

**Structure Elucidation of Selectin Antagonists in
Solution and Synthesis of Sialyl Lewis^x Mimics Pre-
Organized in their Bioactive Conformation**

Inauguraldissertation

zur

Erlangung der Würde eines Doktors der Philosophie

vorgelegt der

Philosophisch-Naturwissenschaftlichen Fakultät

der Universität Basel

von

Mirko Zierke

aus Frankfurt am Main, Deutschland

Konstanz am Bodensee, Mai 2015

Originaldokument gespeichert auf dem Dokumentenserver der Universität Basel

edoc.unibas.ch

Genehmigt von der Philosophisch-Naturwissenschaftlichen Fakultät
auf Antrag von:

Prof. Dr. Beat Ernst, Institut für Molekulare Pharmazie, Universität Basel
Prof. Dr. Valentin Wittmann, Fachbereich Chemie, Universität Konstanz

Basel, den 21.5.2013

Prof. Dr. Jörg Schibler

Acknowledgement

First of all, I would like to thank Prof. Dr. Beat Ernst for giving me the opportunity to perform my doctoral studies in his group and for his confidence and support during my work on two fascinating projects.

Furthermore, I would like to thank Prof. Dr. Valentin Wittmann for accepting to be the co-referee of my thesis.

Special thanks to our collaborators Prof. Dr. Frédéric Allain and Dr. Thomas Aeschbacher of the ‘Glycocode’ team and especially to Dr. Mario Schubert who introduced me to the field of structural biology.

I am grateful to Bea Wagner, Claudia Huber, Gabi Lichtenhahn and Dr. Oliver Schwardt for their administrative and technical support.

Many thanks to all current and former members of the IMP for the great working atmosphere and the outstanding, interdisciplinary research environment. It was a pleasure to work with many IMP members in the E-selectin project, thanks to Dr. Martin Smieško, Dr. Roland Preston, Dr. Brian Cutting, Dr. Norbert Varga, Dr. Florian Binder, Bea Wagner and Dr. Said Rabbani. Furthermore, I would like to thank the members of the ‘007-lab’ Kathi, Lijuan, Xiahua, Flo and Norbert and also the rest of the IMP for the unique time spend together in and outside the institute, especially Kathi, Arjan, Jacqueline, Lijuan, Steffi, Meike, Xiahua, Adam, Fan, Florian, Jonas, Giulio, Matthias, Norbert, Roland, Simon and Wojtek.

The Chinese Academic Basketball team I would like to thank to accommodate a ‘long nose’ in their team with remarkable friendliness and also to give me a small insight into their culture.

I want to thank my friends Claudia, Sarah, Sebastian, Tobias, my sister Lisa Marie and my brothers Florian and Niklas for the spare time spend together at ‘the Lake’, and finally my parents who always believe in me and support me in every respect.

Abbreviations

AcOH	Acetic acid	LPS	Lipopolysaccharides
ADMET	Acyclic diene metathesis polymerization	Lys	Lysine
aq.	Aqueous	MD	Molecular dynamic
BME	Basal Medium Eagle	Me	Methyl
BSA	Bovine serum albumin	MP2	Second-order Møller-Plesset perturbation theory
BSSE	Basis set superposition error	MW	Molecular weight
CM	Crossmetathesis	NaOH	Sodium hydroxid
CP	Counterpoise correction	NMR	Nuclear magnetic resonance
CRD	Carbohydrate recognition domain	NOE	Nuclear Overhauser effect
d	Days	NOESY	Nuclear Overhauser effect correlation spectroscopy
D-Gal	D-Galactose	NRB	Non rotatable bonds
D-GlcNAc	N-Acetylglucosamine	PCC	Pyridinium chlorochromate
D-Neu5NAc	N-Acetyl neuraminic acid, sialic acid	PPTS	Pyridinium p-toluenesulfonate
DCM	Dichloromethane	PSA	Polar surface area
DDQ	2,3-Dichloro-5,6-dicyano-1,4-benzoquinone	PSGL-1	P-selectin glycoprotein ligand 1
DFT	Density functional theory	py	Pyridine
DIBAL-H	Diisobutylaluminium hydride	RCM	Ring closing metathesis
DMAP	4-Dimethylaminopyridine	RDC	Residual dipolar coupling
DME	Dimethoxyethane	rIC ₅₀	Relative IC ₅₀
DMF	N,N-Dimethylformamide	RO5	Lipinski rule of five
DMSO	Dimethylsulfoxide	ROESY	Rotating frame nuclear Overhauser effect correlation spectroscopy
DMTST	Dimethyl(methylthio)sulfonium triflate	ROMP	Ring-opening metathesis polymerisation
DTBMP	2,6-Di-tert-butyl-4-methylpyridine	RT	Room temperature
ee	Enantiomeric excess	s	Second
EGF	Epidermal growth factor domain	SASA	Solvent accessible surface area
ESL-1	E-selectin ligand 1	Satd.	Saturated
Glu	Glutamic acid	SCRs	Short consensus repeats
GlyCAM-1	Glycosylation-dependent cell adhesion molecule-1	sLe ^a	Sialyl Lewis ^a
HEPES	4-(2-Hydroxyethyl)-1-piperazineethanesulfonic acid	sLe ^x	Sialyl Lewis ^x
HPLC	High performance / pressure liquid chromatography	TBAF	<i>Tert</i> -butylammoniumfluorid
Hz	Hertz	TBS	<i>Tert</i> -butyldimethylsilyl
L-Fuc	L-Fucose	Tf	Triflate, (trifluoromethanesulfonate)
Le ^a	Lewis ^a	TFA	Trifluoroacetic acid
Le ^x	Lewis ^x	THF	Tetrahydrofuran
LogP	Octanol-water partition coefficient	TLC	Thin-layer chromatography
		TMS	Trimethylsilyl
		TNF- α	Tumor necrosis factor- α
		UV	Ultraviolet

Table of Contents

CHAPTER 1. – ABSTRACT	1
CHAPTER 2. – INTRODUCTION	5
2.1 Selectins	7
2.1.1 Structure	7
2.1.2 Physiological and pathophysiological role	8
2.1.3 Natural selectin binding epitop sialyl Lewis ^x	10
2.2 Selectin antagonists	11
2.2.1 Evolution of sLe ^x mimics	12
2.3 Macrocycles in drug design	14
2.3.1 Olefin metathesis	15
2.4 Structure determination of carbohydrates by NMR spectroscopy	17
CHAPTER 3. – RESULTS AND DISCUSSION	27
3.1. Evaluating the solution conformation of Lewis^x, core of lead structure sialyl Lewis^x	29
3.1.1 Stabilization of branched oligosaccharides: Lewis ^x benefits from a nonconventional C-H \cdots hydrogen bond (Publication)	29
3.1.1.1 Supporting Information	39
3.1.2 A secondary structure in a wide range of fucosylated glyco-epitopes (Manuscript)	65
3.1.2.1 Supporting Information	77
3.2 Acid pharmacophore orientation in sialyl Lewis^x mimics and synthesis of antagonists pre-organized in their bioactive conformation	111
3.2.1 “Bridging the gap” – Adjusting the acid pharmacophore in sLe ^x mimics by ring closing metathesis (Manuscript)	111
3.2.1.1 Supporting Information	129
3.2.1.2 Additional Experiments: Ring closing metathesis trials for amide and ester series	187
3.2.2 Bioisosteric modification of the acid pharmacophore in cyclic selectin antagonists	203
CHAPTER 4. – SUMMARY AND OUTLOOK	215

Chapter 1. – Abstract

Selectins are an extensively studied class of carbohydrate binding proteins. They mediate the first contact and rolling of leukocytes on endothelia cells, initiating leukocyte infiltration from blood circulation to the diseased or infected tissue. Many inflammatory diseases are associated with an excessive extravasation of leukocytes to the inflamed tissue, and several kinds of metastatic cancer adopt the selectin mediated pathways. Thus, blocking selectins with synthetic antagonists is a promising therapeutic approach.

The common carbohydrate epitope present in all physiological selectin ligands is the tetrasaccharide sialyl Lewis^x. To overcome the problem of structural complexity, nature rigidifies the 3 dimensional Le^x core conformation by several stabilizing elements and thus, pre-organizes the pharmacophores in their bioactive conformation. For the development of potent sLe^x mimics it is crucial to understand the principles of stabilization and to adopt and optimize these structural motives.

Schematically sLe^x can be divided into a rigid Le^x trisaccharide core and a rather flexible neuraminic acid part. The first part of the thesis is focused on the core trisaccharide Le^x, which bears five of six pharmacophores. Since there are contradictory reports about the conformation of Lewis antigens in solution, i.e. do they adopt a single conformation or are they flexible, the conformations of Le^x and related oligosaccharides in solution were investigated.

- An universal approach to analyze conformations of small molecules at room temperature in solution was developed. By converting the small Le^x trisaccharide in a high-molecular-weight glycoconjugate the tumbling time was drastically increased, which is essential to obtain sufficient structural information by NMR spectroscopy. Thus, we achieved a well-defined solution conformation of Le^x, which disclosed a nonconventional CH \cdots O hydrogen bond as a major stabilizing element (Chapter 3.1.1 /**Publication**).
- It could be shown that nonconventional CH \cdots O hydrogen bonds are a common structural element stabilizing the conformation of various branched oligosaccharides. A widespread database search revealed numerous fucosylated carbohydrate structures that fulfill the requirements of nonconventional CH \cdots O hydrogen bonds. Furthermore, the structures of six representative fucosylated carbohydrates were elucidated in solution. All of them are stabilized by nonconventional CH \cdots O hydrogen bonds (Chapter 3.1.2/**Manuscript**).

In the second part of this thesis, the effects of neuraminic acid replacements in sLe^x mimics on conformational flexibility were evaluated, and antagonists with an acid pharmacophore pre-organized in the bioactive conformation were synthesized.

- By solving the solution conformation of a potent selectin antagonist, it was shown that the acid pharmacophore is pre-organized prior to binding by intramolecular interactions of hydrophobic residues. This antagonist showed stronger binding affinities over mimics with a flexible acid moiety. Based on these results, a series of selectin antagonists was designed and synthesized, where the acid pharmacophore is incorporated in a ring system and therefore, locked in the bioactive conformation (Chapter 3.2.1/**Manuscript**).

The synthesis of two additional cyclic selectin antagonist series was not successful. Possible reasons are discussed in chapter 3.2.1.2.

- In chapter 3.2.2, the potential for a bioisosteric replacement of the acid pharmacophore in cyclic selectin antagonists was evaluated by synthesis, biological assays and *ab initio* calculations (**Manuscript**).

