinary tract infections, uropathogenic *Escherichia coli*, FimH antagonists, π−π stacking.

Abbreviations CRD, carbohydrate recognition domain; log*D*, octanol/water distribution coefficient; $FimH_{FL}$, full-length FimH stabilized by peptide complementation; $FimH_{LD}$, FimH lectin domain; $FimH_{PD}$, FimH pilin domain; FP, fluorescence polarization; ITC, isothermal titration calorimetry; K_D, dissociation constant; NMR, nuclear magnetic resonance; PAMPA, parallel artificial membrane permeability assay; P_e, effective permeability; UPEC, uropathogenic *E. coli*; UTI, urinary tract infection.

Abstract

Antimicrobial resistance arising from a frequent use of antibiotics as a first-line treatment in urinary tract infections (UTI) has become a serious concern demanding for alternative treatment strategies. In this context, an anti-adhesive approach targeting FimH, a bacterial lectin used by uropathogenic *E. coli* (UPEC) to attach to urothelial cells of the host and thus initiate the infective process, attracted considerable interest. FimH can adopt different conformations, i.e. a low/medium-affinity state in the absence and a high-affinity state in the presence of shear forces. Until recently, however, only the high-affinity state has been thoroughly investigated, despite the fact that a therapeutically successful FimH antagonist should bind efficaciously to the low-affinity conformation as well. Entry to the mannose binding site of FimH is controlled by two tyrosines. In this communication we show that π–π stacking interactions between the aglycone of mannosidic FimH antagonists and the so-called tyrosine gate can be enhanced by electron poor substituents, which increases affinity for both affinity states. For biphenyl mannosides equipped with polyfluorinated outer aromatic rings, low nanomolar K_D values for the low-affinity state and picomolar K_D values for the high-affinity state were identified. Crystal structures of the lectin domain of FimH co-crystallized with polyfluorinated antagonists show a perfect face-to-face alignment of the two aromatic rings. Substantial effects on the binding thermodynamics suggest that improved stacking interactions have a considerable impact on the dynamics and conformation of the low affinity construct. Moreover, fluorination improves pharmacokinetic parameters as predictive indicators for oral availability.

Introduction

Urinary tract infections (UTI) caused by uropathogenic *E. coli* (UPEC) are among the most common infections. Particularly often affected are women, who have a 40 – 50% risk to experience at least one symptomatic UTI episode during their life (1, 2). Because of the high morbidity and the resulting associated costs, UTI pose a serious burden on healthcare systems (3, 4). Uropathogenic *E. coli* (UPEC) account for approx. 70% of the reported UTI cases (5). To date, patients with acute uncomplicated lower UTI are mainly treated with antibiotics to relieve them of associated symptoms and to prevent the infection from exacerbating into a pyelonephritis or urosepsis, which can be life threatening (6, 7). However, the repeated use of antibiotics as the first-line treatment for UTI led to increased antibiotic-resistance,[8,9] highlighting the need for new strategies for the prevention and treatment of UTI with oral therapeutics (8-10).

The adhesin FimH, located on the tip of type 1 pili of UPEC, is a bacterial virulence factor (11-13). Its adherence to the highly mannosylated glycoprotein uroplakin 1a (UP1a) on urothelial cells of the host is the first step of the infection but also enables UPEC to evade clearance by the bulk flow of urine (14, 15). FimH consists of two domains; an N-terminal lectin domain ($FimH_{LD}$), being implicated as a mannose-specific carbohydrate recognition domain (CRD), and a C-terminal pilin domain ($FimH_{PD}$) anchoring the adhesion FimH to the pilus (16-18). For an efficient colonization of urothelial cells, bacteria only weakly interact with the cell surface, allowing exploration for optimal nutrition supply. However, to avoid clearance from the bladder when shear forces arise, strong adherence is required. To fulfill these two opposing tasks, FimH relies on a sophisticated allosteric mechanism, allowing a fine-tuning of its mannose binding affinity through conformational adaption. In the low-affinity state, there is a close interaction between $FimH_{PD}$ and $FimH_{LD}$, allosterically attenuating the mannose binding capacity. Shear forces induced by urination, provoke the separation of the two domains, which induces the high affinity state. This conformational switch of FimH is by no means a one-step process. Intermediate states and a high variability of the conformational ensemble have been described (19, 20). Although all conformational states of FimH are relevant for the pathogenicity of UPEC, efforts for the development of therapeutic FimH antagonists have mainly been focused on the high affinity conformation of the protein (21).

Within the last three decades, a large number of monovalent α-D-mannosides containing lipophilic aglycones, such as alkyl, phenyl, , diarylamides, dioxocyclo-butenylaminophenyl, umbelliferyl, biphenyl, indol(in)ylphenyl, triazolyl, and thiazolylamino have been synthesized (22-33). Moreover, a wide range of multivalent ligands has been described (34-40). Generally, FimH antagonists consist of a mannose moiety establishing an extended hydrogen bond network within the CRD and with alpha configuration a lipophilic aglycone interacting with the hydrophobic amino acids Tyr48, Tyr137 and Ile52 lining the entrance to the CRD (41).

In order to optimize complementary electrostatic properties to the electron rich tyrosine side chain, a series of electron deficient biphenyl aglycones were attached to mannose. To avoid unfavorable steric clashes, the substitution with fluorine atoms exhibiting a van der Waals radius comparable to hydrogen was explored (42). In case of hexafluorobenzene, the high electronegativity of fluorine induces an inductive polarization that leads 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 (-29.0×10^{-40} C·m^2). Such a complementary charge distribution results in a particularly strong electrostatic attraction (43-47).

The attractive interaction between electron-poor and electron-rich phenyl rings was demonstrated in experimental studies on the rotational barrier in 1,8-diarylnaphthalenes. Fluorine substitution lead to an increase of the energy barrier of approximately 2.1 kJ mol^{-1} per fluorine atom (48-50). Furthermore, studies on arene-arene interactions also indicated the beneficial influence of fluorine substitutions on favoring T-shaped or parallel displaced orientations rather than a sandwich conformation (51-53). The influence of fluoroarenes on protein binding affinity was demonstrated for different targets, including carbonic anhydrase II and the cysteine protease cathepsin L (54-59).

Based on a series of biphenyl α-D-mannosides (**1**–**3**, Figure 1), we aimed to improve affinity for all relevant conformational states of FimH by introducing different fluorination patterns on the terminal aromatic ring. In addition to perfluorination, as in **4** and **5**, a cyano substituent was introduced in the *para* position (→ **6**). This modification is marked by an increase in the Hammett constant, further increasing the electrostatic attraction to the Tyr48 side chain. The difluorinated compound **7** was designed to enable favorable σ–π interactions in an edge-to-face arrangement with the tyrosine gate residues.

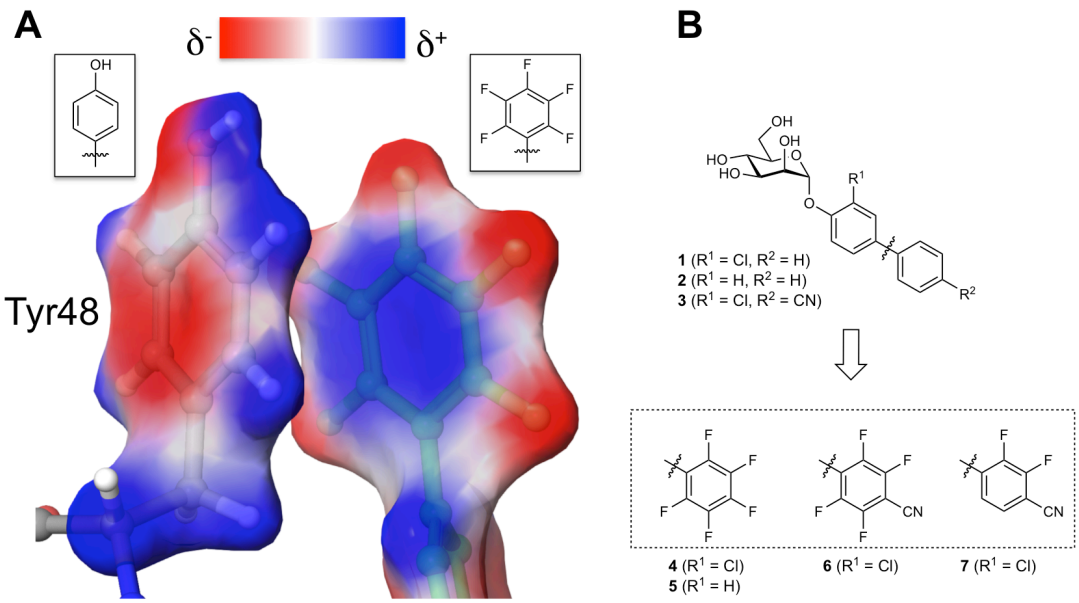
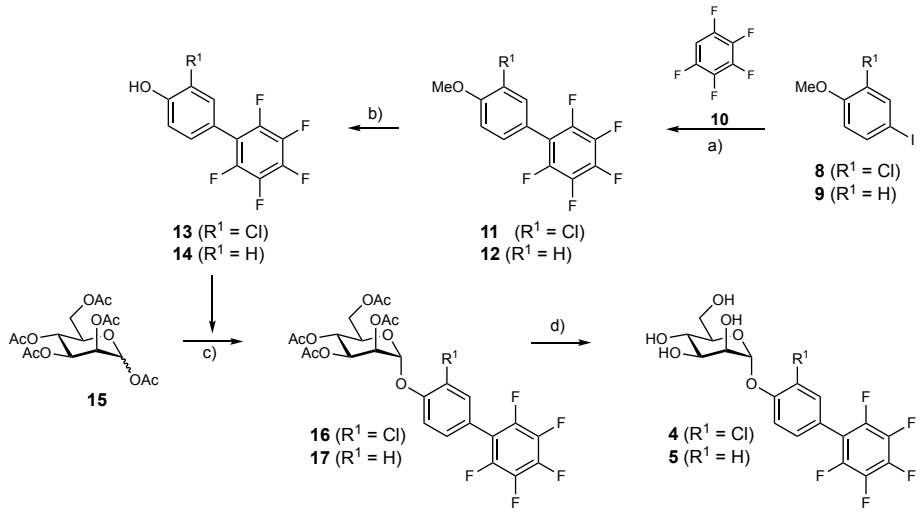


Figure 1 A Electrostatic potential visualized on the interaction of perfluorinated aglycone of mannoside **5** and Tyr48 of FimH. Regions of negative potential (red) of the perfluorinated aglycone overlap with regions of positive potential (blue) of the Tyr48 side chain and *vice versa*. **B** Four biphenyl α-D-mannosides with different fluorination patterns were synthesized.

Results and Discussion

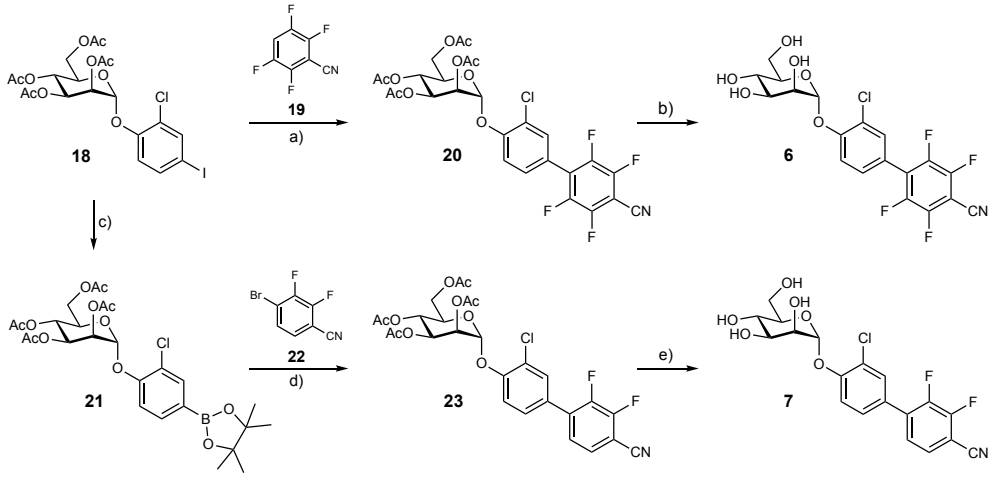
For the exploitation of π–π stacking interactions between the outer aromatic ring of biphenyl α-D-mannopyranosides and Tyr48 four representatives with different fluorination patterns were synthesized.

Synthesis. The synthesis of the biphenyl mannoside derivatives **4** and **5** is depicted in Scheme 1. Pentafluorinated biphenyls **13** and **14** were synthesized starting from anisole derivatives **8** and **9** in a copper-catalyzed arylation of pentafluorobenzene (**10**) followed by demethylation using boron tribromide (60). **13** and **14** were reacted in a $BF_3 \cdot Et_2O$-promoted mannosylation (→ **16** & **17**) followed by deacetylation under Zemplén conditions affording 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, DCM, -78 °C → rt, 5-19 h, 88% for **13**, 90% for **14**; c) $BF_3{\cdot}Et_2O$, DCM, 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 tetrafluorinated derivative **6** (Scheme 2), the previously described mannoside **18** (31) was coupled in a microwave-assisted arylation reaction in the presence of CuI and 1,10-phenantroline with commercially available 2,3,5,6-tetrafluorobenzonitrile (**19)** to give intermediate **20** in 66% yield. As a result of the high electrophilicity of the cyano group, by-products were formed almost exclusively under Zemplén conditions (NaOMe, MeOH) . When milder deacetylation conditions (aq. NH_3 in isopropanol/THF) were applied, by-product formation could not be completely avoided, but the test compound **6** was obtained in reasonable yield



Scheme 2 a) CuI, 1,10-phenantroline, K_3PO_4, *p*-xylene, microwave, 150 → 160 °C, 13 h, 66%; b) 2.0 M $NH_{3(aq)}$, *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 $NH_{3(aq)}$, *i*-PrOH, THF, rt, 4.5 d, 73%.

For the synthesis of the difluorinated derivative **7** (Scheme 2), mannoside **18** was transformed into the boronic acid pinacol ester **21** followed by Suzuki coupling (→ **23**) (61). By mild deprotection with aq. NH_3 compound **7** in good yield without formation of by-products.

FimH constructs. For the determination of binding affinities, two different recombinant FimH constructs were used. The isolated lectin domain of FimH ($FimH_{LD}$), which is locked in the high affinity conformation, was expressed as previously described (25). The full-length FimH variant ($FimH_{FL}$), representing the low-affinity conformational state, consists of the tethered lectin and pilin ($FimH_{PD}$) domains stabilized by a synthetic oligopeptide corresponding to residues 1-14 of the natural FimG donor strand (19).

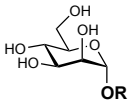
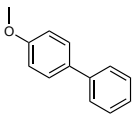
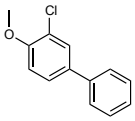
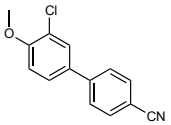
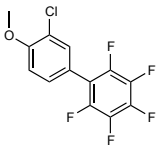
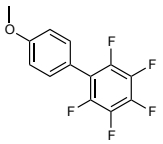
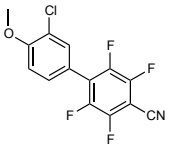
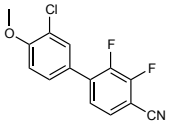
Competitive fluorescence polarization assay. In this assay, the antagonist of interest displaces a fluorescently labeled competitor from the binding site, thereby causing a reduction in fluorescence polarization (27).

***Full-length FimH*.** All fluorinated antagonists exhibited higher affinities to full-length FimH compared to their non-fluorinated counterparts (Table 1). While the introduction of a pentafluorophenyl ring led roughly to a 3-fold improvement of affinity compared to biphenyl α-D-mannoside (**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 improved affinity (**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. The Hammett constant (σ_{para}) reveals that the introduction of a cyano group in the *para*-position of the terminal phenyl ring results in a significantly higher electron-withdrawing potential than for a *para*-fluoro substituent (σ_{para} 0.660 vs. 0.062). Therefore, not surprisingly, the replacement of the *para*-fluoro with a *para*-cyano group led to a further substantial increase in the binding affinity (**4** → **6**, K_D 103 nM vs. 22 nM). Finally, the difluorinated cyano derivative **7** exhibits an activity almost identical to **6**. This compound was designed to capitalize on σ−π interactions resulting from an edge-to-face orientation to Tyr48, which has been shown to be energetically more favorable than the face-to-face conformation in the benzene homodimer (44), However, the strong polarization of the aromatic ring in **7** might create a larger dipole moment, resulting in a stronger interaction with the solvent compared to **6** and, thus, in a higher desolvation penalty. Therefore, the observed similarity of the affinity values between

6 and **7** can be a result of different factors compensating each other. In any case, the mannosides **6** & **7** are the best antagonists of the full-length FimH variant reported to date.

***Lectin domain*.** The fluorination of the terminal phenyl moiety in biphenyl α-D-mannoside (**1**), results in an almost 4-fold affinity gain for antagonist **5**. An additional chloro-substituent (**2** → **4**), however, surprisingly remained without an impact on binding affinity in the FP assay. However, for the corresponding *p*-cyano substituted inhibitors **3** and **6**, a 3-fold gain in potency can be observed. This results in a highly potent inhibitor with picomolar affinity. A similar effect was for the difluorinated derivative **7**. The gain in binding affinity for fluorinated compounds can most likely be attributed to enhanced π–π interactions to the tyrosine gate. It is not clear, however, why this trend is not observed for fluorination of **2**.

Table 1 Affinity of FimH antagonists to $FimH_{FL}$ and $FimH_{LD}$, physicochemical and pharmacokinetic parameters, and quantum mechanical calculation of the π–π stacking energies of Tyr48 with different biphenyl derivatives in the gas phase.

Cpd.		**$FimH_{FL}$** K_D [nM][a]	**$FimH_{LD}$** K_D [nM][a]	**Solubility** [µg/mL]	**log$D_{7.4}$** [d]	**PAMPA** logP_e [e]	**ΔE_{gas}** [kJ mol^{-1}]
1[28]		1'100	15.1	21 ± 1 [b]	2.1 ± 0.1	−4.7 ± 0.1	−19.8
2[64]		458	2.8	5.5 ± 0.2 [b]	2.6 ± 0.1	n.d.	−19.6
3[28]		59.1	1.06	195 ± 5 192 ± 5 [b]	2.1 ± 0.0	−5.2 ± 0.0	−34.3
4		103	4.48	47 ± 7 [c]	2.7 ± 0.2	−4.7 ± 0.2	−35.2
5		346	4.01	131 ± 10 [c]	2.3 ± 0.1	−4.7 ± 0.2	−34.4
6		22.1	0.349	111 ± 7 [c]	2.2 ± 0.1	−4.8 ± 0.1	−42.9
7		22.7	0.577	97 ± 15 [c]	2.2 ± 0.1	−5.0 ± 0.1	−39.8[f]/ −30.7[g]

[*a*] Dissociation constants (K_D) were determined in competitive fluorescence polarization assay towards $FimH_{FL}$ and $FimH_{LD}$ using fluorencent 3'-chloro-N-(2-(3-(3',6'-dihydroxy-3-oxo-3H-spiro- [isobenzofuran-1,9'-xanthen]-5-yl)thioureido)ethyl)-4'-(α-D-mannopyranosyloxy)biphenyl-4-carboxamide as reference compound (27). Experimental errors for measurements with new compounds are given in Table S1. [*b*] Kinetic solubility was measured in a 96-well format using the µSOL Explorer solubility analyzer at the indicated pH in triplicate. [*c*] Thermodynamic solubility was measured in Tris-HCl buffer (pH 7.4, 0.1 M) in duplicate, unless otherwise stated. [*d*] Octanol-water distribution coefficients (log$D_{7.4}$) were determined at pH 7.4 by a miniaturized shake flask procedure in sextuplicate. [*e*] 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. [*f*] ΔE_{gas} obtained with fluorine atoms on the same side as the chlorine atom. [*g*] ΔE_{gas} obtained with fluorine atoms on opposing side of chlorine.

Isothermal Titration Calorimetry (ITC). In order to reveal the influence of fluorination of the aglycone on the thermodynamic signature of binding, ITC experiments were performed (Table 2). In case of antagonists exhibiting an affinity in the low nanomolar or picomolar range (→ **6** & **7**), a competitive format described previously had to be applied (27).

Interaction of antagonists with the lectin domain. Dissociation constants for all antagonists follow the same trend observed in the competitive fluorescence polarization assay. The introduction of an *ortho*-chloro substituent on the phenyl ring adjacent to the anomeric position (**1** → **2**) improved the enthalpy $\Delta\Delta H°$ by −13.1 kJ mol^{-1}. Due to unoptimized π–π stacking of the unsubstituted biphenyl in mannoside **1**, any adjustment of the aglycone has a strong impact on enthalpy. However, the enthalpy gain observed for **2** versus **1** is partially compensated by an unfavorable entropy ($-T\Delta\Delta S°$ = 5.3 kJ mol^{-1}) resulting from a reduced flexibility of the protein–ligand complex (23). The introduction of the cyano group on the terminal aromatic ring (**2** → **3**) leads to a further enthalpy boost of $\Delta\Delta H°$ by −7.8 kJ mol^{-1} by a strengthened π–π interaction to Tyr48. The enthalpy gain $\Delta\Delta H°$ can be further enhanced by the addition of four fluorine substituents (**3** → **6)**, by −6.7 kJ mol^{-1}). Nevertheless, this enthalpy gain is again partly compensated by entropy costs related to a decreased mobility of the ligand and Tyr48. By contrast, the difluorinated cyano derivative **7** shows an enthalpic penalty $\Delta\Delta H°$ of 3.7 kJ mol^{-1} compared to **6**. This could be a consequence of the aforementioned polarization within the phenyl ring leading to a stronger attraction between aglycone and solvent, and thus to higher desolvation costs.

Interaction of antagonists with full-length FimH ($FimH_{FL}$). As observed for $FimH_{LD}$, the K_Ds of ITC measurements are in good agreement with the results from the fluorescent polarization assay. As earlier reported, the affinity of antagonists to $FimH_{FL}$ is again reduced by approximately two orders of magnitude (62). In contrast to measurements with $FimH_{LD}$, an astonishing compensation of enthalpic and entropic contributions is observed. In general, the enthalpy of binding is highly negative, while an entropic penalty mitigates the strength of the interaction. While effects like these are commonly observed for carbohydrate–lectin binding events, the magnitude of the recorded thermodynamic contributions exceeds well beyond the range that is usually reported for such interactions (63). A possible explanation is that the

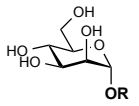
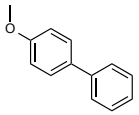
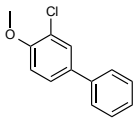
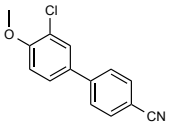
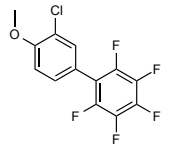
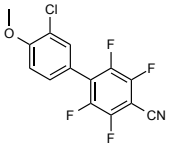
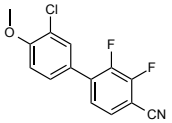
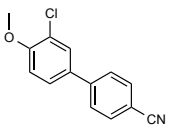
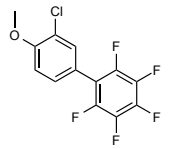
binding site of the inhibitor has a profound impact on the global conformational dynamics of the receptor, an assessment supported by large difference of thermodynamic parameters between measurements with the $FimH_{FL}$ and $FimH_{LD}$ constructs. These significant alterations (e.g. $\Delta\Delta H°_{FL\text{-}\mathbf{6}/LD\text{-}\mathbf{6}}$ = −27.4 kJ mol^{-1}; $-T\Delta\Delta S°_{FL\text{-}\mathbf{6}/LD\text{-}\mathbf{6}}$ = 39.1 kJ mol^{-1}) have to be attributed to changes distal from the binding site, which is structurally identical in the X-ray model of both constructs (19).

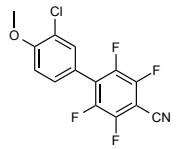
In general, FimH can exist in a flexible low- and a medium-affinity state, characterized by interdomain interactions between lectin and pilin domain. Upon separation of the two domains, the high-affinity state is formed (18-20). It has been reported recently, that binding to FimH is able to alter the equilibrium between the preexisting low- and medium-affinity conformational states and that this transition is likely associated with enthalpy–entropy compensation (20). As documented for inhibitor **6**, the thermodynamic signature of this binding event reveals a significant entropic penalty of $-T\Delta S°_{FL\text{-}\mathbf{4b}}$ = 53.9 kJ mol^{-1}, originating from a restriction in the mobility of the lectin domain as a whole and especially from a fixation of the clamp loop (residues 1–16) forming the mannose binding site. Only when shear forces separate lectin and pilin domains the high-affinity state is formed. The thermodynamic signature of binding to the high-affinity state can be studied with the $FimH_{LD}$ construct, which is locked in the high-affinity state. As expected, the entropy penalty is much smaller ($-T\Delta\Delta S°_{LD\text{-}\mathbf{6}\,/FL\text{-}\mathbf{6}}$ = −39.1 kJ mol^{-1}) because conformational changes are no longer necessary. However, the enthalpy gain ($\Delta\Delta H°_{LD\text{-}\mathbf{6}\,/FL\text{-}\mathbf{6}}$= −27.1 kJ mol^{-1}) associated with the conformational changes is much smaller as well.

As a result, the optimized π–π stacking to Tyr48 of fluorinated aglycones in combination with improved interactions of the clamp loop residues culminate in a highly favorable enthalpy term of $\Delta H°_{FL\text{-}\mathbf{6}}$ = −95.0 kJ mol^{-1}. While the effect of the tetrafluorcyano-substitution of the terminal aromatic ring can be rationalized by enhanced π–π stacking interactions for the high-affinity state ($\Delta\Delta H°_{LD\text{-}\mathbf{6}/LD\text{-}\mathbf{3}}$ = −6.7 kJ mol^{-1}; $-T\Delta\Delta S°_{LD\text{-}\mathbf{6}/LD\text{-}\mathbf{3}}$ = 4.2 kJ mol^{-1}), the observed thermodynamic fingerprint is much more pronounced for the full length construct ($\Delta\Delta H°_{FL\text{-}\mathbf{6}/FL\text{-}\mathbf{3}}$ = −19.1 kJ mol^{-1}; $-T\Delta\Delta S°_{FL\text{-}\mathbf{6}/FL\text{-}\mathbf{3}}$ = 17.3 kJ mol^{-1}). For the inhibitors **4** and **6**, the magnitude of enthalpy–entropy compensation correlates well with the electrostatic properties of the terminal aromatic ring (σ_{para} 0.062 vs. 0.660; $\Delta\Delta H°_{FL\text{-}\mathbf{4}/FL\text{-}\mathbf{6}}$ = 25.1 kJ mol^{-1}; $-T\Delta\Delta S°_{FL\text{-}\mathbf{4}/FL\text{-}\mathbf{6}}$ = −23.8 kJ mol^{-1}). In summary, this observation emphasizes that the improved stacking interactions

with Tyr48 translate to variations in the global arrangement of the protein by an allosteric mechanism.

Table 2. Thermodynamic profile of selected FimH antagonists binding to $FimH_{LD}$ and $FimH_{FL}$. The measurements were performed at 25 °C and pH 7.4.

Cpd.		K_D [nM][a]	$\Delta G°$ [kJ mol^{-1}]	$\Delta H°$ [kJ mol^{-1}]	$-T\Delta S°$ [kJ mol^{-1}]
LD-1[28]		17.7 (14.0 – 22.3)	−44.2	−45.0 (−45.6 – −44.5)	0.8
LD-2		5.8 (4.7 – 7.1)	−47.0	−53.1 (−53.5 – −52.7)	6.1
LD-3[28, b]		1.3 (1.1 – 1.6)	−50.7	−60.9 (−61.4 – −60.4)	10.1
LD-4[b]		3.1 (0.8 – 10.1	−48.6	−55.3 (−57.6 – −53.5)	6.8
LD-6[b]		0.5 (0.3 – 0.8)	−53.3	-67.6 (−67.1 – −66.0)	14.3
LD-7[b]		1.0 (0.6 – 1.7)	−51.3	−63.9 (−64.6 – −63.2)	12.7
FL-3[64]		130 (110 – 150)	−39.3	−75.9 (−77.0 – −74.9)	36.6
FL-4		98 (66 – 141)	−40.0	−69.6 (−73.7 – −65.9)	29.6

FL-6		64 (37 – 102)	−41.2	−97.7 (−100.2 - −90.5)	53.4

[*a*] 95% confidence interval from fitting in parentheses; [*b*] Δ*H°* obtained from direct titration; K_D obtained from competitive titration against *n*-heptyl 2-deoxy-α-D-mannoside.

X-ray Crystallography. To gain insight into the binding mode of the fluorinated FimH antagonists regarding the orientation of the aglycones towards the tyrosine gate motif, we co-crystallized the tetrafluorinated cyanide **6** and the pentafluorinated derivative **5** with $FimH_{LD}$ (Figure 2). Both structures were obtained at a resolution of 2.1 Å. The binding mode of the mannose moieties in **5** & **6** is identical to previously reported non-fluorinated analogues **1** (PDB 4X50) (23) & **3** (PDB 4CST) (27). In all crystal structures, the biphenyl aglycone is interacting with Tyr48 and Tyr137. While the orientation of the side chain of Tyr137 is not affected by different modifications of the biphenyl aglycone, substantial effects can be seen for the orientation of Tyr48. In the structure of **1** (Figure 2B), the Tyr48 side chain adopts a slightly tilted orientation, resulting in a suboptimal arrangement for π–π stacking interactions to the aglycone. Although the introduction of the electron-withdrawing cyano group (→ **3**, Figure 2D) 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 flexibility implicated from the observation of two distinct rotamers (27). By contrast, the terminal tetrafluorocyano phenyl ring in analogue **6** (Figure 2C) is in an optimal parallel alignment with Tyr48, 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** (Figure 2C). It is noteworthy that the aglycones of the co-crystallized antagonists **5** and **6** are involved in crystal packing contacts potentially influencing the binding modes. Therefore the binding poses in the crystal structures have to be validated by solution NMR spectroscopy.

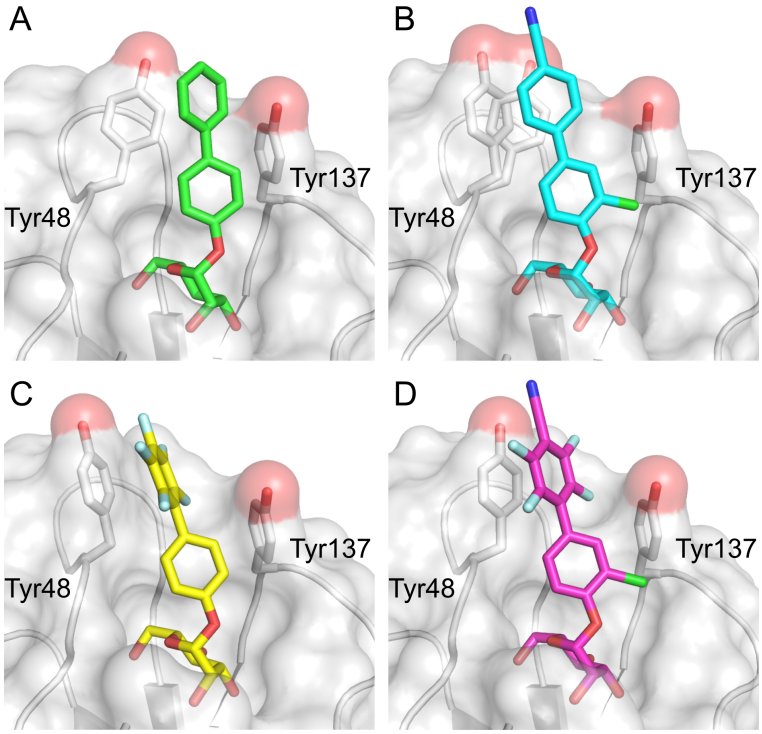


Figure 2 Comparison of the crystal structures of $FimH_{LD}$ co-crystallized with fluorinated antagonists. (A) Binding mode of non-fluorinated analogue **1** (green sticks, PDB 4X50). (B) Binding mode of non-fluorinated analogue **3** (cyan sticks, PDB 4CST). (C) Binding mode of **5** (yellow sticks). (D) Binding mode of **6** (magenta sticks).

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 $FimH_{LD}$ were measured upon addition of the fluorinated compounds **4**, **5** and **6** and their non-fluorinated counterparts **1** and **3** (23). Residues in the mannose binding pocket exhibit nearly identical chemical shifts for all five antagonists (Figure 3A) confirming the similar orientation and hydrogen bond network of their mannosyl moieties as observed in the X-ray structure shown in Figure 2. 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 (23). The formation of a T-shaped NH-π interaction of Thr51 H^N with the Tyr48 ring in complex with most other mannoside antagonists (e.g. with biphenyl mannoside **1**), thereby 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 3B). The $^{1}H,^{15}N$-HSQC spectra of $FimH_{LD}$ with **4**, **5** and **6** revealed the presence of this Tyr48 conformation in agreement with the X-ray structural data (Figure 2, **C** and **D**). A comparable signal shift in the $^{1}H,^{15}N$-HSQC spectrum has been observed for a previously published FimH ligand, which also showed strong π–π interactions with Tyr48 in

the X-ray structure (23). It is tempting to 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.

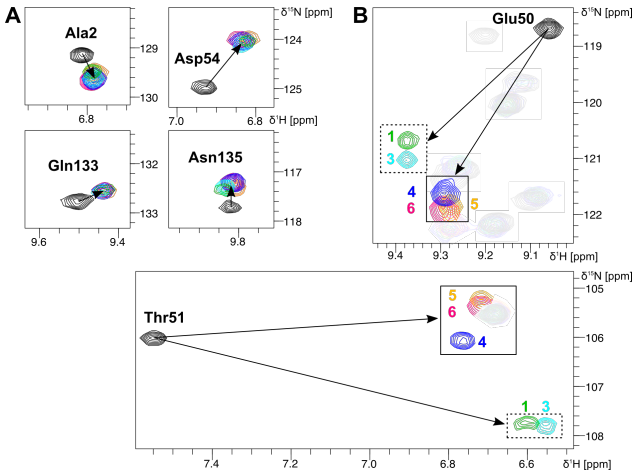


Figure 3 $^{1}H,^{15}N$-HSQC spectral regions of $FimH_{LD}$ with fluorinated compounds **4**, **5** and **6** and with their non-fluorinated counterparts **1** and **3**. A) Signal shift of four representative amino acids in the mannose binding site. B) Signal shift of Glu50 and Thr51.

***Ab initio* calculation of π-π stacking interactions.** The available X-ray structural data was used as input for the quantum mechanical calculation of π–π stacking interaction energies of different aglycones and Tyr48. Analysis of the interaction energies in Table 1 suggest that the π–π stacking is strongest between Tyr48 and the tetrafluorinated cyanide **6** (ΔE_{gas} = −42.9 kJ mol^{-1}) and weakest compounds **1** and **2** exhibiting unsubstituted terminal aromatic rings (ΔE_{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. The addition of the *p*-cyano group or the perfluorinated phenyl ring equally improve the π–π stacking energies as seen for compound **3**, **4** and **5** (ΔE_{gas} between −34.3 and −35.2 kJ mol^{-1}). However, the substitution of the *p*-fluoro with the *p*-cyano group **4a** (→ **6**) and thereby the increase of the Hammett constant of the para substituent, further increases the π–π stacking energy (ΔE_{gas} = −42.9 kJ mol^{-1}). Interestingly, the difluorinated cyanide **7** shows a comparable interaction energy when the two fluorine atoms are on the solvent exposed, coinciding side with the chlorine atom (ΔE_{gas} = −39.8 kJ mol^{-1}).

Physicochemical and *in vitro* pharmacokinetic characterization. To study the influence of polyfluorination of the aglycone on the physicochemical and pharmacokinetic parameters of FimH antagonists, the aqueous solubility and lipophilicity of the new antagonists were determined (64, 65). Furthermore, the permeability through an artificial membrane (PAMPA)

was determined as an estimate for oral bioavailability (66). The results are summarized in Table 1.

Within the non-fluorinated series, introduction of an *o*-chloro substituent (→ **2**) followed by addition of a *p*-cyano group (→ **3**) led to a 4-fold decrease and 9-fold increase, respectively, of the aqueous solubility compared to **1**. While the same *o*-chloro substitution within the fluorinated series (→ **4**) led to a similar 3-fold reduction of the aqueous solubility, no change in this parameter was observed compared to **5** upon replacing the *p*-fluoro by a *p*-cyano group (→ **6**). Since the aqueous solubility of **1** and **2** was determined in a kinetic solubility assay, this parameter cannot be directly compared with the thermodynamic solubility measured for their fluorinated analogues **4** and **5**. By contrast, the thermodynamic solubility was measured for all cyanides (**3**, **6** and **7**) allowing for a detailed comparison. Tetrafluorination of cyanide **3** (→ **6**) had a negative impact on the solubility reducing it 2-fold to 111 µg/mL. Interestingly, the difluorinated analogue **7**, despite a higher predicted polarization of the terminal phenyl ring, had essentially the same solubility as **6**. Nevertheless, the solubility of antagonists **6** and **7** lies above the minimal acceptable value for complete absorption of a 1 mg/kg dose of moderately permeable compounds, estimated by Lipinski *et al.* to be 52 µg/mL (67).

All fluorinated antagonists showed slightly higher lipophilicity than their non-fluorinated equivalents. Interestingly, difluorinated cyanide **7** had the same log$D_{7.4}$ value as the tetrafluorinated **6**. With a lipophilicity between 2.2 and 2.7, these molecules are expected to be well absorbed by means of passive diffusion upon oral administration (68). Moreover, a high lipophilicity supports tubular reabsorption from the glomerular filtrate favoring slower renal clearance and thus a prolonged therapeutic effect in the bladder upon a single dose (69). The increased lipophilicity of the fluorinated antagonists is also reflected in higher effective permeabilities (log P_e) determined in PAMPA. Since a logP_e values above -5.7 indicates a high potential for intestinal absorption, all tested compounds are envisaged to easily permeate the membranes lining the small intestine (70).

Conclusions

Over the years, the high-affinity conformation of FimH present in the isolated $FimH_{LD}$ was the main target in the development of anti-adhesive medicine for UTI neglecting the equally relevant low-affinity state observed in the full-length FimH. A recently published study addressed that issue and showed that even the most potent FimH antagonists bind with around two orders of magnitude lower affinity to $FimH_{FL}$ than to $FimH_{LD}$ (62). With such a need for improvement of efficacy against the low-affinity conformation of FimH (71), we designed and synthesized a series of new FimH antagonists bearing fluorinated aromatic aglycone with the goal to enhance the π−π stacking with Tyr48 located at the entrance of the binding site. The pentafluorinated analogue of biphenyl α-D-mannoside (→ **5**) shows improved binding to full-length FimH. An additional introduction of a *p*-cyano group (→ **6**) results in a dissociation constant of 22.1 nM, the highest reported affinity of an FimH antagonist for $FimH_{FL}$ to date. In case of $FimH_{LD}$, a similar improvement was observed in almost all fluorinated antagonists, with **6** & **7** even reaching subnanomolar affinities. ITC measurements of fluorinated antagonists with $FimH_{LD}$ reveal a marked increase in enthalpy with the highest gain observed for the tetrafluorinated derivative **6**. The thermodynamic signature of the interaction of **6** with the full-length construct reveals important details about the global protein dynamics upon ligand binding. Given that $FimH_{FL}$ is predicted to exist in an equilibrium of low- and high-affinity state in solution (20), the distinct loss of entropy of **6** may indicate a stabilization of a conformation that is characterized by extensive interactions between the lectin and pilin domains and hence fewer degrees of conformational freedom. Furthermore, the strong stacking interaction of the aglycone with Tyr48 leads to a restriction of the mobility of the clamp loop of FimH. Finally, the crystal structures of $FimH_{LD}$ with **5** and **6** reveal a parallel alignment of the aromatic rings of Tyr48 and the aglycone, suggesting particularly strong π−π stacking. This orientation was confirmed by chemical shift perturbation NMR experiments of $FimH_{LD}$ with **4**, **5** and **6** in solution monitoring the backbone signals of the loop residues Glu50 and Thr51 as reporters for the tyrosine gate conformation. This allowed for the quantum mechanical calculation of π−π stacking energies that clearly separate the different substitutions on the terminal ring and revealed its importance for the stacking interaction energies.

The physicochemical and pharmacokinetic parameters solubility, lipophilicity, and permeability were determined to estimate the oral bioavailability of the new FimH antagonists. While the fluorination of the terminal phenyl group had a negative effect on the

aqueous solubility, most inhibitors are still soluble enough (> 52 µg/mL) for oral administration. By contrast, lipophilic modifications introduced on the aglycone were advantageous for permeability indicating high intestinal absorption of all new antagonists. Based on the experimentally determined log$D_{7.4}$ values, a slow renal elimination due to tubular reabsorption and thus, a longer lasting therapeutic effect of a single dose of FimH antagonist, can be expected.

In summary, the introduction of multiple fluorine substituents to the terminal phenyl rings of biphenyl mannoside-type inhibitors of FimH results in improved π–π stacking interactions to Tyr48. Antagonists **6** & **7** display the best pharmacodynamic profile towards full-length FimH as well as $FimH_{LD}$ and nearly optimal pharmacokinetic parameters for oral bioavailability. In future work, we aim to evaluate *in vivo* pharmacological parameters of these promising drug candidates in a mouse model of UTI.

Experimental Part

Synthesis

General methods: NMR spectra were recorded on a Bruker Avance III 500 MHz NMR spectrometer. Assignment of ^{1}H and ^{13}C NMR spectra was achieved using 2D methods (COSY, HSQC, HMBC). Chemical shifts are referenced to residual $CHCl_3$, CHD_2OD or HDO. Optical rotations were measured using Perkin-Elmer Polarimeter 341. Electron spray ionization mass spectra (ESI-MS) were obtained on a Waters micromass ZQ Mass Spectrometer. The LC-HRMS analysis were carried out using a Agilent 1100 LC equipped with a photodiode array detector and a Micromass QTOF I equipped with a 4 GHz digital-time converter. Microwave-assisted reactions were carried out with CEM Discover and Explorer. Reactions were monitored by TLC using glass plates coated with silica gel 60 F_{254} (Merck) 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). MPLC separations were carried out on a CombiFlash Companion or R_f from Teledyne Isco equipped with RediSep normal-phase or RP-18 reversed-phase flash columns. LC–MS separations were carried out on a Waters system equipped with sample manager 2767, pump 2525, PDA 2996, column SunFire™ Prep C18 OBD™ (5 μm, 19 x 150 mm), and Micromass ZQ. All compounds used for biological assays are of 100% purity based on HPLC analytical results. Commercially available reagents were purchased from Sigma-Aldrich, Alfa Aesar and Acros Organics. Dry DMF was purchased from Acros Organics. All the other solvents were dried prior to use where indicated. Methanol (MeOH) was dried by storing over activated molecular sieves 3Å for at least one day. Dichloromethane (DCM) was dried by filtration over Al_2O_3 (Fluka, type 5016 A basic) and stored over activated molecular sieves 4Å. Molecular sieves 3Å and 4Å were activated in vacuo at 200 °C for 30 min immediately before use.

3-Chloro-2',3',4',5',6'-pentafluoro-4-methoxy-biphenyl (11). To a solution of **8** (320 mg, 1.192 mmol) and **10** (200 μL, 1.788 mmol, 1.5 eq) in dry *p*-xylene/DMF (4 mL, 1:1) was added 1,10-phenanthroline (71 mg, 0.358 mmol, 0.3 eq) followed by CuI (68 mg, 0.358 mmol, 0.3 eq) and dry K_3PO_4 (506 mg, 2.334 mmol, 2.0 eq). The mixture was degassed in an ultrasonic bath and flushed with argon. After stirring for 10 h at 130 °C, the mixture was diluted with EtOAc (50 mL), washed with H_2O (2 x 20 mL) and brine (20 mL). The organic layer was dried over Na_2SO_4 and concentrated in vacuo. The residue was purified by MPLC on silica gel (petroleum ether) to yield **11** (132 mg, 36%) as a white solid. ^{1}H NMR (500 MHz, $CDCl_3$): δ = 7.46 (m, 1H, Ar-H), 7.31 (m, 1H, Ar-H), 7.04 (d, *J* = 8.6 Hz, 1H, Ar-H), 3.96 ppm (s, 3H, OCH_3); ^{13}C NMR (125 MHz, $CDCl_3$): δ = 155.89 (Ar-C), 144.15, 140.69, 137.79 (m, 5C, 5 Ar-CF), 131.92, 129.89, 123.04, 119.30 (4 Ar-C), 114.64 (dt, *J* = 3.9, 16.8 Hz, Ar-C), 56.37 ppm (OCH_3).

2',3',4',5',6'-Pentafluoro-4-methoxy-biphenyl (12). To a solution of **9** (187 mg, 0.799 mmol) and **10** (133 μL, 1.199 mmol, 1.5 eq) in dry DMF (0.5 mL) under argon was added 1,10-phenanthroline (16 mg, 0.080 mmol, 0.1 eq) followed by CuI (15 mg, 0.080 mmol, 0.1 eq) and dry K_3PO_4 (339 mg, 1.598 mmol, 2.0 eq). After stirring for 2 h at 130 °C, DMF was co-evaporated with xylene to dryness in vacuo. The residue was purified by MPLC on silica gel (petroleum ether) to yield **12** (100 mg, 46%) as a white solid. Analytical data are with accordance with the literature data (72).

3-Chloro-2',3',4',5',6'-pentafluoro-4-hydroxy-biphenyl (13). To a solution of **11** (132 mg, 0.427 mmol) in CH_2Cl_2 (6 mL) was added dropwise a 1.0 M solution of BBr_3 in CH_2Cl_2 (1.00 mL, 1.0 mmol, 2.3 eq) at -78 °C under argon. The content of the flask was slowly warmed to ambient temperature and stirred for 19 h. Then, the mixture was cooled to 0 °C and carefully quenched by slow addition of H_2O. The mixture was diluted with EtOAc (50 mL), washed with H_2O (2 x 20 mL) and brine (20 mL). The organic

layer was dried over Na_2SO_4 and concentrated in vacuo. The residue was purified by MPLC on silica gel (petroleum ether/EtOAc, 1:0-9:1) to yield **13** (110 mg, 88%) as a white solid. ^{1}H NMR (500 MHz, $CDCl_3$): δ = 7.42 (s, 1H, Ar-H), 7.27 (m, 1H, Ar-H), 7.14 (d, *J* = 8.5 Hz, 1H, Ar-H), 5.76 ppm (s, 1H, OH); ^{13}C NMR (125 MHz, $CDCl_3$): δ = 152.35 (Ar-C), 144.08, 140.69, 137.85 (m, 5C, 5 Ar-CF), 130.78, 130.58, 120.43, 119.54, 116.72 (5 Ar-C), 114.57 ppm (dt, *J* = 4.1, 17.0 Hz, Ar-C); ESI-MS: *m*/*z*: Calcd for $C_{12}H_3ClF_5O$ [M-H]$^-$: 293.0, found: 293.0.

2',3',4',5',6'-Pentafluoro-4-hydroxy-biphenyl (14). To a solution of **12** (100 mg, 0.365 mmol) in CH_2Cl_2 (3 mL) was added dropwise a 1.0 M solution of BBr_3 in CH_2Cl_2 (0.910 mL, 0.910 mmol, 2.5 eq) at -78 °C under argon. The content of the flask was slowly warmed to ambient temperature and stirred for 5 h. Then, the mixture was cooled to 0 °C and carefully quenched by slow addition of H_2O. The mixture was extracted with CH_2Cl_2 (3 x 20 mL). The organic layer was dried over Na_2SO_4 and concentrated in vacuo. The residue was purified by MPLC on silica gel (hexane/EtOAc, 1:0-9:1) to yield **14** (85 mg, 90%) as a white solid. Analytical data are with accordance with the literature data (73).

3-Chloro-2',3',4',5',6'-pentafluoro-biphenyl-4-yl 2,3,4,6-tetra-*O*-acetyl-α-D-mannopyranoside (16). To a suspension of **15** (83 mg, 0.212 mmol) and **13** (0.149 mmol, 0.7 eq) in dry CH_2Cl_2 (4 mL), $BF_3{\cdot}Et_2O$ (80 µL, 0.636 mmol, 3.0 eq) was added dropwise under argon. After stirring for 3 h at 40 °C, the reaction mixture was diluted with EtOAc (50 mL), washed with satd aq $NaHCO_3$ (20 mL), H_2O (20 mL) and brine (20 mL). The organic layer was dried over Na_2SO_4 and concentrated in vacuo. The residue was purified by MPLC on silica gel (petroleum ether/EtOAc, 1:0-3:2) to yield **16** (37 mg, 40%) as colorless oil. $[\alpha]_D^{20}$ +57.3 (*c* 0.74, $CHCl_3$); ^{1}H NMR (500 MHz, $CDCl_3$): δ = 7.49 (br s, 1H, Ar-H), 7.31-7.28 (m, 2H, Ar-H), 5.64-5.61 (m, 2H, H-1, H-3), 5.54 (dd, *J* = 1.9, 3.4 Hz, 1H, H-2), 5.39 (t, *J* = 10.1 Hz, 1H, H-4), 4.28 (dd, *J* = 5.5, 12.2 Hz, 1H, H-6a), 4.15 (ddd, *J* = 2.2, 5.5, 10.1 Hz, 1H, H-5), 4.11 (dd, *J* = 2.2, 12.2 Hz, 1H, H-6b), 2.21, 2.07, 2.04, 2.02 ppm (4 s, 12H, 4 $COCH_3$); ^{13}C NMR (125 MHz, $CDCl_3$): δ = 170.57, 170.07, 169.89 (4C, 4 CO), 151.96 (Ar-C), 144.15, 140.67, 137.65 (m, 5C, 5 Ar-CF), 132.32, 129.76, 124.77, 122.09, 116.78 (5 Ar-C), 114.19 (dt, *J* = 4.1, 17.0 Hz, Ar-C), 96.53 (C-1), 70.07 (C-5), 69.36 (C-2), 68.82 (C-3), 65.90 (C-4), 62.20 (C-6), 20.97, 20.82, 20.79, 20.72 ppm (4 CO*C*H$_3$); ESI-MS: *m*/*z*: Calcd for $C_{26}H_{22}ClF_5NaO_{10}$ [M+Na]$^+$: 647.1, found: 647.2.