Chapter 2. – Introduction

2.1 Selectins

Selectins are a family of Ca^{2+} dependent C-type transmembrane glycoproteins.¹ They are the most intensively studied mammalian carbohydrate binding lectins,² and can be divided in the three subclasses E-selectin (CD62E, ELAM-1, LECAM-2), P-selectin (CD62P, LECAM-3), and L-selectin (CD62L, LAM-1, LECAM-1).³

2.1.1 Structure

Selectins share a similar topology of a short cytoplasmatic tail, a transmembrane domain, several short consensus repeats (SCRs), an epidermal growth factor domain (EGF) and a N-terminal carbohydrate recognition domain (CRD).⁴

The N-terminal CRD is responsible for ligand binding and coordinates the Ca^{2+} ion.⁴ The function of the adjacent EGF domain is not solved completely, but the domain has an influence on ligand specificity and affinity.¹ The SCRs extend the selectins beyond the glycocalix and differ in number and length between the three selectin subclasses and species. Humans P-selectin consists of 9, L-selectin of 2 and E-selectin of 6 SCR molecules.³ The transmembrane domain anchors the selectins in the membrane and the cytoplasmatic tail is involved in signaltransduction.^{5,6}

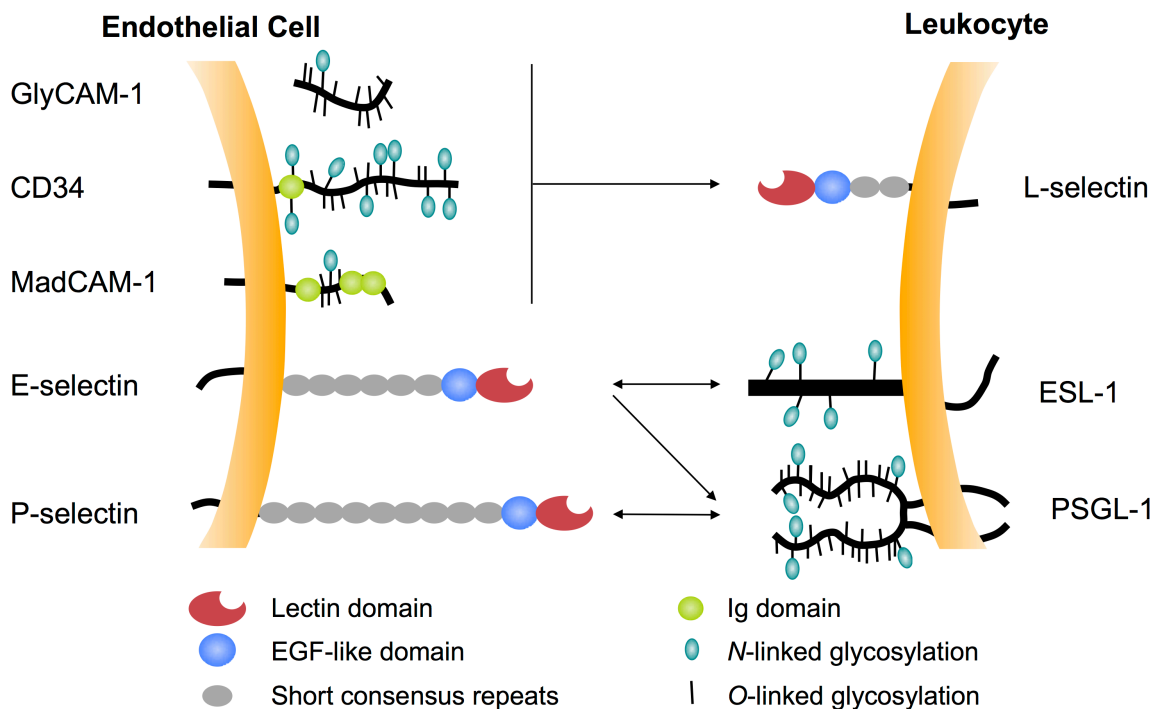


Figure 1⁷. Schematic overview of the three selectin subclasses and binding partners.

The natural ligands of the three selectins are glycolipids and glycoproteins bearing the terminal tetrasaccharide epitops sialyl Lewis^x (sLe^x) or sialyl Lewis^a (sLe^a).^{8,9} They show relatively weak entropic driven binding and fast binding kinetics to their natural ligands. Ligands that bind L- and P-selectin, requires additional negatively charged groups like sulfates or carboxylates for tight binding.

E-selectin is expressed on vascular endothelial cells after an inflammatory stimulus, and binds E-selectin ligand-1 (ESL-1), that is not recognized by P- and L-selectin.¹⁰⁻¹² The dissociation constant for E-selectin ESL-1 binding is 62 μM , the dissociation rate constant k_{off} is 4.6 s^{-1} and the association rate constant k_{on} $7.4 \times 10^4 \text{ M}^{-1} \text{ s}^{-1}$.¹²

P-selectin is expressed on platelets¹³ and on vascular endothelia cells,¹⁴ and binds the P-selectin glycoprotein ligand 1 (PSGL-1) that is stored in Weible palade bodies in leukocytes.¹⁵ Three sulfated tyrosin groups ensure tight binding of PSGL-1.¹⁶ PSGL-1 is also recognized by E-selectin, but the binding is weaker, due to the lack of the secondary binding site for the sulfated tyrosine groups.²⁶ The K_D for P-selectin PSGL-1 binding is 0.3 μM , the binding kinetic is faster than the E-selectin ESL-1 interaction, k_{off} is 1.4 s^{-1} and the k_{on} $4.4 \times 10^6 \text{ M}^{-1} \text{ s}^{-1}$.¹⁶

L-selectin is expressed on monocytes, blood neutrophils, and on T- and B-cells¹⁷ and binds sulfated ligands, *e.g.* addressin (MAdCAM-1),^{18,19} CD34,²⁰ endomucin,²¹ endoglycan²² and glycosylation-dependent cell adhesion molecule-1 (GlyCAM-1).²³

2.1.2 Physiological and pathophysiological role

Selectins are key players in the adaptive and innate immune response, more precisely in the first steps of the inflammation cascade (Figure 2).²⁴ They mediate the initial contact and rolling of leucocytes on blood vessels, a crucial step for the infiltration into inflamed and infected tissue. After an inflammatory stimulus P- and E-selectin close to the side of lesion are displayed on endothelia cells in blood vessels and interact with selectin ligands located on leucocytes, leading to rolling along the endothelia vessels.²⁵ Cytokine mediated activation of $\beta 2$ -integrins on the surface of the leucocytes and binding to ICAM-1 and VCAM-1 on the endothelial surface lead to firm adhesion and finally extravasation of leucocytes to the inflamed tissue.^{24,25}

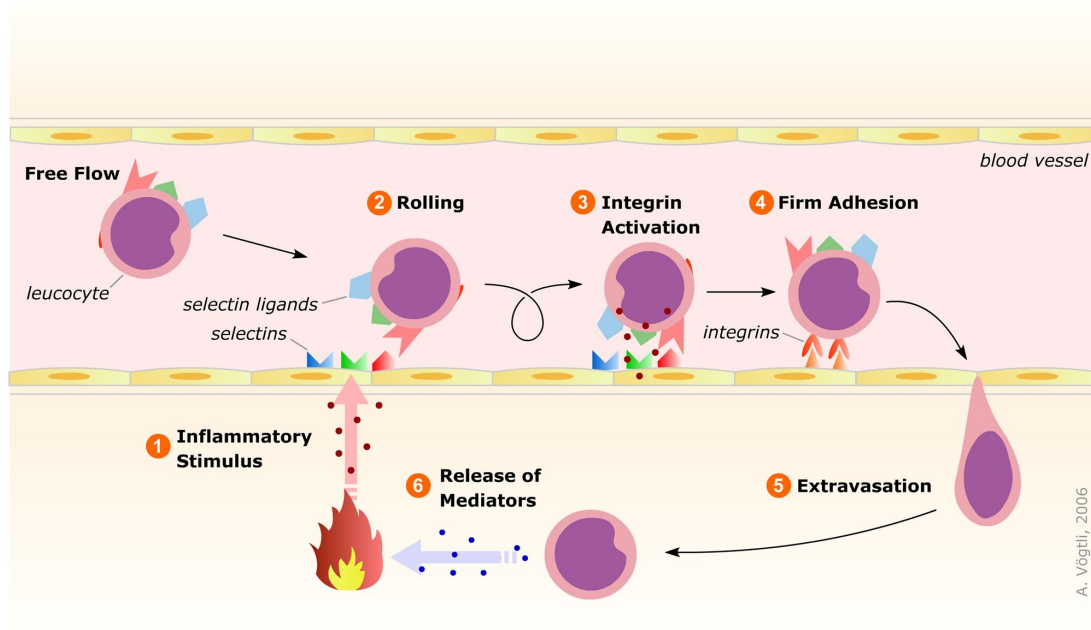


Figure 2. Schematic overview of the inflammatory cascade.²⁶

All three selectins mediate the rolling of leucocytes,³ but they differ in spatially and temporally expression pattern.¹

P-selectins are pre-stored in Weibel palade bodies in the endothelia cells, are released within minutes after injury or inflammation and activate these cells to become adhesive for leucocytes.²⁷ Furthermore, lipopolysaccharides (LPS), tumor necrosis factor- α (TNF- α) and various interleukins can induce *de novo* synthesis and expression on the endothelial surface after 2 - 4 hours.^{28,29}

E-selectin is expressed on endothelia cells 2 - 4 hours after stimulation with immunolatory modulators like TNF- α or interleukin-1, by *de novo* protein synthesis.³⁰ While P-selectin mediates first contacts to the leucocytes and fast rolling, E-selectin allows slow rolling, which seems to be a prerequisite for firm adhesions.³¹

L-selectin is constantly expressed on leucocytes and is involved in trafficking and homing between the lymphatic system and blood.¹ Furthermore, free flowing leucocytes can interact with already adherent ones by L-selectin PSGL-1 interaction, and thus increase the number of adherent leucocytes by this secondary tethering.³²

An aberrant recruitment or hyperactive functioning of effector immune cells characterizes inflammation and causes destruction of healthy tissue.³³ Inappropriate activity of selectins can be associated with a number of acute and chronic inflammatory disorders. Examples are rheumatoid arthritis,³⁴ reperfusion injury,³⁵⁻³⁷ asthma,³⁸ diabetes³⁹ and atherosclerosis.⁴⁰ Furthermore, it was shown, that selectins mediate aggregation of erythrocytes and

leukocytes during vaso-occlusive crisis in a mouse model of sickle cell disease⁴¹ and that tumor cells exploit selectin pathways to extravasate out of the bloodstream.^{42,43} Thus, selectin inhibition can be considered as a promising approach to treat these diseases.

2.1.3 Natural selectin binding epitop sLe^x

The binding affinities of the sLe^x tetrasaccharide is low, with 0.3 to 1.1 mM for E-selectin, 6.8 to 8.8 mM for P-selectin and 3.3 to 4.5 mM for L-selectin.⁴⁴ Nature overcomes this problem of low affinity ligands by multivalency.^{45,46}

The binding mode of sLe^x to E-selectin was intensively studied by NMR spectroscopy⁴⁷⁻⁴⁹ and confirmed by crystallography.⁵⁰ The pharmacophores of sLe^x involved in binding were identified (Figure 3). The L-fucose moiety coordinates the Ca²⁺ ion with two hydroxy groups in 3- and 4-position and the 2-hydroxy group forms an additional hydrogen bond mediated by a water molecule. The 4- and 6-hydroxy groups of the D-galactose are involved in hydrogen bonding and the carboxylate of the neuraminic acid form a salt bridge to the protein. *N*-acetyl-glucosamine is not directly involved in binding and shows no interaction to the protein.

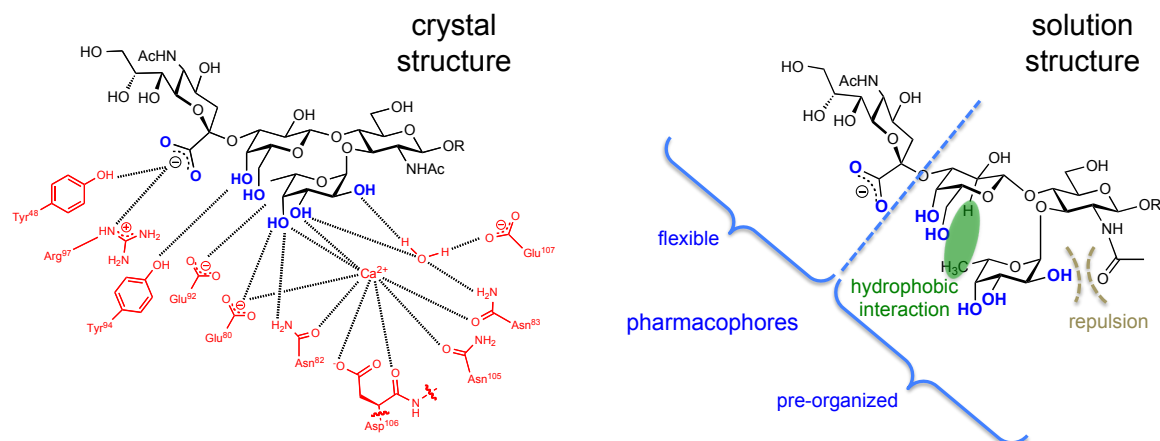


Figure 3. Schematic representation of sLe^x in crystal with E-selectin (amino acids involved in binding are highlighted in red, pharmacophores in blue) and in solution.

The sLe^x/selectin interactions take place under flow conditions, thus, the time to organize the pharmacophores in the bioactive conformation is limited. To overcome this problem, the tetrasaccharide sLe^x adopts a certain degree of pre-organization already in solution. Several NMR and MD studies explored the conformation of sLe^x in solution.^{44,48,51} The tetrasaccharide can be roughly divided in two parts, a rather flexible neuraminic acid

(Neu5NAc) and a relative rigid Le^x (Galβ(1-4)[Fucα(1-3)]GlcNAc) trisaccharide core (Figure 3).

It is hypothesized that the neuraminic acid residue adopts two main conformations in solution, whereof one is identical with the bioactive conformation.⁵¹ In contrast, the Le^x conformation in solution is almost identical to the conformation observed in the complex with E-selectin, *i.e.* the 5 pharmacophoric groups of the L-fucose and D-galactose residue are pre-organized in the bioactive conformation. The Le^x trisaccharide core is mainly stabilized by an exo-anomeric effect,^{52,53} hydrophobic interactions between the L-fucose and β-face of the D-galactose as well as the spatial proximity of the *N*-acetyl residue that restricts the movements of the L-fucose residue.⁵⁴⁻⁵⁶ The GlcNAc residue, which is not involved in protein binding, provides an ideal 3D scaffold for the L-fucose and D-galactose. However, it was discussed extensively in the past, if the previous reported stabilizing elements are sufficient for stabilizing the Le^x trisaccharide in solution. The dihedral angles deviations of the glycosidic bonds between reported Le^x conformations in solution cover a rather large range.⁵⁴⁻⁵⁸ But this could also be justified with drawbacks of NMR spectroscopy elucidating small to midsize molecules (chapter 2.4).

2.2 Selectin antagonists

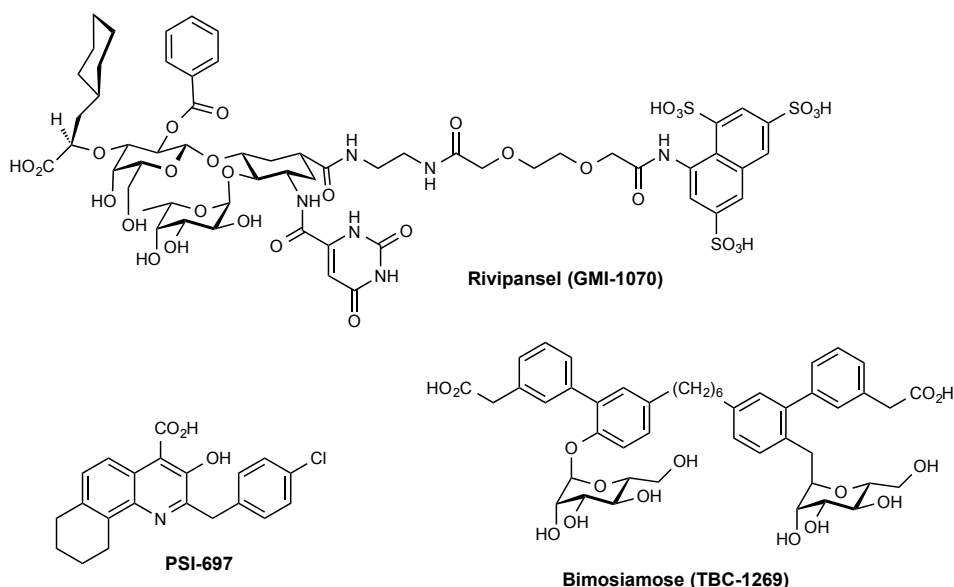


Figure 4: Selected selectin antagonists in clinical trials.

As a result of their high polarity, carbohydrates are unable to passively pass through the enterocyte layer in the small intestine, and if administered parenterally, they suffer from fast

renal excretion.² When interactions with blood plasma components are possible, the plasma half-life required for a successful therapeutic application can be achieved.²

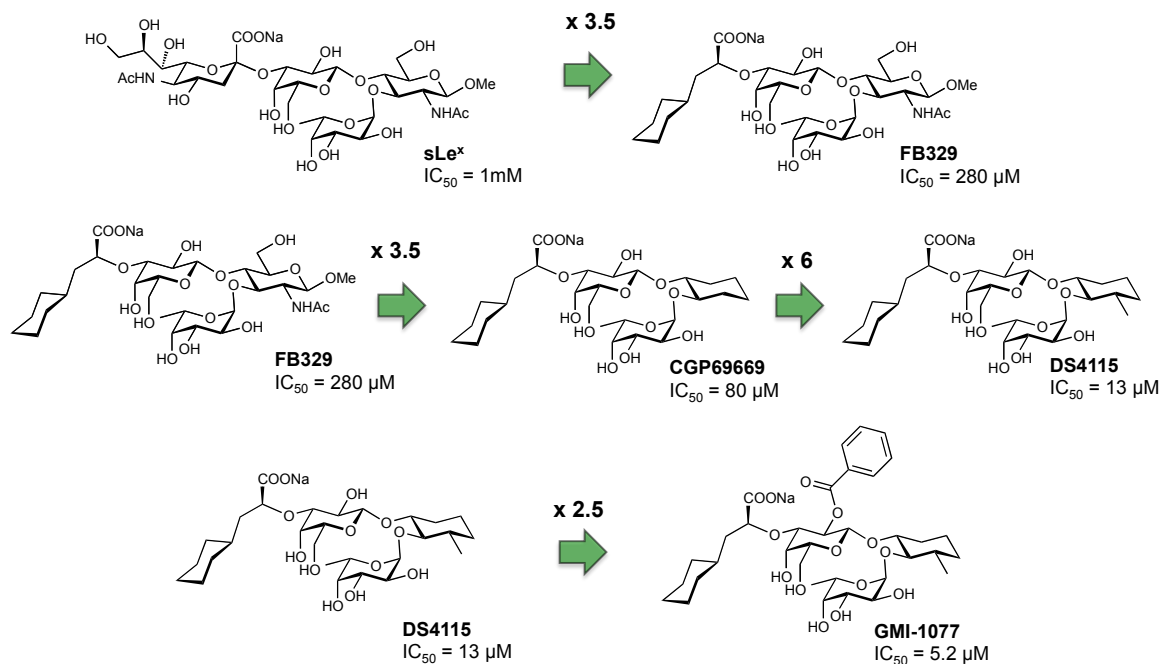
Several synthetic selectin antagonists are currently in clinical evaluation, *i.e.* the pan-selectin antagonist **Rivipansel (GMI-1070)**, which successfully finished clinical phase II to treat vaso-occlusive crisis in sickle cell disease,^{59,60} or the D-mannose based selectin antagonist **TBC-1269**.⁶¹ Furthermore, the non-carbohydrate antagonist **PSI-697**, which shows more drug like properties, was developed and is in clinical phases to treat atherothrombotic and venous thrombotic diseases.^{62,63}

As an alternative approach, biologicals (antibodies^{64,65}, glycoproteins⁶⁶) are currently explored for the treatment of selectin-mediated diseases.

2.2.1 Evolution of sLe^x mimics

The high polarity, complexity and low affinity of the lead structure are challenges for the development of sLe^x mimics. The goal is to mimic structural information of a functional carbohydrate on the one hand and to improve the physicochemical properties on the other hand.² Starting from sLe^x, several compounds were synthesized to mimic the neuraminic acid moiety. (*S*)-cyclohexyl lactic acid turned out to be the most potent substitution (Scheme1; **FB329**²⁶). The binding affinity could be increased by a factor of 3.5.

The (1*R*, 2*R*)-cyclohexane diol mimic the GlcNAc chair conformation adjusting the L-fucose and D-galactose moieties in a position similar as in sLe^x and lead to a 3.5 fold increase in affinity (**FB329**→**CGP69669**). An additional methyl group in 3-position of the (*S*)-cyclohexane residue gave a six-fold improvement (**CGP69669**→**DS4115**) of affinity by forcing the adjacent L-fucose moiety into the bioactive conformation.⁶⁷ A benzoate substituent in 2'-position of the D-galactose, led to an improvement in binding affinity by a factor of 2.5 (**DS4115**→**GMI-1077**).⁶⁷ Thoma *et al.* firstly described these positive effect of 2' benzylation on binding affinity and explained it with a stabilizing effect on the Le^x core.⁶⁸ STD experiments indicated an interaction of the benzoate residue to the protein.⁶⁹ However, this observation could not be confirmed by a recently solved crystal structure.⁷⁰



Scheme 1: Design of sLe^x antagonists.

The benzoate in the 2'-position of the D-galactose moiety is pointing to solution and intramolecular σ - π stacking between the cyclohexane and the aromatic residue was observed.⁷⁰ This benzoate-cyclohexane interaction could pre-organize the structure already before binding, and therefore lead to an entropic gain over flexible ligand **DS4115**.

Recently, E-selectin antagonists with a affinity up to $0.03\ \mu\text{M}$ could be developed by a fragment based drug discovery approach.⁷¹

Although the affinity could be increased significantly and the high polarity could be decreased compared to lead structure sLe^x, druglikeness is still far from being reached. For the prediction of oral availability, several physicochemical rules and filters are available. The most common one is the Lipinski rule of 5 (RO5).^{72,73} Lipinski analyzed a dataset of orally available drugs and clinical candidates for their physicochemical properties, and defined cutoffs for 4 different parameters to be important for solubility and permeability. Oral available drugs that are actively transported were excluded in this study. The rule states that an oral available drug has usually less than 10 hydrogen bond donors, less than 5 hydrogen bond donors, a molecular weight smaller than 500 and an octanol-water partition coefficient ($\log P$) smaller than 5.⁷³

Another rule associated oral availability of drugs with less than 10 rotatable bonds (NRB) and a polar surface area (PSA) smaller than $140\ \text{\AA}^2$.⁷⁴ However, subsequent studies showed, that this limit of 10 rotatable bonds is less stringent in several therapeutic areas,⁷⁵ and that

the upper limits of free rotatable bonds in drugs in humans is 13.⁷⁶ In summary, these rules are just indicators and exceptions will always exist. The estimated success of such predictivity tools is approximately 65 %.⁷⁷

An overview of the RO5, PSA and NRB values of the selectin antagonists **GMI-1070**, **sLe^x**, **GMI-1077**, **FB329** and **DS4115** reveal several violations of the rules mentioned above (Table 1). Strategies to improve the bioavailability are obviously directed towards replacement of the pharmacophores by bioisosteres,⁷⁸ or the reduction of free rotatable bonds by intramolecular cyclisations.

Table 1. Pharmacokinetic and structural parameters for selectin antagonists.