2',3',4',5',6'-Pentafluoro-biphenyl-4-yl 2,3,4,6-tetra-*O*-acetyl-α-D-mannopyranoside (17). To a suspension of **15** (30 mg, 0.077 mmol) and **14** (0.077 mmol, 1.0 eq) in dry CH_2Cl_2 (1 mL), $BF_3{\cdot}Et_2O$ (10 µL, 0.077 mmol, 1.0 eq) was added dropwise under argon. After stirring for 1 h at 40 °C, another portion of $BF_3{\cdot}Et_2O$ (10 µL, 0.077 mmol, 1.0 eq) was added followed by another 2 h later (1.0 eq). After stirring for 3 h, the reaction mixture was diluted with EtOAc (50 mL), washed with satd aq $NaHCO_3$ (20 mL) and H_2O (20 mL). The organic layer was dried over Na_2SO_4 and concentrated in vacuo. The residue was purified by MPLC on silica gel (petroleum ether/EtOAc, 1:0-7:3) to yield **17** (19 mg, 42%) as colorless oil. $[\alpha]_D^{20}$ +68.6 (*c* 0.95, $CHCl_3$); ^{1}H NMR (500 MHz, $CDCl_3$): δ = 7.38 (d, *J* = 8.6 Hz, 2H, Ar-H), 7.21 (d, *J* = 8.8 Hz, 2H, Ar-H), 5.60 (d, *J* = 1.2 Hz, 1H, H-1), 5.57 (dd, *J* = 3.5, 10.1 Hz, 1H, H-3), 5.46 (dd, *J* = 1.7, 3.3 Hz, 1H, H-2), 5.38 (t, *J* = 10.0 Hz, 1H, H-4), 4.28 (dd, *J* = 6.1, 12.7 Hz, 1H, H-6a), 4.10 (m, 2H, H-5, H-6b), 2.21, 2.06, 2.05, 2.02 ppm (4 s, 12H, 4 $COCH_3$); ^{13}C NMR (125 MHz, $CDCl_3$): δ = 170.64, 170.10, 170.09, 169.87 (4 CO), 156.28 (Ar-C), 145.25-136.85 (m, 5C, 5 Ar-CF), 131.74, 121.03, 116.85 (6C, 6 Ar-C), 95.80 (C-1), 69.55 (C-5), 69.44 (C-2), 68.94 (C-3), 66.05 (C-4), 62.24 (C-6), 21.00, 20.83, 20.82, 20.73 ppm (4 CO*C*H$_3$); ESI-MS: *m*/*z*: Calcd for $C_{26}H_{23}F_5NaO_{10}$ [M+Na]$^+$: 613.1, found: 613.1.

3-Chloro-2',3',4',5',6'-pentafluoro-biphenyl-4-yl α-D-mannopyranoside (4). To a solution of **16** (37 mg, 0.059 mmol) in dry MeOH (6 mL) was added 0.3 M MeONa/MeOH (500 µL, 2.5 eq) at rt under argon. The mixture was stirred for 2 h and then neutralized with AcOH (20 µL) and evaporated to

dryness in vacuo. The crude product was purified by MPLC on silica gel (DCM/MeOH, 1:0-9:1) to afford **4** (20 mg, 74%) as a white solid. $[\alpha]_D^{20}$ +79.1 (*c* 1.00, MeOH); ^{1}H NMR (500 MHz, CD_3OD): δ = 7.54-7.52 (m, 2H, Ar-H), 7.39 (m, 1H, Ar-H), 5.65 (d, *J* = 1.7 Hz, 1H, H-1), 4.12 (dd, *J* = 1.9, 3.4 Hz, 1H, H-2), 4.00 (dd, *J* = 3.4, 9.5 Hz, 1H, H-3), 3.81-3.70 (m, 3H, H-4, H-6a, H-6b), 3.63 ppm (ddd, *J* = 2.3, 5.6, 9.7 Hz, 1H, H-5); ^{13}C NMR (125 MHz, CD_3OD): δ = 154.04 (Ar-C), 145.44, 142.03, 139.09 (m, 5C, 5 Ar-CF), 132.87, 131.19, 124.96, 122.26, 118.09 (5 Ar-C), 115.86 (dt, *J* = 3.8, 17.4 Hz, Ar-C), 100.64 (C-1), 76.13 (C-5), 72.37 (C-3), 71.76 (C-2), 68.20 (C-4), 62.66 ppm (C-6)

2',3',4',5',6'-Pentafluoro-biphenyl-4-yl α-D-mannopyranoside (5). To a solution of **17** (19 mg, 0.032 mmol) in dry MeOH (2 mL) was added 1.0 M MeONa/MeOH (50 μL, 1.6 eq) at rt under argon. The mixture was stirred for 2 h and then neutralized with AcOH (10 μL) and evaporated to dryness in vacuo. The crude product was purified by MPLC on silica gel (DCM/MeOH, 1:0-9:3) to afford **5** (12 mg, 89%) as a white solid. $[\alpha]_D^{20}$ +97.3 (*c* 0.60, MeOH); ^{1}H NMR (500 MHz, CD_3OD): δ = 7.41 (d, *J* = 8.5 Hz, 2H, Ar-H), 7.27 (d, *J* = 8.8 Hz, 2H, Ar-H), 5.57 (d, *J* = 1.2 Hz, 1H, H-1), 4.04 (dd, *J* = 1.7, 3.2 Hz, 1H, H-2), 3.92 (dd, *J* = 3.4, 9.4 Hz, 1H, H-3), 3.80-3.71 (m, 3H, H-4, H-6a, H-6b), 3.61 ppm (ddd, *J* = 2.2, 5.2, 9.5 Hz, 1H, H-5); ^{13}C NMR (125 MHz, CD_3OD): δ = 158.71 (Ar-C), 146.48-138.12 (m, 5C, 5 Ar-CF), 132.66, 121.23, 117.92 (5C, 5 Ar-C), 117.19 (m, Ar-C), 100.10 (C-1), 75.61 (C-5), 72.38 (C-3), 71.90 (C-2), 68.33 (C-4), 62.70 ppm (C-6);

4'-(2,3,4,6-Tetra-*O*-acetyl-α-D-mannopyranosyloxy)-3'-chloro-2,3,5,6-tetrafluoro-biphenyl-4-carbonitrile (20). To a microwave tube with **18** (300 mg, 0.513 mmol) and **19** (135 mg, 0.770 mmol, 1.5 eq), dry *p*-xylene (4 mL) was added under argon followed by the addition of 1,10-phenanthroline (51 mg, 0.257 mmol, 0.5 eq), CuI (49 mg, 0.257 mmol, 0.5 eq) and dry K_3PO_4 (218 mg, 1.027 mmol, 2.0 eq). The tube was sealed with a Teflon septum, evacuated through a needle, and flushed with argon. Then, the mixture was degassed in an ultrasonic bath for 10 min, flushed again with argon and exposed to microwave irradiation at 150 °C for 8 h. Then, additional potions of **19** (45 mg, 0.257 mmol, 0.5 eq), 1,10-phenanthroline (17 mg, 0.086 mmol, 0.2 eq), CuI (16 mg, 0.084 mmol, 0.2 eq) and dry K_3PO_4 (72 mg, 1.027 mmol, 0.3 eq) were added. Microwave irradiation was continued for another 5 h at 160 °C. The mixture was diluted with EtOAc (50 mL), washed with H_2O (2 x 30 mL) and brine (20 mL). The organic layer was dried over Na_2SO_4 and concentrated in vacuo. The residue was purified by MPLC on silica gel (petroleum ether/EtOAc, 1:0-7:3) to yield **20** (215 mg, 66%) as colorless oil. $[\alpha]_D^{20}$ +76.6 (*c* 1.05, $CHCl_3$); ^{1}H NMR (500 MHz, $CDCl_3$): δ = 7.53 (br s, 1H, Ar-H), 7.33 (s, 2H, Ar-H), 5.64 (d, *J* = 1.6 Hz, 1H, H-1), 5.59 (dd, *J* = 3.5, 10.0 Hz, 1H, H-3), 5.51 (dd, *J* = 1.8, 3.4 Hz, 1H, H-2), 5.38 (t, *J* = 10.0 Hz, 1H, H-4), 4.25 (dd, *J* = 5.2, 12.0 Hz, 1H, H-6a), 4.13-4.07 (m, 2H, H-5, H-6b), 2.19, 2.05, 2.03, 2.00 ppm (4 s, 12H, 4 $COCH_3$); ^{13}C NMR (125 MHz, $CDCl_3$): δ = 170.45, 169.99, 169.82, 169.78 (4 CO), 152.69 (Ar-C), 147.21 (m, 2C, 2 Ar-CF), 144.01 (m, 2C, 2 Ar-CF), 132.10, 129.65 (2 Ar-C), 125.20 (t, *J* = 16.1 Hz, Ar-C), 124.93, 121.16, 116.68 (3 Ar-C), 107.40 (t, *J* = 3.3 Hz, CN), 96.40 (C-1), 93.39 (t, *J* = 17.1 Hz, Ar-C), 70.09 (C-5), 69.19 (C-2), 68.69 (C-3), 65.74 (C-4), 62.08 (C-6), 20.87, 20.73, 20.71, 20.63 ppm (4 $COCH_3$); IR (neat): ν = 2247 (C≡N) cm^{-1}; ESI-MS: *m/z*: Calcd for $C_{27}H_{22}ClF_4NNaO_{10}$ $[M+Na]^+$: 654.1, found: 654.2.

3'-Chloro-2,3,5,6-tetrafluoro-4'-(α-D-mannopyranosyloxy)-biphenyl-4-carbonitrile (6). To a suspension of **20** in *i*-PrOH (3.0 mL) was added 2.0 M $NH_{3(aq)}$ (3.0 mL) and THF (1.5 mL). After 24 h of stirring at rt, the solvent was evaporated in vacuo. The residue was purified by MPLC on silica gel (DCM/MeOH, 1:0-19:1) followed by MPLC on RP-18 (H_2O/MeOH, 1:0-0:1) to yield **6** (10,8 mg, 39%) as a white solid. $[\alpha]_D^{20}$ +77.0 (*c* 1.08, MeOH); ^{1}H NMR (500 MHz, CD_3OD): δ = 7.63 (br s, 1H, Ar-H), 7.56 (d, *J* = 8.7 Hz, 1H, Ar-H), 7.47 (br d, *J* = 8.6 Hz, 1H, Ar-H), 5.68 (d, *J* = 1.6 Hz, 1H, H-1), 4.12 (dd, *J* = 1.9, 3.4 Hz, 1H, H-2), 4.00 (dd, *J* = 3.4, 9.5 Hz, 1H, H-3), 3.80-3.70 (m, 3H, H-4, H-6a, H-6b), 3.61 ppm (ddd, *J* =

2.3, 5.6, 9.7 Hz, 1H, H-5); ^{13}C NMR (125 MHz, CD_3OD): δ = 154.67 (Ar-C), 148.95 (m, 2C, 2 Ar-CF), 145.31 (m, 2C, 2 Ar-CF), 132.79, 131.17 (2 Ar-C), 126.69 (m, Ar-C), 125.07, 121.75, 118.00 (3 Ar-C), 108.53 (m, CN), 100.58 (C-1), 94.12 (m, Ar-C), 76.21 (C-5), 72.36 (C-3), 71.71 (C-2), 68.18 (C-4), 62.65 ppm (C-6); IR (KBr): ν = 2247 (C≡N) cm^{-1}.

2-Chloro-4-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)phenyl 2,3,4,6-tetra-*O*-acetyl-α-D-mannopyranosyloxy (21). A microwave tube was charged with **18** (100 mg, 0.171 mmol), bis(pinacolato)diborone (52 mg, 0.205 mmol, 1.2 eq) and KOAc (68 mg, 0.688 mmol, 3.0 eq). The tube was sealed with a Teflon septum, evacuated through a needle, and flushed with argon. Then anhydrous DMF (2 mL) was added and the mixture was degassed in an ultrasonic bath for 10 min and flushed again with argon. Then, $PdCl_2(dppf)\cdot CH_2Cl_2$ (5.6 mg, 6.9 μmol, 0,07 eq) was added and the mixture was exposed to microwave irradiation at 120 °C for 2 h. The mixture was diluted with EtOAc (50 mL), washed with H_2O (2 x 20 mL) and brine (20 mL). The organic layer was dried over Na_2SO_4 and concentrated in vacuo. The residue was purified by MPLC on silica gel (petroleum ether/EtOAc, 1:0-2:3) to yield **21** (61 mg, 61%) as colorless oil. $[\alpha]_D^{20}$ +63.4 (*c* 0.56, $CHCl_3$); ^{1}H NMR (500 MHz, $CDCl_3$): δ = 7.83 (d, *J* = 1.5 Hz, 1H, Ar-H), 7.63 (dd, *J* = 1.5, 8.2 Hz, 1H, Ar-H), 7.15 (d, *J* = 8.2 Hz, 1H, Ar-H), 5.63-5.59 (m, 2H, H-1, H-3), 5.53 (dd, *J* = 1.9, 3.4 Hz, 1H, H-2), 5.38 (t, *J* = 10.1 Hz, 1H, H-4), 4.27 (dd, *J* = 5.3, 12.3 Hz, 1H, H-6a), 4.13 (ddd, *J* = 2.3, 5.3, 10.2 Hz, 1H, H-5), 4.04 (dd, *J* = 2.3, 12.3 Hz, 1H, H-6b), 2.20, 2.06, 2.03 (4 s, 12H, 4 $COCH_3$), 1.33 ppm (s, 12H, 4 CCH_3); ^{13}C NMR (125 MHz, $CDCl_3$): δ = 170.63, 170.06, 169.92, 169.86 (4C, 4 CO), 153.48, 137.11, 134.52, 124.02, 116.09 (6C, Ar-C), 96.34 (C-1), 84.26 (2C, 2 *C*$(CH_3)_2$), 69.87 (C-5), 69.44 (C-2), 68.92 (C-3), 65.93 (C-4), 62.17 (C-6), 24.99, 24.97 (4C, 2 C(*C*$H_3)_2$), 21.00, 20.85, 20.81, 20.80 ppm (4 CO*C*H_3); ESI-MS: *m*/*z*: Calcd for $C_{26}H_{34}BClNaO_{12}$ $[M+Na]^+$: 607.2, found: 607.2.

4'-(2,3,4,6-Tetra-*O*-acetyl-α-D-mannopyranosyloxy)-3'-chloro-2,3-difluoro-biphenyl-4-carbonitrile (23). A round-bottom flask was charged with **22** (61 mg, 0.104 mmol), boronate **21** (30 mg, 0.136 mmol, 1.3 eq) and K_3PO_4 (66 mg, 0.312 mmol, 3.0 eq), then evacuated and flushed with argon. Anhydrous DMF (4 mL) was added and the mixture was degassed in an ultrasonic bath for 10 min and flushed with argon followed by the addition of $PdCl_2(dppf)\cdot CH_2Cl_2$ (4 mg, 5.2 μmol, 0.03 eq). After stirring for 2.5 h at 90 °C, the mixture was diluted with EtOAc (50 mL) and washed with H_2O (20 mL) and brine (20 mL). The organic layer was dried over Na_2SO_4 and concentrated in vacuo. The residue was purified by MPLC on silica gel (petroleum ether/EtOAc, 1:0-1:1) to yield **23** (36 mg, 58%) as colorless oil. $[\alpha]_D^{20}$ +70.8 (*c* 0.90, $CHCl_3$); ^{1}H NMR (500 MHz, $CDCl_3$): δ = 7.60 (br s, 1H, Ar-H), 7.45 (m, 1H, Ar-H), 7.41 (br d, *J* = 8.6 Hz, 1H, Ar-H), 7.30-7.28 (m, 2H, Ar-H), 5.63-5.60 (m, 2H, H-1, H-3), 5.53 (m, 1H, H-2), 5.40 (t, *J* = 10.1 Hz, 1H, H-4), 4.28 (dd, *J* = 5.2, 12.2 Hz, 1H, H-6a), 4.14 (ddd, *J* = 2.0, 5.2, 10.0 Hz, 1H, H-5), 4.10 (dd, *J* = 2.0, 12.2 Hz, 1H, H-6b), 2.21, 2.07, 2.04, 2.03 ppm (4 s, 12H, 4 $COCH_3$); ^{13}C NMR (125 MHz, $CDCl_3$): δ = 170.54, 170.08, 169.90, 169.84 (4 CO), 152.39 (dd, *J* = 15.7, 262.2 Hz, Ar-CF), 152.17 (Ar-C), 147.92 (dd, *J* = 11.8, 253.6 Hz, Ar-CF), 134.60 (d, *J* = 10.0 Hz, Ar-C), 131.06 (d, *J* = 3.1 Hz, Ar-C), 128.71 (m, Ar-C), 128.49 (d, *J* = 3.7 Hz, Ar-C), 127.99 (d, *J* = 4.7 Hz, Ar-C), 125.61 (m, Ar-C), 125.02, 116.94 (2 Ar-C), 112.86 (d, *J* = 3.6 Hz, CN), 102.27 (d, *J* = 11.4 Hz, Ar-C), 96.58 (C-1), 70.08 (C-5), 69.33 (C-2), 68.79 (C-3), 65.84 (C-4), 62.15 (C-6), 20.97, 20.81, 20.79, 20.75 ppm (4 CO*C*H_3); IR (neat): ν = 2234 (C≡N) cm^{-1}; ESI-MS: *m*/*z*: Calcd for $C_{27}H_{24}ClF_2NNaO_{10}$ $[M+Na]^+$: 618.1, found: 618.1.

3'-Chloro-2,3-difluoro-4'-(α-D-mannopyranosyloxy)-biphenyl-4-carbonitrile (4c). To a suspension of **21c** in *i*-PrOH (0.5 mL) was added 1.3 M $NH_{3(aq)}$ (2 mL) and THF (1.0 mL). After 4.5 d of stirring at rt, the solvent was evaporated in vacuo. The residue was purified by MPLC on silica gel (DCM/MeOH, 1:0-9:1) to yield **7** (19 mg, 73%) as a white solid. $[\alpha]_D^{20}$ +87.4 (*c* 0.95, MeOH); ^{1}H NMR (500 MHz, CD_3OD): δ =

7.69 (m, 1H, Ar-H), 7.61 (ddd, *J* = 1.7, 5.9, 7.8 Hz, 1H, Ar-H), 7.55-7.51 (m, 2H, Ar-H), 7.47 (ddd, *J* = 1.7, 6.8, 8.3 Hz, 1H, Ar-H), 5.65 (d, *J* = 1.6 Hz, 1H, H-1), 4.12 (dd, *J* = 1.8, 3.4 Hz, 1H, H-2), 4.00 (dd, *J* = 3.4, 9.5 Hz, 1H, H-3), 3.79-3.70 (m, 3H, H-4, H-6a, H-6b), 3.61 ppm (ddd, *J* = 2.3, 5.6, 9.7 Hz, 1H, H-5); ^{13}C NMR (125 MHz, CD_3OD): δ = 154.13 (Ar-C), 153.42 (dd, *J* = 15.8, 258.7 Hz, Ar-CF), 148.95 (dd, *J* = 11.8, 251.2 Hz, Ar-CF), 136.13 (d, *J* = 9.9 Hz, Ar-C), 131.68 (d, *J* = 3.6 Hz, Ar-C), 129.91 (d, *J* = 3.4 Hz, Ar-C), 129.40 (d, *J* = 4.7 Hz, Ar-C), 129.07 (Ar-C), 127.28 (m, Ar-C), 125.14, 118.20 (2 Ar-C), 113.73 (d, *J* = 3.6 Hz, CN), 102.73 (d, *J* = 12.5 Hz, Ar-C), 100.59 (C-1), 76.14 (C-5), 72.37 (C-3), 71.75 (C-2), 68.18 (C-4), 62.65 ppm (C-6); IR (KBr): ν = 2236 (C≡N) cm^{-1};

Protein production and purification

$FimH_{LD}$ and $FimH_{FL}$ were produced and purified as reported earlier (19, 74).

Fluorescence polarization (FP) assay

FP assays were essentially performed as previously described (27). A detailed experimental protocol is given in the Supporting Information.

Isothermal Titration Calorimetry (ITC)

All ITC experiments with $FimH_{LD}$ were performed using a VP-ITC instrument from MicroCal, Inc. (Malvern Instruments, Worcestershire, U.K.) with a sample cell volume of 1.4523 mL. For experiments with $FimH_{FL}$, an ITC_{200} device from MicroCal, Inc. (Malvern Instruments, Worcestershire, U.K.) with a sample cell volume of 203.7 µL was used. The measurements were performed with 2% DMSO at 25 °C, a stirring speed of 307 rpm (VP-ITC) or 750 rpm (ITC_{200}), and 10 µcal s^{-1} (VP-ITC) / 6 µcal s^{-1} (ITC_{200}) reference power. The protein samples were dialyzed in assay buffer (10 mM HEPES, 150 mM NaCl, pH 7.4) prior to all experiments. For experiments with $FimH_{LD}$, the *c*-values [$c = M_t(0)\, K_D^{-1}$, where $M_t(0)$ is the initial protein concentration] for compounds **1** and **2** were in a reliable range between 5 and 1000. For compounds **3**, **4**, **6**, and **7**, the *c*-values of the direct titrations were above 1000, so additional competitive ITC experiments were performed (75). The ligands were titrated into protein preincubated with a nine- to 70-fold excess of *n*-heptyl 2-deoxy-α-D-mannoside, resulting in sigmoidal titration curves. For measurements with $FimH_{FL}$, all compounds were analyzed by direct titration. Baseline correction and peak integration was performed with Origin 7.0 (OriginLab, USA) or NITPIC version 1.2.0 (76). Baseline subtraction and curve fitting with the three variables *N* (stoichiometry), K_D (dissociation constant), and $\Delta H°$ (change in enthalpy) were performed with SEDPHAT versions 10.40 and 12.1b (National Institutes of Health) (77). Global fitting analyses were performed for the competitive titrations (**3**, **4**, **6**, and **7** competing for the binding site) and for the direct titration of the competitor to obtain K_D values. The $\Delta H°$ and *N* values were then obtained by fitting of the direct titrations of **3**, **4**, **6**, and **7**. For compounds **1** and **2**, as well as for all experiments with $FimH_{FL}$, all variables could be determined from a global analysis of two independent direct titrations. The 95% confidence intervals of K_D and $\Delta H°$ were calculated with the 1-dimensional error surface projection. The $\Delta G°$ (Gibb's free energy of binding) and $-T\Delta S°$ (change in entropy) values were calculated from Eq. (1)

$$\Delta G° = \Delta H° - T\Delta S° = -RT \ln K_A \qquad \text{(Eq. 1)}$$

with *T* being the absolute temperature and *R* the universal gas constant (8.314 Jmol^{-1} K^{-1}).

FimH complex crystallization and structure refinement

For crystallization, $FimH_{LD}$ (residues 1–158) at a final concentration of 15 mg/mL (ca. 0.9 mM) with a

3-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 2 months and grow up to their final size within 3 months. They were cryopreserved by addition of 20% ethylene glycol (v/v) and flash-cooled with liquid nitrogen. $FimH_{LD}$/**6** was crystallized in 0.1 M Na acetate 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 (78). Data was collected at the X06SA beamline of the Swiss Light Source (Paul Scherrer Institute, Switzerland) and indexed, integrated and scaled with XDS (79, 80). The structures were solved by molecular replacement with PHASER (81) using the $FimH_{LD}$ / biphenyl α-mannopyranoside complex (PDB code 4X50)(23) as search model. The structures were built using the COOT software[85] and periodically refined with the PHENIX software (82). Geometric restraints for the ligands were generated with PRODRG (83). The final protein models were validated with Molprobity and deposited in the Protein Data Bank with the PDB codes 5MCA and (To be announced). Data collection and refinement statistics are given in the Supporting Information.

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 TSP-*d4* (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 analysed with CcpNmr Analysis 2.4 (84). The backbone resonance assignment as well as CSP data with the non-fluorinated reference compounds **1** and **3** were available from a previous study (23). As all tested mannoside ligands bound in the slow exchange regime, peak assignment was performed based on chemical shift proximity.

***Ab initio* calculations**

Input structures were prepared from the co-crystal structures of **5** and **6** and from the available structures in the Protein Data Bank for **1** (PDB 4X50) & **3** (PDB 4CST). The input structure for compound **2** was generated from **1**, by replacing the appropriate hydrogen with a chlorine substituent. Compound **4** and **7** were equally prepared, starting from **5** and **6**, respectively. The aglycone–Tyr48 complexes were optimized by Jaguar (85) DFT calculations using the B3LYP-MM functional in combination with the cc-pVDZ++ basis-set in the gas phase. The same level of theory was used to compute the π-π stacking interaction energies from the optimized complexes.

Physicochemical and in Vitro Pharmacokinetic 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). ACN and MeOH were ordered from Acros Organics (Geel, Belgium).

The solubility, permeability and lipophilicity of **3** was reported previously (27).

Log$D_{7.4}$ determination (shake-flask method). Equal amounts of Tris-HCl buffer (0.1 M, pH 7.4) and 1-octanol were mixed and vigorously shaken for 5 minutes 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 to 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 aluminium 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 2:

$$logD_{7.4} = log\left(\frac{10\ \mu M - c_b}{c_b} \times \frac{1}{o:b}\right) \quad \text{(Eq. 2)}$$

Aqueous Solubility. Lyophilized compound was added to 200 µL Tris-HCl buffer (pH 7.4, 0.1 M) in duplicate until a precipitation was formed. The vials were put into super-sonic 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 filtrated (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 (66, 70).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), 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. Afterwards, 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 110 163). 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 GutBoxTM (pION), and left undisturbed for 16 h. Then, 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 C18 column (2.1 x 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: ACN containing 0.1% formic acid. The flow rate was maintained at 0.6 mL/min. The gradient was ramped from 95% A/5% B to 5% A/95% B over 1 min, and then hold at 5% A/95% B for 0.1 min. The system was then brought back to 95% A/5% B, resulting in a total duration of 4 min. Fragmentor voltage and collision energy were optimized for the analysis of compounds in multiple reaction monitoring mode in negative mode. The concentration of the analytes was quantified by the Agilent Mass Hunter Quantitative Analysis software (version B.04.00).

References

1. Foxman B, Barlow R, D'Arcy H, Gillespie B, Sobel JD. Urinary tract infection: self-reported incidence and associated costs. *Ann Epidemiol*. **2000**;*10(8)*:509-15.
2. Ronald A. The etiology of urinary tract infection: traditional and emerging pathogens. *Am J Med*. **2002**;*113 Suppl 1A*:14S-9S.
3. Cegelski L, Marshall GR, Eldridge GR, Hultgren SJ. The biology and future prospects of antivirulence therapies. *Nat Rev Microbiol*. **2008**;*6(1)*:17-27.
4. Flores-Mireles AL, Walker JN, Caparon M, Hultgren SJ. Urinary tract infections: epidemiology, mechanisms of infection and treatment options. *Nat Rev Microbiol*. **2015**;*13(5)*:269-84.
5. Wiles TJ, Kulesus RR, Mulvey MA. Origins and virulence mechanisms of uropathogenic Escherichia coli. *Exp Mol Pathol*. **2008**;*85(1)*:11-9.
6. Fihn SD. Clinical practice. Acute uncomplicated urinary tract infection in women. *N Engl J Med*. **2003**;*349(3)*:259-66.
7. Hooton TM, Besser R, Foxman B, Fritsche TR, Nicolle LE. Acute uncomplicated cystitis in an era of increasing antibiotic resistance: a proposed approach to empirical therapy. *Clin Infect Dis*. **2004**;*39(1)*:75-80.
8. Sanchez GV, Master RN, Karlowsky JA, Bordon JM. In vitro antimicrobial resistance of urinary Escherichia coli isolates among U.S. outpatients from 2000 to 2010. *Antimicrob Agents Chemother*. **2012**;*56(4)*:2181-3.
9. Hooton TM. Clinical practice. Uncomplicated urinary tract infection. *N Engl J Med*. **2012**;*366(11)*:1028-37.
10. Clatworthy AE, Pierson E, Hung DT. Targeting virulence: a new paradigm for antimicrobial therapy. *Nat Chem Biol*. **2007**;*3(9)*:541-8.
11. Mulvey MA, Schilling JD, Martinez JJ, Hultgren SJ. Bad bugs and beleaguered bladders: interplay between uropathogenic Escherichia coli and innate host defenses. *Proc Natl Acad Sci U S A*. **2000**;*97(16)*:8829-35.
12. Schilling JD, Mulvey MA, Hultgren SJ. Structure and function of Escherichia coli type 1 pili: new insight into the pathogenesis of urinary tract infections. *J Infect Dis*. **2001**;*183 Suppl 1*:S36-40.
13. Jones CH, Pinkner JS, Roth R, Heuser J, Nicholes AV, Abraham SN, et al. FimH adhesin of type 1 pili is assembled into a fibrillar tip structure in the Enterobacteriaceae. *Proc Natl Acad Sci U S A*. **1995**;*92(6)*:2081-5.
14. Pak J, Pu Y, Zhang ZT, Hasty DL, Wu XR. Tamm-Horsfall protein binds to type 1 fimbriated Escherichia coli and prevents E. coli from binding to uroplakin Ia and Ib receptors. *J Biol Chem*. **2001**;*276(13)*:9924-30.
15. Zhou G, Mo WJ, Sebbel P, Min G, Neubert TA, Glockshuber R, et al. Uroplakin Ia is the urothelial receptor for uropathogenic Escherichia coli: evidence from in vitro FimH binding. *J Cell Sci*. **2001**;*114(Pt 22)*:4095-103.
16. Choudhury D, Thompson A, Stojanoff V, Langermann S, Pinkner J, Hultgren SJ, et al. X-ray structure of the FimC-FimH chaperone-adhesin complex from uropathogenic Escherichia coli. *Science*. **1999**;*285(5430)*:1061-6.
17. Hung CS, Bouckaert J, Hung D, Pinkner J, Widberg C, DeFusco A, et al. Structural basis of tropism of Escherichia coli to the bladder during urinary tract infection. *Mol Microbiol*. **2002**;*44(4)*:903-15.
18. Le Trong I, Aprikian P, Kidd BA, Forero-Shelton M, Tchesnokova V, Rajagopal P, et al. Structural basis for mechanical force regulation of the adhesin FimH via finger trap-like beta sheet twisting. *Cell*. **2010**;*141(4)*:645-55.
19. Sauer MM, Jakob RP, Eras J, Baday S, Eris D, Navarra G, et al. Catch-bond mechanism of the bacterial adhesin FimH. *Nat Commun*. **2016**;*7*:10738.
20. Kalas V, Pinkner JS, Hannan TJ, Hibbing ME, Dodson KW, Holehouse AS, et al. Evolutionary fine-tuning of conformational ensembles in FimH during host-pathogen interactions. *Sci Adv*. **2017**;*3(2)*:e1601944.

21. Eris D, Preston RC, Scharenberg M, Hulliger F, Abgottspon D, Pang L, et al. The Conformational Variability of FimH: Which Conformation Represents the Therapeutic Target? *Chembiochem*. **2016**;*17(11)*:1012-20.
22. Firon N, Ashkenazi S, Mirelman D, Ofek I, Sharon N. Aromatic alpha-glycosides of mannose are powerful inhibitors of the adherence of type 1 fimbriated Escherichia coli to yeast and intestinal epithelial cells. *Infect Immun*. **1987**;*55(2)*:472-6.
23. Fiege B, Rabbani S, Preston RC, Jakob RP, Zihlmann P, Schwardt O, et al. The tyrosine gate of the bacterial lectin FimH: a conformational analysis by NMR spectroscopy and X-ray crystallography. *Chembiochem*. **2015**;*16(8)*:1235-46.
24. Grabosch C, Hartmann M, Schmidt-Lassen J, Lindhorst TK. Squaric acid monoamide mannosides as ligands for the bacterial lectin FimH: covalent inhibition or not? *Chembiochem*. **2011**;*12(7)*:1066-74.
25. Klein T, Abgottspon D, Wittwer M, Rabbani S, Herold J, Jiang X, et al. FimH antagonists for the oral treatment of urinary tract infections: from design and synthesis to in vitro and in vivo evaluation. *J Med Chem*. **2010**;*53(24)*:8627-41.
26. Han Z, Pinkner JS, Ford B, Obermann R, Nolan W, Wildman SA, et al. Structure-based drug design and optimization of mannoside bacterial FimH antagonists. *J Med Chem*. **2010**;*53(12)*:4779-92.
27. Kleeb S, Pang L, Mayer K, Eris D, Sigl A, Preston RC, et al. FimH antagonists: bioisosteres to improve the in vitro and in vivo PK/PD profile. *J Med Chem*. **2015**;*58(5)*:2221-39.
28. Cusumano CK, Pinkner JS, Han Z, Greene SE, Ford BA, Crowley JR, et al. Treatment and prevention of urinary tract infection with orally active FimH inhibitors. *Sci Transl Med*. **2011**;*3(109)*:109ra15.
29. Han Z, Pinkner JS, Ford B, Chorell E, Crowley JM, Cusumano CK, et al. Lead optimization studies on FimH antagonists: discovery of potent and orally bioavailable ortho-substituted biphenyl mannosides. *J Med Chem*. **2012**;*55(8)*:3945-59.
30. Pang L, Kleeb S, Lemme K, Rabbani S, Scharenberg M, Zalewski A, et al. FimH antagonists: structure-activity and structure-property relationships for biphenyl alpha-D-mannopyranosides. *ChemMedChem*. **2012**;*7(8)*:1404-22.
31. Jiang X, Abgottspon D, Kleeb S, Rabbani S, Scharenberg M, Wittwer M, et al. Antiadhesion therapy for urinary tract infections--a balanced PK/PD profile proved to be key for success. *J Med Chem*. **2012**;*55(10)*:4700-13.
32. Schwardt O, Rabbani S, Hartmann M, Abgottspon D, Wittwer M, Kleeb S, et al. Design, synthesis and biological evaluation of mannosyl triazoles as FimH antagonists. *Bioorg Med Chem*. **2011**;*19(21)*:6454-73.
33. Brument S, Sivignon A, Dumych TI, Moreau N, Roos G, Guerardel Y, et al. Thiazolylaminomannosides as potent antiadhesives of type 1 piliated Escherichia coli isolated from Crohn's disease patients. *J Med Chem*. **2013**;*56(13)*:5395-406.
34. Lindhorst TK, Kieburg C, Krallmann-Wenzel U. Inhibition of the type 1 fimbriae-mediated adhesion of Escherichia coli to erythrocytes by multiantennary alpha-mannosyl clusters: the effect of multivalency. *Glycoconj J*. **1998**;*15(6)*:605-13.
35. Nagahori N, Lee RT, Nishimura S, Page D, Roy R, Lee YC. Inhibition of adhesion of type 1 fimbriated Escherichia coli to highly mannosylated ligands. *Chembiochem*. **2002**;*3(9)*:836-44.
36. Appeldoorn CCM, Joosten JAF, Ait el Maate F, Dobrindt U, Hacker J, Liskamp RMJ, et al. Novel multivalent mannose compounds and their inhibition of the adhesion of type 1 fimbriated uropathogenic E. coli. *Tetrahedron: Asymmetry*. **2005**;*16(2)*:361-72.
37. Patel A, Lindhorst TK. A modular approach for the synthesis of oligosaccharide mimetics. *J Org Chem*. **2001**;*66(8)*:2674-80.
38. Touaibia M, Wellens A, Shiao TC, Wang Q, Sirois S, Bouckaert J, et al. Mannosylated G(0) dendrimers with nanomolar affinities to Escherichia coli FimH. *ChemMedChem*. **2007**;*2(8)*:1190-201.

39. Durka M, Buffet K, Iehl J, Holler M, Nierengarten JF, Taganna J, et al. The functional valency of dodecamannosylated fullerenes with Escherichia coli FimH--towards novel bacterial antiadhesives. *Chem Commun (Camb)*. **2011**;*47(4)*:1321-3.
40. Bouckaert J, Li Z, Xavier C, Almant M, Caveliers V, Lahoutte T, et al. Heptyl alpha-D-mannosides grafted on a beta-cyclodextrin core to interfere with Escherichia coli adhesion: an in vivo multivalent effect. *Chemistry*. **2013**;*19(24)*:7847-55.
41. Wellens A, Lahmann M, Touaibia M, Vaucher J, Oscarson S, Roy R, et al. The tyrosine gate as a potential entropic lever in the receptor-binding site of the bacterial adhesin FimH. *Biochemistry*. **2012**;*51(24)*:4790-9.
42. Bondi A. van der Waals Volumes and Radii. *The Journal of Physical Chemistry*. **1964**;*68(3)*:441-51.
43. Battaglia MR, Buckingham AD, Williams JH. The electric quadrupole moments of benzene and hexafluorobenzene. *Chemical Physics Letters*. **1981**;*78(3)*:421-3.
44. Hunter CA, Sanders JKM. The nature of .pi.-.pi. interactions. *Journal of the American Chemical Society*. **1990**;*112(14)*:5525-34.
45. Hunter CA, Lawson KR, Perkins J, Urch CJ. Aromatic interactions. *Journal of the Chemical Society, Perkin Transactions 2*. **2001***(5)*:651-69.
46. Cockroft SL, Hunter CA, Lawson KR, Perkins J, Urch CJ. Electrostatic control of aromatic stacking interactions. *J Am Chem Soc*. **2005**;*127(24)*:8594-5.
47. Riwar LJ, Trapp N, Kuhn B, Diederich F. Substituent Effects in Parallel-Displaced pi-pi Stacking Interactions: Distance Matters. *Angew Chem Int Ed Engl*. **2017**;*56(37)*:11252-7.
48. Cozzi F, Cinquini M, Annunziata R, Dwyer T, Siegel JS. Polar/.pi. interactions between stacked aryls in 1,8-diarylnaphthalenes. *Journal of the American Chemical Society*. **1992**;*114(14)*:5729-33.
49. Cozzi F, Ponzini F, Annunziata R, Cinquini M, Siegel JS. Polar Interactions between Stackedπ Systems in Fluorinated 1,8-Diarylnaphthalenes: Importance of Quadrupole Moments in Molecular Recognition. *Angewandte Chemie International Edition in English*. **1995**;*34(9)*:1019-20.
50. Cozzi F, Siegel JS. Interaction between stacked aryl groups in 1,8-diarylnaphthalenes: Dominance of polar/π over charge-transfer effects. *Pure and Applied Chemistry*. **1995**;*67(5)*.
51. Gung BW, Amicangelo JC. Substituent effects in C6F6-C6H5X stacking interactions. *J Org Chem*. **2006**;*71(25)*:9261-70.
52. Gung BW, Xue X, Zou Y. Enthalpy (DeltaH) and entropy (DeltaS) for pi-stacking interactions in near-sandwich configurations: relative importance of electrostatic, dispersive, and charge-transfer effects. *J Org Chem*. **2007**;*72(7)*:2469-75.
53. Sinnokrot MO, Sherrill CD. Substituent effects in pi-pi interactions: sandwich and T-shaped configurations. *J Am Chem Soc*. **2004**;*126(24)*:7690-7.
54. Doyon JB, Jain A. The pattern of fluorine substitution affects binding affinity in a small library of fluoroaromatic inhibitors for carbonic anhydrase. *Org Lett*. **1999**;*1(2)*:183-5.
55. Kim C-Y, Chang JS, Doyon JB, Baird TT, Fierke CA, Jain A, et al. Contribution of Fluorine to Protein−Ligand Affinity in the Binding of Fluoroaromatic Inhibitors to Carbonic Anhydrase II. *Journal of the American Chemical Society*. **2000**;*122(49)*:12125-34.
56. Abbate F, Casini A, Scozzafava A, Supuran CT. Carbonic anhydrase inhibitors: X-ray crystallographic structure of the adduct of human isozyme II with the perfluorobenzoyl analogue of methazolamide. Implications for the drug design of fluorinated inhibitors. *J Enzyme Inhib Med Chem*. **2003**;*18(4)*:303-8.
57. Giroud M, Harder M, Kuhn B, Haap W, Trapp N, Schweizer WB, et al. Fluorine Scan of Inhibitors of the Cysteine Protease Human Cathepsin L: Dipolar and Quadrupolar Effects in the pi-Stacking of Fluorinated Phenyl Rings on Peptide Amide Bonds. *ChemMedChem*. **2016**;*11(10)*:1042-7.
58. Meyer EA, Castellano RK, Diederich F. Interactions with aromatic rings in chemical and biological recognition. *Angew Chem Int Ed Engl*. **2003**;*42(11)*:1210-50.

59. Salonen LM, Ellermann M, Diederich F. Aromatic rings in chemical and biological recognition: energetics and structures. *Angew Chem Int Ed Engl*. **2011**;*50(21)*:4808-42.
60. Do HQ, Daugulis O. Copper-catalyzed arylation of heterocycle C-H bonds. *J Am Chem Soc*. **2007**;*129(41)*:12404-5.
61. Prieto M, Zurita E, Rosa E, Munoz L, Lloyd-Williams P, Giralt E. Arylboronic acids and arylpinacolboronate esters in Suzuki coupling reactions involving indoles. Partner role swapping and heterocycle protection. *J Org Chem*. **2004**;*69(20)*:6812-20.
62. Mayer K, Eris D, Schwardt O, Sager CP, Rabbani S, Kleeb S, et al. Urinary Tract Infection: Which Conformation of the Bacterial Lectin FimH Is Therapeutically Relevant? *J Med Chem*. **2017**;*60(13)*:5646-62.
63. Daranas AH, Shimizu H, Homans SW. Thermodynamics of binding of D-galactose and deoxy derivatives thereof to the L-arabinose-binding protein. *J Am Chem Soc*. **2004**;*126(38)*:11870-6.
64. Alsenz J, Kansy M. High throughput solubility measurement in drug discovery and development. *Adv Drug Deliv Rev*. **2007**;*59(7)*:546-67.
65. Dearden JC, Bresnen GM. The Measurement of Partition Coefficients. *Quantitative Structure-Activity Relationships*. **1988**;*7(3)*:133-44.
66. Kansy M, Senner F, Gubernator K. Physicochemical high throughput screening: parallel artificial membrane permeation assay in the description of passive absorption processes. *J Med Chem*. **1998**;*41(7)*:1007-10.
67. Lipinski CA. Drug-like properties and the causes of poor solubility and poor permeability. *J Pharmacol Toxicol Methods*. **2000**;*44(1)*:235-49.
68. Waring MJ. Lipophilicity in drug discovery. *Expert Opin Drug Discov*. **2010**;*5(3)*:235-48.
69. Feng B, LaPerle JL, Chang G, Varma MV. Renal clearance in drug discovery and development: molecular descriptors, drug transporters and disease state. *Expert Opin Drug Metab Toxicol*. **2010**;*6(8)*:939-52.
70. Avdeef A, Bendels S, Di L, Faller B, Kansy M, Sugano K, et al. PAMPA--critical factors for better predictions of absorption. *J Pharm Sci*. **2007**;*96(11)*:2893-909.
71. Rodriguez VB, Kidd BA, Interlandi G, Tchesnokova V, Sokurenko EV, Thomas WE. Allosteric coupling in the bacterial adhesive protein FimH. *J Biol Chem*. **2013**;*288(33)*:24128-39.
72. Wei Y, Kan J, Wang M, Su W, Hong M. Palladium-catalyzed direct arylation of electron-deficient polyfluoroarenes with arylboronic acids. *Org Lett*. **2009**;*11(15)*:3346-9.
73. Chen Q-Y, Li Z-T. Photoinduced electron transfer reactions of pentafluoroiodobenzene with aromatic compounds. *Journal of the Chemical Society, Perkin Transactions 1*. **1993***(14)*.
74. Rabbani S, Jiang X, Schwardt O, Ernst B. Expression of the carbohydrate recognition domain of FimH and development of a competitive binding assay. *Anal Biochem*. **2010**;*407(2)*:188-95.
75. Sigurskjold BW. Exact analysis of competition ligand binding by displacement isothermal titration calorimetry. *Anal Biochem.* **2000**;*277(2)*:260-6.
76. Keller S, Vargas C, Zhao H, Piszczek G, Brautigam CA, Schuck P. High-precision isothermal titration calorimetry with automated peak-shape analysis. *Anal Chem*. **2012**;*84(11)*:5066-73.
77. Houtman JC, Brown PH, Bowden B, Yamaguchi H, Appella E, Samelson LE, et al. Studying multisite binary and ternary protein interactions by global analysis of isothermal titration calorimetry data in SEDPHAT: application to adaptor protein complexes in cell signaling. *Protein Sci*. **2007**;*16(1)*:30-42.
78. Rubinson KA, Ladner JE, Tordova M, Gilliland GL. Cryosalts: suppression of ice formation in macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr*. **2000**;*56(Pt 8)*:996-1001.
79. Kabsch W. Integration, scaling, space-group assignment and post-refinement. *Acta Crystallogr D Biol Crystallogr*. **2010**;*66(Pt 2)*:133-44.
80. Kabsch W. Xds. *Acta Crystallogr D Biol Crystallogr*. **2010**;*66(Pt 2)*:125-32.
81. McCoy AJ. Solving structures of protein complexes by molecular replacement with Phaser. *Acta Crystallogr D Biol Crystallogr*. **2007**;*63(Pt 1)*:32-41.

82. Adams PD, Afonine PV, Bunkoczi G, Chen VB, Davis IW, Echols N, et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr*. **2010**;*66(Pt 2)*:213-21.
83. van Aalten DM, Bywater R, Findlay JB, Hendlich M, Hooft RW, Vriend G. PRODRG, a program for generating molecular topologies and unique molecular descriptors from coordinates of small molecules. *J Comput Aided Mol Des*. **1996**;*10(3)*:255-62.
84. Vranken WF, Boucher W, Stevens TJ, Fogh RH, Pajon A, Llinas M, et al. The CCPN data model for NMR spectroscopy: development of a software pipeline. *Proteins*. **2005**;*59(4)*:687-96.
85. Bochevarov AD, Harder E, Hughes TF, Greenwood JR, Braden DA, Philipp DM, et al. Jaguar: A high-performance quantum chemistry software program with strengths in life and materials sciences. *International Journal of Quantum Chemistry*. **2013**;*113(18)*:2110-42.

4.6 Publication 1: FimH ester prodrug approach

This publication describes the synthesis and experimental determination of physicochemical and pharmacokinetic properties of ester prodrugs starting from the bioisostere **5e**.

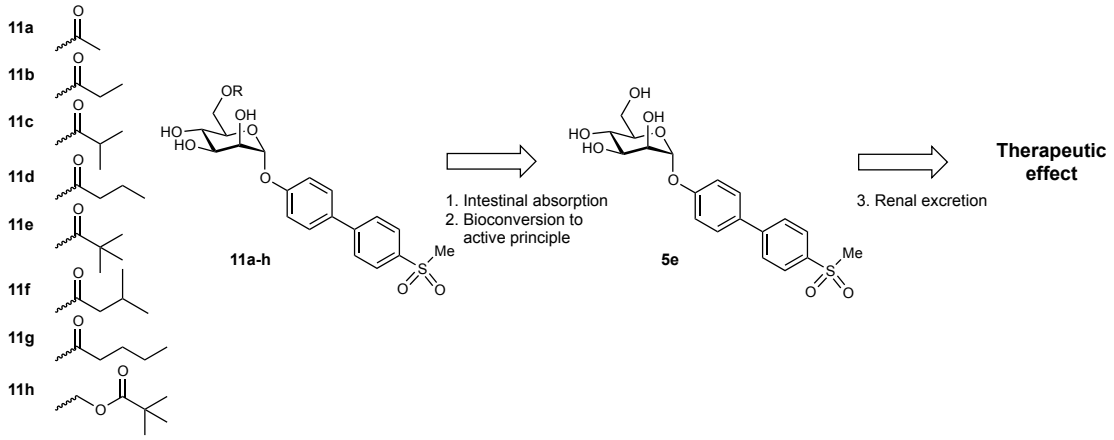


Figure 1 Proposed mode of action of ester prodrugs.

Contribution to the P**roject:**

Philipp Dätwyler was involved in the synthesis of the molecules as well as in the physicochemical and pharmacokinetic evaluation of the ester prodrugs during his master thesis under the supervision of Dr. Wojciech Schönemann and Dr. Simon Kleeb. He further contributed to the writing and arrangement of the publication.

The publication was published in the peer-reviewed *Canadian Journal of Chemistry* in 2016.

Wojciech Schönemann*, Simon Kleeb*, **Philipp Dätwyler**, Oliver Schwardt, and Beat Ernst

DOI: 10.1139/cjc-2015-0582

* contributed equally.

NRC Research Press

ARTICLE

Prodruggability of carbohydrates — oral FimH antagonists

Wojciech Schönemann, Simon Kleeb, Philipp Dätwyler, Oliver Schwardt, and Beat Ernst

Abstract: The bacterial lectin FimH is a promising therapeutic target for the nonantibiotic prevention and treatment of urinary tract infections. In this communication, an ester prodrug approach is described to achieve oral bioavailability for FimH antagonists. By introducing short-chain acyl promoieties at the C-6 position of a biphenyl α-D-mannopyranoside, prodrugs with an excellent absorption potential were obtained. The human carboxylesterase 2 was identified as a main enzyme mediating rapid bioconversion to the active principle. Despite their propensity to hydrolysis within the enterocytes during absorption, these ester prodrugs present a considerable progress in the development of orally available FimH antagonists.

Key words: FimH antagonists, oral bioavailability, ester prodrugs, enzyme-mediated bioconversion, renal excretion.

Résumé : La lectine bactérienne FimH est une cible thérapeutique prometteuse pour la prévention et le traitement sans antibiotiques des infections des voies urinaires. Dans la présente communication, nous décrivons une approche faisant appel à un promédicament à base d'ester pour permettre l'obtention d'une biodisponibilité orale à l'administration d'antagonistes de la protéine FimH. En ajoutant des groupements labiles acylés à courtes chaînes en position C-6 d'un biphényl-α-D-mannopyranoside, nous avons obtenu des promédicaments dotés d'un excellent potentiel d'absorption. Nous avons établi que la carboxylestérase 2 humaine était la principale enzyme assurant une bioconversion rapide de ces promédicaments en leur principe actif. Malgré leur propension à l'hydrolyse à l'intérieur des entérocytes au cours de l'absorption, ces promédicaments à base d'ester représentent un progrès considérable sur le plan du développement d'antagonistes de la protéine FimH biodisponibles après administration par voie orale. [Traduit par la Rédaction]

Mots-clés : antagonistes de la protéine FimH, biodisponibilité orale, promédicaments à base d'ester, bioconversion enzymatique, excrétion rénale.

Introduction

Urinary tract infections (UTIs) primarily caused by uropathogenic *Escherichia coli* (UPEC) are among the most prevalent infectious diseases worldwide and are treated with antibiotics as first line therapy.[1] However, frequent and repeated use of antibiotics can lead to antimicrobial resistance and treatment failure, corroborating the need for alternative therapeutic strategies.[2] Bacterial adhesion to urothelial cells is a crucial first step of the infection cycle, preventing UPEC from being washed out by urine flow and enabling the bacteria to infect urothelial cells.[3] This initial adhesion is mediated by the mannose-specific lectin FimH localized at the tip of bacterial type 1 pili.[4] FimH consists of a lectin domain hosting the carbohydrate recognition domain (CRD) and a pilin domain regulating the switch between the low- and high-affinity states of the CRD.[5] The CRD interacts with the mannosylated glycoprotein uroplakin 1a on the surface of urothelial cells and thereby initiates the bacterial infection.

More than three decades ago, Firon et al. introduced FimH antagonists as a novel class of therapeutics for prevention and treatment of UTI.[6] Since then, several alkyl and aryl α-D-mannopyranosides with nanomolar affinities towards the FimH-CRD were reported.[7] In vivo pharmacokinetic studies with orally bioavailable biphenyl α-D-mannopyranosides in a mouse model were first performed in 2010,[7d] and since then, additional reports describing orally active FimH antagonists have been published.[7f,7g] In either case, high oral dosages (≥50 mg/kg) were necessary to achieve the minimal effective concentrations required for anti-adhesive effects in the bladder. The antagonists were furthermore rapidly eliminated from circulation, such that the therapeutic effect — upon a single dose — could be maintained only for a few hours.

When despite high oral dosages only low antagonist concentration can be detected in the urine, the reasons are either low intestinal absorption (i.e., due to low aqueous solubility or low membrane permeability) or extensive nonrenal elimination.[8] Moreover, undesirably fast renal excretion of the systemically available fraction is due to high glomerular filtration, i.e., low plasma protein binding (PPB), or poor reabsorption in the renal tubules.[9] We recently reported ester prodrugs of antagonist **1** (Fig. 1) where the carboxylic acid moiety on the terminal ring of the biphenyl aglycone is masked as ester, leading to increased intestinal absorption.[7d] In a second approach, we replaced the carboxylic acid moiety by bioisosteres to raise lipophilicity and PPB and, consequently, to slow down excretion of the systemically available antagonist into the bladder.[10] In the case of the cyanide **3**, the physicochemical profile allowed for excellent oral absorption and sustained renal excretion, whereas other bioisosteres, e.g., the *p*-methylsulfonyl-biphenyl mannoside **4a**, were too polar for absorption.

A common feature of carbohydrate derivatives is their high polarity, resulting from their many hydroxyl groups. By peracylation the hydrophilic character can be markedly reduced, leading

Received 30 November 2015. Accepted 10 February 2016.

W. Schönemann,* S. Kleeb,* P. Dätwyler, O. Schwardt, and B. Ernst. Institute of Molecular Pharmacy, Pharmacenter, University of Basel, Klingelbergstrasse 50, CH-4056 Basel, Switzerland.

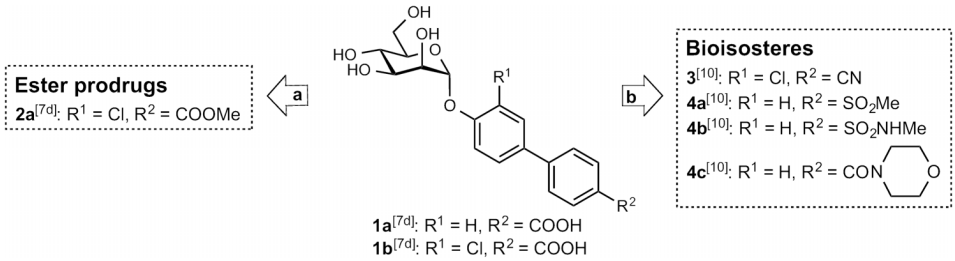
Corresponding author: Beat Ernst (email: beat.ernst@unibas.ch).

*These authors contributed equally to the project.

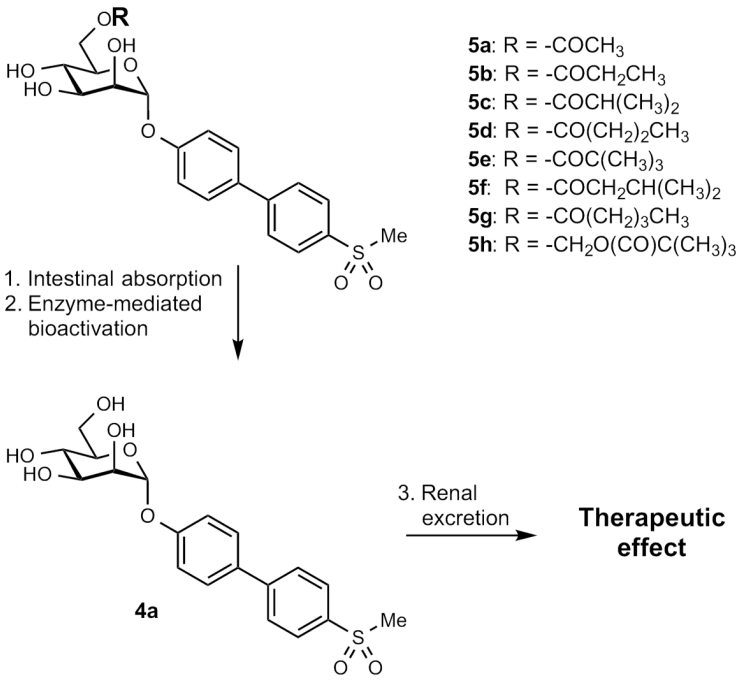
This article is part of a Special Issue dedicated to Professor David Bundle in recognition of his seminal contributions and lifetime achievements in the fields of carbohydrate chemistry and glycobiology.

Can. J. Chem. **94**: 909–919 (2016) dx.doi.org/10.1139/cjc-2015-0582 Published at www.nrcresearchpress.com/cjc on 21 March 2016.

Fig. 1. Pharmacokinetics of biphenyl α-D-mannopyranosides improved by (*a*) an ester prodrug approach[7d] and (*b*) bioisosteric modifications.[10] Compared to the parent carboxylates **1a** and **1b**, the ester prodrugs **2a–2c** exhibit oral availability; with bioisosteres of the carboxylates **1a** and **1b**, a prolonged renal excretion (→ **4a–4c**) and improved oral bioavailablity (→ **3**) could be achieved.



Scheme 1. Proposed mode of action. Ester prodrugs of the biphenyl α-D-mannopyranoside **4a** are hydrolyzed upon absorption by esterases. The active principle **4a** is excreted renally and binds to FimH lectins located at the tip of type 1 pili of UPECs in the bladder, resulting in a therapeutic effect.[7d]



to an increased cellular uptake.[11a,11b] However, polyacylated carbohydrate derivatives may also exhibit disadvantages, like low solubility and complex in vivo pharmacokinetic profiles due to different rates for deacylation leading to numerous metabolic intermediates, whereof each metabolite could be renally excreted individually.

In the present communication, we describe an ester prodrug strategy applied for the optimization of the pharmacokinetic parameters relevant for oral absorption of methylsulfonyl bioisostere **4a**. As opposed to peracylation, only one of the hydroxyl groups was acylated. For the adjustment of lipophilicity and metabolic hydrolysis, the aliphatic esters were elongated or branched, thereby paving the way for orally available prodrugs with sustained excretion of the pharmacologically active principle **4a** into the bladder (Scheme 1).

Results and discussion

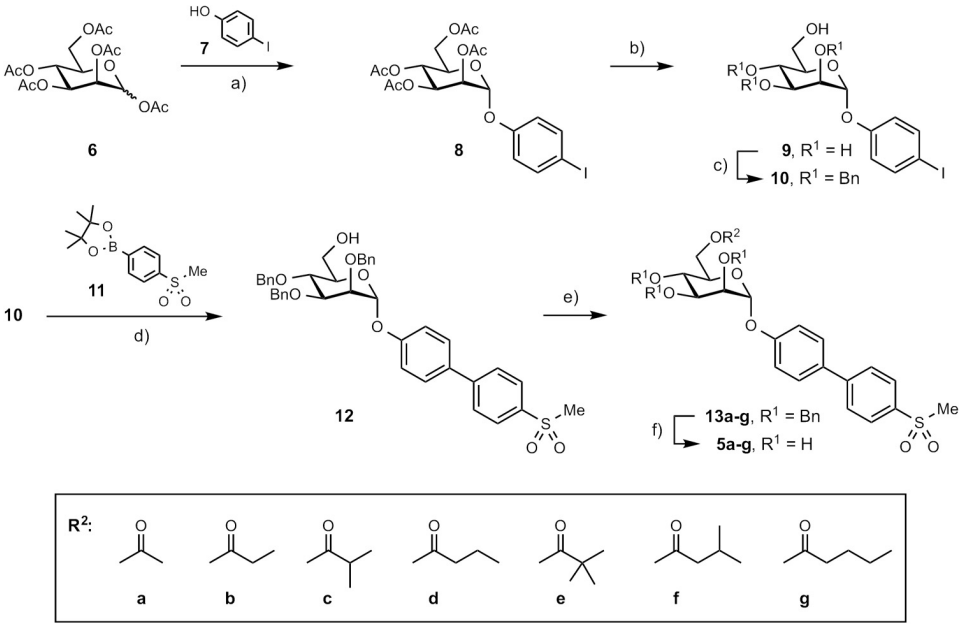
Synthesis of prodrugs 5a–5h

For the acylation, we selected the C-6 hydroxyl group of the mannose moiety of antagonist **4a** because the steric accessibility of the 6-position guarantees successful enzymatic hydrolysis. In the prodrugs **5a–5g**, the C-6 hydroxyl group is directly acylated, whereas in prodrug **5h**, an additional acetal linker was introduced to further improve the accessibility of the metabolic cleavage site. Upon hydrolysis of the ester, the resulting hemiacetal intermediate is expected to collapse spontaneously, releasing the active principal as well as formaldehyde.

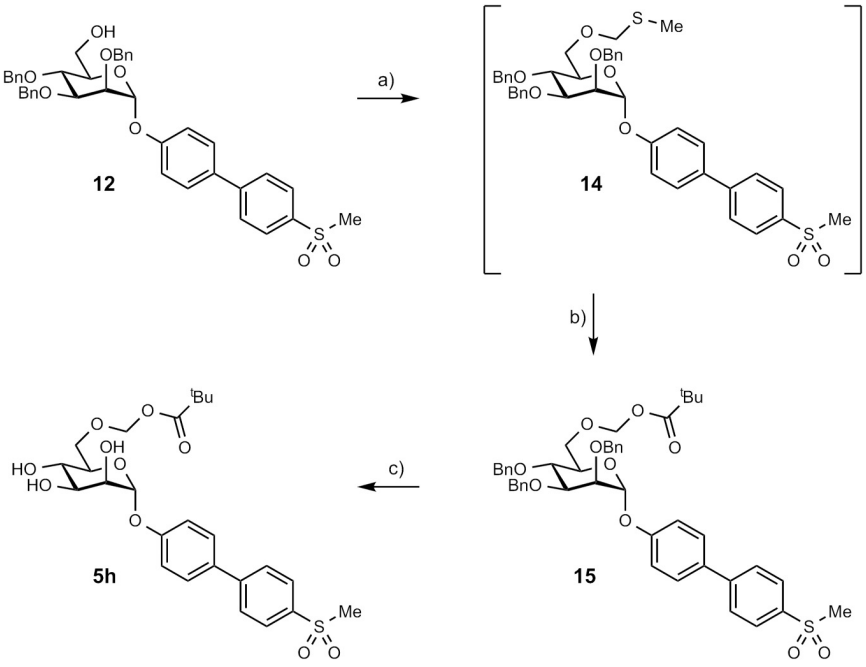
The synthetic route applied for the synthesis of the prodrugs **5a–5g** is depicted in Scheme 2. BF_3·Et_2O-promoted mannosylation of *p*-iodophenol **7** (→ **8**) followed by deacetylation under standard Zemplén conditions yielded mannoside **9**. Selective protection of the primary alcohol in **9** with *tert*-butyldimethylsilyl chloride followed by benzylation of the remaining three hydroxyl groups and, finally, removal of the TBDMS group under acidic conditions was performed according to one of the representative procedures to give intermediate **10**.[12] Suzuki cross-coupling with 4-(methanesulfonyl)phenylboronic acid pinacol ester (**11**) afforded mannoside **12**.[13] Then the promoiety in the 6-position was introduced with acetic anhydrate (→ **13a**) or corresponding acid chlorides (→ **13b–13g**). The prodrugs **5a–5g** were finally obtained after hydrogenolysis in the presence of palladium hydroxide on carbon.

For the synthesis of **5h** (Scheme 3), a modified final acylation strategy was necessary, since a methylene linker had to be introduced between the 6-OH of the mannose moiety and the acyl group. A convenient approach involves methylthiomethyl ether

Scheme 2. (*a*) $BF_3{\cdot}Et_2O$, CH_2Cl_2, 4 Å MS, 40 °C, 30 h, 67%; (*b*) MeONa–MeOH, rt, 27 h, 61%; (*c*) (*i*) TBDMSCl, imidazole, DMF, 0 °C → rt, 18 h; (*ii*) BnBr, NaH, TBAI, DMF, 0 °C → rt, 5 h. (*iii*) H_2SO_4 (1 M), MeOH, 0 °C, 18 h, 58% (for three steps); (*d*) $PdCl_2(dppf){\cdot}CH_2Cl_2$, K_3PO_4, DMF, 80 °C, 19 h, 86%; (*e*) Ac_2O or R^2-Cl, DMAP, pyridine, 0 °C → 60 °C, 3–24 h, 58%–91%; (*f*) $Pd(OH)_2/C$, H_2 (1 bar), EtOH, rt.



Scheme 3. (*a*) Ac_2O, AcOH, DMSO, 4 Å MS, rt, 23 h; (*b*) PivOH, NIS, 0 °C → rt, 15 h, 33%; (*c*) $Pd(OH)_2/C$, H_2 (1 bar), EtOH, rt.



for a subsequent introduction of diverse acyl groups. Starting from **12**, the methylthiomethyl ether **14** was obtained according to a previously described procedure.[14] After removal of the main impurities by flash chromatography, this intermediate was coupled with pivalic acid in presence of *N*-iodosuccinimide yielding the pivaloyloxymethyl ester **15**.[15] Finally, debenzylation by hydrogenolysis gave prodrug **5h**.

Physicochemical and pharmacokinetic properties

Table 1 summarizes the physicochemical and pharmacokinetic properties of the ester prodrugs **5a–5h**, i.e., their aqueous solubility,[16] lipophilicity as quantified by the octanol–water partition coefficient (log *P*),[17] and permeability determined with the parallel artificial membrane permeability assay (PAMPA)[18] as well as with colorectal adenocarcinoma cells (Caco-2).[19] Table 1 also includes metabolic stability data, describing the susceptibility of the ester prodrugs to hydrolysis by rat and human liver associated esterases.[20]

Lipophilicity and permeability

Our primary goal, namely to increase lipophilicity and permeability of the biphenyl mannoside **4a**, could clearly be achieved. The log *P* coefficients increased in parallel with the number of carbons of the acyl promoiety. Whereas acetate **5a** was only

Table 1. Physicochemical and pharmacokinetic parameters of different ester prodrugs **5a–5h**.

Compound	log P^a	Solubility (μg/mL)^b	PAMPA log P_e (cm/s)^c	Caco-2 P_{app} (10^{-6} cm/s)^d a → b (absorptive)	b → a (secretory)	b → a/a → b	RLM $t_{1/2}$ (min)^e	HLM $t_{1/2}$ (min)
4a[10]	0.4±0.0	246±17	−7.2±0.0	0.4±0.0	1.8±0.1	5.0		
5a	0.9±0.1	146±6	−5.4±0.1	1.8±0.7	17.7±1.1	10	33	nd
5b	1.5±0.1	253±10	−5.0±0.0	4.0±0.6	15.2±0.7	3.8	6.5	3.0
5c	1.8±0.1	61±1	−4.6±0.0	10.5±0.9	19.5±0.1	1.9	3.7	nd
5d	1.8±0.0	145±9	−4.7±0.1	17.3±1.9	23.5±1.2	1.4	1.8	1.1
5e	2.3±0.1	58±7	−4.6±0.1	14.1±1.2	19.8±3.3	1.4	15	nd
5f	2.1±0.1	65±4	−4.4±0.0	17.8±2.4	24.3±3.0	1.4	2.0	nd
5g	2.2±0.1	149±5	−4.5±0.1	18.1±0.2	29.4±4.0	1.6	<1	<1
5h	2.1±0.1	154±12	−4.5±0.1	9.4±1.3	30.3±3.2	3.2	44	nd

Note: The indicated values represent the mean ± SD of replicate determinations.

[a]Octanol–water partition coefficients (log *P*) were determined by a miniaturized shake-flask procedure in sextuplicate.[17]

[b]Kinetic aqueous solubility was measured in triplicate.[16]

[c]P_e = effective permeability: diffusion through an artificial membrane was determined by the parallel artificial membrane permeability assay (PAMPA) in quadruplicate.[18]

[d]P_{app} = apparent permeability: permeation through a Caco-2 cell monolayer was assessed in the absorptive (a → b) and secretory (b → a) direction in triplicate. The initial compound concentration (c_0) in the donor chamber was 62.5 μM.[19]

[e]Microsomal stability was determined with pooled male rat liver microsomes (RLM) (0.125 mg/mL) and pooled human liver microsomes (HLM) (0.125 mg/mL) at pH 7.4 and 37 °C. The initial compound concentration was 2 μM. The concentration of the prodrug in the incubation was monitored by LC-MS and $t_{1/2}$ was calculated from the slope of the linear regression from the log percentage compound remaining versus incubation time relationship.[20]

slightly more lipophilic than the parent compound **4a**, propionate **5b**, isobutyrate **5c,** and butyrate **5d** exhibited markedly elevated lipophilicity. The aspired log *P* values above 2 could be reached in the case of pivaloate **5e,** pivaloyloxymethyl **5h**, isovalerate **5f**, and valerate **5g**. Furthermore, the effective permeability (log P_e) deduced from PAMPA rose proportionally to log *P*. Log P_e values above −5.7 are a strong indication for intestinal absorption, whereas those above −6.3 propose only a moderate potential. The PAMPA values for prodrugs **5a** and **5b** (log P_e −5.4 and −5.0) indicate a relevant improvement in membrane permeability compared to parent compound **4a** (log P_e −7.2) and optimal permeability for the more lipophilic esters **5c–5h** (log P_e −4.7 to −4.4).[21]

In addition to PAMPA, bi-directional permeation studies through a monolayer of Caco-2 cells were performed to reveal passive permeation as well as carrier-mediated transport through the cell membranes lining the small intestines.[19] By treatment with bis(4-nitrophenyl) phosphate (BNPP), the esterases expressed by Caco-2 cells were inhibited,[20] enabling the study of membrane permeation independent of enzyme-mediated hydrolysis.[22] Apparent permeability (P_{app}) derived from the experiments in the absorptive direction (apical → basal) paralleled the trends observed in lipophilicity (log *P*) and PAMPA (log P_e). Whereas **5a** and **5b** exhibited low permeability ($P_{app} < 5 \times 10^{-6}$ cm/s), high permeability ($P_{app} > 9 \times 10^{-6}$ cm/s) was detected for the remaining prodrugs. For the most polar prodrugs **5a** and **5b,** high P_{app} in the secretory direction (basal → apical) is leading to unfavorable efflux ratios (b → a/a → b). In these cases, efflux carrier activity probably outbalanced the slow diffusion in the absorptive direction.[23] Otherwise, the more lipophilic prodrugs **5c–5g** diffused more rapidly and therefore appeared only as weak efflux transporter substrates. Finally, the performance of **5h** remains in the high-permeability range. However, its overall evaluation is worsened when the efflux ratio of 3.2 is taken into account.

Solubility

For achieving oral bioavailability, quantitative dissolution of the orally administered prodrug in the intestine is an additional requirement.[24] Regarding aqueous solubility, the prodrugs listed in Table 1 can be divided into two categories: esters with branched acyl promoieties (isobutyrate **5c**, pivaloate **5e**, and isovalerate **5f**) were sparsely soluble in aqueous medium (around 60 μg/mL), whereas the linear esters (acetate **5a**, propionate **5b**, butyrate **5d**, and valerate **5g**) and the extended pivaloyloxymethyl ester in **5h** showed solubility values of at least 145 μg/mL. Provided that the prodrugs are applied at a therapeutic dose of at most 1 mg/kg body weight, aqueous solubility of 52 μg/mL should be aspired,[24] which could barely be achieved with the branched-chain derivatives. By contrast, the prodrugs **5a**, **5b**, **5d**, **5g**, and **5h** markedly exceeded this minimum solubility criterion for quantitative intestinal absorption.

Considering the two pivotal criteria for oral absorption — aqueous solubility and membrane permeability — the prodrug **5b** (high solubility, moderate permeability) as well as the derivatives **5d** and **5g** (moderate solubility, high permeability) showed the most promising profiles for a successful absorption.

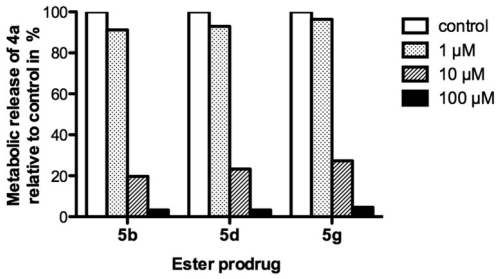
Gastrointestinal stability

Other prerequisites for developing orally bioavailable prodrugs are chemical stability under the conditions encountered in the gastrointestinal tract as well as resistance to hydrolysis during the absorption phase.[25] To assess the stability of the esters against hydrolysis under acidic and physiological conditions, the butyrate **5d** was dissolved in phosphate buffer (20 mM, pH 2.5 and 7.4) and acetate buffer (20 mM, pH 5.0) at 37 °C and stirred for 3 h.[26] At physiological or acidic pH, **5d** proved to be stable. We therefore expect only marginal prodrug loss in the strongly acidic environment of the stomach and the slightly acidic environment of the proximal small bowel.[27]

Enzymatic hydrolysis

Whereas the prodrug has to fulfill stability requirements prior to absorption, it should be rapidly cleaved once in circulation.[25] Enzymatic ester hydrolysis is undesirable during absorption but necessary once the prodrug has reached circulation. It is mediated by plasma-borne esterases or, as in case of many ester prodrugs, by the carboxylesterase (CES) localized in the endoplasmic reticulum of different tissues.[28] The CES superfamily encloses various isozymes classified into five subfamilies. The isozymes human carboxylesterase isotype 1 (hCE1), highly expressed in the liver but scarcely observed in the gastrointestinal tract, and human carboxylesterase isotype 2 (hCE2), present in both liver and small intestine, have been identified as major human CES.[20,29] To estimate the prodrugs' propensity to CES-mediated hydrolysis irrespective of the type of isozyme involved, we incubated the prodrugs **5a–5h** (initial concentration = 2 μM) with rat liver microsomes (RLM)

Fig. 2. Human liver microsome mediated hydrolysis of ester prodrugs **5b**, **5d**, and **5g** in presence of loperamide (1, 10, and 100 μM), a specific inhibitor of the human carboxylesterase isotype 2. The bars represent the metabolic release of the parent compound **4a** in the presence of different concentrations of inhibitor (1, 10, and 100 μM) relative to the accumulation in the control experiment without loperamide.



(0.125 mg/mL in TRIS–HCl 0.1 M, pH 7.4 at 37 °C). In a further step, we assessed the most promising esters **5b**, **5d**, and **5g** using human liver microsomes (HLM) (0.125 mg/mL). The metabolic half-lives ($t_{1/2}$) derived from the microsomal incubations revealed two major trends: first, an increasing susceptibility to hydrolysis along with increased lipophilicity of the linear esters and, second, sterically hindered branched acyl moieties, such as present in pivaloate **5e** and pivaloyloxymethyl **5h**, hampering the enzymatic turnover. Moreover, exposing the pivaloate ester by an acetal linker (→ **5h**) actually did not increase the rate of enzyme-mediated hydrolysis.

Rapid hydrolysis as detected for the propionate, butyrate, and valerate esters (**5b**, **5d**, and **5g**) is essential for the release of the pharmacologically active principle. However, the prodrug approach might only be successful when the cleavage takes place in liver and plasma and not in the small intestines during absorption or, more precisely, when the hydrolysis is mediated by the isozyme hCE1 rather than by hCE2.[25] Since liver microsomes prepared by differential centrifugation from a crude liver homogenate contain both CES, selective inhibition of only one reveals which isozyme is mainly involved. Therefore, loperamide (a specific inhibitor of hCE2) was added during the incubation of **5b**, **5d**, and **5g** with HLM.[30] Figure 2 summarizes the hydrolytic activity in presence of the hCE2 specific inhibitor at ascending concentrations (1, 10, and 100 μM). The metabolic turnover of the esters **5b**, **5d**, and **5g** was in fact inhibited by loperamide, which attributes the hydrolysis to the hCE2 isozyme.[30] Recognition of the employed promoieties by isotype 2 correlates with the reported substrate specificity of the two isozymes. Accordingly, hCE2 prefers esters with a relatively small acyl moiety and a large alcohol group, whereas hCE1 primarily catalyzes the hydrolysis of esters with a large acyl but small alcohol moiety.[20] Bioconversion mediated by isotype 2 might interfere with the intestinal absorption, even in the case of the well-soluble and permeable prodrugs **5b**, **5d**, and **5g**. By hydrolysis within the enterocytes, the polar active principle **4a** is formed and likely effluxed back into the gut lumen instead into the portal blood.[25]

Conclusions

Several short-chain fatty acids (propionic acid, butyric acid, and valeric acid) were identified as useful acyl promoieties for optimizing the intestinal absorption potential of the biphenyl mannoside **4a**. We showed that acylation of only one hydroxyl group on the sugar moiety was sufficient to improve lipophilicity into the range required for membrane permeability, whilst sufficient solubility could be sustained. Moreover, the introduced ester promoieties are stable in acidic conditions of the gastrointestinal tract; however, they are fast hydrolyzed by esterases upon absorption.

The downside of acylation of a hydroxyl group of the sugar moiety could be premature cleavage of the promoiety by hCE2 located in the enterocytes. Therefore, choosing the appropriate prodrug moiety should be guided by enzymatic stability studies. Furthermore, the half-life of the prodrug should allow that the majority of the prodrug is absorbed unchanged.

For proving the benefits of the prodrug approach on oral bioavailability and for assessing whether the intestinal uptake is affected by concomitant hydrolysis, in vivo pharmacokinetic studies in a mouse model are currently performed.

Experimental section

Synthesis

General methods

NMR spectra were recorded on a Bruker Avance DMX-500 (500 MHz) spectrometer. Assignment of ^{1}H and ^{13}C NMR spectra was achieved using 2D methods (COSY, HSQC, and HMBC). Chemical shifts are expressed in ppm using residual $CHCl_3$, CHD_2OD, or HDO as references. Optical rotations were measured using a Perkin-Elmer Polarimeter 341. Electron spray ionization mass spectra (ESI-MS) were obtained on a Waters micromass ZQ mass spectrometer. The LC-HRMS analyses 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 TLC using glass plates coated with silica gel 60 F_{254} (Merck) 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). MPLC separations were carried out on a CombiFlash Companion or Rf from Teledyne Isco equipped with RediSep normal-phase columns. All compounds used for biological assays are at least of 97% purity based on HPLC analytical results. Commercially available reagents were purchased from Aldrich, Alfa Aesar, ABCR, or Acros Organics. Solvents were purchased from Sigma-Aldrich or Acros and were dried prior to use where indicated. Methanol (MeOH), pyridine, and dimethyl sulfoxide (DMSO) were dried by storing with activated molecular sieves of 3 or 4 Å for at least 1 day. Dichloromethane (DCM) was dried by filtration over Al_2O_3 (Fluka, type 5016 A basic). Molecular sieves of 3 and 4 Å were activated in vacuo at 500 °C for 1 h immediately before use.

General procedure A for esterification

To a solution of **12** in dry pyridine (2 mL) were added Ac_2O or the corresponding acyl chloride and a catalytic amount of dimethylaminopyridine (DMAP). The mixture was stirred at rt under argon until the reaction was complete (monitored by TLC) and then diluted with ethyl acetate (EtOAc) and washed with H_2O and brine. The organic layer was dried over Na_2SO_4, concentrated in vacuo, and co-evaporated with xylene. The residue was purified by MPLC on silica gel (petroleum ether – EtOAc, 7:3) to afford **13a–13g**.

General procedure B for hydrogenolysis

A solution of **13a–13g** or **15** in EtOH was stirred under hydrogen (1 bar) in the presence of $Pd(OH)_2$/C (E101 NE/W, 20% Pd) at rt until the reaction was complete (monitored by TLC) and then filtered through Celite, washed with methanol (MeOH), and concentrated in vacuo. The residue was purified by MPLC on silica gel (DCM–MeOH) to give **5a–5h**.

4-Iodophenyl 2,3,4,6-tetra-O-acetyl-α-D-mannopyranoside (8)

Penta-*O*-acetyl-D-mannopyranose (15.0 g, 45.4 mmol), 4-iodophenol (11.2 g, 50.0 mmol, 1.1 equiv.), and activated molecular sieves of 4 Å (2.00 g) were stirred in dry DCM (60 mL) under argon. After 1 h, a first

portion of $BF_3{\cdot}Et_2O$ (10 mL, 81 mmol, 1.8 equiv.) was added dropwise followed by the addition of a second portion (6.8 mL, 55.2 mmol, 1.2 equiv.) 4 h later. The reaction was stirred at 40 °C for 30 h. The reaction mixture was filtered through Celite and the filtrate was diluted with EtOAc (250 mL), washed with satd aq $NaHCO_3$ (3 × 100 mL) and brine (100 mL). The organic layer was dried over Na_2SO_4 and concentrated in vacuo. The residue was crystallized from Et_2O–hexane (1:1) to give **8**. The mother liquor was concentrated and **8** was crystallized again from Et_2O–hexane (2:1). The filtrate was concentrated and purified by MPLC on silica gel (petroleum ether – EtOAc, 9:1). Compound **8** was obtained in an overall yield of 67% (16.8 g). Analytical data were in accordance with literature data.[31]

4-Iodophenyl α-D-mannopyranoside (9)

To a solution of **8** (16.8 g, 30.5 mmol) in dry MeOH (100 mL) was added freshly prepared 1 M MeONa–MeOH (2 mL) under argon. The reaction mixture was stirred overnight at rt and then neutralized with Amberlyst-15 (H^+) ion-exchange resin, filtered, and concentrated in vacuo. Recrystallization from ethanol (250 mL) afforded white crystals of **9**. The mother liquor was concentrated and purified by MPLC on silica gel (DCM–MeOH, 9:1). Compound **9** was obtained in an overall yield of 61% (7.10 g). $[\alpha]_D^{20}$ + 106.5 (*c* 1.00, MeOH). ^{1}H NMR (500 MHz, CD_3OD): δ = 7.61 (d, *J* = 8.9 Hz, 2H, Ar-H), 6.96 (d, *J* = 8.9 Hz, 2H, Ar-H), 5.48 (d, *J* = 1.4 Hz, 1H, H-1), 4.01 (dd, *J* = 1.8, 3.2 Hz, 1H, H-2), 3.89 (dd, *J* = 3.4, 9.4 Hz, 1H, H-3), 3.80–3.69 (m, 3H, H-4, H-6a, H-6b), 3.57 ppm (ddd, *J* = 2.4, 5.4, 9.7 Hz, 1H, H-5). ^{13}C NMR (125 MHz, CD_3OD): δ = 157.84, 139.52, 120.15 (5C, Ar-C), 100.13 (C-1), 85.31 (Ar-C), 75.52 (C-5), 72.32 (C-3), 71.84 (C-2), 68.28 (C-4), 62.65 ppm (C-6). HRMS: *m*/*z*: calcd. for $C_{12}H_{15}INaO_6$ $[M+Na]^+$: 404.9811; found: 404.9808.

4-Iodophenyl 2,3,4-tri-O-benzyl-α-D-mannopyranoside (10)

To a stirred solution of **9** (6.61 g, 17.3 mmol) in dry DMF (17 mL) were added TBDMSCl (2.61 g, 17.3 mmol, 1.0 equiv.) and imidazole (2.35 g, 34.6 mmol, 2.0 equiv.) under argon at 0 °C. After 1 h, the reaction mixture was removed from the ice bath and allowed to reach rt. Another portion of TBDMSCl (0.26 g, 1.73 mmol, 0.1 equiv.) was added after 15 h. The reaction was stirred for 3 h until **9** was completely consumed. The reaction mixture was diluted with DCM (200 mL) and washed with satd aq $NaHCO_3$ (2 × 150 mL) and brine (150 mL). The organic layer was dried over Na_2SO_4, concentrated in vacuo, and co-evaporated with toluene (2 × 100 mL) to afford 9.60 g of crude product. The obtained compound (6.01 g) was dissolved in dry DMF (28 mL) under argon. Sodium hydride (1.75 g, 43.6 mmol, 60% in mineral oil) was added to the stirred solution together with an additional portion of DMF (16 mL) at 0 °C followed by the addition of BnBr (6.47 mL, 54.5 mmol). The reaction mixture was removed from the ice bath and allowed to reach rt. Bu_4NI (0.88 g, 2.42 mmol) was added after 2 h. When the reaction was complete after another 2 h, the mixture was diluted with EtOAc (250 mL) and washed with satd aq $NaHCO_3$ (150 mL), H_2O (150 mL), and brine (100 mL). The organic layer was dried over Na_2SO_4, concentrated in vacuo, and co-evaporated with toluene (75 mL) to afford 10.8 g of a yellowish product. The crude compound (10.8 g) was then dissolved in MeOH (70 mL) under argon and a solution of H_2SO_4 in MeOH (1 M, 560 μL) was added dropwise. The reaction was stirred at 0 °C until completion (18 h, monitored by TLC) and then the mixture was diluted with EtOAc (100 mL) and washed with satd aq $NaHCO_3$ (100 mL), H_2O (100 mL), and brine (100 mL). The aqueous layer was extracted with EtOAc (100 mL). The combined organic layers were dried over Na_2SO_4, concentrated in vacuo, and the major impurities were removed by MPLC on silica gel (petroleum ether – EtOAc, 85:15). The crude **10** was crystallized from petroleum ether – EtOAc (3:1). The mother liquor was concentrated and the residue was recrystallized from MeOH (20 mL). Compound **10** was obtained in an overall yield of 58% (4.32 g) over three steps. $[\alpha]_D^{20}$ + 55.9 (*c* 1.00, EtOAc). ^{1}H NMR (500 MHz, $CDCl_3$): δ = 7.48 (d, *J* = 8.7 Hz, 2H, Ar-H), 7.35–7.24 (m, 15H, Ar-H), 6.67 (d, *J* = 8.7 Hz, 2H, Ar-H), 5.49 (d, *J* = 1.7 Hz, 1H, H-1), 4.90 (d, *J* = 10.9 Hz, 1H, PhCH_2O), 4.78 (d, *J* = 12.2 Hz, 1H, PhCH_2O), 4.70–4.62 (m, 4H, PhCH_2O), 4.07–4.02 (m, 2H, H-4, H-3), 3.89 (br s, 1H, H-2), 3.70–3.69 (m, 2H, H-6a, H-6b), 3.62 (m, 1H, H-5), 1.81 ppm (s, 1H, OH). ^{13}C NMR (125 MHz, $CDCl_3$): δ = 155.93, 138.53, 138.42, 138.35, 138.09, 128.61, 128.58, 128.19, 128.03, 127.95, 127.86, 127.80, 118.73 (23C, Ar-C), 96.66 (C-1), 85.12 (Ar-C), 79.88 (C-3 or C-4), 75.40 (PhCH_2O), 74.74 (C-2), 74.43 (C-3 or C-4), 73.38 (PhCH_2O), 73.28 (C-5), 72.69 (PhCH_2O), 62.06 ppm (C-6). ESI-MS: *m*/*z*: calcd. for $C_{33}H_{33}INaO_6$ $[M+Na]^+$: 675.12, found: 675.18.

4′-(Methylsulfonyl)-biphenyl-4-yl 2,3,4-tri-O-benzyl-α-D-mannopyranoside (12)

Compounds **10** (2.50 g, 3.83 mmol) and **11** (843 mg, 4.21 mmol) were dissolved in dry DMF (20 mL) under argon. The mixture was degassed in an ultrasonic bath and flushed with argon for 5 min followed by the addition of K_3PO_4 (2.44 g, 11.5 mmol, 3.0 equiv.) and $Pd(dppf)Cl_2{\cdot}CH_2Cl_2$ (156 mg, 0.19 mmol, 0.05 equiv.). The mixture was stirred at 80 °C for 19 h. The reaction mixture was diluted with EtOAc (200 mL), washed with H_2O (2 × 120 mL) and brine (120 mL), dried over Na_2SO_4, and concentrated in vacuo. The residue was purified by MPLC on silica gel (petroleum ether – EtOAc, 1:1) to give **12** (2.24 g, 86%) as a white solid. $[\alpha]_D^{20}$ + 70.8 (*c* 1.01, EtOAc). ^{1}H NMR (500 MHz, $CDCl_3$): δ = 7.99–7.97 (m, 2H, Ar-H), 7.72–7.70 (m, 2H, Ar-H), 7.53–7.52 (m, 2H, Ar-H), 7.42–7.29 (m, 15H, Ar-H), 7.08–7.06 (m, 2H, Ar-H), 5.58 (d, *J* = 1.9 Hz, 1H, H-1), 4.97 (d, *J* = 10.9 Hz, 1H, PhCH_2O), 4.87 (d, *J* = 12.3 Hz, 1H, PhCH_2O), 4.78–4.69 (m, 4H, PhCH_2O), 4.16–4.09 (m, 2H, H-3, H-4), 4.00 (d, *J* = 2.2 Hz, 1H, H-2), 3.79–3.73 (m, 3H, H-5, H-6a, H-6b,), 3.09 ppm (s, 3H, SO_2CH_3). ^{13}C NMR (125 MHz, $CDCl_3$): δ = 156.68, 146.14, 138.87, 138.46, 138.35, 138.15, 133.34, 128.78, 128.65, 128.63, 128.24, 128.09, 128.06, 128.02, 127.91, 127.84, 127.67, 116.99 (30C, Ar-C), 96.67 (C-1), 80.00 (C-3 or C-4), 75.46 (PhCH_2O), 74.79 (C-2), 74.59 (C-3 or C-4), 73.41 (PhCH_2O), 73.31 (C-5), 72.74 (PhCH_2O), 62.25 (C-6), 44.79 ppm (SO_2CH_3). ESI-MS: *m*/*z*: calcd. for $C_{40}H_{40}NaO_8S$ $[M+Na]^+$: 703.23, found: 703.26.

4′-(Methylsulfonyl)-biphenyl-4-yl 6-O-acetyl-2,3,4-tri-O-benzyl-α-D-mannopyranoside (13a)

Prepared according to the general procedure A from **12** (100 mg, 0.147 mmol), Ac_2O (27 μL, 0.294 mmol, 2.0 equiv.), and DMAP (1 mg, 0.008 mmol, 0.05 eq). The reaction was started at 0 °C and allowed to warm up to rt. Yield: 104 mg (96%) as a colorless oil. $[\alpha]_D^{20}$ + 70.0 (*c* 1.00, $CHCl_3$). ^{1}H NMR (500 MHz, $CDCl_3$): δ = 7.90 (d, *J* = 8.5 Hz, 2H, Ar-H), 7.63 (d, *J* = 8.4 Hz, 2H, Ar-H), 7.46 (d, *J* = 8.8 Hz, 2H, Ar-H), 7.33–7.17 (m, 15H, Ar-H), 7.03 (d, *J* = 8.7 Hz, 2H, Ar-H), 5.52 (d, *J* = 1.7 Hz, 1H, H-1), 4.88 (d, *J* = 10.8 Hz, 1H, PhCH_2O), 4.77–4.63 (m, 4H, PhCH_2O), 4.54 (d, *J* = 10.8 Hz, 1H, PhCH_2O), 4.26–4.18 (m, 2H, H-6a, H-6b), 4.07 (dd, *J* = 3.0, 9.2 Hz, 1H, H-3), 3.98–3.92 (m, 2H, H-2, H-4), 3.81 (m, 1H, H-5), 2.99 (s, 3H, SO_2CH_3), 1.92 ppm (s, 3H, $COCH_3$). ^{13}C NMR (125 MHz, $CDCl_3$): δ = 170.83 (CO), 156.71, 146.07, 138.84, 138.29, 138.10, 138.09, 133.31, 128.64, 128.59, 128.57, 128.55, 128.24, 128.05, 127.99, 127.95, 127.93, 127.90, 127.84, 127.61, 117.12 (30C, Ar-C), 96.43 (C-1), 79.97 (C-3), 75.34 (PhCH_2O), 74.48 (C-2), 74.32 (C-4), 73.05, 72.56 (2 PhCH_2O), 71.01 (C-5), 63.21 (C-6), 44.74 (SO_2CH_3), 20.95 ppm ($COCH_3$). ESI-MS: *m*/*z*: calcd. for $C_{42}H_{42}NaO_9S$ $[M+Na]^+$: 745.24, found: 745.31.

4′-(Methylsulfonyl)-biphenyl-4-yl 2,3,4-tri-O-benzyl–6-O-propionyl-α-D-mannopyranoside (13b)

Prepared according to the general procedure A from **12** (70 mg, 0.103 mmol), propionyl chloride (33 μL, 0.309 mmol, 3.0 equiv.), and DMAP (1 mg, 0.008 mmol, 0.08 equiv.). The reaction was started at rt and then warmed up to 60 °C. Yield: 38 mg (58%) as a colorless oil. $[\alpha]_D^{20}$ + 63.0 (*c* 1.84, $CHCl_3$). ^{1}H NMR (500 MHz, $CDCl_3$): δ = 7.97 (d, *J* = 8.4 Hz, 2H, Ar-H), 7.70 (d, *J* = 8.4 Hz, 2H, Ar-H), 7.51 (d, *J* = 8.7 Hz, 2H, Ar-H), 7.41–7.27 (m, 15H, Ar-H), 7.10 (d, *J* = 8.7 Hz, 2H, Ar-H), 5.58 (d, *J* = 1.5 Hz, 1H, H-1), 4.95 (d, *J* = 10.7 Hz, 1H, PhCH_2O), 4.83 (d, *J* = 12.3 Hz, 1H, PhCH_2O), 4.77–4.72 (m, 3H, PhCH_2O), 4.60 (d,

J = 10.7 Hz, 1H, PhCH_2O), 4.34–4.28 (m, 2H, H-6a, H-6b), 4.14 (dd, J = 3.0, 9.2 Hz, 1H, H-3), 4.03 (t, J = 9.5 Hz, 1H, H-4), 3.99 (m, 1H, H-2), 3.88 (ddd, J = 2.4, 4.4, 9.8 Hz, 1H, H-5), 3.08 (s, 3H, SO_2CH_3), 2.27 (q, J = 7.5 Hz, 2H, CH_2COO), 1.07 ppm (t, J = 7.6 Hz, 3H, CH_3). ^{13}C NMR (125 MHz, $CDCl_3$): δ = 174.29 (CO), 156.77, 146.18, 138.89, 138.36, 138.17, 138.15, 137.46, 137.16, 133.37, 128.70, 128.65, 128.63, 128.61, 128.29, 128.12, 128.01, 127.99, 127.97, 127.92, 127.68, 118.42, 117.18 (30C, Ar-C), 96.44 (C-1), 80.02 (C-3), 75.44 (PhCH_2O), 74.61 (C-2), 74.50 (C-4), 73.15, 72.65 (2 PhCH_2O), 71.19 (C-5), 63.14 (C-6), 44.80 (SO_2CH_3), 27.66 (CH_2COO), 9.21 ppm (CH_3). ESI-MS: *m/z*: calcd. for $C_{43}H_{44}NaO_9S$ [M+Na]$^+$: 759.26, found: 759.13.

4′-(Methylsulfonyl)-biphenyl-4-yl 2,3,4-tri-O-benzyl-6-O-isobutyryl-α-D-mannopyranoside (13c)

Prepared according to the general procedure A from **12** (70 mg, 0.103 mmol), isobutyryl chloride (33 μL, 0.309 mmol, 3.0 equiv.), and DMAP (2 mg, 0.016 mmol, 0.16 equiv.). The reaction was started at rt and then warmed up to 60 °C. Yield: 63 mg (82%) as a colorless oil. $[\alpha]_D^{20}$ + 60.0 (*c* 3.17, $CHCl_3$). ^{1}H NMR (500 MHz, $CDCl_3$): δ = 7.96 (d, J = 8.4 Hz, 2H, Ar-H), 7.68 (d, J = 8.4 Hz, 2H, Ar-H), 7.49 (d, J = 8.7 Hz, 2H, Ar-H), 7.39–7.26 (m, 15H, Ar-H), 7.09 (d, J = 8.7 Hz, 2H, Ar-H), 5.57 (d, J = 1.6 Hz, 1H, H-1), 4.94 (d, J = 10.6 Hz, 1H, PhCH_2O), 4.81 (d, J = 12.2 Hz, 1H, PhCH_2O), 4.74–4.71 (m, 3H, PhCH_2O), 4.58 (d, J = 10.6 Hz, 1H, PhCH_2O), 4.35 (dd, J = 1.7, 11.9 Hz, 1H, H-6a), 4.25 (dd, J = 5.2, 11.9 Hz, 1H, H-6b), 4.12 (dd, J = 3.0, 9.2 Hz, 1H, H-3), 4.00 (t, J = 9.5 Hz, 1H, H-4), 3.97 (m, 1H, H-2), 3.86 (ddd, J = 1.6, 5.0, 9.8 Hz, 1H, H-5), 3.06 (s, 3H, SO_2CH_3), 2.49 (hept, J = 7.0 Hz, 1H, CH(CH_3)$_2$), 1.09 (d, J = 7.0 Hz, 3H, CH_3), 1.06 ppm (d, J = 7.0 Hz, 3H, CH_3). ^{13}C NMR (125 MHz, $CDCl_3$): δ = 176.85 (CO), 156.73, 146.17, 138.85, 138.33, 138.14, 133.31, 128.67, 128.62, 128.56, 128.22, 128.06, 128.01, 127.95, 127.90, 127.64, 117.14 (30C, Ar-C), 96.34 (C-1), 79.96 (C-3), 75.48 (PhCH_2O), 74.67 (C-2), 74.61 (C-4), 73.16, 72.63 (2 PhCH_2O), 71.32 (C-5), 62.98 (C-6), 44.77 (SO_2CH_3), 34.10 (CH), 19.12, 18.92 ppm (2 CH_3). ESI-MS: *m/z*: calcd. for $C_{44}H_{46}NaO_9S$ [M+Na]$^+$: 773.28, found: 773.28.

4′-(Methylsulfonyl)-biphenyl-4-yl 2,3,4-tri-O-benzyl-6-O-butyryl-α-D-mannopyranoside (13d)

Prepared according to the general procedure A from **12** (54 mg, 0.080 mmol), butyryl chloride (10 μL, 0.094 mmol, 1.2 equiv.), and DMAP (1 mg, 0.008 mmol, 0.1 equiv.). The reaction mixture was stirred at rt. Yield: 45 mg (75%) as a colorless oil. $[\alpha]_D^{20}$ + 68.6 (*c* 1.13, $CHCl_3$). ^{1}H NMR (500 MHz, $CDCl_3$): δ = 8.02 (d, J = 8.4 Hz, 2H, Ar-H), 7.75 (d, J = 8.5 Hz, 2H, Ar-H), 7.56 (d, J = 8.8 Hz, 2H, Ar-H), 7.45–7.32 (m, 15H, Ar-H), 7.15 (d, J = 8.8 Hz, 2H, Ar-H), 5.64 (d, J = 1.7 Hz, 1H, H-1), 5.00 (d, J = 10.7 Hz, 1H, PhCH_2O), 4.87 (d, J = 12.3 Hz, 1H, PhCH_2O), 4.81–4.77 (m, 3H, PhCH_2O), 4.65 (d, J = 10.7 Hz, 1H, PhCH_2O), 4.40–4.33 (m, 2H, H-6a, H-6b), 4.19 (dd, J = 3.0, 9.2 Hz, 1H, H-3), 4.07 (t, J = 9.5 Hz, 1H, H-4), 4.04 (m, 1H, H-2), 3.92 (ddd, J = 2.2, 4.6, 9.8 Hz, 1H, H-5), 3.12 (s, 3H, SO_2CH_3), 2.29 (t, J = 7.5 Hz, 2H, CH_2COO), 1.67–1.59 (m, 2H, CH_2), 0.91 ppm (t, J = 7.4 Hz, 3H, CH_3). ^{13}C NMR (125 MHz, $CDCl_3$): δ = 173.42 (CO), 156.72, 146.10, 138.83, 138.31, 138.12, 138.10, 133.27, 128.64, 128.59, 128.57, 128.55, 128.21, 128.05, 127.98, 127.95, 127.93, 127.91, 127.86, 127.60, 117.10 (30C, Ar-C), 96.37 (C-1), 79.95 (C-3), 75.40 (PhCH_2O), 74.58, 74.49 (C-2, C-4), 73.10, 72.59 (2 PhCH_2O), 71.17 (C-5), 62.94 (C-6), 44.73 (SO_2CH_3), 36.23 (CH_2COO), 18.47 (CH_2), 13.80 ppm (CH_3). ESI-MS: *m/z*: calcd. for $C_{44}H_{46}NaO_9S$ [M+Na]$^+$: 773.28, found: 773.44.

4′-(Methylsulfonyl)-biphenyl-4-yl 2,3,4-tri-O-benzyl-6-O-pivaloyl-α-D-mannopyranoside (13e)

Prepared according to the general procedure A from **12** (100 mg, 0.147 mmol), pivaloyl chloride (36 μL, 0.294 mmol, 2.0 equiv.), and DMAP (1 mg, 0.008 mmol, 0.05 equiv.). The reaction was started at rt and then warmed up to 60 °C. Yield: 101 mg (90%) as a colorless oil. $[\alpha]_D^{20}$ + 88.7 (*c* 1.02, $CHCl_3$). ^{1}H NMR (500 MHz, $CDCl_3$): δ = 7.90 (d, J = 8.4 Hz, 2H, Ar-H), 7.62 (d, J = 8.4 Hz, 2H, Ar-H), 7.44 (d, J = 8.7 Hz, 2H, Ar-H), 7.35–7.20 (m, 15H, Ar-H), 7.05 (d, J = 8.7 Hz, 2H, Ar-H), 5.53 (d, J = 1.6 Hz, 1H, H-1), 4.90 (d, J = 10.6 Hz, 1H, PhCH_2O), 4.77 (d, J = 12.2 Hz, 1H, PhCH_2O), 4.68–4.65 (m, 3H, PhCH_2O), 4.55 (d, J = 10.7 Hz, 1H, PhCH_2O), 4.36 (dd, J = 1.5, 11.8 Hz, 1H, H-6a), 4.15–4.08 (m, 2H, H-3, H-6b), 3.98–3.93 (m, 2H, H-2, H-4), 3.82 (ddd, J = 1.3, 5.3, 9.8 Hz, 1H, H-5), 2.99 (s, 3H, SO_2CH_3), 1.05 ppm (s, 9H, C(CH_3)$_3$). ^{13}C NMR (125 MHz, $CDCl_3$): δ = 178.19 (CO), 156.61, 146.05, 138.75, 138.25, 138.07, 133.16, 128.59, 128.52, 128.46, 128.11, 127.97, 127.91, 127.84, 127.83, 127.52, 127.59, 117.08 (30C, Ar-C), 96.13 (C-1), 79.84 (C-3), 75.40 (PhCH_2O), 74.71, 74.64 (C-2, C-4), 73.11, 72.50 (2 PhCH_2O), 71.33 (C-5), 63.07 (C-6), 44.62 (SO_2CH_3), 38.82 (*C*(CH_3)$_3$), 27.17 ppm (C(*C*H_3)$_3$). ESI-MS: *m/z*: calcd. for $C_{45}H_{48}NaO_9S$ [M+Na]$^+$: 787.29, found: 787.36.