Compound	sLe ^x	GMI-1070	FB329	CGP69669	DS4115	GMI-1077	
R	MW	793.7	1447.4	697.7	578.7	592.7	696.8
O	HD	24	13	9	6	6	5
5	HA	25	37	18	13	13	14
	logP	-6.7*	-5.6*	-3.1*	-0.2*	-0.1*	2.6
	free rot. bond	25	36	20	16	15	17
	PSA	330*	627*	255*	194*	189*	207*
	Affinity (IC ₅₀)	1	0.0043	0.28	0.08	0.019	0.006

*calculated (Schrödinger)

Color code	Out of range (Deviation)			In range
	> 100%	100% - 20%	< 20%	

2.3 Macrocycles in drug design

The introduction of cyclic scaffolds can improve biological and physicochemical properties of drug candidates.⁷⁹ Ring structures can be divided into macrocyclic (≥ 12 atoms), medium (8-11 atoms) and small rings (< 8 atoms), which display qualitative differences in behavior.⁸⁰ The medium rings are rather dominated by transannular interactions and conformational strains that are not present in macrocycles. However, macrocycles contribute a particular degree of pre-organization due to restricted rotations.^{80,81}

A conformational pre-organization can result in improved affinity and selectivity for protein targets, while preserving sufficient bioavailability to reach intracellular targets. However, despite the proven success of more than 100 marketed drugs, this class has been poorly explored within drug discovery.⁸² The marketed drugs mainly derive from natural products or derivatives closely related to naturally occurring macrocycles.⁸² Examples are the macrolide antibiotic erythromycin, antituberculosis compound rifampicin and the immunosuppressant cyclosporine A.⁸² Reasons that macrocycles are under-explored and

poorly exploited for development of novel drug molecules are their structural complexity and accompanied difficulties in the synthesis of analogues and that the preferential adoption of rule of 5-compliant compounds for screening has become widespread.^{73,82}

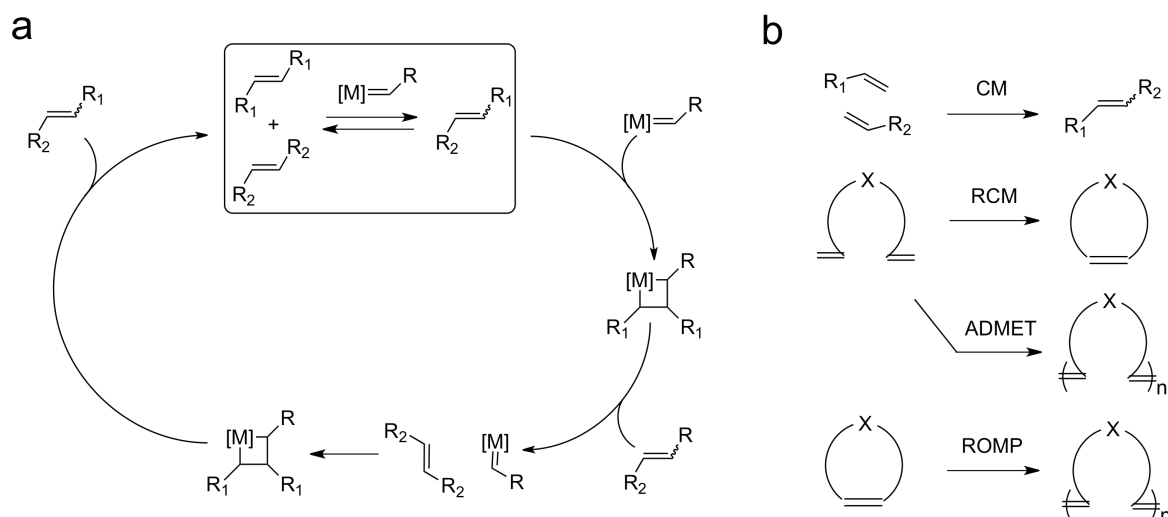
However, rigidified ligands would have an advantageous association rate of binding (k_{on}), since they do not have to reorganize to adopt the bioactive conformation before binding. Furthermore, cyclisation will minimize the unfavorable loss of internal degrees of freedom, and therefore have an entropic gain upon binding. But the enthalpic contribution to the binding event may compensate the entropic term,⁸³ and the assignment and predictions of the energetics are often complex.^{80,84} Several factors like water displacement upon binding and a change of the overall motion of the protein-ligand complex are influencing the thermodynamic fingerprint.⁸⁵

There are several synthetic strategies to insert cyclic scaffolds into drug molecules. The most common is to bridge functionalized termini of linear drug molecules by simple linkers. Examples for common synthetic methods for ring closure are lactonization by Yamaguchi, Mukaiyama and Mitsunobu conditions,⁸⁶ Heck coupling,⁸⁷ copper catalyzed azide-alkyne cycloaddition,^{88,89} and ring closing metathesis (RCM).⁹⁰

2.3.1 Olefin metathesis

The olefin metathesis represents a powerful tool for C-C bond formation and is attracting a vast amount of interest in academia and industry.⁹¹⁻⁹⁵ The mechanism for olefin metathesis was originally proposed by Hérisson and Chauvin and is generally accepted.⁹⁶ According to this mechanism, olefin metathesis proceeds through metallobutane intermediates, generated by the coordination of an olefin to a molybdenum⁹⁷ or ruthenium⁹⁸⁻¹⁰⁰ based alkylidene catalyst, via a series of alternating [2+2]-cycloadditions and cycloreversions (Scheme 2a).^{95,101-103} As all steps in this catalytic cycle are reversible, the main goal is to shift the equilibrium to the desired products. This equilibrium can be influenced by varying the reaction conditions, *e.g.* by choice of catalyst, solvent, temperature or dilution. However, many details of the olefin metathesis reaction still remain unclear.⁹⁵

The most common subtypes of olefin metathesis are illustrated in scheme 2b. The ring-opening metathesis polymerization (ROMP) is the polymerization of monomers with unsaturated strained rings and is driven by ring-strain release.^{94,104}



Scheme 2. Proposed mechanism of the catalytic cycle of olefin metathesis (a), and most common types of olefin metathesis (b).^{91,95}

Olefin cross metathesis (CM) has no enthalpic driving force (like ring strain release) and also no strong entropic benefit. By varying the types of catalysts and the properties of the olefins (steric, electronic) it is possible to reach good product selectivity and stereoselectivity.¹⁰⁵ Chatterjee *et al.* developed a classification system of olefins, in which the olefins can be categorized by their relative abilities to undergo homodimerization via cross metathesis and the susceptibility of their homodimers toward secondary metathesis reactions (Figure 5).¹⁰⁵ Terminal alkenes for example are type I olefins, while acrylates and acrylamides are type II olefins, due to the electron withdrawing groups in proximity to the olefins. The classification differs slightly for different catalysts.¹⁰⁵

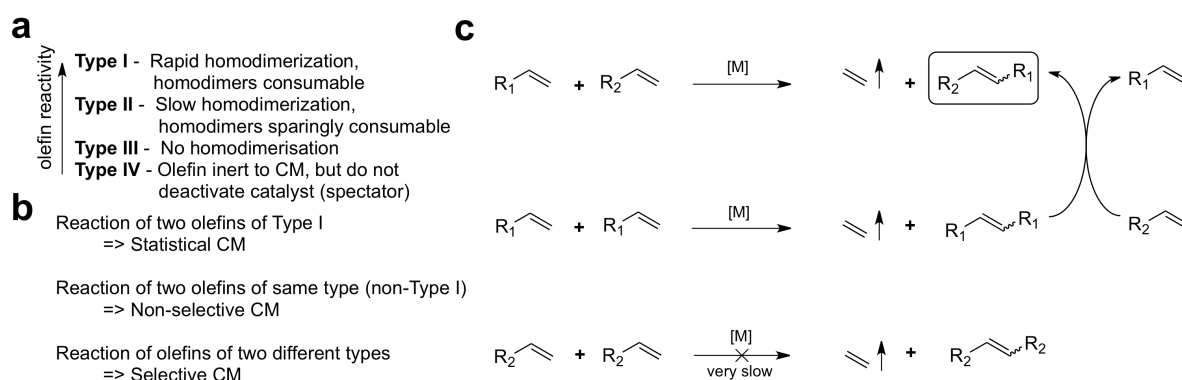


Figure 5. Olefin categorization (a) and rules for selectivity (b). Primary metathesis reactions of Type I (R_1) and Type II (R_2) olefins (c).¹⁰⁵

The ring closing metathesis (RCM) is widely used in organic chemistry^{106,107} and the driving force is predominantly entropic, since one substrate molecule reacts to two product molecules, whereof one is volatile and can leave the equilibrium and thus, the reaction is irreversible and can proceed to completion. The efficiency of the metathesis reaction is influenced by several factors like type of catalyst, concentration, steric congestion near the double bond, presentation of functional groups which serves as a relay entity that assembles the reacting side, and the distance between this polar group and the alkenes to be metathesized.¹⁰⁷

RCM is a common method to introduce a medium- or macrocyclic scaffold in medicinal chemistry. The resulting alkene is in shorter olefin chain almost exclusive the *cis* isomer. With increasing chain length the ratio in favor of *trans* isomer is increasing. The olefin can be further modified,^{108,109} or reduced to the unmodified alkyl chain. By altering the length of the olefin metathesis substrate the size and conformation of the cycle can be easily modified.^{110,111}

2.4 Structure determination of carbohydrates by NMR spectroscopy

The characterization of the conformational properties of carbohydrates is a significant challenge, as they are flexible, and populate multiple (defined) states under physiological conditions.¹¹² For flexible molecules, NMR measurements reflect an average contribution of all conformations. The individual conformations can be discussed by MD simulations.¹¹²

However, for selected oligosaccharides, structures were reported, which have a defined conformation in solution.^{54-58,113}

But even for the probably most extensively studied oligosaccharide, the Le^x trisaccharide epitope, which is assumed to exhibit a rigid conformation in solution (stabilized by different factors like the exo-anomeric effect,^{52,53} hydrophobic interactions of the L-fucose to the β -face of the D-galactose moiety,¹¹⁴ and steric compression of the *N*-acetyl group of the D-GlcNAc residue¹¹⁵) a significant variation in glycosidic torsion angles is reported.⁵⁴⁻⁵⁸

The elucidation of carbohydrate conformations in solution by NMR spectroscopy is challenging due to the tight range of proton resonances between 3 and 4 ppm, which causes several overlaps for oligosaccharides. Furthermore, small oligosaccharides show an unfavorable tumbling time under physiological conditions, which leads to a NOE enhancement factor which is close to zero in nuclear Overhauser effect correlation spectroscopy (NOESY) experiments (Figure 6).¹¹⁷ To overcome the problem of signal

overlay the use of ultra high field spectrometers with their high resolution are useful.¹¹⁸ The unfavorable tumbling time and the resulting poor NOE enhancement factor (Figure 6; Me Le^x trisaccharide (544 Da) is illustrated as a brown line) can be solved by using rotating frame nuclear Overhauser effect correlation spectroscopy (ROESY) experiments, but it has the drawback of smaller theoretical ROEs,¹¹⁹ difficulties in qualifying ROESY cross peaks¹¹⁷ and chemical shift degeneracy.