4′-(Methylsulfonyl)-biphenyl-4-yl 2,3,4-tri-O-benzyl-6-O-isovaleryl-α-D-mannopyranoside (13f)

Prepared according to the general procedure A from **12** (38 mg, 0.056 mmol), isovaleryl chloride (27 μL, 0.223 mmol, 4.0 equiv.), and DMAP (1 mg, 0.008 mmol, 0.14 equiv.). The reaction mixture was stirred at rt. Yield: 35 mg (81%) as a colorless oil. $[\alpha]_D^{20}$ + 59.0 (*c* 0.93, $CHCl_3$). ^{1}H NMR (500 MHz, $CDCl_3$): δ = 7.99 (d, J = 8.4 Hz, 2H, Ar-H), 7.72 (d, J = 8.4 Hz, 2H, Ar-H), 7.53 (d, J = 8.7 Hz, 2H, Ar-H), 7.42–7.29 (m, 15H, Ar-H), 7.11 (d, J = 8.7 Hz, 2H, Ar-H), 5.59 (d, J = 1.6 Hz, 1H, H-1), 4.97 (d, J = 10.7 Hz, 1H, PhCH_2O), 4.84 (d, J = 12.3 Hz, 1H, PhCH_2O), 4.77–4.74 (m, 3H, PhCH_2O), 4.62 (d, J = 10.7 Hz, 1H, PhCH_2O), 4.38 (dd, J = 1.9, 11.9 Hz, 1H, H-6a), 4.29 (dd, J = 4.9, 12.0 Hz, 1H, H-6b), 4.15 (dd, J = 3.0, 9.2 Hz, 1H, H-3), 4.04 (t, J = 9.5 Hz, 1H, H-4), 4.00 (m, 1H, H-2), 3.88 (ddd, J = 1.8, 4.8, 9.8 Hz, 1H, H-5), 3.09 (s, 3H, SO_2CH_3), 2.16 (d, J = 7.1 Hz, 2H, CH_2COO), 2.04 (m, 1H, CH(CH_3)$_2$), 0.89 (d, J = 3.0 Hz, 3H, CH_3), 0.88 ppm (d, J = 3.0 Hz, 3H, CH_3). ^{13}C NMR (125 MHz, $CDCl_3$): δ = 172.93 (CO), 156.76, 146.14, 138.85, 138.33, 138.15, 138.13, 133.29, 128.66, 128.62, 128.60, 128.58, 128.22, 128.09, 127.98, 127.96, 127.94, 127.89, 127.62, 117.11 (30C, Ar-C), 96.39 (C-1), 79.97 (C-3), 75.45 (PhCH_2O), 74.61, 74.57 (C-2, C-4), 73.13, 72.62 (2 PhCH_2O), 71.22 (C-5), 62.89 (C-6), 44.77 (SO_2CH_3), 43.48 (CH_2COO), 25.70 (CH(CH_3)$_2$), 22.56 ppm (2 CH_3). ESI-MS: *m/z*: calcd. for $C_{45}H_{48}NaO_9S$ [M+Na]$^+$: 787.29, found: 787.46.

4′-(Methylsulfonyl)-biphenyl-4-yl 2,3,4-tri-O-benzyl-6-O-valeryl-α-D-mannopyranoside (13g)

Prepared according to the general procedure A from **12** (69 mg, 0.101 mmol), valeryl chloride (49 μL, 0.404 mmol, 4.0 equiv.), and DMAP (1 mg, 0.008 mmol, 0.08 equiv.). The reaction was started at rt and then warmed up 60 °C. Yield: 70 mg (91%) as a yellowish oil. $[\alpha]_D^{20}$ + 63.9 (*c* 1.00, $CHCl_3$). ^{1}H NMR (500 MHz, $CDCl_3$): δ = 7.90 (d, J = 8.4 Hz, 2H, Ar-H), 7.64 (d, J = 8.4 Hz, 2H, Ar-H), 7.45 (d, J = 8.7 Hz, 2H, Ar-H), 7.33–7.20 (m, 15H, Ar-H), 7.03 (d, J = 8.7 Hz, 2H, Ar-H), 5.52 (d, J = 1.5 Hz, 1H, H-1), 4.88 (d, J = 10.7 Hz, 1H, PhCH_2O), 4.76 (d, J = 12.3 Hz, 1H, PhCH_2O), 4.70–4.63 (m, 3H, PhCH_2O), 4.53 (d, J = 10.7 Hz, 1H, PhCH_2O), 4.28–4.21 (m, 2H, H-6a, H-6b), 4.07 (dd, J = 3.0, 9.2 Hz, 1H, H-3), 3.95 (t, J = 9.5 Hz, 1H, H-4), 3.92 (m, 1H, H-2), 3.81 (ddd, J = 2.3, 4.4, 9.8 Hz, 1H, H-5), 3.01 (s, 3H, SO_2CH_3), 2.19 (t, J = 9.5 Hz, 2H, CH_2COO), 1.50–1.44 (m, 2H, CH_2), 1.24–1.17 (m, 2H, CH_2), 0.77 ppm (t, J = 7.4 Hz, 3H, CH_3). ^{13}C NMR (125 MHz, $CDCl_3$): δ = 173.63 (CO), 156.76, 146.12, 138.84, 138.32, 138.14, 138.12, 133.29, 128.65, 128.60, 128.59, 128.57, 128.22, 128.06, 127.99, 127.97, 127.94, 127.92, 127.88, 127.61, 117.12 (30C, Ar-C), 96.43 (C-1), 79.97 (C-3), 75.21 (PhCH_2O), 74.59, 74.52 (C-2, C-4), 73.11, 72.60 (2 PhCH_2O), 71.17 (C-5), 63.01 (C-6), 44.75 (SO_2CH_3), 34.04 (CH_2COO), 27.00, 22.35 (2 CH_2), 13.78 ppm (CH_3). ESI-MS: *m/z*: calcd. for $C_{45}H_{48}NaO_9S$ [M+Na]$^+$: 787.29, found: 787.56.

4′-(Methylsulfonyl)-biphenyl-4-yl 2,3,4-tri-O-benzyl-6-O-pivaloyloxymethyl-α-D-mannopyranoside (15)

Compound **12** (250 mg, 0.367 mmol) was dissolved in dry DMSO under argon followed by the addition of Ac_2O (2 mL) and AcOH (0.2 mL). The reaction was stirred at rt for 19 h. The mixture was diluted with EtOAc (50 mL) and washed with satd aq $NaHCO_3$ (2 × 50 mL) and brine (50 mL). The organic layer was dried over Na_2SO_4,

concentrated in vacuo, and purified by MPLC on silica gel (petroleum ether – EtOAc, 7:3) to afford 153 mg of **14**. The intermediate **14** (113 mg) was dissolved in dry DMF (2 mL) under argon and PivOH (47 mg, 0.460 mmol) and NIS (52 mg, 0.230 mmol) were added. The reaction was quenched after 17 h with Et_3N (0.100 mL). Then, the mixture was diluted with EtOAc (50 mL) and washed with satd aq $NaHCO_3$ (50 mL), H_2O (50 mL) and brine (50 mL). The organic layer was dried over Na_2SO_4, concentrated in vacuo and purified by MPLC on silica gel (petroleum ether – EtOAc, 7:3). Compound **15** was obtained in an overall yield of 33% (72 mg) over two steps. $[\alpha]_D^{20}$ + 64.3 (*c* 0.58, $CHCl_3$). ^{1}H NMR (500 MHz, $CDCl_3$): δ = 7.84 (d, *J* = 8.4 Hz, 2H, Ar-H), 7.63 (d, *J* = 8.4 Hz, 2H, Ar-H), 7.45 (d, *J* = 8.7 Hz, 2H, Ar-H), 7.32–7.21 (m, 15H, Ar-H), 7.02 (d, *J* = 8.7 Hz, 2H, Ar-H), 5.50 (d, *J* = 1.6 Hz, 1H, H-1), 5.22 (s, 2H, OCH_2O), 4.88 (d, *J* = 10.9 Hz, 1H, PhCH_2O), 4.75 (d, *J* = 12.4 Hz, 1H, PhCH_2O), 4.70 (d, *J* = 12.4 Hz, 1H, PhCH_2O), 4.65–4.62 (m, 2H, PhCH_2O), 4.59 (d, *J* = 11.0 Hz, 1H, PhCH_2O), 4.05–3.97 (m, 2H, H-3, H-4), 3.90 (m, 1H, H-2), 3.83 (dd, *J* = 4.4, 10.7 Hz, 1H, H-6a), 3.78–3.72 (m, 2H, H-5, H-6b), 2.99 (s, 3H, SO_2CH_3), 1.08 ppm (s, 9H, $C(CH_3)_3$). ^{13}C NMR (125 MHz, $CDCl_3$): δ = 178.10 (CO), 156.90, 146.19, 138.83, 138.47, 138.19, 133.28, 128.74, 128.60, 128.59, 128.55, 128.08, 128.00, 127.98, 127.86, 127.83, 127.67, 117.16 (30C, Ar-*C*), 96.74 (C-1), 89.94 (OCH_2O), 80.00 (C-3), 75.25 (PhCH_2O), 74.55, 74.50 (C-2, C-4), 73.16, 72.62 (2 PhCH_2O), 72.38 (C-5), 69.04 (C-6), 44.78 (SO_2CH_3), 39.04 ($C(CH_3)_3$), 27.17 ppm ($C(CH_3)_3$). ESI-MS: *m*/*z*: calcd. for $C_{46}H_{50}NaO_{10}S$ [M+Na]$^+$: 817.30, found: 817.39.

4′-(Methylsulfonyl)-biphenyl-4-yl 6-O-acetyl-α-D-mannopyranoside (5a)

Prepared according to the general procedure B from **13a** (50 mg, 0.069 mmol) with 15 mg of $Pd(OH)_2$/C (E101 NE/W, 20% Pd) in EtOH (5 mL). Purified by MPLC on silica gel (DCM–MeOH, 9:1). Yield: 24 mg (75%) as a white solid. $[\alpha]_D^{20}$ + 101.2 (*c* 0.19, MeOH). ^{1}H NMR (500 MHz, CD_3OD): δ = 8.02 (d, *J* = 8.5 Hz, 2H, Ar-H), 7.88 (d, *J* = 8.5 Hz, 2H, Ar-H), 7.71 (d, *J* = 8.8 Hz, 2H, Ar-H), 7.26 (d, *J* = 8.6 Hz, 2H, Ar-H), 5.57 (d, *J* = 1.5 Hz, 1H, H-1), 4.39 (dd, *J* = 1.9, 11.8 Hz, 1H, H-6a), 4.24 (dd, *J* = 6.4, 11.8 Hz, 1H, H-6b), 4.08 (dd, *J* = 1.8, 3.4 Hz, 1H, H-2), 3.94 (dd, *J* = 3.4, 9.0 Hz, 1H, H-3), 3.82–3.73 (m, 2H, H-4, H-5), 3.18 (s, 3H, SO_2CH_3), 1.93 ppm (s, 3H, CH_3). ^{13}C NMR (125 MHz, CD_3OD): δ = 172.68 (CO), 158.24, 147.29, 140.20, 134.42, 129.59, 129.03, 128.47, 118.37 (12C, Ar-C), 99.90 (C-1), 73.02 (C-5), 72.39 (C-3), 71.77 (C-2), 68.52 (C-4), 64.81 (C-6), 44.48 (SO_2CH_3), 20.69 ppm (CH_3). HRMS: *m*/*z*: calcd. for $C_{21}H_{24}NaO_9S$ [M+Na]$^+$: 475.1039, found: 475.1037.

4′-(Methylsulfonyl)-biphenyl-4-yl 6-O-propionyl-α-D-mannopyranoside (5b)

Prepared according to the general procedure B from **13b** (28 mg, 0.038 mmol) with 20 mg of $Pd(OH)_2$/C (E101 NE/W, 20% Pd) in EtOH (5 mL). Purified by MPLC on silica gel (DCM–MeOH, 1:0 to 9:1). Yield: 11 mg (59%) as a white oil. $[\alpha]_D^{20}$ + 93.8 (*c* 0.50, MeOH). ^{1}H NMR (500 MHz, CD_3OD): δ = 8.02 (d, *J* = 8.4 Hz, 2H, Ar-H), 7.87 (d, *J* = 8.4 Hz, 2H, Ar-H), 7.70 (d, *J* = 8.7 Hz, 2H, Ar-H), 7.25 (d, *J* = 8.7 Hz, 2H, Ar-H), 5.58 (d, *J* = 1.0 Hz, 1H, H-1), 4.42 (dd, *J* = 1.6, 11.7 Hz, 1H, H-6a), 4.23 (dd, *J* = 6.7, 11.7 Hz, 1H, H-6b), 4.09 (d, *J* = 1.6 Hz, 1H, H-2), 3.95 (dd, *J* = 3.4, 9.1 Hz, 1H, H-3), 3.82–3.73 (m, 2H, H-4, H-5), 3.18 (s, 3H, SO_2CH_3), 2.25 (dq, *J* = 2.8, 7.6 Hz, 2H, CH_2COO), 1.02 ppm (t, *J* = 7.6 Hz, 3H, CH_3). ^{13}C NMR (125 MHz, CD_3OD): δ = 175.92 (CO), 158.18, 147.26, 140.19, 134.37, 129.57, 129.03, 128.45, 118.34 (12C, Ar-C), 99.75 (C-1), 73.11 (C-5), 72.39 (C-3), 71.74 (C-2), 68.57 (C-4), 64.75 (C-6), 44.48 (SO_2CH_3), 28.20 (CH_2COO), 9.31 ppm (CH_3). HRMS: *m*/*z*: calcd. for $C_{22}H_{26}NaO_9S$ [M+Na]$^+$: 489.1195, found: 489.1192.

4′-(Methylsulfonyl)-biphenyl-4-yl 6-O-isobutyryl-α-D-mannopyranoside (5c)

Prepared according to the general procedure B from **13c** (40 mg, 0.053 mmol) with 25 mg of $Pd(OH)_2$/C (E101 NE/W, 20% Pd) in EtOH (4 mL). Purified by MPLC on silica gel (DCM–MeOH, 1:0 to 9:1). Yield: 18 mg (70%) as a white oil. $[\alpha]_D^{20}$ + 109.8 (*c* 0.85, MeOH). ^{1}H NMR (500 MHz, CD_3OD): δ = 8.02 (d, *J* = 8.4 Hz, 2H, Ar-H), 7.86 (d, *J* = 8.5 Hz, 2H, Ar-H), 7.68 (d, *J* = 8.7 Hz, 2H, Ar-H), 7.25 (d, *J* = 8.7 Hz, 2H, Ar-H), 5.60 (d, *J* = 1.1 Hz, 1H, H-1), 4.45 (dd, *J* = 1.6, 11.7 Hz, 1H, H-6a), 4.20 (dd, *J* = 7.0, 11.7 Hz, 1H, H-6b), 4.09 (dd, *J* = 1.7, 3.2 Hz, 1H, H-2), 3.95 (dd, *J* = 3.4, 9.1 Hz, 1H, H-3), 3.82–3.72 (m, 2H, H-4, H-5), 3.18 (s, 3H, SO_2CH_3), 2.45 (hept, *J* = 7.0 Hz, 1H, C*H*$(CH_3)_2$), 1.06 (d, *J* = 7.0 Hz, 3H, CH_3), 1.02 ppm (d, *J* = 7.0 Hz, 3H, CH_3). ^{13}C NMR (125 MHz, CD_3OD): δ = 178.50 (CO), 158.14, 147.26, 140.17, 134.34, 129.60, 129.02, 128.43, 118.29 (12C, Ar-C), 99.61 (C-1), 73.27 (C-5), 72.37 (C-3), 71.71 (C-2), 68.57 (C-4), 64.84 (C-6), 44.48 (SO_2CH_3), 35.12 (CH), 19.23, 19.14 ppm (2 CH_3). HRMS: *m*/*z*: calcd. for $C_{23}H_{28}NaO_9S$ [M+Na]$^+$: 503.1352, found: 503.1353.

4′-(Methylsulfonyl)-biphenyl-4-yl 6-O-butyryl-α-D-mannopyranoside (5d)

Prepared according to the general procedure B from **13d** (45 mg, 0.060 mmol) with 40 mg of $Pd(OH)_2$/C (E101 NE/W, 20% Pd) in EtOH (5 mL). Purified by MPLC on silica gel (DCM–MeOH, 9:1). Yield: 21 mg (72%) as a colorless oil. $[\alpha]_D^{20}$ + 97.3 (*c* 1.05, MeOH). ^{1}H NMR (500 MHz, CD_3OD): δ = 8.01 (d, *J* = 8.5 Hz, 2H, Ar-H), 7.87 (d, *J* = 8.5 Hz, 2H, Ar-H), 7.69 (d, *J* = 8.8 Hz, 2H, Ar-H), 7.24 (d, *J* = 8.8 Hz, 2H, Ar-H), 5.58 (d, *J* = 1.4 Hz, 1H, H-1), 4.43 (dd, *J* = 1.8, 11.7 Hz, 1H, H-6a), 4.20 (dd, *J* = 6.9, 11.7 Hz, 1H, H-6b), 4.08 (dd, *J* = 1.7, 3.3 Hz, 1H, H-2), 3.93 (dd, *J* = 3.4, 9.1 Hz, 1H, H-3), 3.79 (m, 1H, H-5), 3.73 (m, 1H, H-4), 3.17 (s, 3H, SO_2CH_3), 2.21–2.17 (m, 2H, CH_2COO), 1.55–1.47 (m, 2H, CH_2), 0.83 ppm (t, *J* = 7.4 Hz, 3H, CH_3). ^{13}C NMR (125 MHz, CD_3OD): δ = 175.09 (CO), 158.18, 147.24, 140.19, 134.31, 129.56, 129.03, 128.43, 118.31 (12C, Ar-C), 99.70 (C-1), 73.17 (C-5), 72.39 (C-3), 71.74 (C-2), 68.59 (C-4), 64.72 (C-6), 44.47 (SO_2CH_3), 36.88 (CH_2COO), 19.29 (CH_2), 13.92 ppm (CH_3). HRMS: *m*/*z*: calcd. for $C_{23}H_{28}NaO_9S$ [M+Na]$^+$: 503.1352, found: 503.1350.

4′-(Methylsulfonyl)-biphenyl-4-yl 6-O-pivaloyl-α-D-mannopyranoside (5e)

Prepared according to the general procedure B from **13e** (50 mg, 0.065 mmol) with 15 mg of $Pd(OH)_2$/C (E101 NE/W, 20% Pd) in EtOH (5 mL). Purified by MPLC on silica gel (DCM–MeOH, 9:1). Yield: 24 mg (73%) as a white oil. $[\alpha]_D^{20}$ + 96.0 (*c* 1.18, MeOH). ^{1}H NMR (500 MHz, CD_3OD): δ = 8.02 (d, *J* = 8.5 Hz, 2H, Ar-H), 7.85 (d, *J* = 8.6 Hz, 2H, Ar-H), 7.68 (d, *J* = 8.7 Hz, 2H, Ar-H), 7.26 (d, *J* = 8.9 Hz, 2H, Ar-H), 5.62 (d, *J* = 1.5 Hz, 1H, H-1), 4.45 (dd, *J* = 1.7, 11.7 Hz, 1H. H-6a), 4.15 (dd, *J* = 7.4, 11.7 Hz, 1H, H-6b), 4.09 (dd, *J* = 1.8, 3.4 Hz, 1H, H-2), 3.96 (dd, *J* = 3.4, 9.2 Hz, 1H, H-3), 3.81 (ddd, *J* = 1.6, 7.4, 9.2 Hz, 1H, H-5), 3.73 (t, *J* = 9.6 Hz, 1H, H-4), 3.18 (s, 3H, SO_2CH_3), 1.07 ppm (s, 9H, $C(CH_3)_3$). ^{13}C NMR (125 MHz, CD_3OD): δ = 179.92 (CO), 158.09, 147.28, 140.16, 134.31, 129.65, 129.03, 128.43, 118.26 (12C, Ar-C), 99.43 (C-1), 73.38 (C-5), 72.35 (C-3), 71.68 (C-2), 68.59 (C-4), 65.14 (C-6), 44.48 (SO_2CH_3), 39.72 ($C(CH_3)_3$), 27.43 ppm ($C(CH_3)_3$). HRMS: *m*/*z*: calcd. for $C_{24}H_{30}NaO_9S$ [M+Na]$^+$: 517.1508, found: 517.1507.

4′-(Methylsulfonyl)-biphenyl-4-yl 6-O-isovaleryl-α-D-mannopyranoside (5f)

Prepared according to the general procedure B from **13f** (34 mg, 0.044 mmol) with 40 mg of $Pd(OH)_2$/C (E101 NE/W, 20% Pd) in EtOH (7 mL). Purified by MPLC on silica gel (DCM–MeOH, 9:1). Yield: 12 mg (55%) as a colorless oil. $[\alpha]_D^{20}$ + 120.3 (*c* 0.60, MeOH). ^{1}H NMR (500 MHz, CD_3OD): δ = 8.02 (d, *J* = 8.5 Hz, 2H, Ar-H), 7.88 (d, *J* = 8.5 Hz, 2H, Ar-H), 7.70 (d, *J* = 8.8 Hz, 2H, Ar-H), 7.25 (d, *J* = 8.8 Hz, 2H, Ar-H), 5.58 (d, *J* = 1.3 Hz, 1H, H-1), 4.45 (dd, *J* = 1.6, 11.7 Hz, 1H, H-6a), 4.18 (dd, *J* = 7.0, 11.7 Hz, 1H, H-6b), 4.07 (dd, *J* = 1.7, 3.3 Hz, 1H, H-2), 3.94 (dd, *J* = 3.4, 9.2 Hz, 1H, H-3), 3.79 (m, 1H, H-5), 3.72 (m, 1H, H-4), 3.17 (s, 3H, SO_2CH_3), 2.09 (dd, *J* = 3.0, 7.2 Hz, 2H, CH_2COO), 1.95 (m, 1H, C*H*$(CH_3)_2$), 0.83 ppm (app-t, *J* = 6.3 Hz, 6H, 2 CH_3). ^{13}C NMR (125 MHz, CD_3OD): δ = 174.55 (CO), 158.20, 147.24, 140.19, 134.28, 129.57, 129.03, 128.42, 118.30 (12C, Ar-C), 99.69 (C-1), 73.25 (C-5), 72.39 (C-3), 71.74 (C-2), 68.60 (C-4), 64.72 (C-6), 44.47 (SO_2CH_3), 44.16 (CH_2COO), 26.67 (*C*H$(CH_3)_2$), 22.68, 22.66 ppm (2 CH_3). HRMS: *m*/*z*: calcd. for $C_{24}H_{30}NaO_9S$ [M+Na]$^+$: 517.1508, found: 517.1499.

4′-(Methylsulfonyl)-biphenyl-4-yl 6-O-valeryl-α-D-mannopyranoside (5g)

Prepared according to the general procedure B from **13g** (51 mg, 0.067 mmol) with 40 mg of $Pd(OH)_2/C$ (E101 NE/W, 20% Pd) in EtOH (15 mL). Purified by MPLC on silica gel (DCM–MeOH, 95:5). Yield: 25 mg (76%) as a white solid. $[\alpha]_D^{20}$ + 111.4 (*c* 1.00, MeOH). ^{1}H NMR (500 MHz, CD_3OD): δ = 7.99 (d, *J* = 8.5 Hz, 2H, Ar-H), 7.85 (d, *J* = 8.5 Hz, 2H, Ar-H), 7.68 (d, *J* = 8.8 Hz, 2H, Ar-H), 7.23 (d, *J* = 8.8 Hz, 2H, Ar-H), 5.56 (d, *J* = 1.4 Hz, 1H, H-1), 4.41 (dd, *J* = 1.8, 11.7 Hz, 1H, H-6a), 4.18 (dd, *J* = 7.0, 11.7 Hz, 1H, H-6b), 4.05 (dd, *J* = 1.7, 3.3 Hz, 1H, H-2), 3.91 (dd, *J* = 3.4, 9.1 Hz, 1H, H-3), 3.77 (m, 1H, H-5), 3.70 (t, *J* = 9.6 Hz, 1H, H-4), 3.15 (s, 3H, SO_2CH_3), 2.21–2.17 (m, 2H, CH_2COO), 1.48–1.42 (m, 2H, CH_2), 1.24–1.17 (m, 2H, CH_2), 0.79 ppm (t, *J* = 7.4 Hz, 3H, CH_3). ^{13}C NMR (125 MHz, CD_3OD): δ = 175.26 (CO), 163.63, 158.20, 147.22, 140.20, 134.28, 129.56, 129.02, 128.41, 118.31 (12C, Ar-C), 99.73 (C-1), 73.18 (C-5), 72.40 (C-3), 71.74 (C-2), 68.60 (C-4), 64.75 (C-6), 44.48 (SO_2CH_3), 34.71 (CH_2COO), 27.96, 23.20 (2 CH_2), 13.99 ppm (CH_3). HRMS: *m/z*: calcd. for $C_{24}H_{30}NaO_9S$ $[M+Na]^+$: 517.1508, found: 517.1499.

4′-(Methylsulfonyl)-biphenyl-4-yl 6-O-pivaloyloxymethyl-α-D-mannopyranoside (5h)

Prepared according to the general procedure B from **15** (36 mg, 0.045 mmol) with 30 mg of $Pd(OH)_2/C$ (E101 NE/W, 20% Pd) in EtOH (5 mL). Purified by MPLC on silica gel (DCM–MeOH, 95:5). Yield: 10 mg (24%) as a white solid. $[\alpha]_D^{20}$ + 111.4 (*c* 1.00, MeOH). ^{1}H. NMR (500 MHz, CD_3OD): δ = 7.99 (d, *J* = 8.5 Hz, 2H, Ar-H), 7.86 (d, *J* = 8.5 Hz, 2H, Ar-H), 7.67 (d, *J* = 8.8 Hz, 2H, Ar-H), 7.23 (d, *J* = 8.8 Hz, 2H, Ar-H), 5.52 (d, *J* = 1.6 Hz, 1H, H-1), 5.26 (s, 2H, OCH_2O), 4.03 (dd, *J* = 1.8, 3.3 Hz, 1H, H-2), 3.91–3.90 (m, 2H, H-3, H-6a), 3.85 (m, 1H, H-6b), 3.75–3.71 (m, 2H, H-4, H-5), 3.15 (s, 3H, SO_2CH_3), 1.19 ppm (s, 9H, $C(CH_3)_3$). ^{13}C. NMR (125 MHz, CD_3OD): δ = 179.44 (CO), 158.49, 147.36, 140.16, 134.45, 129.65, 129.01, 128.50, 118.44 (12C, Ar-C), 100.20 (C-1), 90.45 (OCH_2O), 74.36 (C-5), 72.42 (C-2), 71.83 (C-3), 70.53 (C-6), 68.32 (C-4), 44.48 (SO_2CH_3), 39.92 ($C(CH_3)_3$), 27.39 ppm ($C(CH_3)_3$). HRMS: *m/z*: calcd. for $C_{25}H_{32}NaO_{10}S$ $[M + Na]^+$: 547.1614, found: 547.1607.

Pharmacokinetic assays

Materials

DMSO, 1-propanol, 1-octanol, Dulbecco's modified Eagle's medium (DMEM) high glucose, penicillin–streptomycin (solution stabilized, with 10 000 units of penicillin/mL and 10 mg of streptomycin/mL), L-glutamine solution (200 mM), magnesium chloride, ammonium acetate, BNPP, and loperamide hydrochloride were purchased from Sigma-Aldrich (St. Louis, Missouri). PRISMA HT universal buffer, GIT-0 Lipid Solution, and Acceptor Sink Buffer were ordered from pIon (Woburn, Massachusetts). MEM nonessential amino acids solution 10 mM (100×), fetal bovine serum (FBS), and DMEM without sodium pyruvate and phenol red were bought from Invitrogen (Carlsbad, California). Acetonitrile (MeCN) and MeOH were ordered from Acros Organics (Geel, Belgium). Pooled male RLM (Sprague–Dawley) and pooled HLM were ordered from BD Bioscience (Franklin Lakes, New Jersey). The Caco-2 cells were kindly provided by Prof G. Imanidis, FHNW, Muttenz, Switzerland, and originated from the American Type Culture Collection (Rockville, Maryland).

Log P determination

The in silico prediction tool ALOGPS[32] was used to estimate the octanol–water partition coefficients (log *P*) of the compounds. Depending on these values, the compounds were classified into three categories: hydrophilic compounds (log *P* below zero), moderately lipophilic compounds (log *P* between zero and one) and lipophilic compounds (log *P* above 1). For each category, two different ratios (volume of 1-octanol to volume of buffer) were defined as experimental parameters (Table 2).

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

Table 2. Compound classification based on estimated log *P* values.

Compound type	log *P*	Ratios (1-octanol:buffer)
Hydrophilic	<0	30:140,40:130
Moderately lipophilic	0–1	70:110, 110:70
Lipophilic	>1	3:180, 4:180

The log *P* coefficients were calculated from the 1-octanol:buffer ratio (o:b), the initial concentration of the analyte in buffer (1 μM), and the concentration of the analyte in buffer (c_B) with eq. 1:

$$\log P = \log\left(\frac{1\ \mu M - c_B}{c_B} \times \frac{1}{\text{o:b}}\right) \quad (1)$$

The average of the three log *P* values per 1-octanol:buffer ratio was calculated. If the two means obtained for a compound did not differ by more than 0.1 unit, the results were accepted.

PAMPA

Effective permeability (log P_e)was determined in a 96-well format with PAMPA.[18] 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, 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 (190–500 nm) (SpectraMax 190, Molecular Devices, Silicon Valley, California). Then, analyte dissolved in DMSO (10 mM) was added to the remaining buffer to yield 50 μM solutions. To exclude precipitation, the optical density (OD) was measured at 650 nm, and solutions exceeding OD 0.01 were filtrated. Afterwards, samples (150 μL) were withdrawn to determine the reference spectra. A further 200 μL was transferred to each well of the donor plate of the PAMPA sandwich (pIon, P/N 110 163). 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™, and left undisturbed for 16 h. Then, 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 (log P_e) was calculated from the compound flux deduced from the spectra, the filter area, and the initial sample concentration in the donor well with the aid of the PAMPA Explorer Software (pIon, version 3.5).

Caco-2 cell permeation assay

Caco-2 cells were cultivated in tissue culture flasks (BD Biosciences, Franklin Lakes, New Jersey) with DMEM high glucose medium containing L-glutamine (2 mM), nonessential amino acids (0.1 mM), penicillin (100 U/mL), streptomycin (100 μg/mL), and FBS (10%). The cells were kept at 37 °C in humidified air containing 5% CO_2, and the medium was changed every second day. When ap-

proximately 90% confluence was reached, the cells were split in a 1:10 ratio and distributed to new tissue culture flasks. At passage numbers between 60 and 65, they were seeded at a density of 5.3×10^5 cells/well to Transwell 6-well plates (Corning Inc., Corning, New York) with 2.5 mL of culture medium in the basolateral and 2 mL in the apical compartment. The medium was renewed on alternate days. Permeation experiments were performed between days 19 and 21 postseeding. Previously to the experiment, the integrity of the Caco-2 monolayers was evaluated by measuring the transepithelial electrical resistance (TEER) with an Endohm tissue resistance instrument (World Precision Instruments Inc., Sarasota, Florida). Only wells with TEER values higher than 250 Ω·cm² were used. To inhibit CES activity, the Caco-2 cell monolayers were pre-incubated with BNPP (200 μM) dissolved in transport medium (DMEM without sodium pyruvate and phenol red) for 40 min.[22b] Experiments were performed in the apical-to-basolateral (absorptive) and basolateral-to-apical (secretory) directions in triplicate. Transport medium was withdrawn from the donor compartments and replaced by the same volume of compound stock solution (10 mM in DMSO) to reach an initial sample concentration of 62.5 μM. The Transwell plate was shaken (600 rpm, 37 °C) on a Heidolph Titramax 1000 plate-shaker. Samples (40 μL) were withdrawn from the donor and acceptor compartments 30 min after initiation of the experiment and the concentrations were determined by LC-MS (see below). Apparent permeability (P_{app}) was calculated according to eq. 2:

$$(2) \qquad P_{app} = \frac{dQ}{dt} \times \frac{1}{A \times c_0}$$

where dQ/dt is the compound flux (mol s^{-1}), A is the surface area of the monolayer (cm^2), and c_0 is the initial concentration in the donor compartment (mol cm^{-3}).[33] After the experiment, TEER values were measured again and results from wells with values below 250 Ω·cm² were discarded.

Aqueous solubility

Solubility was determined in a 96-well format using the μSOL Explorer solubility analyzer (pIon, version 3.4.0.5). For each compound, measurements were performed in triplicate. Three wells of a deep well plate were filled with 300 μL of PRISMA HT universal buffer, adjusted to pH 7.4 by adding the requested amount of NaOH (0.5 M). Aliquots (3 μL) of a compound stock solution (40–100 mM in DMSO) were added and thoroughly mixed. The final sample concentration was 0.4–1.0 mM and the residual DMSO concentration was 1.0% (*v*/*v*). Fifteen hours after initiation of the experiment, the solutions were filtrated (0.2 μm 96-well filter plates) using a vacuum to collect manifold (Whatman Ltd., Maidstone, UK) to remove any precipitates. Equal amounts of filtrate and 1-propanol were mixed and transferred to a 96-well plate for UV detection (190–500 nm). The amount of material dissolved was calculated by comparison with UV spectra obtained from reference samples, which were prepared by dissolving compound stock solution in a 1:1 mixture of buffer and 1-propanol (final concentrations 0.067–0.167 mM).

Enzymatic hydrolysis by liver microsome associated CES

Incubations were performed in triplicate in a 96-well format on an Eppendorf Thermomixer Comfort. The reaction mixture (270 μL) consisting of liver microsomes (0.139 μg/mL), TRIS–HCl buffer (0.1 M, pH 7.4), and $MgCl_2$ (2 mM) was preheated (37 °C, 500 rpm, 10 min), and the incubation was initiated by adding 30 μL of compound solution (20 μM) in TRIS–HCl buffer. The final concentration of the compound was 2 μM, and the microsomal concentration was 0.125 mg/mL. At the beginning of the experiment (t = 0 min) and after an incubation time of 2, 5, 10, 20, and 30 min, samples (40 μL) were transferred to 120 μL of ice-cooled MeOH and centrifuged (3700 rpm, 4 °C, 10 min). Then, 80 μL of supernatant was transferred to a 96-well plate for analysis by LC-MS (see below). The metabolic half-life ($t_{1/2}$) was calculated from the slope of the linear regression from the log percentage remaining compound versus incubation time relationship. Control experiments were performed in parallel by preincubating the microsomes with the specific CES inhibitor BNPP (1 mM) for 5 min before addition of the compound solution.[20]

Isozyme specific inhibition of CES-mediated hydrolysis

Test compounds were dissolved in DMSO to 1 mM and then diluted with TRIS–HCl buffer (0.1 M, pH 7.4) containing $MgCl_2$ (2 mM) to a concentration of 6 μM. Loperamide hydrochloride was dissolved in DMSO to 20, 2, and 0.2 mM and then diluted with TRIS–HCl buffer containing $MgCl_2$ to a concentration of 750, 75, and 7.5 μM. HLM were suspended in TRIS–HCl buffer containing $MgCl_2$ to a concentration of 30 μg/mL. Compound solution (100 μL) and microsomal suspension (200 μL) mixed with loperamide solution or blank buffer (50 μL) were preheated (37 °C, 500 rpm, 15 min) in separate wells of a 96-well plate. The incubation was initiated by transferring 200 μL of microsome suspension containing loperamide to the compound solution. The final compound concentration was 2 μM, the microsomal concentration was 0.02 mg/mL, and the loperamide concentration was 100, 10, 1, and 0 μM (blank). At the beginning of the experiment (t = 0 min) and after an incubation time of 10, 20, 30, 45, and 60 min, samples (20 μL) were transferred to 60 μL of ice-cooled MeOH and analyzed by LC-MS (see below). The metabolic turnover was assessed as accumulation of product **4a** versus incubation time.[30]

LC-MS measurements

Analyses were performed using a 1100/1200 Series HPLC system coupled to a 6410 Triple Quadrupole mass detector (Agilent Technologies, Inc., Santa Clara, California) equipped with electrospray ionization. The system was controlled with the Agilent MassHunter Workstation Data Acquisition software (version B.01.04). The column used was an Atlantis® T3 C18 column (2.1 × 50 mm) with a 3 μm particle size (Waters Corp., Milford, Massachusetts). The mobile phase consisted of eluent A (10 mM ammonium acetate, pH 5.0 in 95:5 H_2O–MeCN) and eluent B (MeCN containing 0.1% formic acid). The flow rate was maintained at 0.6 mL/min. The gradient was ramped from 95% A – 5% B to 5% A – 95% B over 1 min and then held at 5% A – 95% B for 0.1 min. The system was then brought back to 95% A – 5% B, resulting in a total duration of 4 min. MS parameters such as fragmentor voltage, collision energy, and polarity were optimized individually for each drug, and the molecular ion was followed for each compound in the multiple reaction monitoring mode. The concentrations of the analytes were quantified by the Agilent Mass Hunter Quantitative Analysis software (version B.01.04).

Supplementary material

Supplementary material is available with the article through the journal Web site at http://nrcresearchpress.com/doi/suppl/10.1139/cjc-2015-0582. NMR spectra and HPLC traces to document purity of the test compounds.

Acknowledgement

Financial support from the Swiss National Science Foundation (SNF 200020-129935) is gratefully acknowledged.

References

(1) (a) Fihn, S. D. *N. Engl. J. Med.* **2003**, *349*, 259. doi:10.1056/NEJMcp030027; (b) Hooton, T. M.; Besser, R.; Foxman, B.; Fritsche, T. R.; Nicolle, L. E. *Clin. Infect. Dis.* **2004**, *39*, 75. doi:10.1086/422145.
(2) Sanchez, G. V.; Master, R. N.; Karlowsky, J. A.; Bordon, J. M. *Antimicrob. Agents Chemother.* **2012**, *56*, 2181. doi:10.1128/AAC.06060-11.
(3) Schilling, J. D.; Mulvey, M. A.; Hultgren, S. J. *J. Infect. Dis.* **2001**, *183* (Suppl. 1), S36. doi:10.1086/318855.

(4) Capitani, G.; Eidam, O.; Glockshuber, R.; Grütter, M. G. *Microbes Infect.* **2006**, *8*, 2284. doi:10.1016/j.micinf.2006.03.013.
(5) (a) Sharon, N. *Biochim. Biophys. Acta* **2006**, *1760*, 527. doi:10.1016/j.bbagen.2005.12.008; (b) Le Trong, I.; Aprikian, P.; Kidd, B. A.; Forero-Shelton, M.; Tchesnokova, V.; Rajagopal, P.; Rodriguez, V.; Interlandi, G.; Klevit, R.; Vogel, V.; Stenkamp, R. E.; Sokurenko, E. V.; Thomas, W. E. *Cell* **2010**, *141*, 645. doi:10.1016/j.cell.2010.03.038.
(6) (a) Firon, N.; Ofek, I.; Sharon, N. *Biochem. Biophys. Res. Commun.* **1982**, *105*, 1426. doi:10.1016/0006-291X(82)90947-0; (b) Firon, N.; Ofek, I.; Sharon, N. *Carbohydr. Res.* **1983**, *120*, 235. doi:10.1016/0008-6215(83)88019-7; (c) Firon, N.; Ashkenazi, S.; Mirelman, D.; Ofek, I.; Sharon, N. *Infect. Immun.* **1987**, *55*, 472.
(7) (a) Bouckaert, J.; Berglund, J.; Schembri, M.; Genst, E. D.; Cools, L.; Wuhrer, M.; Hung, C. S.; Pinkner, J.; Slättergard, R.; Zavialov, A.; Choudhury, D.; Langermann, S.; Hultgren, S. J.; Wyns, L.; Klemm, P.; Oscarson, S.; Knight, S. D.; Greve, H. D. *Mol. Microbiol.* **2005**, *55*, 441. doi:10.1111/j.1365-2958.2004.04415.x; (b) Sperling, O.; Fuchs, A.; Lindhorst, T. K. *Org. Biomol. Chem.* **2006**, *4*, 3913. doi:10.1039/b610745a; (c) Han, Z.; Pinkner, J. S.; Ford, B.; Obermann, R.; Nolan, W.; Wildman, S. A.; Hobbs, D.; Ellenberger, T.; Cusumano, C. K.; Hultgren, S. J.; Janetka, J. W. *J. Med. Chem.* **2010**, *53*, 4779. doi:10.1021/jm100438s; (d) Klein, T.; Abgottspon, D.; Wittwer, M.; Rabbani, S.; Herold, J.; Jiang, X.; Kleeb, S.; Lüthi, C.; Scharenberg, M.; Bezençon, J.; Gubler, E.; Pang, L.; Smiesko, M.; Cutting, B.; Schwardt, O.; Ernst, B. *J. Med. Chem.* **2010**, *53*, 8627. doi:10.1021/jm101011y; (e) Schwardt, O.; Rabbani, S.; Hartmann, M.; Abgottspon, D.; Wittwer, M.; Kleeb, S.; Zalewski, A.; Smiesko, M.; Cutting, B.; Ernst, B. *Bioorg. Med. Chem.* **2011**, *19*, 6454. doi:10.1016/j.bmc.2011.08.057; (f) Cusumano, C. K.; Pinkner, J. S.; Han, Z.; Greene, S. E.; Ford, B. A.; Crowley, J. R.; Henderson, J. P.; Janetka, J. W.; Hultgren, S. J. *Sci. Transl. Med.* **2011**, *3*, 109ra115. doi:10.1126/scitranslmed.3003021; (g) Han, Z.; Pinkner, J. S.; Ford, B.; Chorell, E.; Crowley, J. M.; Cusumano, C. K.; Campbell, S.; Henderson, J. P.; Hultgren, S. J.; Janetka, J. W. *J. Med. Chem.* **2012**, *55*, 3945. doi:10.1021/jm300165m; (h) Jiang, X.; Abgottspon, D.; Kleeb, S.; Rabbani, S.; Scharenberg, M.; Wittwer, M.; Haug, M.; Schwardt, O.; Ernst, B. *J. Med. Chem.* **2012**, *55*, 4700. doi:10.1021/jm300192x; (i) Pang, L.; Kleeb, S.; Lemme, K.; Rabbani, S.; Scharenberg, M.; Zalewski, A.; Schädler, F.; Schwardt, O.; Ernst, B. *ChemMedChem* **2012**, *7*, 1404. doi:10.1002/cmdc.201200125.
(8) (a) Smith, D. A.; Jones, B. C.; Walker, D. K. *Med. Res. Rev.* **1996**, *16*, 243. doi:10.1002/(SICI)1098-1128(199605)16:3<243::AID-MED2>3.3.CO;2-R; (b) van de Waterbeemd, H.; Smith, D. A.; Beaumont, K.; Walker, D. K. *J. Med. Chem.* **2001**, *44*, 1313. doi:10.1021/jm000407e.
(9) Feng, B.; LaPerle, J. L.; Chang, G.; Varma, M. V. *Expert Opin. Drug Metab. Toxicol.* **2010**, *6*, 939. doi:10.1517/17425255.2010.482930.
(10) 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. doi:10.1021/jm501524q.
(11) (a) Saxon, E.; Bertozzi, C. R. *Science* **2000**, *287*, 2007. doi:10.1126/science.287.5460.2007; (b) Sarkar, A. K.; Fritz, T. A.; Taylor, W. H.; Esko, J. D. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 3323. doi:10.1073/pnas.92.8.3323.
(12) Fokt, I.; Skora, S.; Conrad, C.; Madden, T.; Emmett, M.; Priebe, W. *Carbohydr. Res.* **2013**, *368*, 111. doi:10.1016/j.carres.2012.11.021.
(13) Prieto, M.; Zurita, E.; Rosa, E.; Muñoz, L.; Lloyd-Williams, P.; Giralt, E. *J. Org. Chem.* **2004**, *69*, 6812. doi:10.1021/jo0491612.
(14) Pojer, P. M.; Angyal, S. J. *Aust. J. Chem.* **1978**, *31*, 1031. doi:10.1071/CH9781031.
(15) Ali, A.; van den Berg, R. J. B. H. N.; Overkleeft, H. S.; van der Marel, G. A.; Codee, J. D. C. *Tetrahedron* **2010**, *32*, 6121. doi:10.1016/j.tet.2010.06.007.
(16) Avdeef, A. In *Pharmacokinetic Optimization in Drug Research: Biological, Physicochemical and Computational Strategies*; Testa, B., van de Waterbeemd, H., Folkers, G., Guy, R., Eds.; Helvetica Chimica Acta: Zurich, 2001; pp. 305–326.
(17) Dearden, J. C.; Bresnen, G. M. *QSAR Comb. Sci.* **1988**, *7*, 133.
(18) Kansy, M.; Senner, F.; Gubernator, K. *J. Med. Chem.* **1998**, *41*, 1007. doi:10.1021/jm970530e.
(19) Artursson, P.; Karlsson, J. *Biochem. Biophys. Res. Commun.* **1991**, *175*, 880. doi:10.1016/0006-291X(91)91647-U.
(20) (a) Imai, T.; Taketani, M.; Shii, M.; Hosokawa, M.; Chiba, K. *Drug Metab. Dispos.* **2006**, *34*, 1734. doi:10.1124/dmd.106.009381; (b) Taketani, M.; Shii, M.; Ohura, K.; Ninomiya, S.; Imai, T. *Life Sci.* **2007**, *81*, 924. doi:10.1016/j.lfs.2007.07.026.
(21) Avdeef, A.; Bendels, S.; Di, L.; Faller, B.; Kansy, M.; Sugano, K.; Yamauchi, Y. *J. Pharm. Sci.* **2007**, *96*, 2893. doi:10.1002/jps.21068.
(22) (a) Imai, T.; Imoto, M.; Sakamoto, H.; Hashimoto, M. *Drug Metab. Dispos.* **2005**, *33*, 1185. doi:10.1124/dmd.105.004226.(b) Ohura, K.; Sakamoto, H.; Ninomiya, S.; Imai, T. *Drug Metab. Dispos.* **2010**, *38*, 323. doi:10.1124/dmd.109.029413.
(23) Seelig, A.; Gerebtzoff, G. *Expert Opin. Drug. Metab. Toxicol.* **2006**, *2*, 733. doi:10.1517/17425255.2.5.733.
(24) Lipinski, C. A. *J. Pharmacol. Toxicol. Methods* **2000**, *44*, 235. doi:10.1016/S1056-8719(00)00107-6.
(25) (a) Beaumont, K.; Webster, R.; Gardner, I.; Dack, K. *Curr. Drug Metab.* **2003**, *4*, 461. doi:10.2174/1389200033489253; (b) Ettmayer, P.; Amidon, G. L.; Clement, B.; Testa, B. *J. Med. Chem.* **2004**, *47*, 2393. doi:10.1021/jm0303812.
(26) Nielsen, A. B.; Buur, A.; Larsen, C. *Eur. J. Pharm. Sci.* **2005**, *24*, 433. doi:10.1016/j.ejps.2004.12.007.
(27) Evans, D. F.; Pye, G.; Bramley, R.; Clark, A. G.; Dyson, T. J.; Hardcastel, J. D. *Gut* **1988**, *29*, 1035. doi:10.1136/gut.29.8.1035.
(28) (a) Li, B.; Sedlacek, M.; Manoharan, I.; Boopathy, R.; Duysen, E. G.; Masson, P.; Lockridge, O. *Biochem. Pharmacol.* **2005**, *70*, 1673. doi:10.1016/j.bcp.2005.09.002; (b) Liederer, B. M.; Borchardt, R. T. *J. Pharm. Sci.* **2006**, *95*, 1177. doi:10.1002/jps.20542.
(29) Satoh, T.; Hosokawa, M. *Chem. Biol. Interact.* **2006**, *162*, 195. doi:10.1016/j.cbi.2006.07.001.
(30) (a) Quinney, S. K.; Sanghani, S. P.; Davis, W. I.; Hurley, T. D.; Sun, Z.; Murry, D. J.; Bosron, W. F. *J. Pharmacol. Exp. Ther.* **2005**, *313*, 1011. doi:10.1124/jpet.104.081265; (b) Wang, J.; Williams, E. T.; Bourgea, J.; Wong, Y. N.; Patten, C. J. *Drug Metab. Dispos.* **2011**, *39*, 1329. doi:10.1124/dmd.111.039628.
(31) Roy, R.; Das, S. K.; Santoyo-González, F.; Hernández-Mateo, F.; Dam, T. K.; Brewer, C. F. *Chem. Eur. J.* **2000**, *6*, 1757. doi:10.1002/(SICI)1521-3765(20000515)6:10<1757::AID-CHEM1757>3.0.CO;2-5.
(32) (a) VCCLAB, Virtual Computational Chemistry Laboratory, 2005; available from http://www.vcclab.org (accessed November 19, 2012); (b) Tetko, I. V.; Gasteiger, J.; Todeschini, R.; Mauri, A.; Livingstone, D.; Ertl, P.; Palyulin, V. A.; Radchenko, E. V.; Zefirov, N. S.; Makarenko, A. S.; Tanchuk, V. Y.; Prokopenko, V. V. *J. Comput. Aided Mol. Des.* **2005**, *19*, 453. doi:10.1007/s10822-005-8694-y.
(33) Hubatsch, I.; Ragnarsson, E. G. E.; Artursson, P. *Nat. Protoc.* **2007**, *2*, 2111. doi:10.1038/nprot.2007.303.

5 Human Neuraminidase 3 inhibitors

5.1 Sialic acids

Sialic acids (or neuramidic acids) are a large family of approximately fifty related acidic monosaccharides sharing a common core structure consisting of a nine-carbon polyhydroxylated α-keto acid (1). They commonly share a carboxylic acid in the 2-position. At physiological pH, the acid is deprotonated, having a strong effect on the physicochemical properties. In the 5-position, sialic acids exhibit the broadest variations. In Figure 5.1, the most common sialic acids are listed, namely *N*-acetylneuraminic acid (**1**, Neu*N*Ac), *N*-glycolylneuraminic acid (**2**, Neu*N*Gc) or deaminoneuraminic acid (**3**, KDN) (2-4). Sialic acids are almost exclusively expressed in higher eukaryotes and in particular groups of eubacteria, and are usually terminal residues of cell surface glycoconjugates (1). There, they mediate a broad range of cell-cell and cell-molecules interactions by stabilizing of cell membranes, acting as chemical messenger in cell-cell regulations and transmembrane receptors, or controlling half-lives of circulating glycoproteins (5,6). The cell-cell regulations play a major role in cell recognition and adhesion processes, such as immune and inflammation responses or nervous system embryogenesis (7,8). The importance of sialic acids is impressively demonstrated by the death of embryotic mutant mice, in which a gene encoding of a key enzyme for sialic acid synthesis was inactivated by gene targeting (9). The synthesis of sialic acids requires the coordinated action of several enzymes catalyzing different steps in biosynthesis and degradation starting from UDP-*N*-acetyl-α-D-glucosamine (10,11). The transfer of sialyl residues is further regulated by neuraminidases (also named sialydases) and sialyltransferases (12). As a result of their broad biological role and presence on cell surfaces, sialic acids are often used as a target for pathogens to modulate and evade the immune response (13-15).

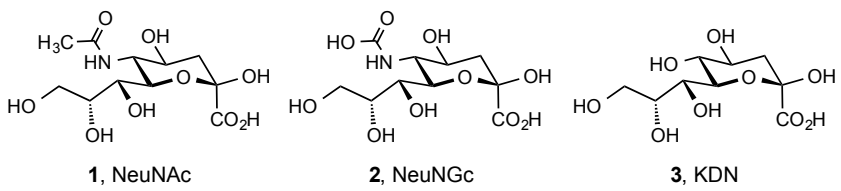


Figure 5.1 Common neuraminic acids.

5.2 Human Neuraminidases

The content of sialic acids in humans bound to glycoproteins and glycolipids and in free circulation is highly regulated by sialyltransferases and neuraminidases (16). Sialyltransferases catalyze the transfer of activated sialic acids (CMP-N-acetylneuraminic acid, CMP-NANA) to different glycoconjugates, almost exclusively to the terminal position (4). The family of human sialyltransferases includes approximately 20 different enzymes, expressed in a different tissue-, cell type-, and stage-specific manner to regulate the sialylation pattern (17). Sialic acids are primary transferred to galactose moieties of glycoconjugates in a α2,3- or α2,6-linked manner (18, 19).

The hydrolytic cleavage of sialic acids from glycoconjugates is facilitated by neuraminidases (NEU). Up to date, four different mammalian NEU (NEU1, NEU2, NEU3 and NEU4) have been identified and characterized. Each neuraminidase differs in subcellular locations, enzymatic properties and substrate specificity to play its unique role in controlling of the level of sialylation (.

Table 5.1) (20).

Table 5.1 Human neuraminidases NEU1-4.

	Localization	**Substrate specificity**	**Major function**
NEU1	Lysosomes	Oligosaccharides Glycopeptides	Degradation in lysosomes Exocytosis Immune function Phagocytosis Elastic fiber assambly
NEU2	Cytosol	Oligosaccharides Glycopeptides Gangliosides	Myoblast differentiation Neural differentiation
NEU3	Plasma membranes	Gangliosides	Neuronal differentiation Apoptosis Adhesion
NEU4	Lysosomes Mitochondria	Oligosaccharides Glycopeptides Gangliosides	Neuronal differentiation Apoptosis Adhesion

5.2.1 Human neuraminidase 3

The human neuraminidase NEU3 is preferably cleaving sialic acids linked in an α2-3, α2-8 or α2-6 manner to galactose being part of gangliosides. The cleavage of the more common sialic acid Neu*N*Ac (**1**) is faster compared to Neu*N*Gc (**2**) (21). NEU3 shares a higher homology with NEU2 and NEU4 (24-30%) than with NEU1 (19-24%) (22). The expression on the surface of plasma membranes implies its important role in regulation of cell surface modulations and expressions, effecting for example neuroblastoma growth, insulin signaling, tumor transformation and invasiveness (23-27).

5.3 Neuraminidase inhibitors

Most synthetic neuraminidase inhibitors are targeted against viral neuraminidases (Figure 5.2A). During the infection cycle, newly formed influenza virus need neuraminidases to be released from human cells by hydrolyzing terminal sialic acids being part of the host cell surface (28). Therefore, viral neuraminidases represent a potent target for the treatment of influenza infections (29). Potent neuraminidase inhibitors are obtained by mimicking the transition state of the hydrolysis by introducing a double bond in the C2/C3 position (→ **4**, DANA) (30). To further improve affinity and specificity for influenza neuraminidases, a guanidinium residue was introduced in the 4-position leading to a subnanomolar inhibitor, marketed as zanamivir (→ **5**)(31). The large hydrophilic surface and charge at physiological pH does not allow an oral administration of zanamivir and therefore it has to be administred by inhalation. By replacing the hydrophilic glycerol side chain by the lipophilic isopentoxy moiety, forming an ester prodrug and further reducing the polar surface area, the orally bioavailable derivative was obtained, marketed as oseltamivir (**6**) (32, 33).

The development of selective inhibitors for the human neuraminidases NEU1-4 are of interest not only for the development of potential therapeutics against different diseases with overexpression of neuraminidases, but also to improve the understanding of the disease-related role of human neuraminidases (Figure 5.2B). Similar to viral neuraminidase inhibitors, high affinity and specificity was achieved by modifications of the glycerol side chain and the 4-position starting from DANA (**4**). Specific NEU1 inhibitor **7** was used to study epithal migration and bacterial adhesion and the specific NEU4 inhibitor **10** was studied in glioblastoma stem cell survival (34-37). The specific NEU2 and NEU3 inhibitors **8** and **9** were developed by Cairo *et al*.

in previous work, however, exhibiting only micromolar affinities (38). Affinities and selectivities are listed in Table 5.2.

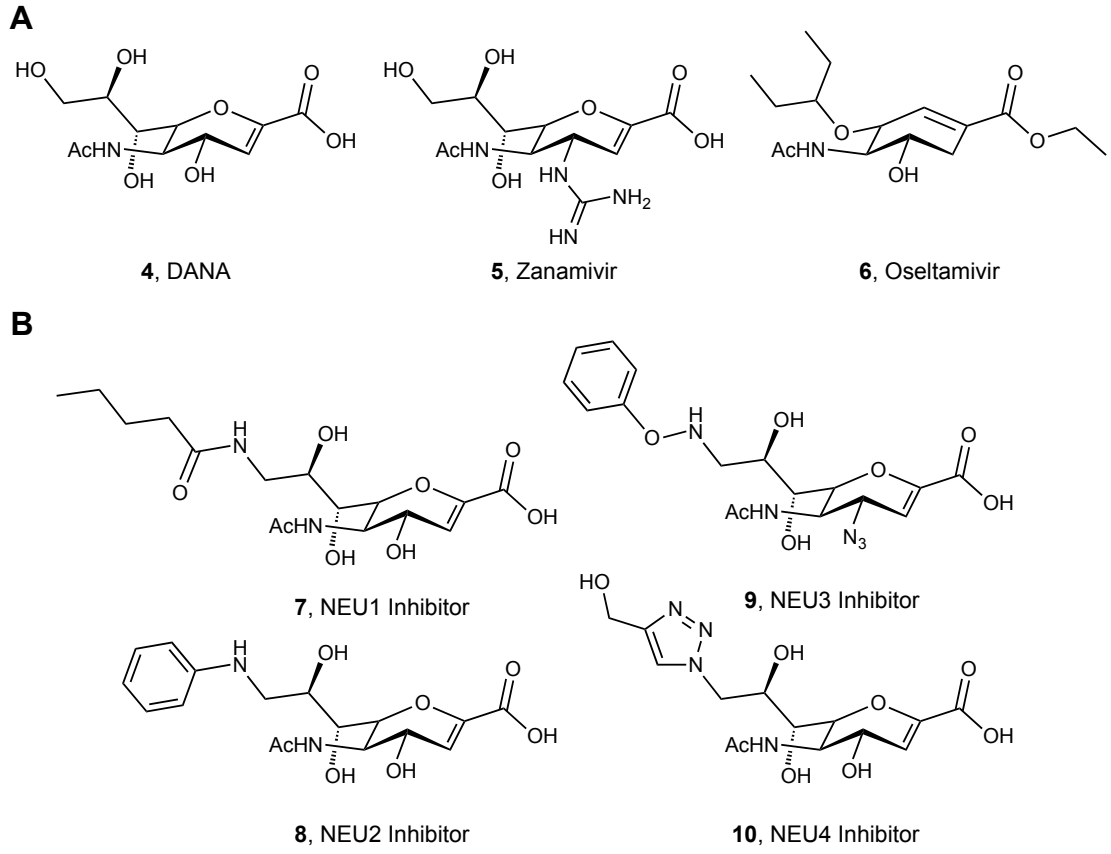


Figure 5.2. **A**) Transition state mimetic **4** and the specific viral neuraminidase inhibitor zanamivir (**5**) and oseltamivir (**6**); **B**) Selective human neuraminidase inhibitors **7**-**10**.

Table 5.2. IC_{50} values and selectivity of human neuraminidase inhibitors.

Human neuraminidase IC_{50} [μM]					Selectivity
Cpd.	NEU1	NEU2	NEU3	NEU4	
7	10	>1000	>1000	>1000	> 100 fold
8	>1000	86	>1000	470	12 fold
9	>1000	920	24	550	38 fold
10	>1000	>1000	80	0.16	500 fold

5.4 References

1. Angata T, Varki A. Chemical diversity in the sialic acids and related alpha-keto acids: an evolutionary perspective. *Chem Rev*. 2002;*102(2)*:439-69.
2. Comb DG, Roseman S. The sialic acids. I. The structure and enzymatic synthesis of N-acetylneuraminic acid. *J Biol Chem*. 1960;*235*:2529-37.
3. Faillard H. The early history of sialic acids. *Trends Biochem Sci*. 1989;*14(6)*:237-41.
4. Buschiazzo A, Alzari PM. Structural insights into sialic acid enzymology. *Curr Opin Chem Biol*. 2008;*12(5)*:565-72.
5. Vimr ER, Kalivoda KA, Deszo EL, Steenbergen SM. Diversity of microbial sialic acid metabolism. *Microbiol Mol Biol Rev*. 2004;*68(1)*:132-53.
6. Schauer R. Achievements and challenges of sialic acid research. *Glycoconj J*. 2000;*17(7-9)*:485-99.
7. Ley K, Gaehtgens P, Fennie C, Singer MS, Lasky LA, Rosen SD. Lectin-like cell adhesion molecule 1 mediates leukocyte rolling in mesenteric venules in vivo. *Blood*. 1991;*77(12)*:2553-5.
8. Gascon E, Vutskits L, Kiss JZ. Polysialic acid-neural cell adhesion molecule in brain plasticity: from synapses to integration of new neurons. *Brain Res Rev*. 2007;*56(1)*:101-18.
9. Schwarzkopf M, Knobeloch KP, Rohde E, Hinderlich S, Wiechens N, Lucka L, et al. Sialylation is essential for early development in mice. *Proc Natl Acad Sci U S A*. 2002;*99(8)*:5267-70.
10. Chou WK, Hinderlich S, Reutter W, Tanner ME. Sialic acid biosynthesis: stereochemistry and mechanism of the reaction catalyzed by the mammalian UDP-N-acetylglucosamine 2-epimerase. *J Am Chem Soc*. 2003;*125(9)*:2455-61.
11. Glaze PA, Watson DC, Young NM, Tanner ME. Biosynthesis of CMP-N,N'-diacetyllegionaminic acid from UDP-N,N'-diacetylbacillosamine in Legionella pneumophila. *Biochemistry*. 2008;*47(10)*:3272-82.
12. Lin CH, Lin CC. Enzymatic and chemical approaches for the synthesis of sialyl glycoconjugates. *Adv Exp Med Biol*. 2001;*491*:215-30.
13. Vimr E, Lichtensteiger C. To sialylate, or not to sialylate: that is the question. *Trends Microbiol*. 2002;*10(6)*:254-7.
14. Schauer R, Reuter G, Muhlpfordt H, Andrade AF, Pereira ME. The occurrence of N-acetyl- and N-glycoloylneuraminic acid in Trypanosoma cruzi. *Hoppe Seylers Z Physiol Chem*. 1983;*364(8)*:1053-7.
15. Preston A, Mandrell RE, Gibson BW, Apicella MA. The lipooligosaccharides of pathogenic gram-negative bacteria. *Crit Rev Microbiol*. 1996;*22(3)*:139-80.
16. Varki A. Sialic acids in human health and disease. *Trends Mol Med*. 2008;*14(8)*:351-60.
17. Harduin-Lepers A, Vallejo-Ruiz V, Krzewinski-Recchi MA, Samyn-Petit B, Julien S, Delannoy P. The human sialyltransferase family. *Biochimie*. 2001;*83(8)*:727-37.
18. Harduin-Lepers A, Recchi MA, Delannoy P. 1994, the year of sialyltransferases. *Glycobiology*. 1995;*5(8)*:741-58.
19. Mondal N, Buffone A, Jr., Stolfa G, Antonopoulos A, Lau JT, Haslam SM, et al. ST3Gal-4 is the primary sialyltransferase regulating the synthesis of E-, P-, and L-selectin ligands on human myeloid leukocytes. *Blood*. 2015;*125(4)*:687-96.
20. Miyagi T, Yamaguchi K. Mammalian sialidases: physiological and pathological roles in cellular functions. *Glycobiology*. 2012;*22(7)*:880-96.
21. Hata K, Wada T, Hasegawa A, Kiso M, Miyagi T. Purification and characterization of a membrane-associated ganglioside sialidase from bovine brain. *J Biochem*. 1998;*123(5)*:899-905.
22. Miyagi T, Wada T, Yamaguchi K, Hata K, Shiozaki K. Plasma membrane-associated sialidase as a crucial regulator of transmembrane signalling. *J Biochem*. 2008;*144(3)*:279-85.
23. Kopitz J, von Reitzenstein C, Andre S, Kaltner H, Uhl J, Ehemann V, et al. Negative regulation of neuroblastoma cell growth by carbohydrate-dependent surface binding of galectin-1 and functional divergence from galectin-3. *J Biol Chem*. 2001;*276(38)*:35917-23.
24. Sasaki A, Hata K, Suzuki S, Sawada M, Wada T, Yamaguchi K, et al. Overexpression of plasma membrane-associated sialidase attenuates insulin signaling in transgenic mice. *J Biol Chem*. 2003;*278(30)*:27896-902.

25. Schengrund CL, Jensen DS, Rosenberg A. Localization of sialidase in the plasma membrane of rat liver cells. *J Biol Chem*. 1972;*247(9)*:2742-6.
26. Yamanami H, Shiozaki K, Wada T, Yamaguchi K, Uemura T, Kakugawa Y, et al. Down-regulation of sialidase NEU4 may contribute to invasive properties of human colon cancers. *Cancer Sci*. 2007;*98(3)*:299-307.
27. Yogeeswaran G, Hakomori S. Cell contact-dependent ganglioside changes in mouse 3T3 gibroblasts and a suppressed sialidase activity on cell contact. *Biochemistry*. 1975;*14(10)*:2151-6.
28. Colman PM, Varghese JN, Laver WG. Structure of the catalytic and antigenic sites in influenza virus neuraminidase. *Nature*. 1983;*303(5912)*:41-4.
29. Moscona A. Neuraminidase inhibitors for influenza. *N Engl J Med*. 2005;*353(13)*:1363-73.
30. Taylor NR, von Itzstein M. Molecular modeling studies on ligand binding to sialidase from influenza virus and the mechanism of catalysis. *J Med Chem*. 1994;*37(5)*:616-24.
31. Taylor NR, Cleasby A, Singh O, Skarzynski T, Wonacott AJ, Smith PW, et al. Dihydropyrancarboxamides related to zanamivir: a new series of inhibitors of influenza virus sialidases. 2. Crystallographic and molecular modeling study of complexes of 4-amino-4H-pyran-6-carboxamides and sialidase from influenza virus types A and B. *J Med Chem*. 1998;*41(6)*:798-807.
32. Kim CU, Lew W, Williams MA, Wu H, Zhang L, Chen X, et al. Structure-activity relationship studies of novel carbocyclic influenza neuraminidase inhibitors. *J Med Chem*. 1998;*41(14)*:2451-60.
33. Lew W, Chen X, Kim CU. Discovery and development of GS 4104 (oseltamivir): an orally active influenza neuraminidase inhibitor. *Curr Med Chem*. 2000;*7(6)*:663-72.
34. Hyun SW, Liu A, Liu Z, Cross AS, Verceles AC, Magesh S, et al. The NEU1-selective sialidase inhibitor, C9-butyl-amide-DANA, blocks sialidase activity and NEU1-mediated bioactivities in human lung in vitro and murine lung in vivo. *Glycobiology*. 2016;*26(8)*:834-49.
35. Silvestri I, Testa F, Zappasodi R, Cairo CW, Zhang Y, Lupo B, et al. Sialidase NEU4 is involved in glioblastoma stem cell survival. *Cell Death Dis*. 2014;*5*:e1381.
36. Albohy A, Zhang Y, Smutova V, Pshezhetsky AV, Cairo CW. Identification of Selective Nanomolar Inhibitors of the Human Neuraminidase, NEU4. *ACS Med Chem Lett*. 2013;*4(6)*:532-7.
37. Magesh S, Moriya S, Suzuki T, Miyagi T, Ishida H, Kiso M. Design, synthesis, and biological evaluation of human sialidase inhibitors. Part 1: selective inhibitors of lysosomal sialidase (NEU1). *Bioorg Med Chem Lett*. 2008;*18(2)*:532-7.
38. Zhang Y, Albohy A, Zou Y, Smutova V, Pshezhetsky AV, Cairo CW. Identification of selective inhibitors for human neuraminidase isoenzymes using C4,C7-modified 2-deoxy-2,3-didehydro-N-acetylneuraminic acid (DANA) analogues. *J Med Chem*. 2013;*56(7)*:2948-58.

5.5 Publication 2: Selective Neuraminidase 3 Inhibitors

This publication describes the synthesis and experimental examination of Neuraminidase 3 (NEU3) inhibitors and ester prodrugs thereof. A selectivity and affinity improvement towards NEU3 was achieved by structural modification of the known neuraminidase inhibitor DANA (**4**). The ester prodrugs did not show a significant improvement in *in vitro* absorption measurements.

Contribution to the Project:

Philipp Dätwyler was evaluating the physicochemical and pharmacokinetic properties of the ester prodrugs and active principle. He further contributed to the writing and arrangement of the publication.

The publication was published in the peer-reviewed *Journal of Medicinal Chemistry* in Februar 2018.

Tianlin Guo[‡], Philipp Dätwyler[§], Ekaterina Demina[†], Michele R. Richards[‡], Peng Ge[‡], Chunxia Zou[‡], Ruixiang Zheng[‡], Alexey V. Pshezhetsky[†], Beat Ernst[§], and Christopher W. Cairo[‡]*

DOI: 10.1021/acs.jmedchem.7b01574

[‡]Alberta Glycomics Centre and Department of Chemistry, University of Alberta, Edmonton Alberta, T6G 2G2, Canada

[§]Department of Pharmaceutical Sciences, Pharmacenter, University of Basel, Klingelbergstrasse 50, CH-4056 Basel, Switzerland

[†]Division of Medical Genetics, Sainte-Justine University Hospital Research Center, University of Montreal, Montréal, Canada

*Corresponding author

Journal of Medicinal Chemistry

Article

Cite This: J. Med. Chem. XXXX, XXX, XXX–XXX

pubs.acs.org/jmc

Selective Inhibitors of Human Neuraminidase 3

Tianlin Guo,[†] Philipp Dätwyler,[‡] Ekaterina Demina,[§] Michele R. Richards,[†] Peng Ge,[†] Chunxia Zou,[†] Ruixiang Zheng,[†] Anne Fougerat,[§] Alexey V. Pshezhetsky,[§] Beat Ernst,[‡] and Christopher W. Cairo*[,†]

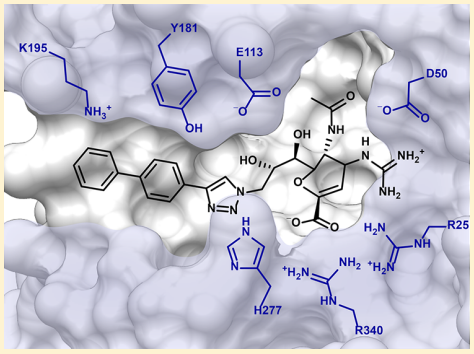
[†]Alberta Glycomics Centre and Department of Chemistry, University of Alberta, Edmonton Alberta T6G 2G2, Canada

[‡]Department of Pharmaceutical Sciences, Pharmacenter, University of Basel, Klingelbergstrasse 50, CH-4056 Basel, Switzerland

[§]Division of Medical Genetics, Sainte-Justine University Hospital Research Center, University of Montreal, Montréal, Quebec H3T 1C5, Canada

Supporting Information

ABSTRACT: Human neuraminidases (NEU) are associated with human diseases including cancer, atherosclerosis, and diabetes. To obtain small molecule inhibitors as research tools for the study of their biological functions, we designed a library of 2-deoxy-2,3-didehydro-*N*-acetylneuraminic acid (DANA) analogues with modifications at C4 and C9 positions. This library allowed us to discover selective inhibitors targeting the human NEU3 isoenzyme. Our most selective inhibitor for NEU3 has a K_i of 320 ± 40 nM and a 15-fold selectivity over other human neuraminidase isoenzymes. This inhibitor blocks glycolipid processing by NEU3 in vitro. To improve their pharmacokinetic properties, various esters of the best inhibitors were synthesized and evaluated. Finally, we confirmed that our best compounds exhibited selective inhibition of NEU orthologues from murine brain.



INTRODUCTION

Neuraminic acids (also known as sialic acids) form a large family of α-keto acid monosaccharides containing a nine-carbon backbone.[1] They are diverse in structure and widely distributed among organisms in nature. The most common form of neuraminic acid in humans is Neu5Ac (5-acetamido-3,5-dideoxy-D-*glycero*-D-*galacto*-non-2-ulosonic acid).[2] Generally, neuraminic acids are the terminal residues in glycan chains, playing crucial roles in molecular recognition.[3] The neuraminic acid content in human cells is controlled by enzymes including sialyltransferases and neuraminidases (NEU; also known as sialidases).[4] In humans, four neuraminidase isoenzymes have been identified: NEU1, NEU2, NEU3, and NEU4.[5] They exhibit different subcellular localization and show different preferences toward glycan substrates.[5] Abnormal activity of human neuraminidases has been linked to diseases including lysosomal storage disorders,[6,7] cardiovascular diseases,[8,9] and cancer.[10−12] Selective inhibitors of NEU4 have been used to investigate the role of NEU4 in glioblastoma,[13] and NEU1 inhibitors have been employed to study epithelial migration and bacterial adhesion.[14] Thus, isoenzyme-selective inhibitors of human neuraminidases are potential tools to study the biological function of these enzymes.[15]

Although there has been significant efforts to develop potent inhibitors of viral neuraminidases,[16] inhibitors of human neuraminidases are not as thoroughly explored,[15] and most viral neuraminidase inhibitors have only weak or limited activity for the human neuraminidases.[15,17,18] The most active viral inhibitor against the human isoenzymes is zanamivir (**6**, Figure 1), which has low micromolar activity for NEU2 and NEU3.[17,19] Most human neuraminidase inhibitors reported to-date are based on 2-deoxy-2,3-didehydro-*N*-acetylneuraminic acid (DANA, **1**; Figure 1).[15,20] Magesh et al. investigated a series of C9-amide analogues of DANA and reported a NEU1-selective inhibitor containing a C9-pentylamide with an IC_{50} of 10 μM (C9-BA-DANA,[21] **2**; Figure 1).[22] Modifications of the C5 and C9 positions of DANA improved potency for DANA analogues against NEU2.[19,23] Selective inhibitors for NEU2 (**3**, Figure 1) and NEU3 (**4**, Figure 1) were identified in a library of DANA analogues when the glycerol side chain was replaced by oxime- and amino-linked aromatic groups.[24] Inhibitor **3** has a reported IC_{50} of 86 μM against NEU2 with 12-fold selectivity. Inhibitor **4** was the most selective compound reported to date for NEU3 with a 38-fold selectivity (IC_{50} = 24 μM). We previously examined triazole modifications at the C5 and C9 positions in DANA and found improved potency of C9-triazolyl DANA analogues toward NEU3.[25] When evaluated against the full panel of NEU isoenzymes, C9-4-hydroxymethyltriazolyl DANA (C9-4HMT-DANA, **5**; 500-fold selective, K_i = 30 nM, Figure 1) turned out to be the most selective inhibitor for NEU4 identified to date. This compound has remained the only reported inhibitor with submicromolar activity for any human NEU.[26] We, therefore, considered that a more extensive library of triazole analogues might take advantage of the hypothesized large C9-pocket of NEU3 and could provide selective inhibitors of NEU3 with further improvements in potency. Furthermore, the activity of

Received: October 23, 2017
Published: February 9, 2018

DOI: 10.1021/acs.jmedchem.7b01574
J. Med. Chem. XXXX, XXX, XXX–XXX

Journal of Medicinal Chemistry Article

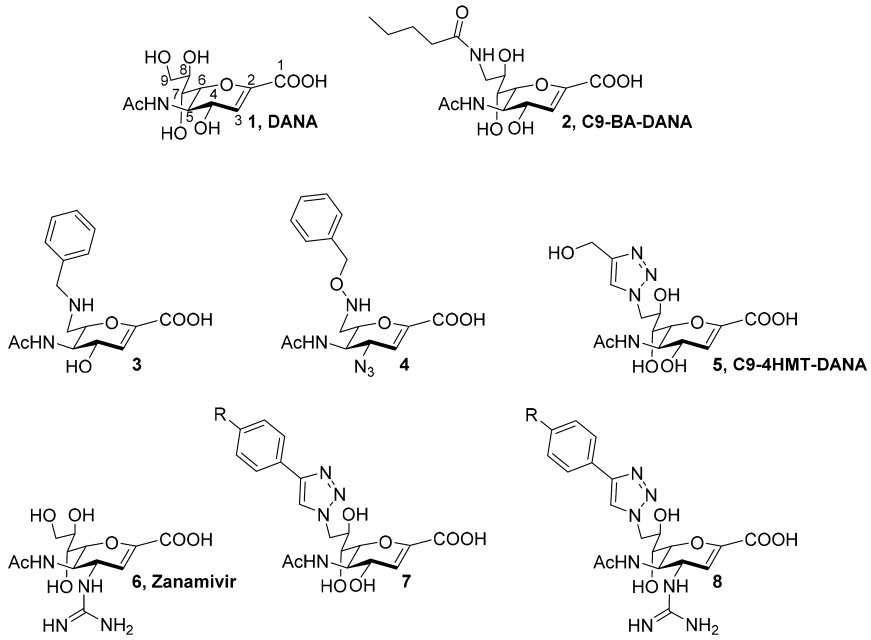
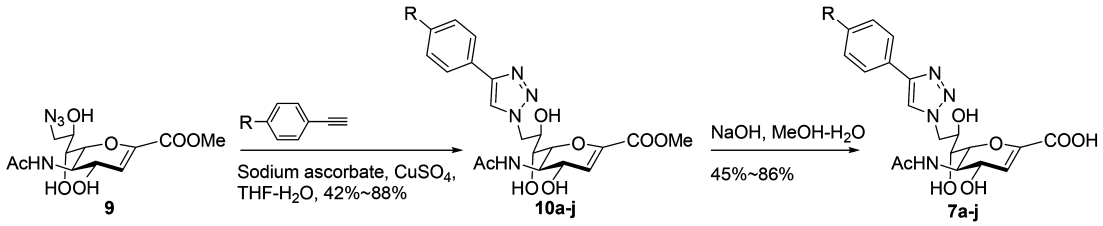


Figure 1. Inhibitor targets.

Scheme 1. Synthetic Route to C9-Phenyltriazole DANA Analogues 7a–j



zanamivir **6** against NEU3 suggested that an appropriate C4 moiety could contribute to increased activity.

In this work, we designed and synthesized a series of compounds featuring a C9-triazolyl group combined with guanidino and other groups at C4. The first series of C9-triazolyl DANA derivatives exhibited remarkably increased potencies relative to the parent compound **1** toward NEU3 and NEU4 (**7**, Figure 1). Although the combination of the best C9-triazolyl fragments with a guanidino group at C4 provided only small gains in potency for NEU3, we found that selectivity for NEU3 over NEU4 was remarkably improved (**8**, Figure 1). The inhibitors presented here constitute the most potent and selective compounds identified for NEU3 so far. In addition, we have established a rationale for the development of future selective inhibitors of NEU3.

RESULTS

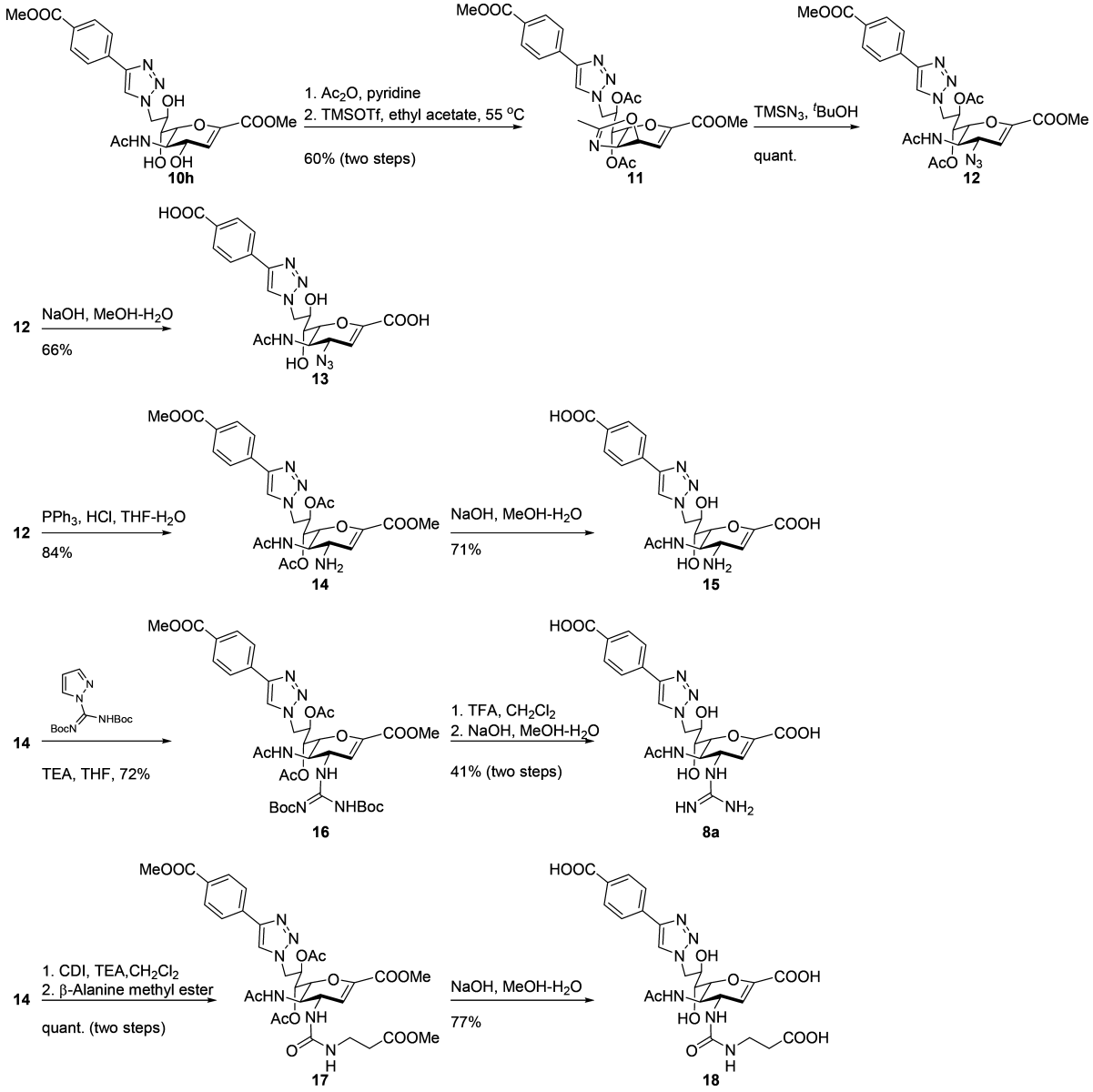
Inhibitor Design and Synthesis. To date, X-ray crystal structure data is only available for human NEU2, including co-crystal data with inhibitors.[27,28] Homology models for NEU3[29−31] have suggested that its glycerol binding pocket is larger than in NEU2. The homology models of NEU3 also indicate that the glycerol side chain interacts with Y179 and Y181. Earlier work reported that large groups, such as a C9-phenyltriazolyl, were tolerated by the C9 pocket of NEU3,[25] but no fragment has been identified which substantially improved potency. We first hypothesized that an appropriate aromatic group at C9 would provide additional interactions and could improve NEU3 potency over other isoenzymes. We therefore explored C9-4-aryl-triazolyl DANA derivatives with variations in electron-donating groups, charge, as well as larger phenyl and phenoxyl groups (**7**, Figure 1). We also considered that a guanidino group at C4 could improve activity for NEU2 and NEU3[17,19] and, therefore, investigated C9-triazolyl DANA derivatives with nitrogen-containing groups at C4, including guanidino, azido, and amino groups.

To prepare analogues with substituted phenyltriazolyl groups at C9, we started from the reported C9-azido-DANA methyl ester (**9**, Scheme 1), which was synthesized from Neu5Ac in six steps.[25] The library of triazolyl derivatives **10a–j** was obtained by CuAAC (Cu-catalyzed azide–alkyne cycloaddition) by following reported conditions (Scheme 1).[25,32,33] Finally, deprotection yielded the test compounds **7a–j**, in moderate to good yields.

For compounds with combined C4,C9-modifications (**8**, Figure 1), two strategies were adopted. Starting from the C9-triazolyl DANA analogue **10h**, a C4-azido group was introduced according to a reported procedure via nucleophilic ring opening of an oxazoline intermediate **11** to yield compound **12** (Scheme 2).[33] Reduction of the C4-azido of compound **12** via Staudinger reduction provided the C4-amino compound **14**.[34,35] After deprotection, the C4-azido and C4-

DOI: 10.1021/acs.jmedchem.7b01574
J. Med. Chem. XXXX, XXX, XXX−XXX

Scheme 2. Synthetic Routes to C4,C9-Modified DANA Analogues 8a, 13, 15, and 18



amino intermediates **12** and **14** gave the test compounds **13** and **15**. Furthermore, the amino group in **14** was converted to a guanidino group using *N*,*N*′-di-Boc-1*H*-pyrazole-1-carboxamidine, yielding the final C4, C9-modified test compound **8a** after deprotection. To obtain an alternative polar group at C4, a C4-*N*-alkyl urea compound was prepared by treatment of **14** with 1,1′-carbonyldiimidazole and *β*-alanine methyl ester. Finally, hydrolysis provided test compound **18**.

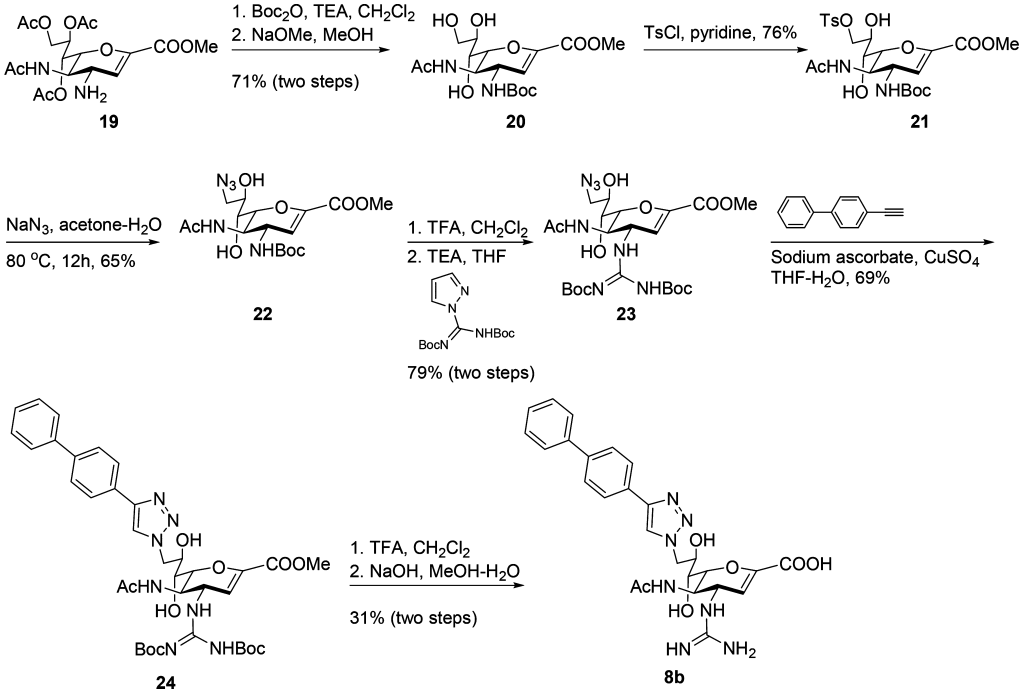
An alternative strategy (Scheme 3) allowed preparation of compound **8b** by reversing the order of the modifications at the 4- and 9-position. First, *O*-acetyl protected C4-amino-DANA **19** was prepared as reported.[34] Protection of the amino group with Boc, followed by acetate deprotection (**20**) and subsequent tosylation at C9 (**21**) allowed for displacement to the C9-azido group (**22**).[25] Next, the guanidino group was introduced at C4 to provide compound **23**. Subsequent CuAAC with a biphenyl ethyne allowed for installation of the C9-triazolyl group, forming compound **24**, which gave test compound **8b** after deprotection.

Additional compounds for testing included a small series of C4-amides, prepared starting from compound **19** by treatment with the appropriate anhydride or acyl chloride to yield the amido compounds **25a**–**d** (Scheme 4) after hydrolysis of the C1-methyl ester with sodium hydroxide. Compounds **1**, **2**, and **6** were also prepared for use in the inhibition assays.

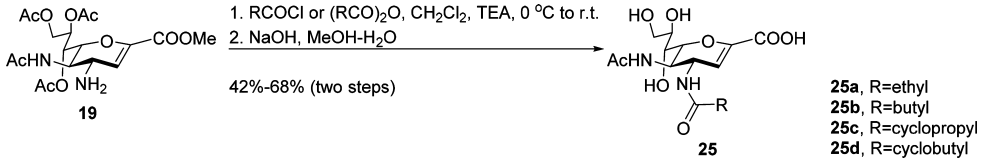
Inhibition Assays. To determine the inhibitory potency of the panel of inhibitors against NEU, isoenzymes were prepared either recombinantly or by purification.[24,26,29] For the determination of enzyme activity (see Table 1), the fluorogenic neuraminidase substrate, 2′-(4-methylumbelliferyl)-*α*-D-*N*-acetylneuraminic acid (4MU-NANA) was used.[36,37] We confirmed that compound **2** acted as a selective inhibitor of NEU1 (IC_{50} of 3.4 ± 0.2 *μ*M).[22] Compounds **7a**–**j** with C9-phenyltriazolyl groups were significantly more potent inhibitors of NEU3 and NEU4 compared to NEU1 and NEU2. For virtually all

DOI: 10.1021/acs.jmedchem.7b01574
J. Med. Chem. XXXX, XXX, XXX–XXX

Scheme 3. Synthetic Route to C4,C9-Modified DANA Analogue 8b



Scheme 4. Synthetic Route to C4-Amide DANA Analogues 25a–d



compounds containing a C9-phenyltriazolyl group, we observed single-digit micromolar or better IC_{50} values toward NEU3. Compounds with an acidic group (**7h**) or larger aromatic groups (**7i** and **7j**) showed improved activity against NEU3 over those with basic groups (**7a** and **7c**) or neutral electron-withdrawing groups (**7f**). Among all C9-modified DANA analogues, the C9-biphenyltriazolyl DANA **7i** showed the highest potency with an IC_{50} of 0.70 ± 0.10 and 0.52 ± 0.10 μM toward NEU3 and NEU4, respectively. The activity of **7i** showed distinct preferences, with 40-fold greater potency against NEU3 and NEU4 than for NEU1 or NEU2.

Our library of C4 modifications revealed the importance of the guanidino group for NEU3 selectivity. The comparison of DANA (**1**) and zanamivir (**6**) was consistent with previous reports[17,19] with potency of 37 ± 6 and 7.8 ± 2.0 μM for NEU2 and 7.7 ± 0.8 vs 4.0 ± 0.6 μM for NEU3. Decreased potency of 49 ± 8 μM and >500 μM was found for NEU1 and 8.3 ± 1.0 vs 47 ± 6 μM for NEU4. Other nitrogen-containing groups at C4, like azido, amino, or amido moieties (**13**, **15**, **18**, and **25**), decreased potency for all isoenzymes. These results are in agreement with previously reported findings as C4-azido- and C4-amino-DANA showed improved potency toward NEU2 and significantly decreased potency toward other isoenzymes.[24] We therefore concluded that a positively charged group of appropriate size at C4 was critical for NEU3 selectivity. Thus, a guanidino group at C4 combined with C9-modification conferred selectivity for NEU3 over NEU4. Selective and potent inhibitors for NEU3 were identified, where compound **8a** (NEU3 IC_{50} of 0.6 ± 0.1 μM) showed 40-fold selectivity, and compound **8b** (NEU3 IC_{50} of 0.58 ± 0.14 μM) had 10-fold selectivity. Although both **8a** and **8b** demonstrated unprecedented potency for NEU3, we selected **8b** for further studies due to its reduced polarity.

Next, we determined the inhibition constants (K_i) of the best inhibitors for their active targets (Table 2). The K_i results showed a trend similar to the IC_{50} values. When zanamivir **6** was tested against NEU2, NEU3, and NEU4, we found K_i values of 5.7 ± 1.5, 0.62 ± 0.09, and 26 ± 4 μM, respectively. Whereas compound **7i** showed similar K_i values for NEU3 and NEU4 (0.28 ± 0.04 and 0.26 ± 0.04 μM, respectively), the 4-guanidino derivative **8b** exhibited 15-fold selectivity for NEU3 (K_i for NEU3 0.32 ± 0.04 μM, and 5 ± 1 μM for NEU4). Because NEU3 is known to favor glycolipid substrates over 4MU-NANA,[38] we also tested the inhibitory activity of the best NEU3 inhibitors to block GM3 hydrolysis by the enzyme (Table 3). In comparison with DANA **1**, **7i** showed a higher inhibitory potency for GM3 cleavage catalyzed by NEU3 (12 ± 2 vs 54 ± 10 μM) and NEU4 (3.7 ± 0.7 vs 26 ± 8 μM); **8b** was 14-fold more potent than DANA (**1**) (3.8 ± 0.5 vs 54 ± 10 μM).

Pharmacokinetic Evaluation of Active Inhibitors. For planning future in vivo experiments with this series of potent NEU3 inhibitors, their pharmacokinetic properties are critical. Carbohydrate mimetics are typically hydrophilic and are, therefore, often delivered as lipophilic ester prodrugs to improve membrane permeability, where the active principle is

DOI: 10.1021/acs.jmedchem.7b01574
J. Med. Chem. XXXX, XXX, XXX–XXX

Table 1. IC_{50} Data for NEU Using 4MU-NANA as the Substrate

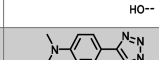
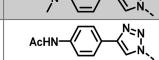
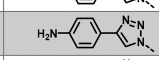
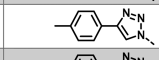
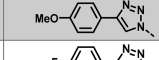
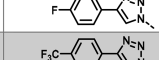
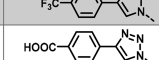
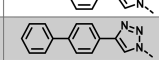
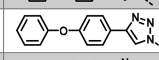
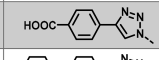
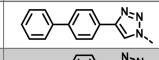
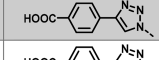
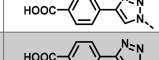
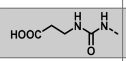
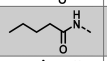
Structure			IC_{50} [μM]			
Compound	**C4**	**C9**	**NEU1**	**NEU2**	**NEU3**	**NEU4**
1 (DANA)			49 ± 8	37 ± 6	7.7 ± 0.8	8.3 ± 1.0
2 (C9-BA-DANA)			3.4 ± 0.2	>500	110 ± 40	220 ± 50
6 (zanamivir)			>500	7.8 ± 2.0	4.0 ± 0.6	47 ± 6
7a			190 ± 40	250 ± 40	9.3 ± 0.8	28 ± 3
7b			59 ± 5	78 ± 20	5.5 ± 0.5	11 ± 1
7c			76 ± 16	350 ± 90	8.7 ± 0.9	12 ± 1
7d			51 ± 7	290 ± 50	8.2 ± 1.3	3.9 ± 0.4
7e			47 ± 6	430 ± 200	5.9 ± 0.4	2.6 ± 0.3
7f			100 ± 40	>500	17 ± 2	1.3 ± 0.2
7g			140 ± 30	360 ± 50	3.3 ± 0.5	5.0 ± 0.9
7h			240 ± 90	110 ± 10	1.1 ± 0.1	3.0 ± 0.3
7i			>500	32 ± 5	0.70 ± 0.10	0.52 ± 0.10
7j			>500	84 ± 24	1.0 ± 0.1	0.97 ± 0.24
8a			>500	45 ± 5	0.61 ± 0.10	24 ± 2
8b			>500	5.9 ± 1.1	0.58 ± 0.14	5.9 ± 1.4
13			>500	>500	20 ± 6	400 ± 70
15			>500	>500	>500	>500
18			>500	430 ± 300	>500	66 ± 16
25a			>500	>500	>500	>500
25b			>500	>500	>500	>500
25c			>500	>500	>500	>500
25d			>500	>500	>500	>500

Table 2. K_i Determinations

Structure			K_i [μM][a]		
Compound	**C4**	**C9**	**NEU2**	**NEU3**	**NEU4**
1 (DANA)			25 ± 4	1.6 ± 0.3	5.8 ± 0.6
6 (zanamivir)			5.7 ± 1.5	0.62 ± 0.09	26 ± 4
7i			48 ± 9	0.28 ± 0.04	0.26 ± 0.04
8b			17 ± 4	0.32 ± 0.04	5 ± 1
5[b]			ND	ND	0.030 ± 0.02

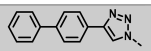
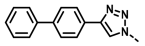
[a]None of these compounds showed significant activity toward NEU1, so only NEU2, NEU3, and NEU4 determinations were made. [b]Previously reported data.[26]

released via hydrolysis by esterases.[39,40] Therefore, C1-esters of inhibitors **1**, **2**, **5**, **7i**, and **8b** were prepared. The methyl ester of each compound was obtained after CuAAC as described earlier, whereas elongated aliphatic esters of **7i** were prepared from the C1 methyl ester of **7i** (**7i-OMe**) via transesterification catalyzed by LiBr and DBU.[41] The octanol−water partition coefficient at pH 7.4 (log $D_{7.4}$), membrane permeation (log P_e), kinetic solubility, and metabolic stability of the selected compounds were determined.[42] The log $D_{7.4}$ values were determined using a shake-flask approach.[43,44] Membrane permeation log P_e was determined using a parallel artificial membrane permeability assay (PAMPA).[45] A summary of the pharmacokinetic parameters is provided in Table 4. A plot of the measured log $D_{7.4}$ against log P_e is shown in Figure 2 with indications for moderate and high absorption potential.[46] Passive permeability (log P_e) of ester prodrugs is dependent on lipophilicity. In general, the ester prodrugs showed consistently low solubility and time-dependent precipitation. While the solution of the dissolved esters appeared clear after the initial dilution of the DMSO stock with water, turbidity developed after several hours, indicating the formation of precipitates. Because PAMPA required 16 h incubation time, measurements of insufficiently soluble compounds were performed with 10% DMSO.

The passive permeation (log P_e) of all compounds was below the threshold of −6.3, predicting moderate absorption potential (Figure 2). Only the largest aliphatic ester prodrug (**7i-OBu**)

DOI: 10.1021/acs.jmedchem.7b01574
J. Med. Chem. XXXX, XXX, XXX−XXX

Table 3. Inhibition of GM3 Cleavage

Structure			IC_{50} [μM]			
Compound	**C4**	**C9**	**NEU3**	**NEU4**	**Rel/NEU3**[a]	**Rel/NEU4**[a]
DANA (1)	HO--	HO--	54 ± 10	26 ± 8	ND	ND
7i	HO--		12 ± 2	3.7 ± 0.7	4.1	7.0
8b			3.8 ± 0.5	ND	14	ND

[a]Relative activity as compared to DANA.

Table 4. Pharmacokinetic Evaluation of C1-Esters

parent compd	C1(O)R	$clogD_{7.4}$[a]	log $D_{7.4}$[b]	log P_e [log cm s^{-1}][c]	solubility [μg mL^{-1}][d]
DANA, **1**	OMe	−3.31	<−1.5	nd	nd
2	OMe	−2.01	−1.0 ± 0.1	−8.3 ± 0.1	nd
5	OMe	−3.49	<−1.5	nd	nd
8b	OMe	−1.95	1.2 ± 0.1	−7.4 ± 0.1	nd
7i	OMe	0.91	2.3 ± 0.1	−7.8 ± 0.4*	Tris pH 7.4:1.5 Prism pH 6.5: >1.0 NaCl 0.9%: >1.0
7i	OEt	1.27	3.4 ± 0.1	−8.3 ± 0.3*	Tris pH 7.4:13.8 Prism pH 6.5:3.2 NaCl 0.9%: 3.7
7i	OPr	1.79	nd	−8.2 ± 0.3*	Tris pH 7.4:2.8 Prism pH 6.5:1.7 NaCl 0.9%: 2.4
7i	OBu	2.23	3.8 ± 0.1	−6.5 ± 0.4*	Tris pH 7.4:3.9 Prism pH 6.5:3.2 NaCl 0.9%: 5.0

[a]$clogD_{7.4}$ was calculated using Marvin Sketch software (version 16.10.10.0) [b]Octanol−water distribution coefficients (log $D_{7.4}$) were determined at pH 7.4 by a miniaturized shake flask procedure in sextuplicate in Tris buffer (0.1 M). [c]*Pe* = effective permeability: diffusion through an artificial membrane was determined by the parallel artificial membrane permeability assay (PAMPA) in quadruplicate at pH 6.5 in Prism buffer. Measurements indicated with * indicate inclusive 10% DMSO. [d]The kinetic solubility was measured in singlet measurements in three different buffer systems (Tris 0.1 M, pH 7.4; Prism pH 6.5; 0.9% NaCl in water).