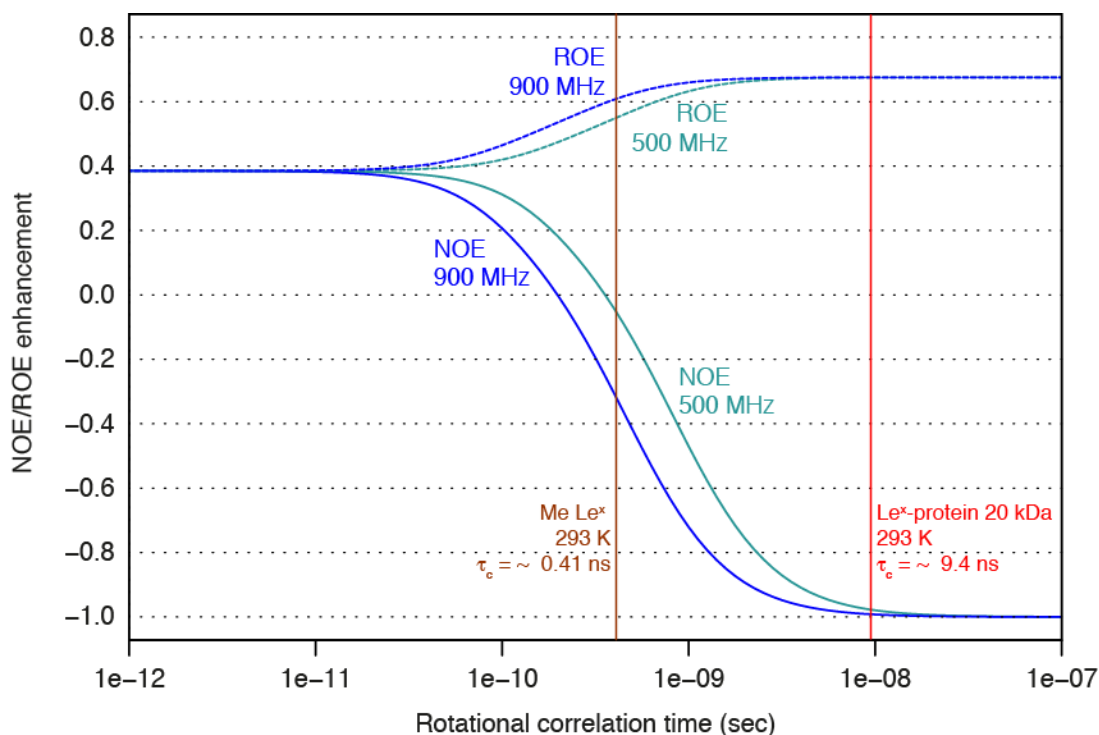


Figure 6¹¹⁶. Maximal NOE and ROE enhancements calculated for a transient NOE experiment at two different field strengths in D₂O.¹¹⁷ The correlations for a NOESY experiment are essentially identical.

When carbohydrates are attached to proteins, an increased NOE transfer is observed, leading to improved NOE cross peaks (Figure 6; Le^x glycan is illustrated by a red line). Slynko *et al.* could elucidate a well-defined structure of a *N*-linked heptasaccharide glycan from a *Camphylobacter jejuni* glycoprotein by attaching an unlabeled heptasaccharide to a ¹³C, ¹⁵N labeled protein using a *in vitro* glycosylation methods.¹¹³ With editing and filtering techniques¹²⁰ they were able to suppress the protein signals and extract numerous NOEs. Using these experimental NOEs, they were able to calculate a well-defined conformation of the oligosaccharide. The drawback of this method is the limitation for substrates, which are accepted by the particular glycosidase. Other carbohydrates and mimics would not be recognized by glycosidases.

The distance restraints that are used for the structural calculations can be translated from the NOE cross peaks assuming a $1/r^6$ dependence of the peak volume (or in practice the peak intensities based on the signal to noise ratios). The distance can be calculated using the equation: $r_{\text{ref}} / r_x = (I_x / I_{\text{ref}})^{1/6}$. r_{ref} is the distance of a defined reference signal (for example the distance between the H61 and H62 protons in GlcNAc which is 1.77 \AA)¹²¹ and I_{ref} the corresponding signal intensity. r_x is the unknown distance and I_x the measured cross peak intensity.

Altogether, NMR spectroscopy can be a useful tool to analyze the solution conformation of oligosaccharides, which display a defined structure and also support the development of carbohydrate-based mimetics.

References

- (1) Kansas, G. *Blood* **1996**, *88*, 3259.
- (2) Ernst, B.; Magnani, J. L. *Nat. Rev. Drug. Discov.* **2009**, *8*, 661.
- (3) Vestweber, D.; Blanks, J. E. *Physiological Reviews* **1999**, *79*, 181.
- (4) Erbe, D. V.; Wolitzky, B. A.; Presta, L. G.; Norton, C. R.; Ramos, R. J.; Burns, D. K.; Rumberger, J. M.; Rao, B. N.; Foxall, C.; Brandley, B. K. *J. Cell Biol.* **1992**, *119*, 215.
- (5) Lorenzon, P.; Vecile, E.; Nardon, E.; Ferrero, E.; Harlan, J. M.; Tedesco, F.; Dobrina, A. *Journal of Cell Biology* **1998**, *142*, 1381.
- (6) Hu, Y. Y.; Szente, B.; Kiely, J. M.; Gimbrone, M. A. *Journal of Biological Chemistry* **2001**, *276*, 48549.
- (7) Kindly provided by Roland C. Preston.
- (8) Phillips, M. L.; Nudelman, E.; Gaeta, F. C. A.; Perez, M.; Singhal, A. K.; Hakomori, S. I.; Paulson, J. C. *Science* **1990**, *250*, 1130.
- (9) Berg, E. L.; Robinson, M. K.; Mansson, O.; Butcher, E. C.; Magnani, J. L. *J. Bio. Chem.* **1991**, *266*, 14869.
- (10) Lenter, M.; Levinovitz, A.; Isenmann, S.; Vestweber, D. *Journal of Cell Biology* **1994**, *125*, 471.
- (11) Levinovitz, A.; Muhlhoff, J.; Isenmann, S.; Vestweber, D. *Journal of Cell Biology* **1993**, *121*, 449.
- (12) Wild, M. K.; Huang, M. C.; Schulze-Horsel, U.; van der Merwe, P. A.; Vestweber, D. *Journal of Biological Chemistry* **2001**, *276*, 31602.

- (13) Hsulín, S. C.; Berman, C. L.; Furie, B. C.; August, D.; Furie, B. *Journal of Biological Chemistry* **1984**, *259*, 9121.
- (14) McEver, R. P.; Beckstead, J. H.; Moore, K. L.; Marshallcarlson, L.; Bainton, D. F. *Journal of Clinical Investigation* **1989**, *84*, 92.
- (15) McEver, R. P. *Thrombosis and Haemostasis* **2002**, *87*, 364.
- (16) Mehta, P.; Cummings, R. D.; McEver, R. P. *Journal of Biological Chemistry* **1998**, *273*, 32506.
- (17) Siegelman, M. H.; Vanderijn, M.; Weissman, I. L. *Science* **1989**, *243*, 1165.
- (18) Streeter, P. R.; Berg, E. L.; Rouse, B. T. N.; Bargatze, R. F.; Butcher, E. C. *Nature* **1988**, *331*, 41.
- (19) Nakache, M.; Berg, E. L.; Streeter, P. R.; Butcher, E. C. *Nature* **1989**, *337*, 179.
- (20) Baumhueter, S.; Singer, M. S.; Henzel, W.; Hemmerich, S.; Renz, M.; Rosen, S. D.; Lasky, L. A. *Science* **1993**, *262*, 436.
- (21) Kanda, H.; Tanaka, T.; Matsumoto, M.; Umemoto, E.; Ebisuno, Y.; Kinoshita, M.; Noda, M.; Kannagi, R.; Hirata, T.; Murai, T.; Fukuda, M.; Miyasaka, M. *International Immunology* **2004**, *16*, 1265.
- (22) Fieger, C. B.; Sasseti, C. M.; Rosen, S. D. *Journal of Biological Chemistry* **2003**, *278*, 27390.
- (23) Imai, Y.; Singer, M. S.; Fennie, C.; Lasky, L. A.; Rosen, S. D. *Journal of Cell Biology* **1991**, *113*, 1213.
- (24) Ley, K.; Laudanna, C.; Cybulsky, M. I.; Nourshargh, S. *Nat Rev Immunol* **2007**, *7*, 678.
- (25) Lawrence, M. B.; Springer, T. A. *Cell* **1991**, *65*, 859.
- (26) Binder, F. P. C., PhD Thesis, Basel, **2011**.
- (27) Wagner, D. D.; Frenette, P. S. *Blood* **2008**, *111*, 5271.
- (28) Bischoff, J.; Brasel, C. *Biochem. Biophys. Res. Commun.* **1995**, *210*, 174.
- (29) Gotsch, U.; Jager, U.; Dominis, M.; Vestweber, D. *Cell Ad. Commun.* **1994**, *2*, 7.
- (30) Bevilacqua, M. P.; Pober, J. S.; Mendrick, D. L.; Cotran, R. S.; Gimbrone, M. A. *PNAS* **1987**, *84*, 9238.
- (31) Springer, T. A. *Annual Review of Physiology* **1995**, *57*, 827.
- (32) Mitchell, D. J.; Li, P.; Reinhardt, P. H.; Kubes, P. *Blood* **2000**, *95*, 2954.
- (33) Ulbrich, H.; Eriksson, E. E.; Lindbom, L. *Trends in Pharmacological Sciences* **2003**, *24*, 640.
- (34) Cronstein, B. N.; Weissmann, G. *Arthritis and Rheumatism* **1993**, *36*, 147.

-
- (35) Shreeniwas, R.; Koga, S.; Karakurum, M.; Pinsky, D.; Kaiser, E.; Brett, J.; Wolitzky, B. A.; Norton, C.; Plocinski, J.; Benjamin, W.; Burns, D. K.; Goldstein, A.; Stern, D. *Journal of Clinical Investigation* **1992**, *90*, 2333.
- (36) Billups, K. L.; Palladino, M. A.; Hinton, B. T.; Sherley, J. L. *Journal of Laboratory and Clinical Medicine* **1995**, *125*, 626.
- (37) Palluy, O.; Morliere, L.; Gris, J. C.; Bonne, C.; Modat, G. *Free Radical Biology and Medicine* **1992**, *13*, 21.
- (38) Gundel, R. H.; Wegner, C. D.; Torcellini, C. A.; Clarke, C. C.; Haynes, N.; Rothlein, R.; Smith, C. W.; Letts, L. G. *Journal of Clinical Investigation* **1991**, *88*, 1407.
- (39) Albertini, J. P.; Valensi, P.; Lormeau, B.; Aourousseau, M. H.; Ferriere, F.; Attali, J. R.; Gattegno, L. *Diabetes Care* **1998**, *21*, 1008.
- (40) Ross, R. *Nature* **1993**, *362*, 801.
- (41) Turhan, A.; Weiss, L. A.; Mohandas, N.; Collier, B. S.; Frenette, P. S. *PNAS* **2002**, *99*, 3047.
- (42) Ludwig, R. J.; Schön, M. P.; Boehncke, W.-H. *Expert Opin. Therap. Targets* **2007**, *11*, 1103.
- (43) Laeubli, H.; Borsig, L. *Seminar Cancer Biol.* **2010**, *20*, 169.
- (44) Poppe, L.; Brown, G. S.; Philo, J. S.; Nikrad, P. V.; Shah, B. H. *J. Am. Chem. Soc.* **1997**, *119*, 1727.
- (45) Lee, R. T.; Lee, Y. C. *Glycoconj. J.* **2000**, *17*, 543.
- (46) Mammen, M.; Choi, S. K.; Whitesides, G. M. *Angew. Chem. Int. Ed.* **1998**, *37*, 2755.
- (47) Scheffler, K.; Ernst, B.; Katopodis, A.; Magnani, J. L.; Wang, W. T.; Weisemann, R.; Peters, T. *Angew. Chem. Int. Ed.* **1995**, *34*, 1841.
- (48) Harris, R.; Kiddle, G. R.; Field, R. A.; Milton, M. J.; Ernst, B.; Magnani, J. L.; Homans, S. W. *J. Am. Chem. Soc.* **1999**, *121*, 2546.
- (49) Scheffler, K.; Brisson, J.-R.; Weisemann, R.; Magnani, J.; Wong, W.; Ernst, B.; Peters, T. *J. Biomol. NMR* **1997**, *9*, 423.
- (50) Somers, W. S.; Tang, J.; Shaw, G. D.; Camphausen, R. T. *Cell* **2000**, *103*, 467.
- (51) Lin, Y. C.; Hummel, C. W.; Huang, D. H.; Ichikawa, Y.; Nicolaou, K. C.; Wong, C. H. *J. Am. Chem. Soc.* **1992**, *114*, 5452.
- (52) Lemieux, R. U. *Molecular rearrangements* Interscience, New York, 1964.
- (53) Wolfe, S.; Whangbo, M. H.; Mitchell, D. J. *Carbohydr. Res.* **1979**, *69*, 1.
- (54) Su, Z.; Wagner, B.; Cocinero, E. J.; Ernst, B.; Simons, J. P. *Chem. Phys. Lett.* **2009**, *477*, 365.