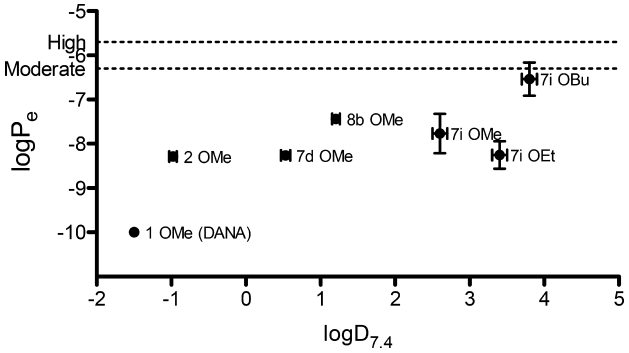


Figure 2. Passive permeability log P_e of ester prodrugs versus lipophilicity. The dotted lines indicate the threshold for moderate and high absorption potential.[46] Error bars show the standard deviation of the measurement in six independent measurements of log $D_{7.4}$ and four independent measurements of log P_e.

reached the lower limit of this threshold. In combination with the low solubility, the compounds are classified as class IV compounds in the biopharmaceutical drug classification system.[47] This class of compounds is highly susceptible to individual transport mechanisms and, therefore,[48] we further explored the permeability of the *n*-butyl ester of **7i** (**7i-OBu**) using a colon carcinoma (Caco2) cell-based permeation assay.[49]

The threshold for a moderate absorption potential is 2.0×10^{-6} cm s^{-1} in the absorptive direction (apical to basolateral: a−b)[42] and was not reached by **7i-OBu**. Additionally, the compound had an efflux ratio of 6.7, predicting active efflux. When the metabolic stability of the prodrug in human liver microsomes was assessed (Table 5), the hydrolysis of the ester was slow, limiting the release of the active principle. A slow release of the active principle can be beneficial, especially for hydrophilic carbohydrate mimetics to overcome the common problem of fast renal excretion, but in our case the active principle is already lipophilic and the biphenyl moiety is likely prone to oxidation by cytochromes. In summary, the lead compound **7i-**

Table 5. Cell Permeability of 7i-OBu

	Caco2 P_{app} [10^{-6} cm s^{-1}][a]			
	a−b	b−a	b−a/a−b	MetStab HLM $T_{1/2}$ [min][b]
7i-OBu	0.32 ± 0.09	2.2 ± 0.2	6.7	>90

[a]P_{app} = apparent permeability: permeation through a Caco-2 cell monolayer was assessed in the absorptive (a → b) and secretory (b → a) direction in triplicate after 30 min. The initial compound concentration (c_0) in the donor chamber was 62.5 μM. [b]The metabolic stability assay was performed in human liver microsomes (HLM 0.5 mg mL^{-1}) in triplicate.

DOI: 10.1021/acs.jmedchem.7b01574
J. Med. Chem. XXXX, XXX, XXX−XXX

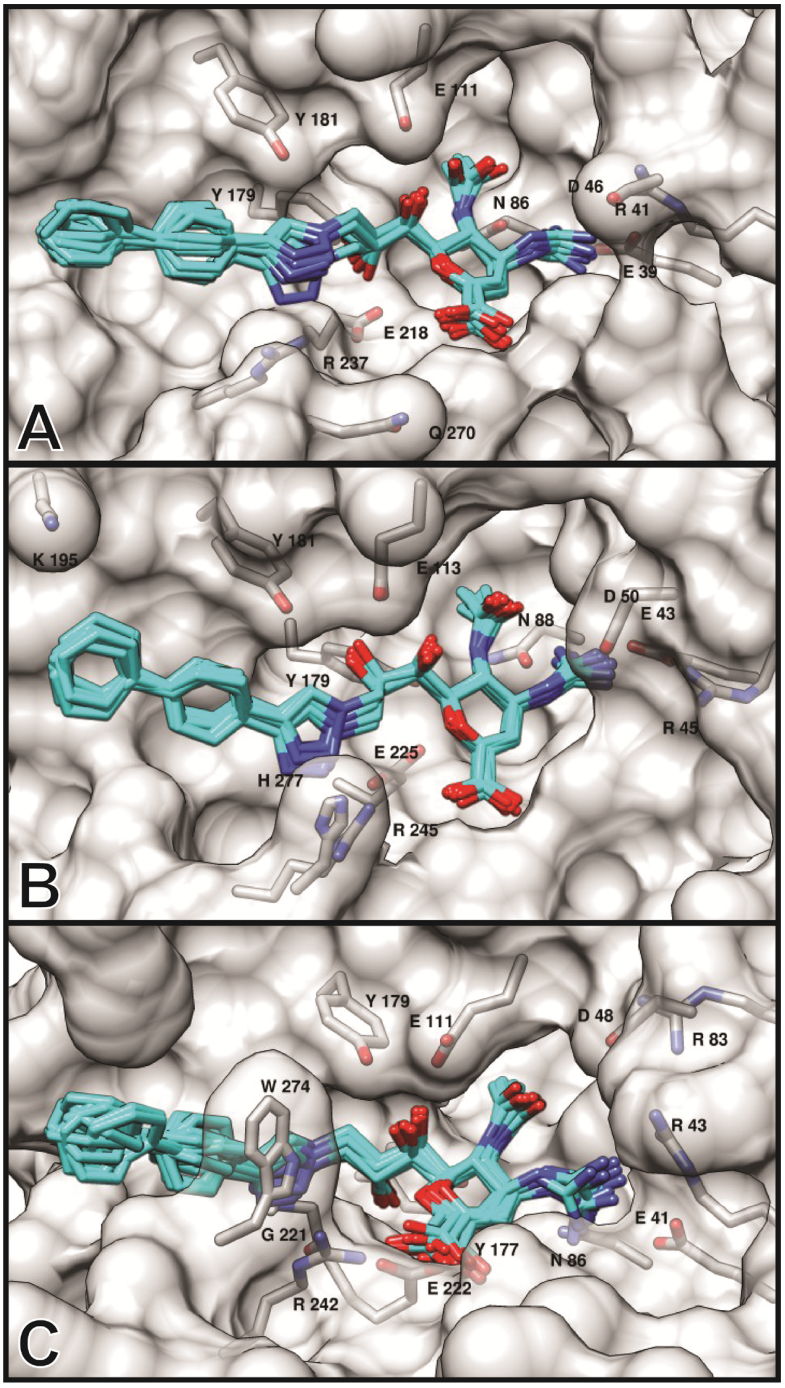


Figure 3. Average enzyme structures for (A) NEU2, (B) NEU3, and (C) NEU4 with **8b** in the active site. The inhibitor is shown as 10 representative structures taken from the last 25 ns of the MD simulation.

OBu showed insufficient solubility, a suboptimal efflux ratio, and slow permeability in a cell permeation assay and slow metabolic release of the active principle. Together, these results predict low oral bioavailability without further modification.

Active Site Models of Human NEUs. To gain structural insight into the selectivity of active inhibitors, we ran 50 ns molecular dynamics (MD) simulations of **8b** in complex with NEU2, NEU3, and NEU4. NEU models were based on reported X-ray or homology models.[27,29−31,50] The starting structures for each simulation are shown in the Supporting Information. Inhibitor **8b** remained bound in the three active sites throughout the MD simulations (Figure 3) and the parent DANA ring in **8b** maintained the ring conformation that mimics the proposed transition state of the oxocarbenium ion intermediate. We analyzed ring conformations of **8b** in the simulations using standard nomenclature for six-membered monosaccharides, with C2 being the anomeric carbon (Figure 1, **1**).[51] The population of ring conformations[51] observed for **8b** were found over E_5−$^6H_5$−^6E and were similar for all three enzymes, with NEU2 centered at 94° ($^6H_5$), NEU3 centered at

DOI: 10.1021/acs.jmedchem.7b01574
J. Med. Chem. XXXX, XXX, XXX−XXX

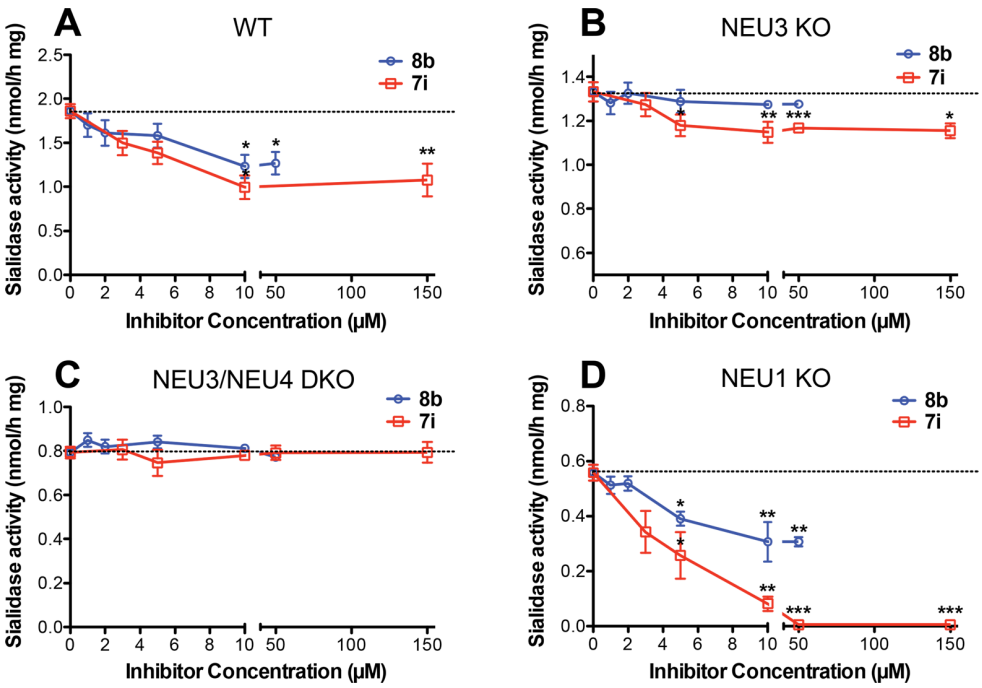


Figure 4. Inhibition of NEU activity in murine brain homogenate. Brains were extracted from WT C57Bl6 mice and those KO for the NEU1 (*neu1*$^{-/-}$), NEU3 (*neu3*$^{-/-}$), NEU4 (*neu4*$^{-/-}$), or double KO (DKO) for the NEU3/4 genes (*neu3*$^{-/-}$; *neu4*$^{-/-}$). Brain homogenates were assayed for inhibition of neuraminidase activity against 4MU-NANA. The WT brain homogenates (A) showed partial inhibition (up to 70% at 10 μM) of neuraminidase activity by both **8b** and **7i**, while the NEU3 KO brain homogenate (B) showed no inhibition with **8b** and NEU3/4 DKO brain homogenate (C) showed no significant inhibition with either compound **8b** nor **7i**. The NEU1 KO brain homogenate (D) was the only sample to show dramatic inhibition for both compounds **8b** and **7i**, consistent with specificity of these compounds for the NEU3/4 isoenzymes. Importantly, all homogenates showed a complete inhibition of neuraminidase activity at 150 μM of DANA **1** (data not shown). Points are labeled as significantly different (*, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$) from the activity in the absence of the inhibitors according to *t* test with a Dunlett post-test. Data are shown as mean values ± SD of 6−12 independent experiments.

106° ($^6H_5/^6E$), and NEU4 centered at 82° ($E_5/^6H_5$).[52] For comparison, the ring conformation of DANA **1** bound to NEU2 is E_5 in the co-crystal structure (PDB 1VCU).[27]

We had predicted that large groups at C9 would provide better inhibitors of NEU3 and NEU4 based on previous inhibitor studies,[25] as well as on previous MD simulations.[53] One reason for this difference in the C9 binding pocket is found in the P267−G275 loop of NEU2, which contains an extra amino acid in comparison to the same loops in NEU3 and NEU4. In Figure 3, a key residue on this loop is shown for each enzyme: Q270 for NEU2, H277 for NEU3, and W274 for NEU4. In NEU2, Q270 forms a H-bond with the triazole of **8b** for 10% of the simulation (Figure 3A, Supporting Information). Furthermore, the triazole is within 7 Å of Q270 for 30% of the simulation (6.4 ± 0.6 Å, average distance between triazole N3 of **8b** and carbonyl oxygen of Q270). To accommodate the large biphenyl-triazole, the P267−G275 loop changes position and moves Q270 in closer contact with the carboxylate of **8b** (5.5 ± 0.8 Å, average distance between carboxylate of **8b** and terminal amide of Q270, as shown in Figure 3A). The comparable residue in NEU3, H277, forms a H-bond with the triazole for 17% of the simulation and is within 7 Å of the triazole for 100% of the time (Figure 3B, 4.7 ± 0.4 Å, average distance between triazole N3 of **8b** and carbonyl oxygen of H277). While W274 in NEU4 does not form any H-bonds with **8b**, it remains within 7 Å of the biphenyl group of **8b** for 100% of the simulation (Figure 3C, 3.8 ± 0.6 Å, average distance between carbonyl oxygen of W274 and the carbon *ortho* to the triazole in **8b**). However, this relatively close contact is not sufficient to prevent rotation of the biphenyl-triazole substituent. The flexibility of **8b** in the active site of NEU4 could be one reason for the higher inhibitory potency for NEU3 over NEU4.

Looking at the IC_{50} values for **1** and **6**, we sought an explanation from our MD simulations for the selectivity the C4 guanidine provides for NEU3 over NEU4. The C4 guanidine of **8b** forms the most highly populated H-bonds between the inhibitor and NEU3. The H-bond to E43 is populated for 92% of the simulation (Figure 3B), whereas the comparable H-bonds in NEU2 and NEU4 are only populated for 47% (to E39, Figure 3A) and 52% (to E41, Figure 3C) of the simulation, respectively (see Supporting Information). Additionally, there is a water bridge between **8b** and D50 of NEU3 that is populated for 88% of the simulation. None of the water bridges in simulations with **8b** and NEU2 or NEU4 are occupied for as much of the simulation time (Supporting Information). Thus, we propose that the guanidine moiety displaces an ordered water from the C4 pocket in the NEU3 active site. In contrast, NEU2 and NEU4 are not able to maintain contacts to this group.

Activity of Inhibitors Against Murine NEU Orthologues. To confirm the selectivity of our active NEU3 inhibitors in vivo against the mouse orthologues, we assayed neuraminidase activity in the tissues of previously described gene-targeted mouse strains deficient in NEU1 (*neu1*$^{-/-}$), NEU3 (*neu3*$^{-/-}$), and NEU4 (*neu4*$^{-/-}$),[54−56] respectively. We also produced a mouse line with a double NEU3/NEU4 deficiency (*neu3*$^{-/-}$ *neu4*$^{-/-}$) by cross-breeding the individual

DOI: 10.1021/acs.jmedchem.7b01574
J. Med. Chem. XXXX, XXX, XXX−XXX

knockouts. As expected, the expression levels of both *neu3* and *neu4* were below the detection limit in the tissues of the *neu3*$^{-/-}$ *neu4*$^{-/-}$ mice.[54]

We assayed neuraminidase activity in the mouse brain tissues, where comparable amounts of NEU1, NEU3, and NEU4 (and only negligible amounts of NEU2) were previously described in the WT mice.[57] Acidic neuraminidase activity in the WT brain homogenate, assayed using the 4MU-NANA substrate, was reduced to ∼85% in the presence of 10 μM **8b** and to ∼60% in the presence of 50 μM **7i**, suggesting that both NEU3 and NEU4 isoenzymes contribute approximately 10−15% of the net brain neuraminidase activity against 4MU-NANA. Further increase of **7i** to 150 μM did not reduce activity, indicating that even at this concentration the compound did not inhibit mouse NEU1 (Figure 4A). In contrast, the residual neuraminidase activity levels measured in the brain tissues of NEU1-deficient mice in the presence of saturating concentrations of **8b** and **7i** were drastically different: **8b** reduced the activity to about ∼60% of that measured in the absence of inhibitors, whereas **7i** completely inhibited the activity (Figure 4D). In the brain of NEU3-deficient mice, only **7i** could reduce the activity by approximately 20%, which corresponded to the NEU4 fraction of the activity, the rest being NEU1 (Figure 4B). Finally, both **8b** and **7i** failed to cause any inhibition of neuraminidase activity in the brain homogenates of NEU3/NEU4 double knockout mice where all activity was presumed to come from the NEU1 isoenzyme (Figure 4C). Together, our data confirm that the **8b** and **7i** can discriminate between isoforms of mammalian neuraminidases in tissues and could be used to target the specific isozymes in vivo.

DISCUSSION

In glycobiology, human neuraminidases constitute an important family of enzymes[58] because they play crucial roles in numerous biological processes through the control of sialoglycoconjugate catabolism. Potent and selective chemical inhibitors represent valuable tools for investigating the biological functions of these enzymes. Viral neuraminidase inhibitors with low nanomolar activities have seen successful clinical application as antiviral drugs.[16] However, there remains a need for strategies to generate selective and potent inhibitors for human neuraminidases.

In this work, we identified selective inhibitors of NEU3 from a small library of DANA analogues with modifications at C4 and C9. On the basis of previously reported structure−activity relationship studies (SAR) and protein structures, we hypothesized that large aromatic groups at C9 could improve potency against NEU3. In our study, C9 modifications not only significantly improved the potency toward NEU3 but also for NEU4. Furthermore, the inhibitory profile of zanamivir **6** toward human neuraminidases encouraged us to also study the effects of C4 modifications. Surprisingly, among all the nitrogen-containing groups we tested, only the guanidino group showed significant improvements. Although compounds bearing a C4-guanidino group and C9-modification only slightly improved potency toward NEU3, this modification substantially decreased the potency toward NEU4. Moreover, compound **8b**, with a biphenyltriazolyl group at C9 and a guandino group at C4, showed slightly improved potency over **7i**, which had only a biphenyltriazolyl group at C9. The selectivity of compound **8a**, which bears a 4-carboxylphenyltriazolyl group at C9 and a guanidino group at C4 for NEU3 may be due to the presence of a carboxylate on the aryl group.

Analysis of the pharmacokinetic properties of our lead compounds **7i** and **8b** suggested that they are unlikely to be orally bioavailable. Permeability assays using artificial membranes (PAMPA) showed that the parent compounds **7i** and **8b**, and most of the C1-ester prodrugs were not sufficiently permeable for oral bioavailability. Only the most lipophilic *n*-butyl ester **7i-OBu** reached this limit with a log $P_e = -6.5 \pm 0.4$. The low solubility of 3−5 μg mL^{-1} further contributes to poor pharmacokinetic properties. Furthermore, cell permeability data for **7i-OBu** suggested that the compound was likely subject to active efflux, which adds to the limitation of its bioavailability. Finally, the examination of the liver microsome stability of **7i-OBu** showed that the ester was only slowly hydrolyzed. In summary, esterification did not improve pharmacokinetic properties sufficiently to allow for oral bioavailability. Future studies should implement modifications of the active principle to reduce the risk of insufficient permeability and solubility.

Molecular modeling studies provided insight into the structural elements responsible for selectivity between NEU3 and NEU4 relative to other isoenzymes. Because only crystal structures for NEU2 were available, we used homology models to study NEU3 and NEU4. Our previous hypothesis of larger C9 pockets available on both NEU3 and NEU4 explained the activity of compounds such as **7i**, which gained substantial activity over DANA (**1**) for both NEU3 and NEU4. Inhibitors of both NEU3 and NEU4 could be useful agents as these enzymes are known to act on similar substrates and may compensate for each other in vivo.[59] The selectivity over NEU1 and NEU2 can be attributed to the limited size of their C9 pockets. Furthermore, our models provided a rationale for the observed selectivity of inhibitors with a C4-guanidino group (**6**, **8a**, **8b**) for NEU3. The inclusion of the C4-guanidino group confers a notable improvement in activity against NEU3, and our data suggested this could be due to the displacement of an ordered water in the active site.[60] The absence of a comparable ordered water molecule prevented these compounds from gaining similar activity against NEU4. In general, the C4-guanidino modification reduced NEU4 activity relative to the parent structures (compare **1** and **6**; **7h** and **8a**; **7i** and **8b**). This finding provides an important design principle for future NEU3 inhibitor design.

CONCLUSION

In this study, we demonstrated that the scaffold of the neuraminidase inhibitor DANA **1** can be used as a starting point for the generation of new potent and selective inhibitors of human neuraminidase isoenzymes. First, aryltriazolyl groups at C9 greatly improved the potency of inhibitors toward NEU3 and NEU4, and second, the inclusion of a guanidino group at C4 provided selectivity between NEU3 and NEU4, revealing a clear strategy for the design of inhibitors with improved selectivity for NEU3 with submicromolar potencies. These inhibitors may be adapted as useful tools to study the biological function of human neuraminidases. In summary, our data confirmed that divergent structural features of the active site of human neuraminidase isoenzymes allowed identification of selective and potent inhibitors of this important family of human glycosidases. Future studies will address the activity of these compounds in vivo and provide new tools for dissecting the specific biological roles of NEU isoenzymes.

DOI: 10.1021/acs.jmedchem.7b01574
J. Med. Chem. XXXX, XXX, XXX−XXX

EXPERIMENTAL SECTION

General Synthetic Procedures. All reagents and solvents were purchased from Sigma-Aldrich unless otherwise noted and used without further purification. Reactions were monitored with TLC (Merck TLC Silica gel 60 F_{254}), and spots were visualized under UV light (254 nm) or by charring with 0.5% H_2SO_4/EtOH. Compounds were purified by flash column chromatography with silica gel (SiliCycle SiliaFlash F60, 40−63 μm particle size) or recrystallization with solvents specified in the corresponding experiments. Proton (^{1}H) and carbon (^{13}C) NMR spectra were recorded on Varian 400 (400 MHz for ^{1}H; 100 MHz for ^{13}C), Varian 500 (500 MHz for ^{1}H; 125 MHz for ^{13}C), or Varian 700 (700 MHz for ^{1}H; 175 MHz for ^{13}C). High-resolution mass spectrometry (HR-MS) analysis was performed on Agilent Technologies 6220 TOF spectrometer. Purity of all final products used for inhibitor assays and pharmacokinetic studies was determined to be ≥95% by HPLC or LC-MS (see Supporting Information for details).

*5-Acetamido-2,6-anhydro-3,5-dideoxy-D-glycero-D-galacto-non-2-enonic Acid (DANA, **1**).* Compound **1** was synthesized as previously reported.[25] ^{1}H NMR (500 MHz, CD_3OD) δ 5.67 (d, *J* = 2.3 Hz, 1H, H-3), 4.36 (dd, *J* = 8.6, 2.3 Hz, 1H, H-4), 4.10 (dd, *J* = 10.9, 1.1 Hz, 1H, H-6), 3.99 (dd, *J* = 10.9, 8.6 Hz, 1H, H-5), 3.87 (ddd, *J* = 9.1, 5.4, 3.1 Hz, 1H, H-8), 3.80 (dd, *J* = 11.4, 3.1 Hz, 1H, H-9), 3.65 (dd, *J* = 11.4, 5.4 Hz, 1H, H-9′), 3.52 (dd, *J* = 9.1, 1.1 Hz, 1H, H-7), 2.02 (s, 3H, $COCH_3$). ^{13}C NMR (125 MHz, CD_3OD) δ 174.68, 170.02 (2 × C═O), 149.95 (C-2), 108.34 (C-3), 77.24 (C-6), 71.29 (C-8), 70.22 (C-7), 68.70 (C-4), 64.94 (C-9), 51.96 (C-5), 22.82 ($COCH_3$). HR-MS (ESI): calcd for $C_{11}H_{16}NO_8$ $[M - H]^-$ 290.0876, found 290.0879

*Methyl-5-acetamido-2,6-anhydro-3,5-dideoxy-D-glycero-D-galacto-non-2-enonate (DANA-OMe, **1**-OMe).* Compound **1**-OMe was synthesized as previously reported.[25] ^{1}H NMR (500 MHz, D_2O) δ 6.13 (d, *J* = 2.4 Hz, 1H, H-3), 4.60 (dd, *J* = 9.0, 2.4 Hz, 1H, H-4), 4.36 (d, *J* = 10.9 Hz, 1H, H-6), 4.17 (dd, *J* = 10.8, 9.0 Hz, 1H, H-5), 4.00 (ddd, *J* = 9.0, 6.0, 2.7 Hz, 1H, H-8), 3.95 (dd, *J* = 11.9, 2.7 Hz, 1H, H-9), 3.90 (s, 3H, $COOCH_3$), 3.77−3.71 (m, 2H, H-7, H-9), 2.15 (s, 3H, $COCH_3$). ^{13}C NMR (125 MHz, D_2O) δ 175.78, 165.30 (2 × C═O), 144.37 (C-2), 113.54 (C-3), 77.09 (C-6), 70.84 (C-8), 68.92 (C-7), 68.02 (C-4), 63.99 (C-9), 53.91 ($COOCH_3$), 50.59 (C-5), 23.07 ($COCH_3$). HR-MS (ESI): calcd for $C_{12}H_{19}NNaO_8$ $[M + Na]^+$ 328.1008, found 328.1006.

*5-Acetamido-9-pentanamido-2,6-anhydro-3,5-dideoxy-D-glycero-D-galacto-non-2-enonic Acid (C9-BA-DANA, **2**).* Compound **2** was synthesized as previously reported.[22] ^{1}H NMR (500 MHz, CD_3OD) δ 5.93 (d, *J* = 1.7 Hz, 1H, H-3), 4.43 (dd, *J* = 8.6, 1.7 Hz, 1H, H-4), 4.18 (d, *J* = 10.8 Hz, 1H, H-6), 4.02−3.95 (m, 1H, H-5), 3.95−3.87 (m, 1H, H-8), 3.59 (dd, *J* = 13.9, 2.9 Hz, 1H, H-9), 3.43 (d, *J* = 9.0 Hz, 1H, H-7), 3.32−3.27 (m, 1H, H-9′), 2.22 (t, *J* = 7.6 Hz, 2H, α-CH_2), 2.03 (s, 3H, $COCH_3$), 1.64−1.54 (m, 2H, β-CH_2), 1.35 (dd, *J* = 15.0, 7.5 Hz, 2H, γ-CH_2), 0.92 (t, *J* = 7.4 Hz, 3H, δ-CH_3). ^{13}C NMR (126 MHz, CD_3OD) δ 177.26, 174.88 (N−C═O), 165.81 (C-1), 145.81 (C-2), 113.17 (C-3), 77.79 (C-6), 71.50 (C-7), 70.24 (C-4), 68.03 (C-8), 51.97 (C-5), 44.40 (C-9), 36.87 (C-α), 29.28 (C-β), 23.44 (C-γ), 22.86 ($COCH_3$), 14.22 (C-δ). HRMS (ESI): calcd for $C_{16}H_{25}N_2O_8$ $[M - H]^-$ 373.1616, found 373.1614.

*Methyl-5-acetamido-9-pentanamido-2,6-anhydro-3,5-dideoxy-D-glycero-D-galacto-non-2-enonate (C9-BA-DANA-OMe, **2**-OMe).* Compound **2**-OMe was synthesized as previously reported.[22] ^{1}H NMR (500 MHz, CD_3OD) δ 5.93 (d, *J* = 2.4 Hz, 1H, H-3), 4.43 (dd, *J* = 8.8, 2.4 Hz, 1H, H-4), 4.19 (d, *J* = 10.6 Hz, 1H, H-6), 3.98 (dd, *J* = 10.6, 8.8 Hz, 1H, H-5), 3.94−3.89 (m, 1H, H-8), 3.77 (s, 3H, $COOCH_3$), 3.59 (dd, *J* = 14.0, 3.1 Hz, 1H, H-9), 3.42 (d, *J* = 9.1 Hz, 1H, H-7), 3.31 (m, 1H, H-9′), 2.23 (t, *J* = 7.6 Hz, 2H, α-CH_2), 2.03 (s, 3H, $COCH_3$), 1.64−1.54 (m, 2H, β-CH_2), 1.40−1.30 (m, 2H, γ-CH_2), 0.93 (t, *J* = 7.4 Hz, 3H, δ-CH_3). ^{13}C NMR (125 MHz, CD_3OD) δ 177.31, 174.90, 164.34 (3 × C═O), 145.18 (C-2), 113.65 (C-3), 77.97 (C-6), 71.50 (C-7), 70.19 (C-4), 67.90 (C-8), 52.90 ($COOCH_3$), 52.01 (C-5), 44.49 (C-9), 36.86 (C-α), 29.27 (C-β), 23.44 (C-γ), 22.85 ($COCH_3$), 14.24 (C-δ). HRMS (ESI): calcd for $C_{17}H_{29}N_2O_8$ $[M + H]^+$ 389.1924, found 389.1909.

*5-Acetamido-2,6-anhydro-4-guanidino-3,4,5-trideoxy-D-glycero-D-galacto-non-2-enonic Acid (Zanamivir, **6**).* Compound **6** was synthesized as previously reported.[61,62] ^{1}H NMR (500 MHz, D_2O) δ 5.70 (d, *J* = 1.9 Hz, 1H, H-3), 4.54 (dd, *J* = 9.3, 1.9 Hz, 1H, H-4), 4.46 (m, 1H, H-6), 4.29 (dd, *J* = 10.5, 9.3 Hz, 1H, H-5), 4.02 (ddd, *J* = 9.1, 6.2, 2.5 Hz, 1H, H-8), 3.96 (dd, *J* = 11.9, 2.5 Hz, 1H, H-9), 3.77−3.69 (m, 2H, H-7, H-9′), 2.11 (s, 3H, $COCH_3$). ^{13}C NMR (125 MHz, D_2O) δ 175.38, 170.10 (2 × C═O), 157.99 (C═N), 150.19 (C-2), 104.79 (C-3), 76.33 (C-6), 70.74 (C-8), 69.11 (C-7), 64.03 (C-9), 52.11 (C-4), 48.71 (C-5), 22.93 ($COCH_3$). HR-MS (ESI): calcd for $C_{12}H_{21}N_4O_7$ $[M + H]^+$ 333.1405, found 333.1400

General Procedure of CuAAC Reaction and Hydrolysis of Methyl Ester. To a solution of methyl 5-acetamido-9-azido-2,6-anhydro-3,5-dideoxy-D-glycero-D-galacto-non-2-enonate (**9**)[25] (1 equiv) and the corresponding alkyne (1.5 equiv) in THF−H_2O (2:1), sodium L-ascorbate (0.5 equiv) and copper(II) sulfate (0.5 equiv) were added sequentially. The reaction mixture was kept stirring at room temperature and monitored by TLC until no azide remained. Silica gel was then added to the reaction mixture, and the solvent was removed under reduced pressure. The residue was separated by flash chromatography to provide the desired products with yields of 42%−88%. To hydrolyze the C1-methyl ester, the product was dissolved in MeOH, and 0.5 M NaOH was added. The mixture was kept stirring at room temperature. After completion, the pH of the solution was adjusted to 2 with Amberlite IR-120 (H^+). The solution was filtered, concentrated, and purified by flash chromatography or recrystallization to provide the desired products with yields of 45%−88%.

*5-Acetamido-9-(4-(dimethylamino)phenyl)-2,6-anhydro-3,5-dideoxy-D-glycero-D-galacto-non-2-enonic Acid (**7a**).* Compound **7a** was prepared as described above in 62% yield (two steps, 86 mg). ^{1}H NMR (500 MHz, CD_3OD) δ 8.11 (s, 1H, triazole-H), 7.64 (d, *J* = 8.9 Hz, 2H, Ar-H), 6.81 (d, *J* = 8.9 Hz, 2H, Ar-H), 5.90 (d, *J* = 2.3 Hz, 1H, H-3), 4.49 (dd, *J* = 14.0, 7.5 Hz, 1H, H-9′), 4.40 (dd, *J* = 8.7, 2.3 Hz, 1H, H-4), 4.31−4.25 (m, 1H, H-8), 4.14 (dd, *J* = 10.9, 1.0 Hz, 1H, H-6), 3.99 (dd, *J* = 10.9, 8.7 Hz, 1H, H-5), 3.41 (dd, *J* = 9.2, 1.0 Hz, 1H, H-7), 2.96 (s, 6H, $N(CH_3)_2$), 1.99 (s, 3H, $COCH_3$). ^{13}C NMR (125 MHz, CD_3OD) δ 175.16 (C═O), 152.12, 127.59, 120.10, 113.93 (Ar-C), 149.18 (triazole-C4), 121.83 (triazole-C5), 77.67 (C-6), 71.23 (C-7), 69.79 (C-4), 68.01 (C-8), 55.13 (C-9), 51.96 (C-5), 40.78 (N-CH_3), 22.65 ($COCH_3$). HR-MS (ESI): calcd for $C_{21}H_{26}N_5O_7$ $[M - H]^-$ 460.1832, found 460.1834.

*5-Acetamido-9-(4-acetamidophenyl)-2,6-anhydro-3,5-dideoxy-D-glycero-D-galacto-non-2-enonic Acid (**7b**).* Compound **7b** was prepared as above in 65% yield (two steps, 90 mg). ^{1}H NMR (500 MHz, CD_3OD) δ 8.28 (s, 1H, triazole-H), 7.76 (d, *J* = 8.2 Hz, 2H, Ar-H), 7.62 (d, *J* = 8.2 Hz, 2H, Ar-H), 5.70 (s, 1H, H-3), 4.50 (dd, *J* = 13.5, 7.7 Hz, 1H, H-9′), 4.35 (d, *J* = 8.6 Hz, 1H, H-4), 4.30−4.27 (m, 1H, H-8), 4.10 (d, *J* = 10.7 Hz, 1H, H-6), 4.04−3.96 (m, 1H, H-5), 3.39 (d, *J* = 7.7 Hz, 1H, H-7), 2.13, 1.98 (2 × s, 3H, 2 × $COCH_3$). ^{13}C NMR (125 MHz, DDMSO-d_6) δ 172.10, 168.34, 168.25, 165.18 (4 × C═O), 147.63 (C-2), 145.74 (triazole-C4), 121.67 (triazole-C5), 138.80, 138.69, 125.72, 125.44, 119.20, 119.11 (Ar-C), 108.60 (C-3), 75.66 (C-6), 69.95 (C-7), 68.10 (C-4), 65.90 (C-8), 53.70(C-9), 50.81 (C-5), 22.97, 22.49 (2 × $COCH_3$). HR-MS (ESI): calcd for $C_{21}H_{24}N_5O_8$ $[M - H]^-$ 474.1625, found 474.1636.

*5-Acetamido-9-(4-amidophenyl)-2,6-anhydro-3,5-dideoxy-D-glycero-D-galacto-non-2-enonic Acid (**7c**).* Compound **7c** was prepared as above in 63% yield (two steps, 80 mg). ^{1}H NMR (500 MHz, CD_3OD) δ 8.09 (s, 1H, triazole-H), 7.55 (d, *J* = 8.6 Hz, 2H, Ar-H), 6.79 (d, *J* = 8.6 Hz, 2H, Ar-H), 5.88 (d, *J* = 2.4 Hz, 1H, H-3), 4.82 (dd, *J* = 14.1, 2.6 Hz, 1H, H-9), 4.48 (dd, *J* = 14.1, 7.6 Hz, 1H, H-9′), 4.41 (dd, *J* = 8.7, 2.4 Hz, 1H, H-4), 4.28 (ddd, *J* = 9.5, 7.6, 2.6 Hz, 1H, H-8), 4.14 (dd, *J* = 10.9, 1.0 Hz, 1H, H-6), 4.00 (dd, *J* = 10.9, 8.7 Hz, 1H, H-5), 3.41 (dd, *J* = 9.5, 1.0 Hz, 1H, H-7), 1.98 (s, 3H, $COCH_3$). ^{13}C NMR (125 MHz, CD_3OD) δ 175.13, 166.94 (2 × C═O), 149.13, 148.42 (Ar-C, triazole-C4), 146.69 (C-2), 127.77, 116.96 (Ar-C), 122.01, 121.94 (Ar-C, triazole-C5), 112.19 (C-3), 77.57 (C-6), 71.24 (C-7), 69.81 (C-4), 68.08 (C-8), 55.14 (C-9), 51.93 (C-5), 22.72 ($COCH_3$). HR-MS (ESI): calcd for $C_{19}H_{22}N_5O_7$ $[M - H]^-$ 432.1529, found 432.1513.

DOI: 10.1021/acs.jmedchem.7b01574
J. Med. Chem. XXXX, XXX, XXX−XXX

*5-Acetamido-9-(4-methylphenyl)-2,6-anhydro-3,5-dideoxy-D-glycero-D-galacto-non-2-enonic Acid (**7d**).* Compound **7d** was prepared as above in 75% yield (two steps, 100 mg). ^{1}H NMR (500 MHz, CD_3OD) δ 8.24 (s, 1H, triazole-H), 7.68 (d, J = 8.0 Hz, 2H, Ar-H), 7.22 (d, J = 8.0 Hz, 2H, Ar-H), 5.79 (s, 1H, H-3), 4.50 (dd, J = 13.9, 7.4 Hz, 1H, H-9′), 4.39 (d, J = 8.2 Hz, 1H, H-4), 4.29 (brs, 1H, H-8), 4.13 (d, J = 10.9 Hz, 1H, H-6), 4.06−3.97 (m, 1H, H-5), 3.42 (d, J = 9.0 Hz, 1H, H-7), 2.34 (s, 3H, $PhCH_3$), 1.98 (s, 3H, $COCH_3$). ^{13}C NMR (125 MHz, CD_3OD) δ 175.00 (C═O), 148.64 (triazole-C4), 123.08 (triazole-C5), 139.31, 130.58, 129.01, 126.62 (Ar-C), 110.40 (C-3), 77.27 (C-6), 71.24 (C-7), 69.84 (C-4), 68.35 (C-8), 55.16 (C-9), 51.95 (C-5), 22.76 ($COCH_3$), 21.31 ($PhCH_3$). HR-MS (ESI): calcd for $C_{20}H_{23}N_4O_7$ $[M - H]^-$ 431.1567, found 431.1568.

*5-Acetamido-9-(4-methoxyphenyl)-2,6-anhydro-3,5-dideoxy-D-glycero-D-galacto-non-2-enonic Acid (**7e**).* Compound **7e** was prepared as above in 59% yield (two steps, 80 mg). ^{1}H NMR (500 MHz, CD_3OD) δ 8.20 (s, 1H, triazole-H), 7.72 (d, J = 8.9 Hz, 2H, Ar-H), 6.97 (d, J = 8.9 Hz, 2H, Ar-H), 5.77 (d, J = 2.0 Hz, 1H, H-3), 4.49 (dd, J = 14.0, 7.6 Hz, 1H, H-9′), 4.39 (dd, J = 8.7, 2.0 Hz, 1H, H-4), 4.32−4.25 (m, 1H, H-8), 4.12 (d, J = 10.8 Hz, 1H, H-6), 4.00 (dd, J = 10.8, 8.7 Hz, 1H, H-5), 3.81 (s, 3H, $PhOCH_3$), 3.41 (d, J = 9.2 Hz, 1H, H-7), 1.98 (s, 3H, $COCH_3$). ^{13}C NMR (125 MHz, CD_3OD) δ 174.96 ($COCH_3$), 161.27, 128.00, 124.40, 115.37 (Ar-C), 148.49 (triazole-C4), 122.59 (triazole-C5), 109.95 (C-3), 77.20 (C-6), 71.26 (C-7), 69.86 (C-4), 68.38 (C-8), 55.80 ($PhOCH_3$), 55.15 (C-9), 51.97 (C-5), 22.74 ($COCH_3$). HR-MS (ESI): calcd for $C_{20}H_{23}N_4O_8$ $[M - H]^-$ 447.1516, found 447.1527.

*5-Acetamido-9-(4-fluorophenyl)-2,6-anhydro-3,5-dideoxy-D-glycero-D-galacto-non-2-enonic Acid (**7f**).* Compound **7f** was prepared as above in 62% yield (two steps, 80 mg). ^{1}H NMR (500 MHz, CD_3OD) δ 8.28 (s, 1H, triazole-H), 7.87−7.80 (m, 2H, Ar-H), 7.20−7.13 (m, 2H, Ar-H), 5.92 (d, J = 2.4 Hz, 1H, H-3), 4.86 (dd, J = 14.0, 2.6 Hz, 1H, H-9), 4.51 (dd, J = 14.0, 7.7 Hz, 1H, H-9′), 4.41 (dd, J = 8.7, 2.4 Hz, 1H, H-4), 4.29 (ddd, J = 9.6, 7.7, 2.6 Hz, 1H, H-8), 4.14 (dd, J = 10.8, 1.1 Hz, 1H, H-6), 3.99 (dd, J = 10.8, 8.7 Hz, 1H, H-5), 3.43 (dd, J = 9.6, 1.1 Hz, 1H, H-7), 1.99 (s, 3H, $COCH_3$). ^{13}C NMR (125 MHz, CD_3OD) δ 175.15 (C═O), 164.09 (d, J = 245.9 Hz, Ar-C), 147.63 (triazole-C4), 123.31 (triazole-C5), 128.61 (d, J = 8.2 Hz, Ar-C), 128.34 (d, J = 3.2 Hz, Ar-C), 116.77 (d, J = 22.0 Hz, Ar-C), 112.86 (C-3), 77.70 (C-6), 71.30 (C-7), 69.80 (C-4), 67.95 (C-8), 55.25 (C-9), 51.96 (C-5), 22.64 ($COCH_3$). HR-MS (ESI): calcd for $C_{19}H_{20}FN_4O_7$ $[M - H]^-$ 435.1316, found 435.1324.

*5-Acetamido-9-(4-(trifluoromethyl)phenyl)-2,6-anhydro-3,5-dideoxy-D-glycero-D-galacto-non-2-enonic Acid (**7g**).* Compound **7g** was prepared as above in 19% yield (two steps, 40 mg). ^{1}H NMR (500 MHz, CD_3OD) δ 8.44 (s, 1H, triazole-H), 8.02 (d, J = 8.2 Hz, 2H, Ar-H), 7.72 (d, J = 8.2 Hz, 2H, Ar-H), 5.95 (d, J = 2.4 Hz, 1H, H-3), 4.90 (dd, J = 14.0, 2.6 Hz, 1H, H-9), 4.53 (dd, J = 14.0, 7.7 Hz, 1H, H-9′), 4.42 (dd, J = 8.7, 2.4 Hz, 1H, H-4), 4.37−4.21 (m, 1H, H-8), 4.16 (dd, J = 10.8, 1.0 Hz, 1H, H-6), 4.00 (dd, J = 10.8, 8.7 Hz, 1H, H-5), 3.45 (dd, J = 9.1, 1.0 Hz, 1H, H-7), 1.99 (s, 3H, $COCH_3$). ^{13}C NMR (125 MHz, CD_3OD) δ 175.20 (C═O), 147.04 (triazole-C4), 124.50 (triazole-C5), 135.81 (Ar-C), 130.90 (q, J = 32.3 Hz, Ar-C), 127.01 (Ar-C), 126.89 (q, J = 3.8 Hz, Ar-C), 113.43 (C-3), 77.81 (C-6), 71.35 (C-7), 69.80 (C-4), 67.88 (C-8), 55.34 (C-9), 51.95 (C-9), 22.65 ($COCH_3$). HR-MS (ESI): calcd for $C_{20}H_{20}F_3N_4O_7$ $[M - H]^-$ 485.1284, found 485.1282.

*5-Acetamido-9-(4-carboxyphenyl)-2,6-anhydro-3,5-dideoxy-D-glycero-D-galacto-non-2-enonic Acid (**7h**).* Compound **7h** was prepared as above in 74% yield (two steps, 100 mg). ^{1}H NMR (500 MHz, CD_3OD) δ 8.42 (s, 1H, triazole-H), 8.06 (d, J = 8.5 Hz, 2H, Ar-H), 7.90 (d, J = 8.5 Hz, 2H, Ar-H), 5.84 (d, J = 1.9 Hz, 1H, H-3), 4.88 (dd, J = 14.0, 2.1 Hz, 1H, H-9), 4.54 (dd, J = 14.0, 7.6 Hz, 1H, H-9′), 4.43 (dd, J = 8.7, 1.9 Hz, 1H, H-4), 4.31 (t, J = 7.6 Hz, 1H, H-8), 4.15 (d, J = 10.9 Hz, 1H, H-6), 4.03 (dd, J = 10.9, 8.7 Hz, 1H, H-5), 3.44 (d, J = 7.6 Hz, 1H, H-7), 1.99 (s, 3H, $COCH_3$). ^{13}C NMR (125 MHz, CD_3OD) δ 175.08, 169.69, 167.90 (3 × C═O), 147.50 (triazole-C4), 124.53 (triazole-C5), 136.19, 131.49, 126.44 (Ar-C), 111.19 (C-3), 77.37 (C-6), 71.31 (C-7), 69.84 (C-4), 68.24 (C-8), 55.29 (C-9), 51.90 (C-5), 22.82 ($COCH_3$). HR-MS (ESI): calcd for $C_{20}H_{21}N_4O_9$ $[M - H]^-$ 461.1309, found 461.1316.

*Methyl-5-acetamido-9-(4-biphenyl)-2,6-anhydro-3,5-dideoxy-D-glycero-D-galacto-non-2-enonate Acid (**7i-OMe**).* C1-Methyl ester of **7i** was obtained via CuAAC as described. 69% (71 mg). ^{1}H NMR (500 MHz, DMSO-d_6) δ 8.46 (s, 1H, triazole-H), 8.19 (d, J = 7.4 Hz, 1H, Ar-H), 7.93 (d, J = 7.4 Hz, 2H, Ar-H), 7.74 (d, J = 7.6 Hz, 2H, Ar-H), 7.70 (d, J = 7.0 Hz, 2H, Ar-H), 7.49−7.42 (m, 2H, Ar-H), 7.38−7.32 (m, 1H, Ar-H), 5.81 (s, 1H, H-3), 5.41 (d, J = 5.4 Hz, 1H, OH), 5.36 (d, J = 4.7 Hz, 1H, OH), 5.10 (s, 1H, NH), 4.77 (d, J = 13.4 Hz, 1H, H-9), 4.38−4.25 (m, 2H, H-9′, H-4), 4.09−3.95 (m, 2H, H-8, H-6), 3.80−3.72 (m, 1H, H-7), 3.69 (s, 3H, $COOCH_3$), 1.90 (s, 3H, $COCH_3$). ^{13}C NMR (126 MHz, DMSO-d_6) δ 172.06, 162.12 (2 × C═O), 145.50 (triazole-C4), 143.15 (C-2), 139.56, 139.21, 130.05, 128.92, 127.47, 127.04, 126.45, 125.59 (Ar-C), 122.56 (triazole-C5), 113.36 (C-3), 76.40 (C-6), 69.85 (C-7), 68.00 (C-4), 65.47 (C-8), 53.82 (C-9), 52.06 (α-CH_2), 50.26 (C-5), 22.52 ($COCH_3$). HR-MS (ESI): calcd for $C_{26}H_{28}N_4O_7$ $[M - H]^-$ 509.2031, found 509.2034.

General Procedure of Transesterification for Elongated Esters of 7i. Compound **7i-OMe** was dissolved in corresponding alcohol, then LiBr (5 equiv) and DBU (0.5 equiv) was added. After stirring at room temperature overnight, solvents were removed under reduced pressure and the residue was purified by flash chromatography to give the desired products (43−62%).

*Ethyl-5-acetamido-9-(4-biphenyl)-2,6-anhydro-3,5-dideoxy-D-glycero-D-galacto-non-2-enonate (**7i-OEt**).* Yield 10 mg (62%). ^{1}H NMR (700 MHz, DMSO-d_6) δ 8.48 (s, 1H, triazole-H), 8.22 (d, J = 8.1 Hz, 1H, Ar-H), 7.95 (d, J = 8.2 Hz, 2H, Ar-H), 7.76 (d, J = 8.2 Hz, 2H, Ar-H), 7.72 (d, J = 7.7 Hz, 2H, Ar-H), 7.48 (t, J = 7.6 Hz, 2H, Ar-H), 7.37 (t, J = 7.4 Hz, 1H, Ar-H), 5.82 (d, J = 2.4 Hz, 1H, H-3), 5.42 (d, J = 6.2 Hz, 1H, OH), 5.38 (d, J = 6.5 Hz, 1H, OH), 5.12 (d, J = 4.8 Hz, 1H, NH), 4.80 (dd, J = 13.9, 2.3 Hz, 1H, H-9), 4.38−4.31 (m, 2H, H-9′, H-4), 4.19−4.13 (m, 2H, α-CH_2), 4.06 (ddd, J = 14.8, 8.7, 2.6 Hz, 1H, H-8), 3.99 (d, J = 10.9 Hz, 1H, H-6), 3.79−3.73 (m, 1H, H-5), 3.35 (dd, J = 8.7, 5.1 Hz, 1H, H-7), 1.92 (s, 3H, $COOCH_3$), 1.21 (t, J = 7.1 Hz, 3H, β-CH_3). ^{13}C NMR (175 MHz, DMSO-d_6) δ 172.09, 161.55 (2 × C═O), 145.49 (triazole-C4), 143.26 (C-2), 139.57, 139.21, 130.08, 128.94, 127.48, 127.06, 126.46, 125.59 (Ar-C), 122.55 (triazole-C5), 113.18 (C-3), 76.47 (C-6), 69.95 (C-7), 68.13 (C-4), 65.49 (C-8), 60.82 (α-CH_2), 53.77 (C-9), 50.31 (C-5), 22.53 ($COCH_3$), 13.98 (β-CH_3). HR-MS (ESI): calcd for $C_{27}H_{31}N_4O_7$ $[M - H]^-$ 523.2187, found 523.2186.

*Propyl-5-acetamido-9-(4-biphenyl)-2,6-anhydro-3,5-dideoxy-D-glycero-D-galacto-non-2-enonate (**7i-OPr**).* Yield 10 mg (63%). ^{1}H NMR (700 MHz, DMSO-d_6) δ 8.47 (d, J = 5.6 Hz, 1H, triazole-H), 8.22 (d, J = 8.2 Hz, 1H, Ar-H), 7.95 (d, J = 8.2 Hz, 2H, Ar-H), 7.76 (d, J = 8.2 Hz, 2H, Ar-H), 7.72 (d, J = 7.5 Hz, 2H, Ar-H), 7.48 (t, J = 7.7 Hz, 2H, Ar-H), 7.37 (t, J = 7.4 Hz, 1H, Ar-H), 5.83 (d, J = 2.3 Hz, 1H, H-3), 5.41 (d, J = 6.2 Hz, 1H, OH), 5.38 (d, J = 6.5 Hz, 1H, OH), 5.12 (d, J = 4.7 Hz, 1H, NH), 4.80 (dd, J = 13.9, 2.1 Hz, 1H, H-9), 4.39−4.31 (m, 2H, H-9′, H-4), 4.10−4.04 (m, 3H, α-CH_2, H-8), 3.99 (d, J = 10.8 Hz, 1H, H-6), 3.77 (dd, J = 10.8, 8.6 Hz, 1H, H-5), 1.92 (s, 3H, $COOCH_3$), 1.61 (dd, J = 14.1, 7.2 Hz, 2H, β-CH_2), 0.88 (t, J = 7.2 Hz, 3H, γ-CH_3). ^{13}C NMR (175 MHz, DMSO-d_6) δ 172.09, 161.59 (2 × C═O), 145.49 (triazole-C4), 143.23 (C-2), 139.57, 139.21, 130.08, 128.94, 127.48, 127.06, 126.46, 125.59 (Ar-C), 122.49 (triazole-C5), 113.14 (C-3), 76.47 (C-6), 69.96 (C-7), 68.20 (C-4), 66.14 (C-8), 65.51 (α-CH_2), 53.76 (C-9), 50.31 (C-5), 22.53 (β-CH_2), 21.42 ($COCH_3$), 10.19 (γ-CH_3). HR-MS (ESI): calcd for $C_{28}H_{33}N_4O_7$ $[M - H]^-$ 537.2344, found 537.2345.