- (55) Azurmendi, H. F.; Martin-Pastor, M.; Bush, C. A. *Biopolymers* **2002**, *63*, 89.
- (56) Imberty, A.; Mikros, E.; Koca, J.; Mollicone, R.; Oriol, R.; Pérez, S. *Glycoconj. J.* **1995**, *12*, 331.
- (57) Wormald, M. R.; Edge, C. J.; Dwek, R. A. *Biochem. Biophys. Res. Commun.* **1991**, *180*, 1214.
- (58) Miller, K. E.; Mukhopadhyay, C.; Cagas, P.; Bush, C. A. *Biochemistry* **1992**, *31*, 6703.
- (59) Chang, J.; Patton, J. T.; Sarkar, A.; Ernst, B.; Magnani, J. L.; Frenette, P. S. *Blood* **2010**, *116*, 1779.
- (60) Telen, M. J.; Wun, T.; McCavit, T. L.; De Castro, L. M.; Krishnamurti, L.; Lanzkron, S.; Hsu, L. L.; Smith, W. R.; Rhee, S.; Magnani, J. L.; Thackray, H. *Blood* **2015**.
- (61) Avila, P. C.; Boushey, H. A.; Wong, H.; Grundland, H.; Liu, J.; Fahy, J. V. *Clin. Exp. Allergy* **2004**, *34*, 77.
- (62) Kaila, N.; Janz, K.; Huang, A.; Moretto, A.; DeBernardo, S.; Bedard, P. W.; Tam, S.; Clerin, V.; Keith, J. C.; Tsao, D. H. H.; Sushkova, N.; Shaw, G. D.; Camphausen, R. T.; Schaub, R. G.; Wang, Q. *J. Med. Chem.* **2006**, *50*, 40.
- (63) Japp, A. G.; Chelliah, R.; Tattersall, L.; Lang, N. N.; Meng, X.; Weisel, K.; Katz, A.; Burt, D.; Fox, K. A. A.; Feuerstein, G. Z.; Connolly, T. M.; Newby, D. E. *Journal of the American Heart Association* **2013**, *2*.
- (64) Friedman, G.; Jankowski, S.; Shahla, M.; Goldman, M.; Rose, R. M.; Jean Kahn, R.; Vincent, J.-L. *Crit. Care. Med.* **1996**, *24*, 229.
- (65) Suzuki, K.; Fukushima, S.; Coppen, S. R.; Yamahara, K.; Varela-Carver, A.; Ermakov, A.; Yacoub, M. H. *J. Mol. Cell. Card.* **2006**, *40*, 975.
- (66) Leppänen, A.; Mehta, P.; Ouyang, Y.-B.; Ju, T.; Helin, J.; Moore, K. L.; van Die, I.; Canfield, W. M.; McEver, R. P.; Cummings, R. D. *J. Bio. Chem.* **1999**, *274*, 24838.
- (67) Schwizer, D.; Patton, J. T.; Cutting, B.; Smieško, M.; Wagner, B.; Kato, A.; Weckerle, C.; Binder, F. P. C.; Rabbani, S.; Schwarzt, O.; Magnani, J. L.; Ernst, B. *Chem. – Eur. J.* **2012**, *18*, 1342.
- (68) Thoma, G.; Banteli, R.; Jahnke, W.; Magnani, J. L.; Patton, J. T. *Angewandte Chemie* **2001**, *113*, 3756.
- (69) Weckerle, C., PhD Thesis, University of Basel, **2010**.
- (70) Preston, R. C., PhD Thesis, University of Basel, **2014**.

-
- (71) Egger, J.; Weckerle, C.; Cutting, B.; Schwardt, O.; Rabbani, S.; Lemme, K.; Ernst, B. *J. Am. Chem. Soc.* **2013**, *135*, 9820.
- (72) Lipinski, C. A. *Drug Discov. Today: Technol.* **2004**, *1*, 337.
- (73) Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J. *Adv. Drug Deliv. Rev.* **2001**, *46*, 3.
- (74) Veber, D. F.; Johnson, S. R.; Cheng, H.-Y.; Smith, B. R.; Ward, K. W.; Kopple, K. *D. J. Med. Chem.* **2002**, *45*, 2615.
- (75) Lu, J. J.; Crimin, K.; Goodwin, J. T.; Crivori, P.; Orrenius, C.; Xing, L.; Tandler, P. J.; Vidmar, T. J.; Amore, B. M.; Wilson, A. G. E.; Stouten, P. F. W.; Burton, P. S. *J. Med. Chem.* **2004**, *47*, 6104.
- (76) Tian, S.; Li, Y.; Wang, J.; Zhang, J.; Hou, T. *Mol. Pharmaceut.* **2011**, *8*, 841.
- (77) Lipinski, C. *10th Swiss Course on Medicinal Chemistry* Leysin, **2012**.
- (78) Meanwell, N. A. *J. Med. Chem.* **2011**, *54*, 2529.
- (79) Aigner, M.; Hartl, M.; Fauster, K.; Steger, J.; Bister, K.; Micura, R. *ChemBioChem* **2011**, *12*, 47.
- (80) Mallinson, J.; Collins, I. *Future Med. Chem.* **2012**, *4*, 1409.
- (81) Smith, B.; March, J. *March's Advanced Organic Chemistry: Reactions, Mechanisms and Structure (5th Edition)* **2001**, 184.
- (82) Driggers, E. M.; Hale, S. P.; Lee, J.; Terrett, N. K. *Nat. Rev. Drug. Discov.* **2008**, *7*, 608.
- (83) Reynolds, C.; de Koning, C. B.; Pelly, S. C.; van Otterlo, W. A. L.; Bode, M. L. *Chem. Soc. Rev.* **2012**, *41*, 4657.
- (84) Bissantz, C.; Kuhn, B.; Stahl, M. *J. Med. Chem.* **2010**, *53*, 5061.
- (85) Williams, D. H.; Calderone, C. T.; O'Brien, D. P.; Zerella, R. *Chem. Commun.* **2002**, *0*, 1266.
- (86) Winssinger, N.; Barluenga, S. *Chem. Commun.* **2007**, *0*, 22.
- (87) Breslin, H. J.; Lane, B. M.; Ott, G. R.; Ghose, A. K.; Angeles, T. S.; Albom, M. S.; Cheng, M.; Wan, W.; Haltiwanger, R. C.; Wells-Knecht, K. J.; Dorsey, B. D. *J. Med. Chem.* **2011**, *55*, 449.
- (88) Bogdan, A. R.; James, K. *Org. Lett.* **2011**, *13*, 4060.
- (89) Chouhan, G.; James, K. *Org. Lett.* **2011**, *13*, 2754.
- (90) Lee, C. W.; Grubbs, R. H. *J. Org. Chem.* **2001**, *66*, 7155.
- (91) Chauvin, Y. *Angew. Chem. Int. Ed.* **2006**, *45*, 3740.
- (92) Grubbs, R. H. *Angew. Chem. Int. Ed.* **2006**, *45*, 3760.

- (93) Schrock, R. R. *Angew. Chem. Int. Ed.* **2006**, *45*, 3748.
- (94) Grubbs, R. H. *Handbook of Metathesis*; Wiley-VCH Verlag GmbH, **2008**.
- (95) Vougioukalakis, G. C.; Grubbs, R. H. *Chem. Rev.* **2010**, *110*, 1746.
- (96) Jean-Louis Hérisson, P.; Chauvin, Y. *Die Makromolekulare Chemie* **1971**, *141*, 161.
- (97) Schrock, R. R.; Murdzek, J. S.; Bazan, G. C.; Robbins, J.; DiMare, M.; O'Regan, M. *J. Am. Chem. Soc.* **1990**, *112*, 3875.
- (98) Schwab, P.; Grubbs, R. H.; Ziller, J. W. *J. Am. Chem. Soc.* **1996**, *118*, 100.
- (99) Scholl, M.; Ding, S.; Lee, C. W.; Grubbs, R. H. *Org. Lett.* **1999**, *1*, 953.
- (100) Garber, S. B.; Kingsbury, J. S.; Gray, B. L.; Hoveyda, A. H. *J. Am. Chem. Soc.* **2000**, *122*, 8168.
- (101) Grubbs, R. H.; Burk, P. L.; Carr, D. D. *J. Am. Chem. Soc.* **1975**, *97*, 3265.
- (102) Katz, T. J.; McGinnis, J. *J. Am. Chem. Soc.* **1975**, *97*, 1592.
- (103) Leconte, M.; Basset, J.-M.; Quignard, F.; Larroche, C. *Mechanistic Aspects of the Olefin Metathesis Reaction. In Reactions of Coordinated Ligands* Plenum: New York, **1986**; Vol. 1.
- (104) Wiberg, K. B. *Angew. Chem. Int. Ed.* **1986**, *25*, 312.
- (105) Chatterjee, A. K.; Choi, T.-L.; Sanders, D. P.; Grubbs, R. H. *J. Am. Chem. Soc.* **2003**, *125*, 11360.
- (106) Nicolaou, K. C.; Bulger, P. G.; Sarlah, D. *Angew. Chem. Int. Ed.* **2005**, *44*, 4490.
- (107) Furstner, A. *Chem. Commun.* **2011**, *47*, 6505.
- (108) Proisy, N.; Sharp, S. Y.; Boxall, K.; Connelly, S.; Roe, S. M.; Prodromou, C.; Slawin, A. M. Z.; Pearl, L. H.; Workman, P.; Moody, Christopher J. *Chem. Biol.* **2006**, *13*, 1203.
- (109) Moulin, E.; Barluenga, S.; Totzke, F.; Winssinger, N. *Chem. – Eur. J.* **2006**, *12*, 8819.
- (110) Tao, Z.-F.; Sowin, T. J.; Lin, N.-H. *Synlett* **2007**, *2007*, 2855.
- (111) Dandapani, S.; Marcaurelle, L. A. *Nat. Chem. Biol.* **2010**, *6*, 861.
- (112) DeMarco, M. L.; Woods, R. J. *Glycobiology* **2008**, *18*, 426.
- (113) Slynko, V.; Schubert, M.; Numao, S.; Kowarik, M.; Aebi, M.; Allain, F. H. T. *J. Am. Chem. Soc.* **2009**, *131*, 1274.
- (114) Titz, A.; Marra, A.; Cutting, B.; Smiesko, M.; Papandreou, G.; Dondoni, A.; Ernst, B. *Eur. J. Org. Chem.* **2012**, 5534.
- (115) Simanek, E. E.; McGarvey, G. J.; Jablonowski, J. A.; Wong, C. H. *Chem. Rev.* **1998**, *98*, 833.

- (116) Kindly provided by Thomas Aeschbacher.
- (117) Neuhaus, D.; Williamson, M. P. *The Nuclear Overhauser Effect in Structural and Conformational Analysis*; Wiley-VCH: New York, **2000**.
- (118) Blundell, C. D.; Reed, M. A. C.; Overduin, M.; Almond, A. *Carbohydr. Res.* **2006**, *341*, 1985.
- (119) Bothner-By, A. A.; Stephens, R. L.; Lee, J.; Warren, C. D.; Jeanloz, R. W. *J. Am. Chem. Soc.* **1984**, *106*, 811.
- (120) Peterson, R.; Theimer, C.; Wu, H.; Feigon, J. *J. Biomol. NMR* **2004**, *28*, 59.
- (121) Brown, G. M.; Levy, H. A. *Acta Crystallogr. Sect. B-Struct. Sci.* **1979**, *35*, 656.