*Butyl-5-acetamido-9-(4-biphenyl)-2,6-anhydro-3,5-dideoxy-D-glycero-D-galacto-non-2-enonate (**7i-OBu**).* Yield 7 mg (43%). ^{1}H NMR (700 MHz, DMSO-d_6) δ 8.47 (s, 1H, triazole-H), 8.22 (d, J = 8.2 Hz, 1H, Ar-H), 7.97−7.93 (m, 2H, Ar-H), 7.78−7.74 (m, 2H, Ar-H), 7.73−7.70 (m, 2H, Ar-H), 7.49−7.46 (m, 2H, Ar-H), 7.39−7.35 (m, 1H, Ar-H), 5.82 (d, J = 2.5 Hz, 1H, H-3), 5.40 (d, J = 6.2 Hz, 1H, OH), 5.38 (d, J = 6.5 Hz, 1H, OH), 5.12 (d, J = 4.8 Hz, 1H, NH), 4.81 (dd, J = 13.9, 2.5 Hz, 1H, H-9), 4.39−4.31 (m, 2H, H-9′, H-4), 4.11 (td, J = 6.5, 1.1 Hz, 2H, α-CH_2), 4.07 (tdd, J = 8.7, 6.3, 2.5 Hz, 1H, H-8), 3.99 (d, J = 10.6 Hz, 1H, H-6), 3.77 (dt, J = 10.6, 8.5 Hz, 1H, H-5),

DOI: 10.1021/acs.jmedchem.7b01574
J. Med. Chem. XXXX, XXX, XXX−XXX

3.36 (dd, J = 8.6, 5.8 Hz, 1H, H-7), 1.92 (s, 3H, $COOCH_3$), 1.60−1.55 (m, 2H, β-CH_2), 1.32 (dq, J = 14.7, 7.4 Hz, 2H, γ-CH_2), 0.86 (t, J = 7.4 Hz, 3H, δ-CH_3). ^{13}C NMR (175 MHz, DMSO-d_6) δ 172.10, 161.58 (2 × C═O), 145.49 (triazole-C4), 143.23 (C-2), 139.84, 139.57, 139.20, 130.09, 128.94, 127.48, 127.05, 126.46, 125.59 (Ar-C), 122.48 (triazole-C5), 113.15 (C-3), 76.50 (C-6), 69.99 (C-7), 68.27 (C-4), 65.49 (C-8), 64.42 (α-CH_2), 53.74 (C-9), 50.32 (C-5), 30.03 (β-CH_2), 22.53 ($COCH_3$), 18.59 (γ-CH_2), 13.48 (δ-CH_3). HR-MS (ESI): calcd for $C_{29}H_{35}N_4O_7$ $[M - H]^-$ 551.2500, found 551.2509.

*5-Acetamido-9-(4-biphenyl)-2,6-anhydro-3,5-dideoxy-D-glycero-D-galacto-non-2-enonic Acid (**7i**).* Compound **7i** was obtained after hydrolysis of **7i-OMe** as described above in 75% (52 mg). ^{1}H NMR (500 MHz, CD_3OD) δ 8.34 (s, 1H, triazole-H), 7.88 (d, J = 8.1 Hz, 2H, Ar-H), 7.67 (d, J = 8.1 Hz, 2H, Ar-H), 7.62 (d, J = 7.8 Hz, 2H, Ar-H), 7.42 (t, J = 7.6 Hz, 2H, Ar-H), 7.32 (t, J = 7.4 Hz, 1H, Ar-H), 5.74 (d, J = 1.9 Hz, 1H, H-3), 4.88 (dd, J = 13.9, 2.2 Hz, 1H, H-9), 4.52 (dd, J = 13.9, 7.7 Hz, H-9′), 4.39 (dd, J = 8.7, 1.9 Hz, H-4), 4.37−4.21 (m, 1H, H-8), 4.13 (d, J = 10.8 Hz, 1H, H-6), 4.02 (dd, J = 10.8, 8.7 Hz, 1H, H-5), 3.42 (d, J = 9.3 Hz, 1H, H-7), 2.00 (s, 3H, $COCH_3$). ^{13}C NMR (125 MHz, CD_3OD) δ 174.93, 169.51 (2 × C═O), 149.28 (triazole-C4), 123.47 (triazole-C5), 148.20 (C-2), 142.24, 141.76, 130.84, 129.94, 128.54, 128.48, 127.86, 127.10 (Ar-C), 77.14 (C-6), 71.34 (C-7), 69.83 (C-4), 68.49 (C-8), 55.25 (C-9), 52.00 (C-5), 22.77 ($COCH_3$). HR-MS (ESI): calcd for $C_{25}H_{25}N_4O_7$ $[M - H]^-$ 493.1723, found 493.1729.

*5-Acetamido-9-(4-phenoxyphenyl)-2,6-anhydro-3,5-dideoxy-D-glycero-D-galacto-non-2-enonic Acid (**7j**).* Compound **7j** was prepared as above in 40% (two steps, 40 mg). ^{1}H NMR (500 MHz, CD_3OD) δ 8.24 (s, 1H, triazole-H), 7.78 (d, J = 8.7 Hz, 2H, Ar-H), 7.35 (dd, J = 8.7, 7.4 Hz, 2H, Ar-H), 7.12 (t, J = 7.4 Hz, 1H, Ar-H), 7.05−6.98 (m, 4H, Ar-H), 5.93 (d, J = 2.4 Hz, 1H, H-3), 4.85 (dd, J = 14.0, 2.5 Hz, 1H, H-9), 4.50 (dd, J = 14.0, 7.7 Hz, 1H, H-9′), 4.42 (dd, J = 8.7, 2.4 Hz, 1H, H-4), 4.30 (ddd, J = 10.0, 7.7, 2.5 Hz, 1H, H-8), 4.16 (m, 1H, H-6), 4.01 (dd, J = 10.8, 8.8 Hz, 1H, H-5), 3.47−3.41 (m, 1H, H-7), 2.00 (s, 3H, $COCH_3$). ^{13}C NMR (125 MHz, CD_3OD) δ 175.18, 166.12 (2 × C═O), 158.95, 158.31, 131.00, 128.28, 126.98, 124.77, 120.19, 119.98 (Ar-C), 148.06 (triazole-C4), 123.10 (triazole-C5), 145.90 (C-2), 113.04 (C-3), 77.73 (C-6), 71.32 (C-7), 69.83 (C-4), 67.95 (C-8), 55.25 (C-9), 51.95 (C-5), 22.70 ($COCH_3$). HR-MS (ESI): calcd for $C_{25}H_{25}N_4O_8$ $[M - H]^-$ 509.1672, found 509.1674.

*5-Acetamido-9-(4-carboxyphenyl)-2,6-anhydro-4-guanidino-3,4,5-trideoxy-D-glycero-D-galacto-non-2-enonic Acid (**8a**).* To a solution of compound **16** (40 mg) in 1 mL of DCM, 100 μL of TFA was added. The solution was then stirred at room temperature for 2 h. After completion, DCM and TFA were removed under reduced pressure. The residue was dissolved in 2 mL of 0.1 N NaOH and stirred at room temperature for 1 h. After completion, Amberlite IR 120 (H^+) was added to the reaction mixture to adjust the pH of the solution to 2. The suspension was then filtered, and the filtrate was concentrated and dissolved in minimum methanol. The desired product was precipitated with ethyl acetate. Yield 10 mg (41%). ^{1}H NMR (500 MHz, D_2O) δ 8.49 (s, 1H, triazole-H), 8.02 (d, J = 8.2 Hz, 2H, Ar-H), 7.93 (d, J = 8.2 Hz, 2H, Ar-H), 5.68 (d, J = 2.1 Hz, 1H, H-3), 4.91 (dd, J = 14.4, 2.8 Hz, 1H, H-9), 4.72 (dd, J = 14.4, 6.7 Hz, 1H, H-9′), 4.48 (dd, J = 9.5, 2.1 Hz, 1H, H-4), 4.46−4.41 (m, 1H, H-8), 4.40 (d, J = 11.0 Hz, 1H, H-7), 4.26 (t, J = 9.5 Hz, 1H, H-5), 3.53 (d, J = 9.5 Hz, 1H, H-6), 1.99 (s, 3H, $COCH_3$). ^{13}C NMR (125 MHz, D_2O) δ 175.37, 169.94, 163.66 (3 × C═O), 157.96 (C═N), 150.14 (C-2), 147.81 (triazole-C4), 124.65 (triazole-C5), 133.00, 130.62, 126.33 (Ar-C), 104.76 (C-3), 76.02 (C-6), 69.79 (C-8), 69.00 (C-7), 54.41 (C-9), 51.83 (C-4), 48.65 (C-5), 22.75 ($COCH_3$). HR-MS (ESI): calcd for $C_{21}H_{24}N_7O_8$ $[M - H]^-$ 502.1686, found 502.1683.

*Methyl 5-Acetamido-9-(4-biphenyl-1H-1,2,3-triazol-1-yl))-4-guanidino-2,6-anhydro-4-[2,3-bis(tert-butoxycarbonyl)guanidino]-3,4,5-trideoxy-D-glycero-D-galacto-non-2-enonate (**8b-OMe**).* A solution of compound **24** (150 mg) in 3 mL of DCM was charged with TFA (300 μL), and the solution was stirred at room temperature for 2 h. After completion, DCM and TFA were removed under reduced pressure. The residue was dissolved in 5 mL of methanol and cooled to 0 °C, followed by addition of NaOMe (54 mg, 5 equiv). The reaction mixture was stirred at 0 °C for 1 h. After completion, Amberlite IR 120 (H^+) was added to neutralize the solution. The suspension was then filtered, and the filtrate was concentrated and purified by flash chromatography to give the desired product. Yield 50 mg (45%). ^{1}H NMR (500 MHz, CD_3OD) δ 8.31 (s, 1H, triazole-H), 7.84 (d, J = 8.1 Hz, 2H, Ar-H), 7.62 (d, J = 8.1 Hz, 2H, Ar-H), 7.57 (d, J = 7.5 Hz, 2H, Ar-H), 7.38 (t, J = 7.5 Hz, 2H, Ar-H), 7.29 (t, J = 7.3 Hz, 1H, Ar-H), 5.87 (d, J = 2.0 Hz, 1H, H-3), 4.58−4.47 (m, 2H, H-9′, H-6), 4.43 (d, J = 10.2 Hz, 1H, H-4), 4.33−4.20 (m, 2H, H-8, H-5), 3.72 (s, 3H, $COOCH_3$), 3.58 (d, J = 9.0 Hz, 1H, H-7), 1.97 (s, 3H, $COCH_3$). ^{13}C NMR (125 MHz, CD_3OD) δ 174.57, 163.86 (2 × C═O), 158.95 (C═N), 148.23 (triazole-C4), 146.33, 142.29, 141.69, 130.72, 129.94, 128.58, 128.50, 127.84, 127.11 (Ar-C), 123.71 (triazole-C5), 109.28 (C-3), 77.93 (C-6), 71.21 (C-8), 70.11 (C-7), 55.15 (C-9), 53.07 ($COOCH_3$), 51.51 (C-5), 22.69 ($COCH_3$). HR-MS (ESI): calcd for $C_{27}H_{32}N_7O_6$ $[M + H]^+$ 550.2414, found 550.2395.

*5-Acetamido-9-(4-biphenyl-1H-1,2,3-triazol-1-yl))-4-guanidino-2,6-anhydro-4-[2,3-bis(tert-butoxycarbonyl)guanidino]-3,4,5-trideoxy-D-glycero-D-galacto-non-2-enonic Acid (**8b**).* A solution of compound **24** (180 mg) in 3 mL DCM was charged with 300 μL of TFA, and the solution was stirred at room temperature for 2 h. After completion, DCM and TFA were removed under reduced pressure. The residue was dissolved in 3 mL of 0.1 N NaOH and stirred at room temperature for 1 h. After completion, Amberlite IR 120 (H^+) was added to adjust the pH of the solution to 2. The suspension was then filtered, and the filtrate was concentrated and dissolved in minimum methanol. The desired product was precipitated with ethyl acetate. Yield 10 mg (31%). ^{1}H NMR (500 MHz, CD_3OD) δ 8.34 (s, 1H, triazole-H), 7.87 (d, J = 8.1 Hz, 2H, Ar-H), 7.66 (d, J = 8.1 Hz, 2H, Ar-H), 7.61 (d, J = 7.5 Hz, 2H, Ar-H), 7.41 (t, J = 7.6 Hz, 2H, Ar-H), 7.32 (t, J = 7.3 Hz, 1H, Ar-H), 5.89 (s, 1H, H-3), 4.88 (d, J = 14.3 Hz, 1H, H-9), 4.60−4.47 (m, 2H, H-9′, H-6), 4.42 (d, J = 10.0 Hz, 1H, H-4), 4.28 (m, 2H, H-5, H-8), 3.57 (d, J = 9.0 Hz, 1H, H-7), 1.98 (s, 3H, $COCH_3$). ^{13}C NMR (125 MHz, CD_3OD) δ 174.56, 165.08 (2 × C═O), 158.96 (C═N), 148.25 (C-2), 146.68 (triazole-C4), 123.70 (triazole-C5), 142.30, 141.70, 130.71, 129.95, 128.58, 128.50, 127.85, 127.12 (Ar-C), 109.14 (C-3), 77.79 (C-6), 71.21 (C-8), 70.06 (C-7), 55.19 (C-9), 51.53 (C-5), 22.72 ($COCH_3$). HR-MS (ESI): calcd for $C_{26}H_{28}N_7O_6$ $[M - H]^-$ 534.2101, found 534.2105.

*2-Methyl-4,5-dihydro-(methyl(7,8-di-O-acetyl-9-(4-(methoxycarbonyl)phenyl)-2,6-anhydro-3,4,5-trideoxy-D-glycero-D-talo-non-2-en)onate)[5,4-d]-1,3-oxazole (**11**).* A solution of compound **10h** (250 mg, 1 equiv) in anhydrous pyridine was cooled to 0 °C, followed by dropwise addition of acetic anhydride (230 μL, 4.5 equiv). The reaction mixture was allowed to warm to room temperature and kept stirring overnight. After completion, the reaction was quenched with methanol and the solvents were removed under reduced pressure. The residue was dissolved in ethyl acetate and carefully washed with 0.05 M HCl, water, and brine sequentially and dried over Na_2SO_4. The solution was then concentrated and purified by flash chromatography, providing a crude fully protected product, which was used in the next step without further purification. 370 mg (quant, crude product). The obtained crude protected product (800 mg, 1 equiv, several batches' product of last step) was dissolved in 10 mL of ethyl acetate. The solution was warmed to 40 °C and TMSOTf (408 μL, 3 equiv) was added dropwise. The resulting solution was kept stirring at 50 °C for 4 h. After completion, the solution was added to a vigorously stirred cold saturated sodium bicarbonate solution. The aqueous phase was separated and extracted with ethyl acetate. The organic phase was combined, dried over Na_2SO_4, concentrated, and purified by flash chromatography to give the desired product (430 mg, 60%). ^{1}H NMR (500 MHz, CD_3OD) δ 8.49 (s, 1H, triazole-H), 8.03−7.97 (d, J = 8.5 Hz, 2H, Ar-H), 7.88 (d, J = 8.5 Hz, 2H, Ar-H), 6.39 (d, J = 4.0 Hz, 1H, H-3), 5.65−5.57 (m, 2H, H-7, H-8), 5.22 (dd, J = 14.8, 2.6 Hz, 1H, H-9), 4.94 (dd, J = 9.5, 4.0 Hz, 1H, H-4), 4.79 (m, 1H, H-9′), 4.02 (t, J = 9.5 Hz, 1H, H-5), 3.87, 3.79 (2 × s, 2 × 3H, 2 × $COOCH_3$), 3.60 (dd, J = 9.5, 2.3 Hz, 1H, H-6), 2.17 (s, 3H, oxazole-CH_3), 1.97, 1.95 (2 × s, 2 × 3H, 2 × $COOCH_3$). ^{13}C NMR (125 MHz, CD_3OD) δ 171.55, 171.32, 168.03, 163.34 (4 × C═O), 170.12 (oxazole-O-C═N), 148.09 (triazole-C4), 124.39 (triazole-C5), 147.66

DOI: 10.1021/acs.jmedchem.7b01574
J. Med. Chem. XXXX, XXX, XXX−XXX

(C-2), 136.30, 131.27, 130.78, 126.55 (Ar-C), 109.05 (C-3), 78.35 (C-6), 74.22 (C-4), 73.52 (C-8), 70.86 (C-7), 62.68 (C-5), 53.27, 52.81 (2 × $COOCH_3$), 51.09 (C-9), 20.73, 20.67 ($COCH_3$), 14.06 (oxazole-CH_3). HR-MS (ESI): calcd for $C_{26}H_{28}N_4NaO_{10}$ $[M + Na]^+$ 579.1703, found 579.1697

*Methyl 5-Acetamido-7,8-di-O-acetyl-9-(4-(methoxycarbonyl)-phenyl)-2,6-anhydro-4-azido-3,4,5-trideoxy-D-glycero-D-galacto-non-2-enonate (**12**).* To a solution of compound **11** (430 mg, 1 equiv) in dry ^{t}BuOH, $TMSN_3$ (507 μL, 5 equiv) was added and the resulting solution was stirred at 80 °C under a nitrogen atmosphere for 12 h. After completion, the solution was cooled to room temperature, concentrated, and purified by flash chromatography to give the desired product. Yield 470 mg (quant). ^{1}H NMR (500 MHz, CD_3OD) δ 8.51 (s, 1H, triazole-H), 8.07−8.02 (m, 2H, Ar-H), 7.92−7.87 (m, 2H, Ar-H), 5.97 (d, *J* = 2.2 Hz, 1H, H-3), 5.57−5.56 (dd, *J* = 3.3, 2.2 Hz, 1H, H-4), 5.52 (dt, *J* = 9.1, 2.9 Hz, 1H, H-8), 5.29 (dd, *J* = 14.8, 2.7 Hz, 1H, H-9), 4.71 (dd, *J* = 14.8, 9.1 Hz, 1H, H-9′), 4.49 (dd, *J* = 10.3, 1.9 Hz, 1H, H-6), 4.27−4.18 (m, 2H, H-5, H-7), 3.89, 3.80 (2 × s, 2× 3H, 2 × $COOCH_3$), 2.13, 1.93, 1.92 (3 × s, 3 × 3H, 3 × $COCH_3$). ^{13}C NMR (125 MHz, CD_3OD) δ 173.49, 171.83, 171.42, 168.09, 163.07 (5 × C═O), 147.64 (triazole-C4), 124.51 (triazole-C5), 146.32 (C-2), 136.25, 131.26, 130.85, 126.54 (Ar-C), 109.74 (C-3), 78.44 (C-6), 73.96 (C-8), 69.60 (C-7), 60.46 (C-4), 53.21, 52.74 (2 × $COOCH_3$), 51.17 (C-9), 48.39 (C-5), 22.89, 20.88, 20.63 (3 × $COCH_3$). HR-MS (ESI): calcd for $C_{26}H_{29}N_7NaO_{10}$ $[M + Na]^+$ 622.1874, found 622.1866

*5-Acetamido-9-(4-(methoxycarbonyl)phenyl)-2,6-anhydro-4-azido-3,4,5-trideoxy-D-glycero-D-galacto-non-2-enonic Acid (**13**).* A sample of 60 mg of compound **12** was dissolved in 2 mL of 0.5 N NaOH, and the solution was stirred under room temperature for 1 h. After completion, Amberlite IR 120 (H^+) was added to adjust the pH of the solution as 2. The suspension was then filtered, and the filtrate was concentrated and purified by flash chromatography to give the desired product. Yield 32 mg (66%). ^{1}H NMR (500 MHz, CD_3OD) δ 8.43 (s, 1H, triazole-H), 8.06 (d, *J* = 8.2 Hz, 2H, Ar-H), 7.91 (d, *J* = 8.2 Hz, 2H, Ar-H), 5.73 (s, 1H), 4.53 (dd, *J* = 14.0, 7.6 Hz, 1H, H-9′), 4.32−4.25 (m, 2H, H-4, H-8), 4.23 (d, *J* = 10.8 Hz, 1H, H-6), 4.18−4.10 (m, 1H, H-5), 3.45 (d, *J* = 9.3 Hz, 1H, H-7), 1.98 (s, 3H, $COCH_3$). ^{13}C NMR (125 MHz, CD_3OD) δ 174.32 (C═O), 147.53 (triazole-C4), 124.47 (triazole-C5), 136.17, 131.48, 126.42 (Ar-C), 104.50 (C-3), 77.02 (C-6), 71.03 (C-8), 69.86 (C-7), 60.23 (C-4), 55.25 (C-9), 48.53 (C-5), 22.78 ($COCH_3$). HR-MS (ESI): calcd for $C_{20}H_{20}N_7O_8$ $[M - H]^-$ 486.1373, found 486.1378

*Methyl 5-Acetamido-7,8-di-O-acetyl-9-(4-(methoxycarbonyl)-phenyl)-2,6-anhydro-4-amino-3,4,5-trideoxy-D-glycero-D-galacto-non-2-enonate (**14**).* To a solution of compound **12** (50 mg, 1 equiv) in THF (2 mL), 0.5 N HCl (200 μL, 2 equiv) was added, followed by triphenylphosphine (29 mg, 1.1 equiv). The resulting mixture was stirred at room temperature overnight. After completion, solvents were removed under reduced pressure and the residue was purified by flash chromatography, providing the desired product. Yield 39 mg (84%). ^{1}H NMR (500 MHz, CD_3OD) δ 8.57 (s, 1H, triazole-H), 8.05 (d, *J* = 8.4 Hz, 2H, Ar-H), 7.92 (d, *J* = 8.4 Hz, 2H, Ar-H), 6.06 (d, *J* = 2.3 Hz, 1H, H-3), 5.63−5.55 (m, 2H, H-7, H-8), 5.29−5.21 (m, 1H, H-9), 4.74 (dd, *J* = 14.7, 8.4 Hz, 1H, H-9′), 4.64 (dd, *J* = 10.0, 1.1 Hz, 1H, H-6), 4.35 (t, *J* = 10.0 Hz, 1H, H-5), 4.15 (dd, *J* = 10.0, 2.3 Hz, 1H, H-4), 3.90, 3.81 (2 × s, 2 × 3H, 2 × $COOCH_3$), 2.12, 1.97, 1.95 (3 × s, 3 × 3H, 3 × $COCH_3$). ^{13}C NMR (125 MHz, CD_3OD) δ 174.28, 171.71, 171.31, 168.15, 162.80 (5 × C═O), 147.66 (triazole-C4), 124.59 (triazole-C5), 147.29, 136.28, 130.84, 126.56 (Ar-C), 107.15 (C-3), 77.96 (C-6), 73.34 (C-8), 69.49 (C-7), 53.34, 52.78 (2 × $COOCH_3$), 51.67 (C-4), 51.31 (C-9), 46.73 (C-5), 23.14, 20.89, 20.66 (3 × $COCH_3$). HR-MS (ESI): calcd for $C_{26}H_{31}N_5NaO_{10}$ $[M + Na]^+$ 596.1969, found 596.1967.

*5-Acetamido-7,8-di-O-acetyl-9-(4-(methoxycarbonyl)phenyl)-2,6-anhydro-4-amino-3,4,5-trideoxy-D-glycero-D-galacto-non-2-enonic Acid (**15**).* A sample of 35 mg of compound **14** was dissolved in 400 μL of 1N NaOH, and the solution was kept stirring at rt for 1 h. After completion, the reaction mixture was adjusted to pH 2 with Amberlite IR 120 (H^+). The suspension was then filtered, and the filtrate was concentrated and dissolved in minimum methanol. The desired product was precipitated with ethyl acetate. Yield 20 mg (71%). ^{1}H NMR (500 MHz, CD_3OD) δ 8.47 (s, 1H, triazole-H), 8.08 (d, *J* = 8.2 Hz, 2H, Ar-H), 7.94 (d, *J* = 8.2 Hz, 2H, Ar-H), 5.84 (s, 1H, H-3), 4.89 (d, *J* = 14.4 Hz, 1H, H-9), 4.56 (dd, *J* = 14.0, 7.5 Hz, 1H, H-9′), 4.41−4.26 (m, 3H, H-8, H-6, H-5), 4.18 (d, *J* = 7.1 Hz, 1H, H-4), 3.56 (d, *J* = 9.1 Hz, 1H, H-7), 2.03 (s, 3H, $COCH_3$). ^{13}C NMR (125 MHz, CD_3OD) δ 174.86, 169.48 (2 × C═O), 147.50 (triazole-C4), 124.66 (triazole-C5), 136.29, 131.48, 131.25, 126.45 (Ar-C), 103.03 (C-3), 77.01 (C-6), 70.71 (C-8), 70.03 (C-7), 55.19 (C-9), 51.29 (C-4), 47.54 (C-5), 23.03 ($COCH_3$). HR-MS (ESI): calcd for $C_{20}H_{22}N_5O_8$ $[M - H]^-$ 460.1468, found 460.1482.

*Methyl 5-Acetamido-9-(4-(methoxycarbonyl)phenyl-1H-1,2,3-triazol-1-yl))-4-[2,3-bis(tert-butoxycarbonyl)guanidino]-7,8-di-O-acetyl-2,6-anhydro-3,4,5-trideoxy-D-glycero-D-galacto-non-2-enonate (**16**).* To a solution of compound **14** (40 mg, 1 equiv) in 2 mL of anhydrous DCM, TEA (40 μL, 4 equiv) was added. The solution was cooled to 0 °C, and *N*,*N*′-Di-Boc-1*H*-pyrazole-1-carboxamidine (42 mg, 2 equiv) was added. The reaction mixture was allowed to warm to room temperature and kept stirring overnight. After completion, the reaction was quenched with water and extracted with ethyl acetate. The organic phase was washed with brine, dried over Na_2SO_4, concentrated, and purified by flash chromatography to give the desired product. Yield 40 mg (crude product, 72%). ^{1}H NMR (500 MHz, CD_3OD) δ 8.55 (s, 1H, triazole-H), 8.06 (d, *J* = 8.3 Hz, 2H, Ar-H), 7.92 (d, *J* = 8.3 Hz, 2H), 6.00 (d, *J* = 2.3 Hz, 1H, H-3), 5.58 (d, *J* = 1.5 Hz, 1H, H-7), 5.56−5.52 (m, 1H, H-8), 5.31 (dd, *J* = 14.8, 2.4 Hz, 1H, H-9), 5.02 (dd, *J* = 10.2, 2.3 Hz, 1H, H-4), 4.74 (dd, *J* = 14.8, 9.0 Hz, 1H, H-9′), 4.53 (dd, *J* = 10.2, 1.5 Hz, 1H, H-6), 4.27 (t, *J* = 10.2 Hz, 1H, H-5), 3.91, 3.80 (2 × s, 2 × 3H, 2 × $COOCH_3$), 2.12, 1.94, 1.85 (3 × s, 3 × 3H, 3 × $COCH_3$), 1.51, 1.46 (2 × s, 2 × 9H, 2 × Boc). ^{13}C NMR (125 MHz, CD_3OD) δ 173.57, 171.77, 171.41, 168.07, 164.32, 163.39, 158.01 (7 × C═O), 153.82 (C═N), 147.64 (triazole-C4), 124.50 (triazole-C5), 145.61 (C-2), 136.34, 131.27, 131.27, 126.55 (Ar-C), 111.83 (C-3), 84.84, 80.57 (2 × ^{t}Boc-C$(CH_3)_3$), 78.89 (C-6), 74.00 (C-8), 69.87 (C-7), 53.07, 52.74 (2 × $COOCH_3$), 51.27 (C-9), 50.84 (C-4), 47.90 (C-5), 28.59, 28.26 (2 × ^{t}Boc-C$(CH_3)_3$), 22.77, 20.86, 20.65 (3 × $COCH_3$). HR-MS (ESI): calcd for $C_{37}H_{49}N_7NaO_{14}$ $[M + Na]^+$ 838.3235, found 838.3226.

*Methyl 5-Acetamido-9-(4-(methoxycarbonyl)phenyl-1H-1,2,3-triazol-1-yl))-4-(3-(3-methoxy-3-oxopropyl)ureido)-7,8-di-O-acetyl-2,6-anhydro-3,4,5-trideoxy-D-glycero-D-galacto-non-2-enonate (**17**).* A solution of compound **14** (100 mg, 1 equiv) and TEA (56 mg, 2 equiv) in anhydrous DCM was cooled to 0 °C and charged with 1,1′-carbonyldiimidazole (39 mg, 1.2 equiv). The reaction mixture was then warmed to room temperature and kept stirring for 2 h until TLC showed consumption of the amine. The solution was then cooled to 0 °C, and the methyl ester of β-alanine (56 mg, 2 equiv) was added. The solution was warmed to room temperature and kept stirring overnight. After completion, the reaction was quenched with water and extracted by ethyl acetate. The organic layer was washed with water and brine, and dried over Na_2SO_4. After subsequent concentration, the residue was purified by flash chromatography to give the desired product. Yield 140 mg (quant). ^{1}H NMR (500 MHz, CD_3OD) δ 8.54 (s, 1H, triazole-H), 8.07 (d, *J* = 8.7 Hz, 2H, Ar-H), 7.92 (d, *J* = 8.7 Hz, 2H, Ar-H), 5.92 (d, *J* = 2.5 Hz, 1H, H-3), 5.55 (m, 2H, H-8, H-4), 5.31 (dd, *J* = 14.8, 2.6 Hz, 1H, H-9), 4.73 (dd, *J* = 14.8, 8.9 Hz, 1H, H-9′), 4.55 (dd, *J* = 9.9, 2.5 Hz, 1H, H-7), 4.46 (dd, *J* = 10.2, 2.0 Hz, 1H, H-6), 4.10 (t, *J* = 10.2 Hz, 1H, H-5), 3.91, 3.78, 3.66 (3 × s, 3 × 3H, 3 × $COOCH_3$), 3.36 (td, *J* = 6.6, 1.9 Hz, 2H, CH_2), 2.48 (t, *J* = 6.5 Hz, 2H, CH_2), 2.10, 1.93, 1.87 (3 × s, 3 × 3H, 3 × $COCH_3$). ^{13}C NMR (125 MHz, CD_3OD) δ 174.12, 173.59, 171.79, 171.41, 168.13, 163.56, 160.42 (7 × C═O), 147.66 (C-2), 145.10 (triazole-C4), 124.46 (triazole-C5), 136.30, 131.26, 130.88, 126.54 (Ar-C), 114.06 (C-1), 79.10 (C-6), 73.99 (C-8), 69.97 (C-7), 52.96, 52.70, 52.15 (3 × $COOCH_3$), 51.26 (C-9), 50.38 (C-5), 36.92, 35.64 (CH_2CH_2), 22.85, 20.84, 20.59 ($COCH_3$). HR-MS (ESI): calcd for $C_{31}H_{39}N_7N_6O_{13}$ $[M + Na]^+$ 703.2575, found 703.2571.

*5-Acetamido-9-(4-carboxyphenyl)-2,6-anhydro-4-(3-(2-carboxyethyl)ureido)-3,4,5-trideoxy-D-glycero-D-galacto-non-2-enonic Acid (**18**).* A sample of 140 mg of compound **17** was dissolved

DOI: 10.1021/acs.jmedchem.7b01574
J. Med. Chem. XXXX, XXX, XXX−XXX

in 5 mL of 0.1 N NaOH and stirred at room temperature for 1 h. After completion, the reaction mixture adjusted to pH 2 with the addition of Amberlite IR 120 (H^+). The suspension was then filtered, and the filtrate was concentrated and dissolved in minimum methanol. The desired product was precipitated with ethyl acetate. 88 mg (77%). ^{1}H NMR (500 MHz, CD_3OD) δ 8.44 (s, 1H, triazole-H), 8.07 (d, J = 8.4 Hz, 2H, Ar-H), 7.93 (d, J = 8.4 Hz, 2H, Ar-H), 5.66 (d, J = 1.9 Hz, 1H, H-3), 4.57 (dd, J = 9.8, 1.9 Hz, 1H, H-4), 4.52 (dd, J = 14.0, 7.7 Hz, 1H, H-9′), 4.30 (dd, J = 12.1, 4.8 Hz, 1H, H-8), 4.19 (d, J = 10.8 Hz, 1H, H-6), 4.02 (t, J = 10.3 Hz, 1H, H-5), 3.44 (d, J = 9.3 Hz, 1H, H-7), 3.37 (t, J = 6.4 Hz, 2H, CH_2), 2.46 (t, J = 6.4 Hz, 2H, CH_2), 1.94 (s, 3H, $COCH_3$). ^{13}C NMR (125 MHz, CD_3OD) δ 175.74, 174.73, 160.76, 169.64 (4 × C═O), 147.51 (triazole-C4), 124.40 (triazole-C5), 136.24, 131.68, 131.45, 126.42 (Ar-C), 109.35 (C-3), 77.81 (C-6), 71.36 (C-8), 69.80 (C-7), 55.31 (C-9), 37.00, 35.78 (CH_2CH_2), 22.74 ($COCH_3$). HR-MS (ESI): calcd for $C_{24}H_{27}N_6O_{11}$ $[M - H]^-$ 575.1738, found 575.1738

*Methyl 5-Acetylamino-4-(tert-butoxycarbonyl)amino-2,6-anhydro-3,5-dideoxy-D-glycero-D-galacto-non-2-enonate (**20**).* To a solution of compound **19** (600 mg, 1 equiv) and TEA (389 μL, 2 equiv) in 20 mL of anhydrous DCM at 0 °C, di-*tert*-butyl dicarbonate (456 mg, 1.5 equiv) was added dropwise. The mixture was them warmed to room temperature and kept stirring overnight. After completion, solvent was removed and the residue was purified by flash chromatography to give the desired compound in 350 mg (crude product, 47%). The crude product (350 mg, 1 equiv) was dissolved in 10 mL of methanol, and cooled to 0 °C, followed by addition of NaOMe (92 mg, 3 equiv). The solution was kept stirring at 0 °C for 1 h until no starting material remained. The solution was adjusted to pH 2 by addition of Amberlite IR 120 (H^+). The suspension was then filtered, and the filtrate was concentrated and purified by flash chromatography to give the desired product. Yield 190 mg (71%). ^{1}H NMR (500 MHz, CD_3OD) δ 5.82 (d, J = 2.2 Hz, 1H, H-3), 4.46 (d, J = 10.1 Hz, 1H, H-4), 4.23 (d, J = 10.1 Hz, 1H, H-6), 4.05 (t, J = 10.1 Hz, 1H, H-5), 3.88 (ddd, J = 9.2, 5.4, 2.9 Hz, 1H, H-8), 3.81 (dd, J = 11.4, 2.9 Hz, 1H, H-9), 3.65 (dd, J = 11.4, 5.4 Hz, 1H, H-9′), 3.58 (dd, J = 9.3, 1.1 Hz, 1H, H-7), 1.98 (s, 3H, $COOCH_3$), 1.44 (s, 9H, ^{t}Boc-$C(CH_3)_3$). ^{13}C NMR (125 MHz, CD_3OD) δ 174. 65, 164.26, 158.35 (3 × C═O), 145.71 (C-2), 112.19 (C-3), 80.56 (^{t}Boc-$C(CH_3)_3$), 78.62 (C-6), 71.13 (C-8), 70.00 (C-7), 64.90 (C-9), 52.79 (C-4), 50.26 (C-5), 28.70 (^{t}Boc-$C(CH_3)_3$), 22.70 ($COCH_3$). HR-MS (ESI): calcd for $C_{17}H_{29}N_2O_9$ $[M + H]^+$ 405.1873, found 405.1875

*Methyl 4-(tert-Butoxycarbonyl)amino-5-acetylamino-9-(4-methylbenzenesulfonate)-2,6-anhydro-3,5-dideoxy-D-glycero-D-galacto-non-2-enonate (**21**).* A solution of compound **20** (190 mg, 1 equiv) in anhydrous pyridine was cooled down to 0 °C, TsCl (98 mg, 1.1 equiv) was then added slowly under stirring. The solution was warmed to room temperature and kept stirring overnight. After completion, the reaction was quenched by methanol. The solution was concentrated and purified by flash chromatography to give the desired product. Yield 200 mg (76%). ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.78 (d, J = 8.1 Hz, 2H, Ar-H), 7.33 (d, J = 8.1 Hz, 2H, Ar-H), 6.93, 5.23, 5.11, 3.59 (4 × d, 4H, 2 × NH, 2 × OH), 5.82 (d, J = 2.3 Hz, 1H, H-3), 4.55 (td, J = 9.6, 2.2 Hz, 1H, H-8), 4.40−4.31 (m, 1H, H-6), 4.21−4.17 (m, 1H, H-5), 4.15−4.11 (m, 2H, H-9′, H-4), 4.01−3.95 (m, 1H, H-9′), 3.71 (s, 3H, $COOCH_3$), 3.56−3.48 (m, 1H, H-7), 2.43 (s, 3H, $PhCH_3$), 2.00 (s, 3H, $COCH_3$), 1.42 (s, 9H, ^{t}Boc-$C(CH_3)_3$). ^{13}C NMR (125 MHz, $CDCl_3$) δ 174.01, 162.21, 156.80 (3 × C═O), 145.07 (C-2), 144.86, 132.59, 129.91, 128.02 (Ar-C), 109.92 (C-3), 80.60 (^{t}Boc-$C(CH_3)_3$), 72.56 (C-9), 68.53 (C-7), 67.90 (C-8), 52.38 (C-6), 50.19 (C-4), 48.68 (C-5), 28.26 (^{t}Boc-$C(CH_3)_3$), 22.86, 21.63 ($COCH_3$, $PhCH_3$). HR-MS (ESI): calcd for $C_{24}H_{35}N_2O_{11}S$ $[M + H]^+$ 559.1962, found 559.1966.

*Methyl 5-Acetamido-7,8-di-O-acetyl-9-azido-2,6-anhydro-4-[2,3-bis(tert-butoxycarbonyl)guanidino]-3,4,5-trideoxy-D-glycero-D-galacto-non-2-enonate (**23**).* Compound **21** (200 mg, 1 equiv) was dissolved in 3 mL of acetone−water (2:1), and NaN_3 (117 mg, 5 equiv) was added. The solution was heated at 67 °C under N_2 for 2 days. After completion, the solution was concentrated and purified by flash chromatography to give 100 mg of compound **22** (crude product, 75%), which was used in the next step without further purification. The crude product was dissolved in 2 mL of anhydrous DCM, and 200 μL of TFA was added. The solution was kept stirring at room temperature until no starting material remained. Solvent was then removed under vacuum, and the residue was dissolved in 2 mL of anhydrous DCM, and TEA (140 μL, 4 equiv) was added. After the solution was cooled down to 0 °C, *N,N′*-di-Boc-1*H*-pyrazole-1-carboxamidine (150 mg, 2 equiv) was added. The reaction mixture was allowed to warm up to room temperature and kept stirring overnight. After completion, the reaction was quenched with water and extracted with ethyl acetate. The organic phase was washed with brine, dried over Na_2SO_4, concentrated, and purified by flash chromatography to give the desired product. Yield 108 mg (82%). ^{1}H NMR (500 MHz, $CDCl_3$) δ 8.63, 8.15, 7.70, 6.43 (2 × d, 2 × brs, 4 H, 2 × NH, 2 × OH), 5.82 (d, J = 2.4 Hz, 1H, H-3), 5.20 (ddd, J = 10.2, 8.1, 2.4 Hz, 1H, H-4), 4.22−4.15 (m, 2H, H-6, H-5), 4.02 (td, J = 10.2, 6.1 Hz, 1H, H-8), 3.72 (dd, J = 12.6, 2.8 Hz, 1H, H-9), 3.60−3.53 (m, 2H, H-9′, H-7), 2.04 (s, 3H, $COCH_3$), 1.52 (2 × s, 2 × 9H, 2 × ^{t}Boc-$C(CH_3)_3$). ^{13}C NMR (125 MHz, $CDCl_3$) δ 174.04, 162.29, 162.08, 157.56 (4 × C═O), 152.70 (C═N), 146.31 (C-2), 107.51 (C-3), 84.39, 80.11 (2 × ^{t}Boc-$C(CH_3)_3$), 69.14 (C-6), 54.89 (C-9), 52.53 (C-7), 51.53 (C-8), 48.33 (C-5), 28.23, 28.04 (2 × ^{t}Boc-$C(CH_3)_3$), 22.96 ($COCH_3$). HR-MS (ESI): calcd for $C_{23}H_{38}N_7O_{10}$ $[M + H]^+$ 572.2680, found 572.2681.

*Methyl 5-Acetamido-9-(4-biphenyl-1H-1,2,3-triazol-1-yl))-4-[2,3-bis(tert-butoxycarbonyl)guanidino]-2,6-anhydro-4-[2,3-bis(tert-butoxycarbonyl)guanidino]-3,4,5-trideoxy-D-glycero-D-galacto-non-2-enonate (**24**).* Compound **23** (200 mg, 1 equiv) and 4-ethynylbiphenyl (32 mg, 1.5 equiv) were reacted following the protocol above to provide the desired product. Yield 180 mg (69%). ^{1}H NMR (700 MHz, $CDCl_3$) δ 8.55, 8.06 (2 × d, 2H, 2 × NH), 7.94 (s, 1H, triazole-H), 7.81 (d, J = 8.3 Hz, 2H, Ar-H), 7.57 (dd, J = 13.1, 7.8 Hz, 4H, Ar-H), 7.39 (t, J = 7.7 Hz, 2H, Ar-H), 7.31 (t, J = 7.4 Hz, 1H, Ar-H), 5.75 (d, J = 2.3 Hz, 1H, H-3), 5.51 (brs, 1H, OH), 5.16−5.10 (m, 1H, H-4), 4.90 (dd, J = 14.0, 1.7 Hz, 1H, H-9), 4.54 (dd, J = 14.0, 6.7 Hz, 1H, H-9′), 4.48−4.43 (m, 1H, H-8), 4.20 (d, J = 10.4 Hz, 1H, H-6), 3.96 (td, J = 10.4, 6.2 Hz, 1H, H-5), 3.69 (s, 3H, $COOCH_3$), 3.34 (d, J = 9.0 Hz, 1H, H-7), 1.90 (s, 3H), 1.47, 1.43 (2 × s, 2 × 9H, 2 × ^{t}Boc-$C(CH_3)_3$). ^{13}C NMR (176 MHz, $CDCl_3$) δ 174.10, 162.26, 162.03, 157.37 (4 × C═O), 152.66 (C═N), 146.98 (C-2), 146.26 (triazole-C4), 121.69 (triazole-C5), 140.63, 140.43, 129.41, 128.77, 127.37, 126.86, 125.99 (Ar-C), 107.54 (C-3), 84.22, 79.96 (2 × ^{t}Boc-$C(CH_3)_3$), 69.33 (C-6), 68.54 (C-4), 53.90 (C-9), 52.42 (C-8), 51.58 (C-7), 48.37 (C-5), 28.17, 27.99 (^{t}Boc-$C(CH_3)_3$), 22.86 ($COCH_3$). HR-MS (ESI): calcd for $C_{37}H_{48}N_7O_{10}$ $[M + H]^+$ 750.3463, found 750.3454.

General Procedure for Synthesis of Compounds 25a−d. A solution of compound **19** (1 equiv) and TEA (3 equiv) in anhydrous DCM was cooled to 0 °C, and the corresponding anhydride or acyl chloride (3 equiv) was added dropwise. The resulting mixture was warmed to room temperature and kept stirring overnight. After completion, the reaction was quenched with water and extracted with ethyl acetate. The organic layer was collected and washed with saturated $NaHCO_3$ and brine, and sequentially and dried with Na_2SO_4. Solvents were removed under reduced pressure, and the residue was separated by flash chromatography to give the desired crude products. For hydrolysis of the C1-methyl ester, the crude product was dissolved in MeOH, and 0.5 M NaOH was added. The mixture was kept stirring at room temperature. After completion, the pH was adjusted to 2 with Amberlite IR-120 (H^+). The solution was then filtered and purified by flash chromatography to provide the desired products with yields of 42%−68% (over two steps).

*5-Acetamido-2,6-anhydro-4-propionamido-3,4,5-trideoxy-D-glycero-D-galacto-non-2-enonic Acid (**25a**).* Yield 28 mg (68%, over two steps). ^{1}H NMR (500 MHz, CD_3OD) δ 5.49 (d, J = 2.0 Hz, 1H, H-3), 4.75 (dd, J = 9.7, 2.0 Hz, 1H, H-4), 4.20 (d, J = 10.8 Hz, 1H, H-6), 4.10−4.06 (m, 1H, H-5), 3.86−3.85 (m, 1H, H-8), 3.79 (dd, J = 11.4, 3.0 Hz, 1H, H-9), 3.63 (dd, J = 11.4, 5.4 Hz, 1H, H-9′), 3.55 (d, J = 9.0 Hz, 1H, H-7), 2.17 (q, J = 7.6 Hz, 2H, α-CH_2), 1.93 (s, 3H, $COCH_3$), 1.09 (t, J = 7.6 Hz, 3H, β-CH_3). ^{13}C NMR (125 MHz,

DOI: 10.1021/acs.jmedchem.7b01574
J. Med. Chem. XXXX, XXX, XXX−XXX

CD_3OD) δ 177.40, 174.17 (2 × C═O), 106.15 (C-3), 77.49 (C-6), 71.46 (C-8), 70.05 (C-7), 64.88 (C-9), 49.64 (C-5), 30.40 (α-CH_2), 22.78 ($COCH_3$), 10.59 (β-CH_3). HR-MS (ESI): calcd for $C_{14}H_{21}N_2O_8$ $[M - H]^-$ 345.1298, found 345.1302

*5-Acetamido-2,6-anhydro-4-pentanamido-3,4,5-trideoxy-D-glycero-D-galacto-non-2-enonic Acid (**25b**).* Yield 25 mg (56%, over two steps). ^1H NMR (500 MHz, CD_3OD) δ 5.47 (d, J = 2.2 Hz, 1H, H-3), 4.76 (dd, J = 9.8, 2.2 Hz, 1H, H-4), 4.19 (d, J = 10.8 Hz, 1H, H-6), 4.09−4.05 (m, 1H, H-5), 3.87−3.85 (m, 1H, H-8), 3.78 (dd, J = 11.5, 3.1 Hz, 1H, H-9), 3.66 (dd, J = 11.5, 5.1 Hz, 1H, H-9′), 3.57 (t, J = 7.8 Hz, 1H, H-7), 2.17 (t, J = 7.5 Hz, 2H, α-CH_2), 1.94 (s, 3H, $COCH_3$), 1.60−1.51 (m, 2H, β-CH_2), 1.34−1.29 (m, 2H, γ-CH_2), 0.90 (t, J = 7.4 Hz, 3H, δ-CH_3). ^{13}C NMR (125 MHz, CD_3OD) δ 176.55, 174.18, 170.01 (3 × C═O), 151.02 (C-2), 105.96 (C-3), 77.51 (C-6), 71.55 (C-8), 69.91 (C-7), 64.71 (C-9), 49.60 (C-4), 49.00 (C-5), 37.02 (α-CH_2), 29.24 (β-CH_2), 23.31 (γ-CH_2), 23.31 ($COCH_3$), 14.19 (δ-CH_3). HR-MS (ESI): calcd for $C_{16}H_{25}N_2O_8$ $[M - H]^-$ 373.1611, found 373.1612.

*5-Acetamido-2,6-anhydro-4-cyclopropanecarboxamido-3,4,5-trideoxy-D-glycero-D-galacto-non-2-enonic Acid (**25c**).* Yield 22 mg (44%, over two steps). ^1H NMR (500 MHz, CD_3OD) δ 5.54 (d, J = 2.1 Hz, 1H, H-3), 4.77 (dd, J = 9.8, 2.1 Hz, 1H, H-4), 4.20 (d, J = 10.7 Hz, 1H, H-6), 4.11−4.07 (m, 1H, H-5), 3.90−3.82 (m, 1H, H-8), 3.79 (dd, J = 11.4, 3.0 Hz, 1H, H-9), 3.66 (dd, J = 11.4, 5.2 Hz, 1H, H-9′), 3.57 (d, J = 9.0 Hz, 1H, H-7), 1.94 (s, 3H, $COCH_3$), 1.58−1.53 (m, 1H, α-CH), 0.88−0.78 (m, 2H, β-CH_2), 0.74−0.72 (m, 2H, β-CH_2). ^{13}C NMR (125 MHz, CD_3OD) δ 176.90, 174.36, 169.54 (3 × C═O), 150.41 (C-2), 106.92 (C-3), 77.64 (C-6), 71.52 (C-8), 69.95 (C-7), 64.77 (C-9), 49.92 (C-4), 49.28 (C-5), 22.87 ($COCH_3$), 15.05 (α-CH), 7.57, 7.49 (2 × β-CH_2). HR-MS (ESI): calcd for $C_{15}H_{21}N_2O_8$ $[M - H]^-$ 357.1298, found 357.1305.

*5-Acetamido-2,6-anhydro-4-cyclobutanecarboxamido-3,4,5-trideoxy-D-glycero-D-galacto-non-2-enonic Acid (**25d**).* Yield 19 mg (42%, over two steps). ^1H NMR (500 MHz, CD_3OD) δ 5.49 (d, J = 2.2 Hz, 1H, H-3), 4.75 (dd, J = 9.8, 2.2 Hz, 1H, H-4), 4.21 (d, J = 10.8 Hz, 1H, H-6), 4.10−4.06 (m, 1H, H-5), 3.88−3.84 (m, *m*, 1H, H-8), 3.78 (dd, J = 11.5, 3.1 Hz, 1H, H-9), 3.66 (dd, J = 11.5, 5.2 Hz, 1H, H-9′), 3.56 (d, J = 9.3 Hz, 1H, H-7), 3.09−3.02 (m, 1H, α-CH), 2.27−1.77 (m, 6H, 3 × CH_2), 1.93 (s, 3H, $COCH_3$). ^{13}C NMR (125 MHz, CD_3OD) δ 178.06, 174.20, 169.63 (3 × C═O), 150.52 (C-2), 106.54 (C-3), 77.53 (C-6), 71.55 (C-8), 69.94 (C-7), 64.77 (C-9), 49.57 (C-4), 49.14 (C-5), 40.90 (α-CH), 26.44, 26.02, 19.09 (3 × CH_2), 22.88 ($COCH_3$). HR-MS (ESI): calcd for $C_{16}H_{23}N_2O_8$ $[M - H]^-$ 371.1454, found 371.1458.

Inhibition Assay. Inhibition assays against 4MU-NANA cleavage and GM3 cleavage were performed using protocols reported previously.[24] NEU3 and NEU2 were expressed as N-terminal MBP fusion proteins in *Escherichia coli* and purified as previously reported.[29] NEU4 was expressed as an MBP fusion protein in *E. coli* and purified (see Supporting Information).[18]

NEU1 was overexpressed as a $(His)_6$ fusion protein in HEK293 cells and used as a crude preparation from cell lysate.[63] The CathA-IRES-NEU1$(His)_6$ biscistronic recombinant plasmid was constructed by removing the hIL12 and CD19tmpk from pDy.hIL12.IRES.CD19tmpk.WS vector kindly provided J. Medin (University of Toronto Health Network) and replaced respectively by human cathepsin A (CathA) and NEU1$(His)_6$ cDNA. The NEU1$(His)_6$ cDNA was PCR-amplified and inserted in the MCS of the pDy.hIL12.IRES.CD19tmpk.WS between the XbaI and *Bam*HI sites. The CathA cDNA was PCR-modified and inserted into the MCS of pDy.hIL12.IRES.NEU1-GFP and pDy.hIL12.IRES.NEU1 between the *Eco*RI and AscI sites. All constructs were verified by PCR, enzymatic digestions, and DNA sequencing. The final CathA-IRES-NEU1$(His)_6$ construct was assessed for sialidase activity in transfected HEK293T cells. The CathA-IRES-NEU1$(His)_6$ biscistronic HIV-1-based recombinant vesicular stomatitis virus glycoprotein-pseudotyped (VSVg) LV was generated by transient transfections of HEK293T cells by LV plasmid construct, packaging plasmid pCMVDR8.91, and the VSV-g envelope-coding plasmid pMD.G as described previously.[64] HEK293T cells were cultured in DMEM supplemented with 10% FBS, 1% streptomycin−penicillin at 37 °C in a humidified 5% CO_2 atmosphere. To overexpress NEU1, cells were transduced with CathA-IRES-NEU1$(His)_6$ lentivirus at 6 MOI in the presence of polybrene (8 μg mL^{-1}). Five, 8, and 10 days after transduction, a portion of the cells was harvested to verify NEU1 expression by measuring acidic neuraminidase activity in the cell lysate. The transduced cells showed significantly higher neuraminidase activity in the cell lysates (76−85 nmol h^{-1} mg^{-1} of protein) as compared with the lysate of wild-type HEK293T cells (1.7−2.0 nmol h^{-1} mg^{-1} of protein).

All assays were conducted in 0.1 M sodium acetate buffer at optimum pH for each enzyme (pH 4.5 for NEU1, NEU3, and NEU4; pH 5.5 for NEU2).[24] To get comparable IC_{50} values among the four isoenzymes, similar activity of each enzyme was used in the assay based on 4MU-NANA activity.

For assays using 4MU-NANA as the substrate, inhibitors in 3-fold serial dilutions were incubated with enzyme at 0 °C for 15 min. 4MU-NANA was then added to the mixture, making the final concentration of 4MU-NANA as 50 μM and the total volume of the reaction mixture as 20 μL. After incubation at 37 °C for 30 min, the reaction was quenched with 100 μL of 0.2 M sodium glycine buffer (pH 10.2). The reaction mixture was transferred to 386-well plate, and the enzyme activity was determined by measuring fluorescence (λ_{ex} = 365 nm; λ_{em} = 445 nm) using a plate reader (Molecular Devices, Sunnyvale CA). Assays were performed with duplicates for each point, and IC_{50} was obtained by plotting the data with Graphpad Prism 7.0. For curves that showed less than a 50% decrease in signal, fits were conducted using maximum inhibition values found for DANA.

For inhibition assays against GM3 cleavage, a method developed by Markely and co-workers was adopted.[65] The assay was conducted in 0.1 M sodium acetate buffer (pH 4.5). After inhibitors in serial dilutions were incubated with enzyme at 0 °C for 15 min, GM3 was added, making the final concentration of GM3 500 μM and the total volume of the reaction mixture 20 μL. The reaction mixture was incubated at 37 °C for 30 min and quenched with 100 μL of freshly made 0.2 M sodium borate buffer (pH 10.2). Then 0.8% malononitrile solution (40 μL) was added to form a fluorescent adduct with the free sialic acid. Fluorescence was obtained (λ_{ex} = 357 nm; λ_{em} = 434 nm), and the data were processed using Graphpad Prism 7.0. For curves that showed less than a 50% decrease in signal, fits were conducted using maximum inhibition values found for DANA.

K_i Determinations. Enzymes were incubated with inhibitors in serial dilutions at 0 °C for 15 min and serial concentrations of 4MU-NANA were added. The reaction mixture was transferred to a 386-well plate immediately, and the rate of product formation was obtained by measuring fluorescence (λ_{ex} = 315 nm; λ_{em} = 450 nm) every 1 min for 30 min. The obtained data were processed with Graphpad Prism 7.0 for K_i determination.

clogD$_{7.4}$ Calculation. The clog$D_{7.4}$ was calculated using MarvinSketch (Chemaxon, version 16.10.10.0). The consensus method was used with 0.1 M Cl^- and 0.1 M Na^+ ions and pK_a correction library.

Log $D_{7.4}$ Determination (Shake-Flask Method). Equal amounts of TRIS-HCl buffer (0.1 M, 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 stock solutions of test compounds (DMSO, 10 mM) were diluted to 10 μM in saturated buffer (final DMSO concentration; 0.1%). The buffer was transferred to a 96-well plate and saturated 1-octanol was added, resulting in a 40/140 and 30/150 1-octanol to water ratio, respectively (140/40 and 150/30 for compounds with clog$D_{7.4}$ below 0). Each ratio was measured in triplicate, and simultaneous measurements were conducted with codeine as 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-

DOI: 10.1021/acs.jmedchem.7b01574
J. Med. Chem. XXXX, XXX, XXX−XXX

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 eq 1:

$$\log D_{7.4} = \log\left(\frac{10\ \mu\text{M} - c_b}{c_b} \times \frac{1}{\text{o:b}}\right) \quad (1)$$

Parallel Artificial Membrane Permeation Assay. Effective permeability ($\log P_e$) was determined in a 96-well format with parallel artificial membrane permeation assay (PAMPA).[45] For each compound, measurements were performed at pH 7.4 in quadruplicate. Four wells of a deep well plate were filled with 500 μL of Prisma HT buffer (pH 7.4, pION P/N 110 151). Then analyte dissolved in DMSO (10 mM) was added to the buffer to yield 50 μM solutions (containing 0.5% DMSO). Afterward, 200 μL was transferred to each well of the donor plate of the PAMPA sandwich (pIon Billerica, USA, P/N 110 163). The filter membranes at the bottom of the acceptor plate were infused with 5 μL of GIT-0 lipid solution (pIon, P/N 110 669), and 200 μL of acceptor buffer (pION, P/N 110 139) was filled into each acceptor well. The experiment was conducted with 10% DMSO in Prisma HT buffer and acceptor buffer for compounds with insufficient solubility. The sandwich was assembled, placed in the GutBoxTM (pION), and left undisturbed for 16 h. Then it was disassembled and the compound concentration of donor, acceptor, and reference compartment was measured by LC-MS (see below). Effective permeability ($\log P_e$) was calculated from the compound flux deduced from the concentration in the donor and acceptor compartment as well as the retention in the filter area with the aid of the PAMPA Explorer Software (pIon, version 3.5).

Metabolic Stability in Human Liver Microsomes. Incubations were performed in triplicate in a 96-well format on an Eppendorf Thermomixer Comfort. The reaction mixture (270 μL) consisting of liver microsomes, TRIS-HCl buffer (0.1 M, pH 7.4), and $MgCl_2$ (2 mM) was preheated (37 °C, 500 rpm, 10 min), and the incubation was initiated by adding 30 μL of compound solution (200 μM) in TRIS-HCl buffer. The final concentration of the compound was 20 μM, including 0.5 mg mL^{-1} human liver microsomes. At the beginning of the experiment (t = 0 min) and after an incubation time of 10, 30, 60, and 90 min, samples (40 μL) were transferred to 120 μL of ice-cold MeOH and centrifuged (3700 rpm, 4 °C, 10 min). Then 80 μL of supernatant was transferred to a 96-well plate for analysis by LC-MS (see below). The metabolic half-life ($t_{1/2}$) was calculated from the slope of the linear regression from the log percentage remaining compound versus incubation−time relationship. Control experiments were performed in parallel with control substance as positive control or containing 1 mM bis(4-nitrophenyl) phosphate (esterase inhibitor, negative control). Compounds with $t_{1/2}$ > 90 min were considered to be metabolically stable.

Colorectal Adenocarcinoma (Caco-2) Cell Permeation Assay. Cultivation, splitting, seeding of Caco-2 cells as well as the assay protocol previously described by Schönemann et al.[39] Permeation assays were conducted in triplicate at a concentration of 62.5 μM containing 0.6% DMSO. Samples (40 μL) were withdrawn after 30 min and analyzed by LC-MS (see below), and the apparent permeation (P_{app}) was calculated according to eq 2. The compound flux (mol/s) is dQ/dt, A is the surface area of the monolayer (cm^2), and c_0 is the initial concentration.

$$P_{app} = \frac{dQ}{dt} \times \frac{1}{A \times c_0} \quad (2)$$

Compounds with a $P_{app} < 2 \times 10^{-6}$ cm/s are considered as low permeable according to Hou et al. (2007).[42,46]

Solubility. A sample of 5 μL of predesolved compounds (DMSO, 10 mM) were added to 500 μL of buffer (Tris-HCl 0.1 M pH 7.4, μL Prisma HT buffer, pH 7.4, pION P/N 110 151) and water containing 0.9% NaCl) ending in 1% DMSO. The vials were put into supersonic 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 MeOH, and analyzed by LC-MS (see below).

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 C18 column (2.1 mm × 50 mm) with a 3 μm particle size (Waters Corp., Milford, MA, USA). The mobile phase consisted of eluent A, 0.1% formic acid in water, and eluent B, MeCN containing 0.1% formic acid. The flow rate was maintained at 0.6 mL/min. The gradient was ramped from 95% A/5% B to 5% A/95% B over 1 min, and then hold at 5% A/95% B for 0.1 min. The system was then brought back to 95% A/5% B, resulting in a total duration of 4 min. 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).

Molecular Dynamics (MD) Simulations. MD simulations were run for the inhibitor **8b** bound to NEU2, NEU3, and NEU4. Inhibitor **8b** was simulated in its zwitterionic form, i.e., as a carboxylate at C1 and as a guanidinium at C4. For simulations with NEU2, we used the crystal structure of NEU2 bound to **1** (PDB 1VCU)[27] as a starting structure for the enzyme. The nonterminal residues G227, E228, S284, G295, P286, and G287 were added to the crystal structure for NEU2 using Modeler in Chimera.[58,66−68] We used our previously reported homology model for NEU3,[29,30] and we used the SWISS-MODEL Web site[69−73] to generate a homology model for NEU4 based on the 1VCU structure.[27] The portion of NEU4 between N289 and F368 is not homologous to NEU2, therefore, residues 290−367 (78 in total) were removed from the NEU4 model before running MD simulations. The coordinates for **8b** were based on those of **6** bound to NEU2 (PDB 2F0Z)[50] with the biphenyl-triazole and hydrogens added in Avogadro.[74,75] To obtain the initial position for **8b** in the active site of Neu2, we aligned it to **6** in the crystal structure 2F0Z,[50] and we kept the active site waters for our simulations. For simulations with NEU3 and NEU4, each enzyme was aligned separately with the starting structure for NEU2 using the MatchMaker tool in Chimera.[58,67,68] The starting positions for **8b** in the active sites of NEU2, NEU3, and NEU4 are shown in the Supporting Information.

All simulations were run in AMBER 15[76] using *pmemd.cuda* (GPU acceleration) on Nvidia GeForce GTX 980 GPUs. The *ff14SB* force field[77] was used for NEU2, NEU3, and NEU4, while the general AMBER force field (GAFF)[78] was used for **8b**. Partial charges for **8b** were assigned using the AM1 with bond charge correction (AM1-BCC) model[79] in the *antechamber* module of AmberTools15.[76] The enzyme−inhibitor complexes with NEU2 and NEU3 were neutralized with the addition of Na^+ ions and that for NEU4 were neutralized with the addition of Cl^- ions. All complexes were solvated in a box of TIP3P water[80] with 10 Å between the solute and the edges of the box in all three dimensions. For all systems, the water was first minimized using 100 steps of steepest descent, followed by 4900 steps of conjugate gradient. Then the entire system was minimized with 100 steps of steepest descent, followed by 4900 steps of conjugate gradient. The systems were further equilibrated by heating from 5 to 300 K over 50 ps, followed by cooling back to 5 K over an additional 50 ps. After the annealing step, the systems were again heated from 5 to 300 K over 100 ps then allowed to run at 300 K for 100 ps before the production simulations were started. Production was run for 50 ns. The time step was 2 fs, bonds to hydrogen were constrained with the SHAKE[81] algorithm, and the cutoff for nonbonded interactions was 8.0 Å. The temperature was maintained with the Berendsen thermostat[82] (*ntt* = 1) with velocities rescaled every 1 ps. The final 25 ns of the simulations were used for analysis with the *cpptraj* module of AmberTools15.[76,83]

Neuraminidase Assay in Mouse Brain Tissue. Mice with targeted disruption of the *neu1* (*neu1*$^{-/-}$), *neu*3 (*neu*3$^{-/-}$), and *neu*4 (*neu*4$^{-/-}$) genes have been previously described.[54,56,84] Mice with a combined deficiency of NEU4 and NEU3 were obtained by intercrossing *neu4* and *neu3* knockout (KO) mouse strains.[84] Mice

DOI: 10.1021/acs.jmedchem.7b01574
J. Med. Chem. XXXX, XXX, XXX−XXX

were housed in an enriched environment with continuous access to food and water, under constant temperature and humidity, on a 12 h light:dark cycle. Approval for the animal care and the use in the experiments was granted by the Animal Care and Use Committee of the Ste. Justine University Hospital Research Center.

At the age of 12 weeks, mice were sacrificed using a CO_2 chamber and their brains extracted, snap-frozen with liquid nitrogen, and kept at −80 °C. For measurement of neuraminidase activity, 100 mg of frozen brain tissue was homogenized in water in a ratio of 100 mg of tissue per 500 μL of water in the 1.5 mL Eppendorf tubes using a Kontes Pellet motorized pestle. Protein concentration in the homogenate was measured by the Bradford method using the Bio-Rad reagent. Acidic α-neuraminidase activity was assayed at pH 4.6 using fluorogenic 4-methylumbelliferyl-*N*-acetyl neuraminic acid (4MU-NANA) substrate as previously described.[54] The reaction mixture contained an aliquot of homogenate corresponding to 300 μg of total protein, neuraminidase inhibitor in a concentration of 0−150 μM and substrate in a final concentration of 200 μM. The reaction was carried on at 37 °C for 30 min, after which it was terminated by the addition of 1900 μL of 0.4 M glycine buffer, pH 10.5. For blank samples, the reaction mixture contained buffer, inhibitor, and substrate only, but the same volume of homogenate was added after the termination of reaction.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jmedchem.7b01574.

^{1}H and ^{13}C NMR data, HR-MS data, and HPLC traces for intermediates and final products, IC_{50} curves and K_i determinations (PDF)

Molecular formula strings (CSV)

PDB files for modeling Neu2, Neu3, and Neu4 with **8b** (ZIP)

■ AUTHOR INFORMATION

Corresponding Author

*Phone: 780-492-0377. Fax: 780-492-8231. E-mail: ccairo@ualberta.ca.

ORCID

Beat Ernst: 0000-0001-5787-2297

Christopher W. Cairo: 0000-0003-3363-8708

Author Contributions

T.G. synthesized and characterized all compounds unless otherwise noted, conducted IC_{50} and K_i measurements, and wrote the manuscript; P.G. synthesized compounds **25a−d**; C.Z. and R.Z. developed and implemented protein purification protocols; M.R.R. performed molecular modeling and analysis and wrote the manuscript; P.D. and B.E. designed and conducted PK analysis and wrote the manuscript; A.F. produced HEK293 cells expressing NEU1; E.D. and A.P. designed and performed inhibition experiments in murine brain homogenate, analyzed data, and wrote the manuscript; C.W.C. designed experiments, interpreted data, and wrote the manuscript.

Notes

The authors declare the following competing financial interest(s): The authors note that a provisional patent application has been filed on a portion of this work.

■ ACKNOWLEDGMENTS

This work was supported by a grant from the Canadian Glycomics Network (GlycoNet), and the Natural Sciences and Engineering Research Council of Canada (NSERC).

■ ABBREVIATIONS USED

NEU, neuraminidase enzyme; Neu5Ac, 5-acetamino-3,5-dideoxy-D-*glycero*-D-*galacto*-non-2-ulosonic acid; DANA, 2-deoxy-2,3-didehydro-*N*-acetylneuraminic acid; 4MU-NANA, 2′-(4-methylumbellifery)-α-D-*N*-acetylneuraminic acid; SAR, structure−activity relationship; CuAAC, copper-catalyzed azide−alkyne cycloaddition; C9-BA-DANA, C9-(butyl-*N*-amide) DANA; C9-4HMT-DANA, C9-(4-hydroxymethyltriazolyl) DANA

■ REFERENCES

(1) Angata, T.; Varki, A. Chemical diversity in the sialic acids and related α-keto acids: An evolutionary perspective. *Chem. Rev.* **2002**, *102*, 439−470.

(2) Chen, X.; Varki, A. Advances in the biology and chemistry of sialic acids. *ACS Chem. Biol.* **2010**, *5*, 163−176.

(3) Varki, A. Sialic acids in human health and disease. *Trends Mol. Med.* **2008**, *14*, 351−360.

(4) Buschiazzo, A.; Alzari, P. M. Structural insights into sialic acid enzymology. *Curr. Opin. Chem. Biol.* **2008**, *12*, 565−572.

(5) Miyagi, T.; Yamaguchi, K. Mammalian sialidases: Physiological and pathological roles in cellular functions. *Glycobiology* **2012**, *22*, 880−896.

(6) Seyrantepe, V.; Poupetova, H.; Froissart, R.; Zabot, M.-T.; Maire, I.; Pshezhetsky, A. V. Molecular pathology of NEU1 gene in sialidosis. *Hum. Mutat.* **2003**, *22*, 343−352.

(7) Kwak, J. E.; Son, M.-Y.; Son, Y. S.; Son, M. J.; Cho, Y. S. Biochemical and molecular characterization of novel mutations in GLB1 and NEU1 in patient cells with lysosomal storage disorders. *Biochem. Biophys. Res. Commun.* **2015**, *457*, 554−560.

(8) Gayral, S.; Garnotel, R.; Castaing-Berthou, A.; Blaise, S.; Fougerat, A.; Berge, E.; Montheil, A.; Malet, N.; Wymann, M. P.; Maurice, P.; Debelle, L.; Martiny, L.; Martinez, L. O.; Pshezhetsky, A. V.; Duca, L.; Laffargue, M. Elastin-derived peptides potentiate atherosclerosis through the immune Neu1-PI3K gamma pathway. *Cardiovasc. Res.* **2014**, *102*, 118−127.

(9) Yang, A.; Gyulay, G.; Mitchell, M.; White, E.; Trigatti, B. L.; Igdoura, S. A. Hypomorphic sialidase expression decreases serum cholesterol by downregulation of VLDL production in mice. *J. Lipid Res.* **2012**, *53*, 2573−2585.

(10) Miyagi, T.; Takahashi, K.; Hata, K.; Shiozaki, K.; Yamaguchi, K. Sialidase significance for cancer progression. *Glycoconjugate J.* **2012**, *29*, 567−577.

(11) Takahashi, K.; Hosono, M.; Sato, I.; Hata, K.; Wada, T.; Yamaguchi, K.; Nitta, K.; Shima, H.; Miyagi, T. Sialidase NEU3 contributes neoplastic potential on colon cancer cells as a key modulator of gangliosides by regulating Wnt signaling. *Int. J. Cancer* **2015**, *137*, 1560−1573.

(12) Shiga, K.; Takahashi, K.; Sato, I.; Kato, K.; Saijo, S.; Moriya, S.; Hosono, M.; Miyagi, T. Upregulation of sialidase NEU3 in head and neck squamous cell carcinoma associated with lymph node metastasis. *Cancer Sci.* **2015**, *106*, 1544−1553.

(13) Silvestri, I.; Testa, F.; Zappasodi, R.; Cairo, C. W.; Zhang, Y.; Lupo, B.; Galli, R.; Di Nicola, M.; Venerando, B.; Tringali, C. Sialidase NEU4 is involved in glioblastoma stem cell survival. *Cell Death Dis.* **2014**, *5*, e1381.

(14) Hyun, S. W.; Liu, A.; Liu, Z.; Cross, A. S.; Verceles, A. C.; Magesh, S.; Kommagalla, Y.; Kona, C.; Ando, H.; Luzina, I. G.; Atamas, S. P.; Piepenbrink, K. H.; Sundberg, E. J.; Guang, W.; Ishida, H.; Lillehoj, E. P.; Goldblum, S. E. The NEU1-selective sialidase inhibitor, C9-butyl-amide-DANA, blocks sialidase activity and NEU1-

mediated bioactivities in human lung in vitro and murine lung in vivo. *Glycobiology* **2016**, *26*, 834−849.
(15) Cairo, C. W. Inhibitors of the human neuraminidase enzymes. *MedChemComm* **2014**, *5*, 1067−1074.
(16) von Itzstein, M. The war against influenza: Discovery and development of sialidase inhibitors. *Nat. Rev. Drug Discovery* **2007**, *6*, 967−974.
(17) Hata, K.; Koseki, K.; Yamaguchi, K.; Moriya, S.; Suzuki, Y.; Yingsakmongkon, S.; Hirai, G.; Sodeoka, M.; von Itzstein, M.; Miyagi, T. Limited inhibitory effects of oseltamivir and zanamivir on human sialidases. *Antimicrob. Agents Chemother.* **2008**, *52*, 3484−3491.
(18) Albohy, A.; Mohan, S.; Zheng, R. B.; Pinto, B. M.; Cairo, C. W. Inhibitor selectivity of a new class of oseltamivir analogs against viral neuraminidase over human neuraminidase enzymes. *Bioorg. Med. Chem.* **2011**, *19*, 2817−2822.
(19) Li, Y.; Cao, H.; Yu, H.; Chen, Y.; Lau, K.; Qu, J.; Thon, V.; Sugiarto, G.; Chen, X. Identifying selective inhibitors against the human cytosolic sialidase NEU2 by substrate specificity studies. *Mol. BioSyst.* **2011**, *7*, 1060−1072.
(20) Magesh, S.; Savita, V.; Moriya, S.; Suzuki, T.; Miyagi, T.; Ishida, H.; Kiso, M. Human sialidase inhibitors: Design, synthesis, and biological evaluation of 4-acetamido-5-acylamido-2-fluoro benzoic acids. *Bioorg. Med. Chem.* **2009**, *17*, 4595−4603.
(21) This compound has been previously named as a *C*9-butyl *N*-amide derivative of DANA (C9-BA-DANA).This could more appropriately be named as a C9-pentylamide, but we maintain the previous nomenclature for consistency with the literature.
(22) Magesh, S.; Moriya, S.; Suzuki, T.; Miyagi, T.; Ishida, H.; Kiso, M. Design, synthesis, and biological evaluation of human sialidase inhibitors. Part 1: Selective inhibitors of lysosomal sialidase (NEU1). *Bioorg. Med. Chem. Lett.* **2008**, *18*, 532−537.
(23) Khedri, Z.; Li, Y.; Cao, H.; Qu, J.; Yu, H.; Muthana, M. M.; Chen, X. Synthesis of selective inhibitors against V. cholerae sialidase and human cytosolic sialidase NEU2. *Org. Biomol. Chem.* **2012**, *10*, 6112−6120.
(24) Zhang, Y.; Albohy, A.; Zou, Y.; Smutova, V.; Pshezhetsky, A. V.; Cairo, C. W. Identification of selective inhibitors for human neuraminidase isoenzymes using C4,C7-modified 2-deoxy-2,3-didehydro-N-acetylneuraminic acid (DANA) analogues. *J. Med. Chem.* **2013**, *56*, 2948−2958.
(25) Zou, Y.; Albohy, A.; Sandbhor, M.; Cairo, C. W. Inhibition of human neuraminidase 3 (NEU3) by C9-triazole derivatives of 2,3-didehydro-N-acetyl-neuraminic acid. *Bioorg. Med. Chem. Lett.* **2010**, *20*, 7529−7533.
(26) Albohy, A.; Zhang, Y.; Smutova, V.; Pshezhetsky, A. V.; Cairo, C. W. Identification of selective nanomolar inhibitors of the human neuraminidase, NEU4. *ACS Med. Chem. Lett.* **2013**, *4*, 532−537.
(27) Chavas, L. M. G.; Tringali, C.; Fusi, P.; Venerando, B.; Tettamanti, G.; Kato, R.; Monti, E.; Wakatsuki, S. Crystal structure of the human cytosolic sialidase Neu2 - Evidence for the dynamic nature of substrate recognition. *J. Biol. Chem.* **2005**, *280*, 469−475.
(28) Buchini, S.; Gallat, F.-X.; Greig, I. R.; Kim, J.-H.; Wakatsuki, S.; Chavas, L. M. G.; Withers, S. G. Tuning mechanism-based inactivators of neuraminidases: Mechanistic and structural Insights. *Angew. Chem.* **2014**, *126*, 3450−3454.
(29) Albohy, A.; Li, M. D.; Zheng, R. B.; Zou, C.; Cairo, C. W. Insight into substrate recognition and catalysis by the human neuraminidase 3 (NEU3) through molecular modeling and site-directed mutagenesis. *Glycobiology* **2010**, *20*, 1127−1138.
(30) Albohy, A.; Richards, M. R.; Cairo, C. W. Mapping substrate interactions of the human membrane-associated neuraminidase, NEU3, using STD NMR. *Glycobiology* **2015**, *25*, 284−293.
(31) Magesh, S.; Suzuki, T.; Miyagi, T.; Ishida, H.; Kiso, M. Homology modeling of human sialidase enzymes NEU1, NEU3 and NEU4 based on the crystal structure of NEU2: Hints for the design of selective NEU3 inhibitors. *J. Mol. Graphics Modell.* **2006**, *25*, 196−207.
(32) Meldal, M.; Tornøe, C. W. Cu-catalyzed azide−alkyne cycloaddition. *Chem. Rev.* **2008**, *108*, 2952−3015.
(33) Lu, Y.; Gervay-Hague, J. Synthesis of C-4 and C-7 triazole analogs of zanamivir as multivalent sialic acid containing scaffolds. *Carbohydr. Res.* **2007**, *342*, 1636−1650.
(34) Shidmoossavee, F. S.; Watson, J. N.; Bennet, A. J. Chemical insight into the emergence of influenza virus strains that are resistant to Relenza. *J. Am. Chem. Soc.* **2013**, *135*, 13254−13257.
(35) Staudinger, H.; Meyer, J. Über neue organische phosphorverbindungen III. Phosphinmethylenderivate und phosphinimine. *Helv. Chim. Acta* **1919**, *2*, 635−646.
(36) Potier, M.-C.; Mameli, L.; Bélisle, M.; Dallaire, L.; Melancon, S. B. Fluorometric assays of neuraminidase with a sodium (4-methylumbellifery-α-D-N-acetylneuraminate) substrate. *Anal. Biochem.* **1979**, *94*, 287−296.
(37) Warner, T. G.; O'Brien, J. S. Synthesis of 2′-(4-methylumbelliferyl)-α-D-N-acetylneuraminic acid and detection of skin fibroblast neuraminidase in normal humans and in sialidosis. *Biochemistry* **1979**, *18*, 2783−2787.
(38) Monti, E.; Bassi, M. T.; Papini, N.; Riboni, M.; Manzoni, M.; Venerando, B.; Croci, G.; Preti, A.; Ballabio, A.; Tettamanti, G.; Borsani, G. Identification and expression of NEU3, a novel human sialidase associated to the plasma membrane. *Biochem. J.* **2000**, *349*, 343−351.
(39) Schönemann, W.; Kleeb, S.; Dätwyler, P.; Schwardt, O.; Ernst, B. Prodruggability of carbohydrates — Oral FimH antagonists. *Can. J. Chem.* **2016**, *94*, 909−919.
(40) Schade, D.; Kotthaus, J.; Riebling, L.; Kotthaus, J.; Müller-Fielitz, H.; Raasch, W.; Hoffmann, A.; Schmidtke, M.; Clement, B. Zanamivir amidoxime- and N-hydroxyguanidine-based prodrug approaches to tackle poor oral bioavailability. *J. Pharm. Sci.* **2015**, *104*, 3208−3219.
(41) Scebach, D.; Thaler, A.; Blaser, D.; Ko, S. Y. Transesterifications with 1,8-diazabicyclo[5.4.0]undec-7-ene/lithium bromide (DBU/LiBr) − Also applicable to cleavage of peptides from resins in merrifleld syntheses. *Helv. Chim. Acta* **1991**, *74*, 1102−1118.
(42) Hou, T.; Wang, J.; Zhang, W.; Xu, X. ADME evaluation in drug discovery. 7. Prediction of oral absorption by correlation and classification. *J. Chem. Inf. Model.* **2007**, *47*, 208−218.
(43) Kah, M.; Brown, C. D. LogD: Lipophilicity for ionisable compounds. *Chemosphere* **2008**, *72*, 1401−1408.
(44) OECD. *Partitions Coefficient (n-Octanol/Water), Shake-Flask Method*; Guideline for Testing of Chemicals, no.107, 1995.
(45) Kansy, M.; Senner, F.; Gubernator, K. Physicochemical high throughput screening: Parallel artificial membrane permeation assay in the description of passive absorption processes. *J. Med. Chem.* **1998**, *41*, 1007−1010.
(46) Avdeef, A.; Bendels, S.; Di, L. i.; Faller, B.; Kansy, M.; Sugano, K.; Yamauchi, Y. PAMPA—Critical factors for better predictions of absorption. *J. Pharm. Sci.* **2007**, *96*, 2893−2909.
(47) Amidon, G. L.; Lennernäs, H.; Shah, V. P.; Crison, J. R. A theoretical basis for a biopharmaceutic drug classification: The correlation of in vitro drug product dissolution and in vivo bioavailability. *Pharm. Res.* **1995**, *12*, 413−420.
(48) Custodio, J. M.; Wu, C.-Y.; Benet, L. Z. Predicting drug disposition, absorption/elimination/transporter interplay and the role of food on drug absorption. *Adv. Drug Delivery Rev.* **2008**, *60*, 717−733.
(49) Hubatsch, I.; Ragnarsson, E. G. E.; Artursson, P. Determination of drug permeability and prediction of drug absorption in Caco-2 monolayers. *Nat. Protoc.* **2007**, *2*, 2111−2119.
(50) Chavas, L. M. G., Kato, R., McKimm-Breschkin, J., Colman, P. M., Fusi, P., Tringali, C., Venerando, B., Tettamanti, G., Monti, E., Wakatsuki, S. *Crystal Structure of the Human Sialidase NEU2 in Complex with Zanamivir Inhibitor (PDB ID: 2F0Z)*; Protein Data Bank: RCSB, 2005.
(51) IUPAC-IUB Joint Commission on Biochemical Nomenclature (JCBN). Conformational nomenclature for five and six-membered ring forms of monosaccharides and their derivatives. *Eur. J. Biochem.* **1980**, *111*, 295−29810.1111/j.1432-1033.1980.tb04941.x.

DOI: 10.1021/acs.jmedchem.7b01574
J. Med. Chem. XXXX, XXX, XXX−XXX

(52) Fushinobu, S. *Cremer−Pople Parameter Calculator*; ISU, 2006; http://enzyme13.bt.a.u-tokyo.ac.jp/CP/ (accessed Jul 6, 2017).
(53) Woods, R. J.; Grant, O. C.; Makeneni, S.; Foley, B. L. The effect of substrate presentation and activation on neuraminidase NEU2 specificity. *FASEB J.* **2017**, *31*, lb111.
(54) Seyrantepe, V.; Canuel, M.; Carpentier, S.; Landry, K.; Durand, S.; Liang, F.; Zeng, J.; Caqueret, A.; Gravel, R. A.; Marchesini, S.; Zwingmann, C.; Michaud, J.; Morales, C. R.; Levade, T.; Pshezhetsky, A. V. Mice deficient in Neu4 sialidase exhibit abnormal ganglioside catabolism and lysosomal storage. *Hum. Mol. Genet.* **2008**, *17*, 1556−1568.
(55) Seyrantepe, V.; Hinek, A.; Peng, J.; Fedjaev, M.; Ernest, S.; Kadota, Y.; Canuel, M.; Itoh, K.; Morales, C. R.; Lavoie, J.; Tremblay, J.; Pshezhetsky, A. V. Enzymatic activity of lysosomal carboxypeptidase (cathepsin) A is required for proper elastic fiber formation and inactivation of endothelin-1. *Circulation* **2008**, *117*, 1973−1981.
(56) Yamaguchi, K.; Shiozaki, K.; Moriya, S.; Koseki, K.; Wada, T.; Tateno, H.; Sato, I.; Asano, M.; Iwakura, Y.; Miyagi, T. Reduced susceptibility to colitis-associated colon carcinogenesis in mice lacking plasma membrane-associated sialidase. *PLoS One* **2012**, *7*, e41132.
(57) Koseki, K.; Wada, T.; Hosono, M.; Hata, K.; Yamaguchi, K.; Nitta, K.; Miyagi, T. Human cytosolic sialidase NEU2-low general tissue expression but involvement in PC-3 prostate cancer cell survival. *Biochem. Biophys. Res. Commun.* **2012**, *428*, 142−149.
(58) Yang, Z.; Lasker, K.; Schneidman-Duhovny, D.; Webb, B.; Huang, C. C.; Pettersen, E. F.; Goddard, T. D.; Meng, E. C.; Šali, A.; Ferrin, T. E. UCSF Chimera, MODELLER, and IMP: An integrated modeling system. *J. Struct. Biol.* **2012**, *179*, 269−278.
(59) Smutova, V.; Albohy, A.; Pan, X.; Korchagina, E.; Miyagi, T.; Bovin, N.; Cairo, C. W.; Pshezhetsky, A. V. Structural basis for substrate specificity of mammalian neuraminidases. *PLoS One* **2014**, *9*, e106320.
(60) Varghese, J. N.; Chandana Epa, V.; Colman, P. M. Three-dimensional structure of the complex of 4-guanidino-Neu5Ac2en and influenza virus neuraminidase. *Protein Sci.* **1995**, *4*, 1081−1087.
(61) von Itzstein, M.; Wu, W.-Y.; Jin, B. The synthesis of 2,3-didehydro-2,4-dideoxy-4-guanidinyl-N-acetylneuraminic acid: A potent influenza virus sialidase inhibitor. *Carbohydr. Res.* **1994**, *259*, 301−305.
(62) von Itzstein, M.; Wu, W.-Y.; Kok, G. B.; Pegg, M. S.; Dyason, J. C.; Jin, B.; Van Phan, T.; Smythe, M. L.; White, H. F.; Oliver, S. W.; Colman, P. M.; Varghese, J. N.; Ryan, D. M.; Woods, J. M.; Bethell, R. C.; Hotham, V. J.; Cameron, J. M.; Penn, C. R. Rational design of potent sialidase-based inhibitors of influenza virus replication. *Nature* **1993**, *363*, 418−423.
(63) Pshezhetsky, A. V.; Potier, M. Association of N-acetylgalactosamine-6-sulfate sulfatase with the multienzyme lysosomal complex of β-galactosidase, cathepsin A, and neuraminidase: Possible implication for intralysosomal catabolism of keratan sulfate. *J. Biol. Chem.* **1996**, *271*, 28359−28365.
(64) Yoshimitsu, M.; Sato, T.; Tao, K.; Walia, J. S.; Rasaiah, V. I.; Sleep, G. T.; Murray, G. J.; Poeppl, A. G.; Underwood, J.; West, L.; Brady, R. O.; Medin, J. A. Bioluminescent imaging of a marking transgene and correction of Fabry mice by neonatal injection of recombinant lentiviral vectors. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101*, 16909−16914.
(65) Markely, L. R. A.; Ong, B. T.; Hoi, K. M.; Teo, G.; Lu, M. Y.; Wang, D. I. C. A high-throughput method for quantification of glycoprotein sialylation. *Anal. Biochem.* **2010**, *407*, 128−133.
(66) Šali, A.; Blundell, T. L. Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.* **1993**, *234*, 779−815.
(67) Resource for Biocomputing, Visualization, and Informatics. *UCSF Chimera*, candidate version 1.11; University of California, San Francisco, 2016 (supported by NIGMS P41-GM103311).
(68) Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E. C.; Ferrin, T. E. UCSF Chimera − A visualization system for exploratory research and analysis. *J. Comput. Chem.* **2004**, *25*, 1605−1612.
(69) *SWISS-MODEL*; Swiss Institute of Bioinformatics, 2016;https://swissmodel.expasy.org/interactive (accessed Jul 8, 2016).
(70) Biasini, M.; Bienert, S.; Waterhouse, A.; Arnold, K.; Studer, G.; Schmidt, T.; Kiefer, F.; Cassarino, T. G.; Bertoni, M.; Bordoli, L.; Schwede, T. SWISS-MODEL: Modelling protein tertiary and quaternary structure using evolutionary information. *Nucleic Acids Res.* **2014**, *42*, W252−W258.
(71) Kiefer, F.; Arnold, K.; Künzli, M.; Bordoli, L.; Schwede, T. The SWISS-MODEL Repository and associated resources. *Nucleic Acids Res.* **2009**, *37*, D387−D392.
(72) Bordoli, L.; Kiefer, F.; Arnold, K.; Benkert, P.; Battey, J.; Schwede, T. Protein structure homology modeling using SWISS-MODEL workspace. *Nat. Protoc.* **2009**, *4*, 1−13.
(73) Arnold, K.; Bordoli, L.; Kopp, J.; Schwede, T. The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. *Bioinformatics* **2006**, *22*, 195−201.
(74) *Avogadro: An Open-Source Molecular Builder and Visualization Tool*, version 1.1.1; 2012; http://avogadro.openmolecules.net/.
(75) Hanwell, M. D.; Curtis, D. E.; Lonie, D. C.; Vandermeersch, T.; Zurek, E.; Hutchison, G. R. Avogadro: An advanced semantic chemical editor, visualization, and analysis platform. *J. Cheminformatics* **2012**, *4*, 17.
(76) Case, D. A.; Berryman, J. T.; Betz, R. M.; Cerutti, D. S.; Cheatham, T.E., III; Darden, T. A.; Duke, R. E.; Giese, T. J.; Gohlke, H.; Goetz, A. W.; Homeyer, N.; Izadi, S.; Janowski, P.; Kaus, J.; Kovalenko, A.; Lee, T. S.; LeGrand, S.; Li, P.; Luchko, T.; Luo, R.; Madej, B.; Merz, K. M.; Monard, G.; Needham, P.; Nguyen, H.; Nguyen, H. T.; Omelyan, I.; Onufriev, A.; Roe, D. R.; Roitberg, A.; Salomon-Ferrer, R.; Simmerling, C. L.; Smith, W.; Swails, J.; Walker, R. C.; Wang, J.; Wolf, R. M.; Wu, X.; York, D. M.; Kollman, P. A. *AMBER 15*; University of California: San Francisco, CA, 2015.
(77) Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins: Struct., Funct., Genet.* **2006**, *65*, 712−725.
(78) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. Development and testing of a general Amber force field. *J. Comput. Chem.* **2004**, *25*, 1157−1174.
(79) Jakalian, A.; Bush, B. L.; Jack, D. B.; Bayly, C. I. Fast, efficient generation of high-quality atomic charges. AM1-BCC model: I. Method. *J. Comput. Chem.* **2000**, *21*, 132−146.
(80) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **1983**, *79*, 926−935.
(81) Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. C. Numerical integration of the cartesian equations of motion of a system with constraints: Molecular dynamics of n-alkanes. *J. Comput. Phys.* **1977**, *23*, 327−341.
(82) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* **1984**, *81*, 3684−3690.
(83) Roe, D. R.; Cheatham, T. E. PTRAJ and CPPTRAJ: Software for processing and analysis of molecular dynamics trajectory data. *J. Chem. Theory Comput.* **2013**, *9*, 3084−3095.
(84) Pan, X.; De Aragao, C. B. P.; Velasco-Martin, J. P.; Priestman, D. A.; Wu, H. Y.; Takahashi, K.; Yamaguchi, K.; Sturiale, L.; Garozzo, D.; Platt, F. M.; Lamarche-Vane, N.; Morales, C. R.; Miyagi, T.; Pshezhetsky, A. V. Neuraminidases 3 and 4 regulate neuronal function by catabolizing brain gangliosides. *FASEB J.* **2017**, *31*, 3467−3483.

6 Conclusion and outlook

Carbohydrate mimetics are a challenging class of molecules for the development of drug candidates. Their large hydrophilic surface and the inherent complexity requires a consequent structural modification to enhance pharmacokinetic properties. The aim of this thesis was to obtain orally bioavailable carbohydrate mimetics and further improve the predictive power of the *in vitro* assays implemented in the PADMET-platform. Specific assays have been advanced (e.g. phase II metabolism, carboxylesterase specificity and metabolite analysis), leading to a PADMET-platform that can address characteristic questions to certain problems. In the case of E-selectin antagonists, the predictive power of the PADMET-platform was further tested in animal studies in mice.

The PK properties of FimH antagonists have been excessively studied previously (1-5). In this thesis, E-selectin antagonists and NEU3 inhibitors are newly assessed by the PADMET-platform to optimize their PK properties. According to the Biopharmaceutics Drug Distribution and Classification System (BDDCS), the currently developed E-selectin and FimH antagonists, as well as the NEU3 inhibitors, can be classified differently, as illustrated in Table 6.1 (6).

Table 6.1 The classification of the different lead structures discussed in this thesis according to the BDDCS (6).

	High solubility	**Low solubility**
High permeability	**Class 1** • Extensive metabolism • Minimal transporter effects **FimH antagonists**	**Class 2** • Extensive metabolism • Efflux transporter effects predominate in the gut, while absorptive and efflux transporter effects occur the in liver
Low permeability	**Class 3** • Poor metabolism • Absorptive transporter effects • Efflux transporter effects possible **E-selectin antagonists**	**Class 4** • Poor metabolism • Absorptive and efflux transporter effects could be important **NEU3 inhibitors**

E-selectin antagonist

The development of E-selectin antagonists is demanding, since they are derived from the highly polar tetrasaccharide sLe^x. Lead development studies done by Norman *et al.* introduced compound GCP69669A, a tetrasaccharide mimetic showing good affinity towards E-selectin with a reduced complexity and polarity (7). GCP69669A, still a highly hydrophilic molecule containing two carbohydrate moieties and a carboxylic acid, was chosen as the starting point for further development of E-selectin antagonists with improved pharmacokinetic properties. Attempts to further reduce the polar surface area and molecular weight in order to achieve the structural properties advocated by Lipinski *et al.* and Veber *et al.* were not successful (Chapter 3.5) (8, 9). Besides the loss in affinity, the molecule did not show sufficient absorption potential in PAMPA and Caco-2 cell based assay.

Therefore, other approaches were followed to improve the pharmacokinetic properties of E-selectin antagonists. Firstly, ester prodrugs were developed (Manuscript 1). The implementation of large aliphatic promoieties increased the passive permeation of the ester prodrugs, reaching high absorption potential in PAMPA. Combining the results of PAMPA, the hydrolysis rate in liver microsomes, and carboxylesterase specificity, a O-methyl-cyclohexyl promoiety was identified as promising starting point. However, all ester prodrugs showed active efflux in a Caco-2 cell based assay. Furthermore, the observed hydrolysis of the ester bond by pancreatic enzymes in a simulated gastrointestinal fluid could be avoided by reducing the size of the lactic acid residue. Additionally, metabolic studies in liver microsomes showed that the promoiety is partly susceptible for oxidation, leading to a metabolite, which no longer can be converted to the active principle by carboxylesterases. Unfortunately, the ester prodrug did not show any oral bioavailability in *in vivo* studies in mice. Nevertheless, the ester prodrug approach was able to show a promising approach to address the short half-life of carbohydrate mimetics. Once in circulation, the slow hydrolysis of the ester prodrug and therefore prolonged release of the active principle led to a longer apparent plasma half-life.

As a different strategy, a bioisosteric replacement of the carboxylic acid (Manuscript 2) was introduced into the ligand scaffold. By forming amides, not only the affinity, but also the pharmacokinetic properties were improved. The amidic replacement enhanced the lipophilicity of the molecules but did not improve the passive permeability. Two different molecules containing a terminal isopropyl moiety on the amide were identified to show active uptake in a Caco-2 cell permeation assay. *In vivo*, the E-selectin antagonist featuring a

cyclobutylamine-2-isoproxy moiety showed the highest oral bioavailability (7.7 %). Furthermore, the molecule has good aqueous solubility and sufficient stability in liver microsomes. Therefore, other structural modifications to improve the affinity were implemented. Even though the active uptake was no longer observable in Caco2 cells, the highly affine compound showed an oral bioavailability of 4.3 % in mice.

The ester prodrug and amide approach showed the importance of the active transport of E-selectin antagonists to reach oral bioavailability. Therefore, currently developed E-selectin antagonists can be classified in the BDDCS as class 3. Even though beneficial for the affinity, the observed drop in solubility of E-selectin antagonists with a CF_3-group in the 6-position of the fucose moiety (Chapter 3.5 and Manuscript 3) has to be monitored carefully.

Overall, the complexity, size, and polarity of the core structure of E-selectin antagonists promotes molecules with low passive permeability. Thus, oral bioavailability most likely has to be achieved by active transport and future work to reach higher oral bioavailability should focus on the identification of the responsible transporter for the active uptake of amidic E-selectin antagonists and further evaluate the potential thereof.

FimH antagonists

Previous work by Kleeb *et al.* showed the potential of biphenyl α-D-manno-pyranosides with a bioisosteric replacement of the terminal carboxylic acid to get orally bioavailable FimH-antagonists with a promising PK/PD profile (3). Further improvements of the arene-arene interactions by forming polyfluorinated FimH antagonists did not alter the physicochemical properties (e.g. solubility, log$D_{7.4}$, passive permeability in PAMPA) of the antagonists (Manuscript 6). Therefore, we assume that those molecules have a similar PK/PD profile as previously reported.

Starting from a well soluble, but insufficiently permeable bioisostere, ester prodrugs attached to the C-6 moiety of the mannose were synthesized (Publication 1) (4). Several short-chain aliphatic acids were identified as useful promoieties for optimizing the intestinal absorption potential, whilst sufficient solubility could be sustained. The downside of acylation of a hydroxyl group of the sugar moiety is the premature cleavage of the promoiety by CES2 located in the enterocytes. Therefore, choosing the appropriate prodrug moiety should be guided by enzymatic stability studies addressing the carboxylesterase specificity.

Overall, the FimH antagonists reported in Manuscript 6 have a high potential to show good pharmacokinetic properties and are classified as class 1 in the BDDCS. As a next step, the primary findings should be confirmed *in vivo* and further ADMET properties, especially the toxicity, should be evaluated.

NEU3 inhibitors

Specific NEU3 inhibitors and ester prodrugs thereof were developed (Publication 2) (10). Analysis of the pharmacokinetic properties suggested that they are unlikely to be orally bioavailable. Permeability assays using artificial membranes (PAMPA) and Caco-2 cells showed that the compounds were not sufficiently permeable and furthermore showed active efflux. The low solubility of 3–5 µg/mL further contributes to poor pharmacokinetic properties. Therefore, the inhibitors have to be classified as class 4 in the BDDCS. Finally, the examination of the liver microsomal stability showed that the ester was only slowly hydrolyzed. In summary, esterification did not improve pharmacokinetic properties sufficiently to allow for oral bioavailability. Future studies should implement modifications of the active principle to enhance permeability and solubility.

References

1. Klein T, Abgottspon D, Wittwer M, Rabbani S, Herold J, Jiang X, et al. FimH antagonists for the oral treatment of urinary tract infections: from design and synthesis to in vitro and in vivo evaluation. *J Med Chem*. **2010**;*53(24)*:8627-41.
2. Jiang X, Abgottspon D, Kleeb S, Rabbani S, Scharenberg M, Wittwer M, et al. Antiadhesion therapy for urinary tract infections--a balanced PK/PD profile proved to be key for success. *J Med Chem*. **2012**;*55(10)*:4700-13.
3. Kleeb S, Pang L, Mayer K, Eris D, Sigl A, Preston RC, et al. FimH antagonists: bioisosteres to improve the in vitro and in vivo PK/PD profile. *J Med Chem*. **2015**;*58(5)*:2221-39.
4. Schonemann W, Kleeb S, Datwyler P, Schwardt O, Ernst B. Prodruggability of carbohydrates — oral FimH antagonists. *Can J Chem*. **2016***(94)*:10.
5. Kleeb S, Jiang X, Frei P, Sigl A, Bezencon J, Bamberger K, et al. FimH Antagonists: Phosphate Prodrugs Improve Oral Bioavailability. *J Med Chem*. **2016**;*59(7)*:3163-82.
6. Shugarts S, Benet LZ. The role of transporters in the pharmacokinetics of orally administered drugs. *Pharm Res*. **2009**;*26(9)*:2039-54.
7. Norman KE, Anderson GP, Kolb HC, Ley K, Ernst B. Sialyl Lewis(x) (sLe(x)) and an sLe(x) mimetic, CGP69669A, disrupt E-selectin-dependent leukocyte rolling in vivo. *Blood*. **1998**;*91(2)*:475-83.
8. Lipinski CA. Drug-like properties and the causes of poor solubility and poor permeability. *J Pharmacol Toxicol Methods*. **2000**;*44(1)*:235-49.
9. Veber DF, Johnson SR, Cheng HY, Smith BR, Ward KW, Kopple KD. Molecular properties that influence the oral bioavailability of drug candidates. *J Med Chem*. **2002**;*45(12)*:2615-23.
10. Guo T, Datwyler P, Demina E, Richards MR, Ge P, Zou C, et al. Selective Inhibitors of Human Neuraminidase 3. *J Med Chem*. **2018**.