Chapter 3. – Results and Discussion

3.1. Evaluating the solution conformation of Lewis^x, core of lead structure sialyl Lewis^x

3.1.1 Stabilization of branched oligosaccharides: Lewis^x benefits from a nonconventional C-H \cdots hydrogen bond

J. Am. Chem. Soc., 2013, 135 (36), pp 13464–13472

DOI: 10.1021/ja4054702

Publication Date (Web): August 12, 2013

Copyright © 2013 American Chemical Society

Contributions:

- Manuscript preparation
- Compound synthesis
- Chemical glycosylation of the FimH mutant
- NMR spectra assignment and structure calculation

Stabilization of Branched Oligosaccharides: Lewis^x Benefits from a Nonconventional C–H···O Hydrogen Bond

Mirko Zierke,[†] Martin Smieško,[†] Said Rabbani,[†] Thomas Aeschbacher,[‡] Brian Cutting,[†] Frédéric H.-T. Allain,[‡] Mario Schubert,^{*,‡} and Beat Ernst^{*,†}

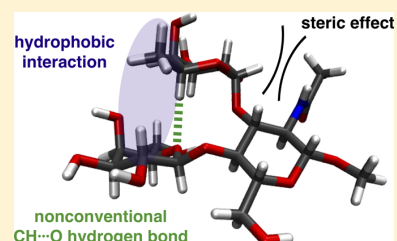
[†]University of Basel, Klingelbergstraße 50, CH-4056 Basel, Basel-City, Switzerland

[‡]Institute of Molecular Biology and Biophysics, ETH Zürich, CH-8093 Zürich, Switzerland

Supporting Information

ABSTRACT: Although animal lectins usually show a high degree of specificity for glycan structures, their single-site binding affinities are typically weak, a drawback which is often compensated in biological systems by an oligovalent presentation of carbohydrate epitopes. For the design of monovalent glycomimetics, structural information regarding solution and bound conformation of the carbohydrate lead represents a valuable starting point. In this paper, we focus on the conformation of the trisaccharide Lewis^x (Gal[Fuca(1–3)]β(1–4)GlcNAc). Mainly because of the unfavorable tumbling regime, the elucidation of the solution conformation of Lewis^x by NMR has only been partially successful so far. Lewis^x was therefore attached to a ¹³C,¹⁵N-labeled protein. ¹³C,¹⁵N-filtered NOESY NMR techniques at ultrahigh field

allowed increasing the maximal NOE enhancement, resulting in a high number of distance restraints per glycosidic bond and, consequently, a well-defined structure. In addition to the known contributors to the conformational restriction of the Lewis^x structure (exoanomeric effect, steric compression induced by the NHAc group adjacent to the linking position of L-fucose, and the hydrophobic interaction of L-fucose with the β-face of D-galactose), a nonconventional C–H···O hydrogen bond between H–C(5) of L-fucose and O(5) of D-galactose was identified. According to quantum mechanical calculations, this C–H···O hydrogen bond is the most prominent factor in stabilization, contributing 40% of the total stabilization energy. We therefore propose that the nonconventional hydrogen bond contributing to a reduction of the conformational flexibility of the Lewis^x core represents a novel element of the glycode. Its relevance to the stabilization of related branched oligosaccharides is currently being studied.



INTRODUCTION

Selectins are probably the most intensely studied mammalian carbohydrate binding proteins. First discovered in 1989,¹ their functions as adhesion molecules in the early stages of inflammation are well understood.² For diseases in which cell adhesion, extravasation of leukocytes from the bloodstream, or migration of specific lymphocytes has been implicated in the pathology, selectins present an attractive therapeutic target.³ The family of selectins consisting of E-, P-, and L-selectin recognizes the common carbohydrate epitope sialyl Lewis^x (Neu5Aca(2–3)Galβ(1–4)[Fuca(1–3)]GlcNAc, sLe^x (1); Figure 1), which is present in all physiological selectin ligands identified so far.⁴ sLe^x (1) was therefore regarded as lead structure for almost 20 years. Countless studies aiming at its optimization into a druglike mimetic have been reported.⁵

Although animal lectins usually display a high degree of specificity for glycan structures, their single-site binding affinities are typically weak. This drawback is often compensated in biological systems by an oligovalent presentation of the carbohydrate epitopes or the carbohydrate recognition domains (CRD) of the lectins.⁶ In addition, the pharmacokinetic properties of carbohydrates such as bioavailability and plasma half-life are typically insufficient for therapeutic applications.³ For the design of druglike mimetics

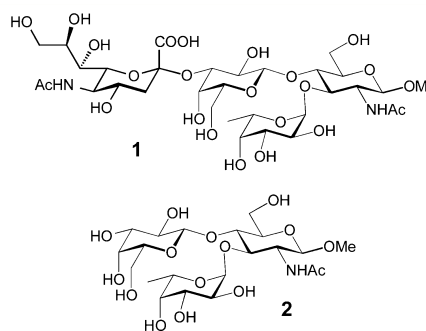


Figure 1. Methyl sialyl Lewis^x (sLe^x, 1) and methyl Lewis^x (Le^x, 2).

structural information regarding the solution and bound conformation of the carbohydrate lead represent a valuable starting point.

The conformation of sLe^x (1) bound to E- and P-selectin was first elucidated by NMR⁷ and later confirmed by X-ray crystallography.⁸ The analysis of the solution conformation of

Received: May 31, 2013

Published: August 12, 2013

sLe^x (1) can be divided into two parts: (i) the conformation of the Le^x core Galβ(1-4)[Fucβ(1-3)]GlcNAc (2) and (ii) the conformation of the glycosidic bond in Neu5Acα(2-3)Gal (Figure 1).

In this paper, we focus on the core conformation of Le^x (2) in solution that is stabilized by two distinct factors. First, the acetyl group of the GlcNAc moiety or equatorial alkyl groups in the 2-position of carbocyclic GlcNAc mimetics restrict the conformational flexibility of the core and therefore entropically improve binding affinities.^{5,9} Second, the methyl group of L-fucose is optimally suited to stabilize the Le^x core through a hydrophobic interaction with the β-face of D-galactose.¹⁰ This structural insight into the solution conformation of Le^x was obtained by molecular dynamics simulation (MD) and NMR spectroscopy,¹¹ as well as X-ray crystallography¹² (Figure 2).

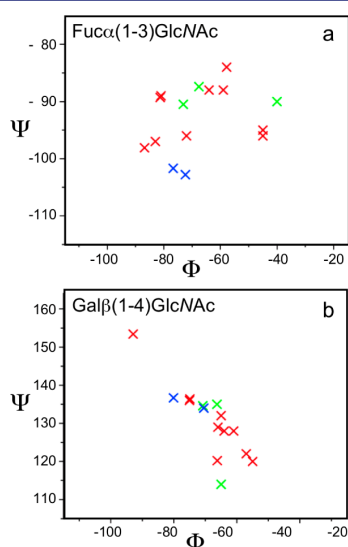


Figure 2. Previously reported Le^x structures/substructures: ϕ/ψ angles of the Fuc α (1,3)GlcNAc linkage (a) and the Gal β (1,4)GlcNAc linkage (b). Torsion angles based on NMR data and MD simulations are shown in red,^{11a,e,g-j} those of structures based on residual dipolar coupling (RDC) data in green,¹¹ and those from the crystal structure of Le^x in blue.¹² The torsion angles are defined as follows: ϕ , O5–C1–O1–C₂; ψ , C1–O1–C₂–C₃. A detailed list containing angles and references of all the displayed structures is given in the Supporting Information (Table S1).

While individual MD^{11a–g} and residual dipolar coupling (RDC)^{11h,i} studies yielded well-defined values for the ϕ/ψ torsion angles of the two glycosidic bonds of Le^x, the total set of ϕ/ψ values ranges e.g. from $-55/120^{\circ}$ ^{11h} to $-93/153^{\circ}$ ^{11g} for the glycosidic bond Gal β (1-4)GlcNAc. Therefore, a single defined solution conformation of Le^x (2) could not be obtained so far. In particular, elucidating the solution conformation of Le^x by NMR was severely hampered by the unfavorable tumbling regime, the small NOEs of usually performed ROESY experiments,¹³ difficulties in quantifying ROESY cross-peaks,¹⁴ and chemical shift degeneracy. In the case of methyl Le^x (2) (MW 544 Da), the rotational correlation time τ_c at 293 K is only 0.41 ns, resulting in a maximal NOE enhancement between 0.0 and -0.4 for a NOESY experiment (Figure S1, Supporting Information).¹⁴ Therefore, only a small number of inter-residual distance restraints are observed that are not

sufficient to deduce a well-defined structure.¹⁴ Finally, for the two X-ray structures of Le^x (2)¹² ϕ/ψ torsion angles with differing values were reported: e.g. -71 and -80° for the ϕ values of Gal β (1-4)GlcNAc. Differences in the crystal packing may explain these deviations.

Slyenko et al.¹⁵ demonstrated that the covalent attachment of an oligosaccharide to a protein has the advantage that the NOE transfer within the carbohydrate is largely enhanced because of the increase of the overall rotational correlation time. By attaching an unlabeled oligosaccharide to a ¹³C,¹⁵N-labeled recombinantly expressed protein, ¹³C,¹⁵N-filtered NOESY NMR techniques at ultrahigh field allowed increasing the maximal NOE enhancement close to -1 (Figure S1, Supporting Information), resulting in a high number of distance restraints per glycosidic bond and consequently a well-defined structure with a single favored conformation.¹⁵

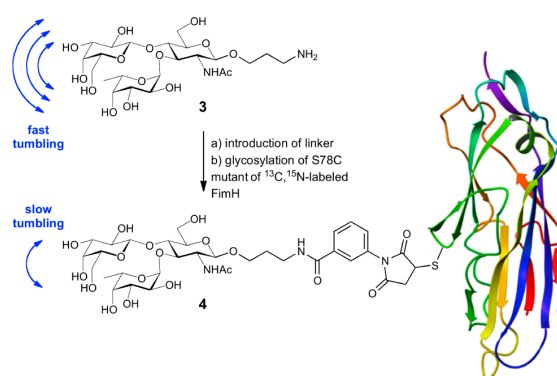
In this paper, we applied this approach to the structural analysis of Le^x chemically linked to a ¹³C,¹⁵N-labeled bacterial lectin (MW of ~ 20 kD). The resulting increase of the correlation time τ_c together with the high resolution obtained at 900 MHz enabled the observation of numerous inter-residual NOEs that could be readily quantified and converted into distance restraints. On the basis of the hereby obtained well-defined solution structure, the stereoelectronic effects responsible for the stabilization of the Le^x core structure were analyzed and are presented within this work.

RESULTS

When Lewis^x is covalently linked to a protein, the low-molecular-weight carbohydrate is converted into a large glycoconjugate with a drastically increased tumbling time and consequently a more favorable range for the detection of NOEs. For this purpose, we developed a generally applicable approach featuring the carbohydrate or a mimic thereof with a linker that can be chemically coupled to a cysteine of a ¹³C,¹⁵N-labeled protein (Scheme 1).

For this work, Le^x was equipped with a 3-propanolamine aglycone ($\rightarrow 3$) and coupled to the carrier protein via a 3-maleimidobenzoic acid linker ($\rightarrow 4$). For the protein component we selected the ¹³C,¹⁵N-labeled bacterial protein FimH with a Ser78Cys mutation.¹⁶

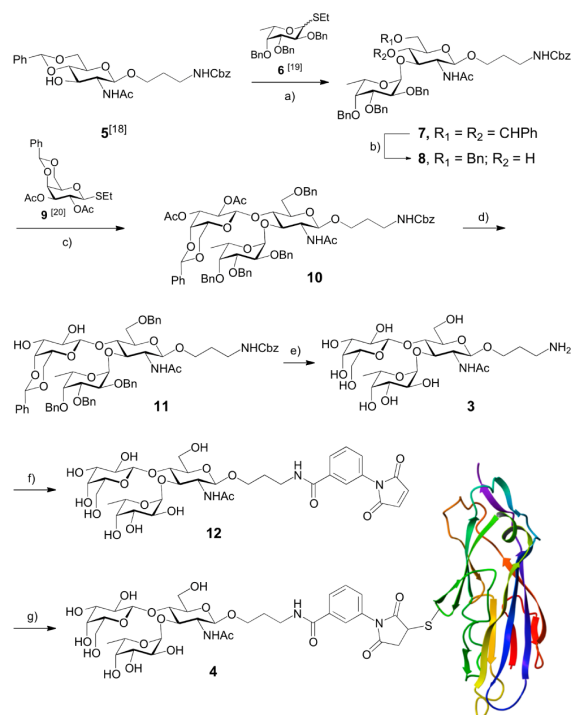
Scheme 1. ^a



^aFor the improvement of the tumbling properties and consequently the extractable NMR spectroscopic information, low-molecular-weight Le^x (3) was linked to the bacterial protein FimH ($\rightarrow 4$).

Ligand Synthesis and Chemical Glycosylation of Carrier Protein. To link oligosaccharides to carriers, 3-propanolamine is typically used.¹⁷ 3-Aminopropyl Le^x (3) was obtained by glycosylating the GlcNAc derivative 5¹⁸ first with the L-fucose building block 6¹⁹ and then, after deprotection of the 4-position, with the thiogalactoside 9²⁰ (Scheme 2). The first glycosylation step was promoted by

Scheme 2. Synthesis of Le^x Equipped with a Linker and Its Coupling to the Carrier Protein, the FimH S78C Mutant^a



^aLegend: (a) TBAB, CuBr₂, 4 Å molecular sieves, DCM/DMF, 77%; (b) NaBH₃(CN), HCl, THF, 85%; (c) DMTST, 4 Å molecular sieves, DCM, 60%; (d) NaOMe, MeOH; (e) Pd(OH)₂, H₂, DCM/MeOH/H₂O/AcOH, 77% over two steps; (f) 3-maleimidobenzoic acid *N*-hydroxy-succinimide ester (MBS), DMSO, H₂O, 62%; (g) ¹³C,¹⁵N-labeled S78C FimH mutant protein, 37 °C, 15 h, sodium phosphate buffer.

tetrabutylammonium bromide and copper(II) bromide, yielding the α -fucoside 7 in 77% yield. After the regioselective cleavage of the benzylidene acetal in 7 using sodium cyanoborohydride and hydrogen chloride in ether (\rightarrow 8), the 4-hydroxy group of the GlcNAc residue was galactosylated, giving the protected trisaccharide 10 in 60% yield.

The acetyl groups and the carbobenzoxy protection were removed by hydrolysis under Zemplén conditions and by catalytic hydrogenolysis, respectively, giving 3-aminopropyl Le^x (3) in 77% yield.

As linker, we chose 3-maleimidobenzoic acid, because its rigidity guarantees favorable tumbling properties and its ¹H NMR resonances are located outside the characteristic carbohydrate ranges. With the bifunctional 3-maleimidobenzoic acid *N*-hydroxysuccinimide ester (MBS), coupling with 3-aminopropyl Le^x (3) was performed in DMSO/water to give maleimide 12 in 62% yield. The final step was the coupling to

the S78C mutant of the ¹³C,¹⁵N-labeled bacterial FimH protein (MW 19.714 kDa). Ser78 was selected for mutation to Cys, because it is positioned in a solvent-exposed loop connecting strands D1 and D' (PDB entry 1TR7).²¹ The S78C mutant was expressed as ¹³C,¹⁵N-labeled protein in *E. coli* BL21(DE3) strain. Under physiological conditions, the nucleophilic thiol group of the cysteine residue was conjugated selectively to the maleimido group of the Le^x derivative 12 (Scheme 2; for a nonreducing SDS-PAGE, MALDI-TOF-MS data and a ¹H,¹⁵N-HSQC spectrum of the glycoconjugate see Figures S2 and S3 in the Supporting Information). Although the Michael addition proceeded only to approximately 50%, separation of the glycoconjugate 4 from the unreacted FimH protein was not necessary, because unreacted protein did not disturb the NMR measurements of the carbohydrate.

Extracting Carbohydrate Distance Restraints by NMR Spectroscopy. Similar to our study with the bacterial glycoprotein AcrA,¹⁵ we used the ¹³C,¹⁵N-labeled glycoprotein (see above) linked to unlabeled carbohydrate to detect distance-related NOE cross-peaks within the carbohydrate by recording 2D ¹³C F1-filtered F2-filtered NOESY²² (in D₂O) and 2D ¹⁵N F1-filtered F2-filtered NOESY¹⁵ spectra (in H₂O) (Figure 3a,b). In these types of NOESY experiments, the

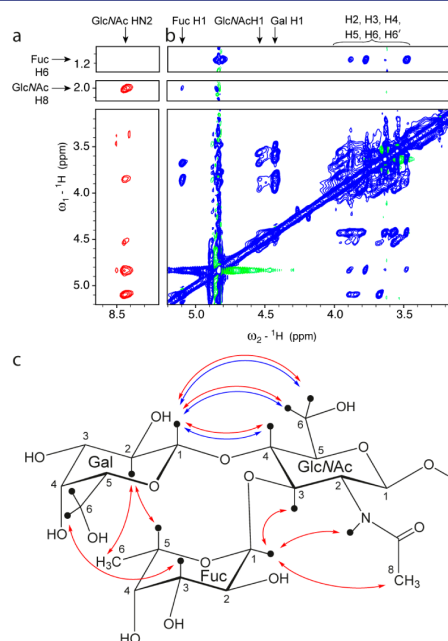


Figure 3. NOE cross-peaks between protons of the Le^x trisaccharide: (a) ¹⁵N filtered-filtered 2D NOESY recorded with 96 scans in 16 h; (b) ¹³C filtered-filtered 2D NOESY recorded with 96 scans in 16 h for Le^x-FimH (4); (c) schematic overview of interresidual NOEs of Le^x-FimH (4) (red arrows) and the free Me Le^x (2) (blue arrows).

signals of the ¹³C,¹⁵N-labeled protein are suppressed, resulting in spectra containing only resonances of the unlabeled carbohydrate and the linker. To obtain maximal resolution, spectra were recorded at 900 MHz. The assignment of Le^x resonances was basically identical with that of free Le^x (3), which was confirmed by 2D ¹³C F1-filtered TOCSY and ¹H,¹³C-HSQC spectra (Figure S4, Supporting Information). Since no carbohydrate–protein NOE cross-peaks were present

in the 2D ^{13}C F2-filtered NOESY experiment, the carbohydrate moiety was assumed not to interact with the protein surface. As a result, we could extract 24 unambiguous NOE cross-peaks between various nonexchangeable protons (CH_x) from a 2D ^{13}C F1-filtered F2-filtered NOESY and 4 unambiguous NOE cross-peaks between the exchangeable H^{N} amide of the acetamido group and proximal nonexchangeable protons from a 2D ^{15}N F1-filtered F2-filtered NOESY (Table S2, Supporting Information).

To evaluate the improvement resulting from the increased tumbling time, we compared 2D filtered-filtered NOESY spectra of Le^x attached to FimH ($\text{Le}^x\text{-FimH}$ (4)) with 2D NOESY spectra of free methyl Le^x (2) measured at 293 K. Due to the unfavorable tumbling time of free Le^x (2) most NOE cross-peaks are either absent or very weak, even though the 2D NOESY pulse sequence lacks the filter elements and hence is more sensitive. Whereas 28 NOEs were observed for $\text{Le}^x\text{-FimH}$ (4), only 9 NOEs were detected for the free Le^x (2) (Table S2, Supporting Information). The inter-residual NOE restraints that are of particular importance for conformational studies are shown schematically in Figure 3c and are summarized in Table 1. For $\text{Le}^x\text{-FimH}$ (4), 9 inter-residual restraints could be detected, in contrast to 3 for the free Le^x (2).

Solution Conformations of 2 and 4. With the help of NOE distance restraints, the structural ensembles of Le^x were calculated using Cyana²³ with subsequent refinement by Amber^{24a} applying the GLYCAM06 force field^{24b} (Figure 4; see Table S3 in the Supporting Information for NMR structure determination statistics). From the 28 NOE restraints for $\text{Le}^x\text{-FimH}$ (4) a well-defined structural ensemble with narrow distributions of the glycosidic torsion angles was obtained (Figure 4a and Figure S5 (Supporting Information)). Figure 4b shows a representative structure of this ensemble and Table S4 (Supporting Information) the corresponding ^1H – ^1H distances. In contrast, the structural ensemble of methyl Le^x (2) calculated from only 9 restraints displayed a considerable scattering of torsion angles (Figure S6 in the Supporting Information). Obviously, the obtained NOE restraints were not sufficient to calculate an ensemble structure with high precision (Figure 4c).

We then compared our structure model of Le^x with those reported from previous studies (Figure S7 in the Supporting Information), namely the crystal structure of methyl Le^x (2),¹² protein structures containing Le^x as ligand or as part of their glycosylation,²⁵ earlier solution structures obtained by residual dipolar couplings, limited NOEs, and molecular modeling (MD).^{11a,e,g–j} We observed a high agreement of the glycosidic torsion angles from our solution structure with those of the crystal structure of methyl Le^x (2)¹² and some Le^x structures determined by NMR spectroscopy in combination with MD,^{11a,h} confirming that the structures are identical and accurate. Deviation among these confirmations are of the same order as that between the two Le^x molecules in the asymmetric unit cell of the methyl Le^x (2) crystal structure.¹²

What Stabilizes the Le^x Structure? We hypothesized that the stabilization of the Le^x conformation originates from the interface between the stacked fucose and galactose moieties. Previously, it was suggested that hydrophobic interactions between the two moieties as well as steric effects of the acetamido group of GlcNAc both contributed to the increased stability of the conformation.^{26,27} Since such interactions should lead to changes of the chemical shifts in comparison to the corresponding, unstacked disaccharides, we measured the

Table 1. Inter-Residual NOEs of $\text{Le}^x\text{-FimH}$ (4) and Me Le^x (2) at 293 K and Their Corresponding Distances

proton pair	$\text{Le}^x\text{-FimH}$ (4)		methyl Le^x (2)	
	S/N of NOE cross-peaks	^1H – ^1H distance (Å)	S/N of NOE cross-peaks	^1H – ^1H distance (Å)
Inter-Residual NOEs				
Gal H1-GlcNAc H4	910	2.3 ^a	206	2.6 ^a
Gal H1-GlcNAc H62	438	2.6 ^a	206 ^c	2.6 ^a
Gal H1-GlcNAc H61	795	2.4 ^a	226	2.5 ^a
Gal H2-Fuc H5	209 ^c	3.0 ^a		
Gal H2-Fuc Q6	939 ^c	2.8 ^{a,b}		
Gal H6-Fuc H3	718 ^c	2.7 ^{a,b}		
GlcNAc H3-Fuc H1	286 ^c	2.8 ^a		
GlcNAc Q8-Fuc H1	142	3.8 ^{a,b}		
GlcNAc HN2-Fuc H1	129	3.2 ^a		
Intra-residual NOEs for Calibration				
GlcNAc H61–H62	1004 ^c	1.77 ^d	3209 ^c	1.77 ^d
GlcNAc H61–H62	4577	1.77 ^e	1837 ^c	1.77 ^d

^aThe ^1H – ^1H distances were calculated from experimentally obtained NOE intensities. The H61–H62 cross-peak of GlcNAc was used as a reference with a distance of 1.77 Å and assuming a r^{-6} dependence of the NOE intensities. For the structure calculations the distances reported in this table were increased by a 0.5 Å tolerance and used as upper limit distance restraints. ^bSignal to noise ratios (S/N) from cross-peaks involving methyl or methylene protons were divided by 3 or 2, respectively. ^cOnly one cross-peak was used because of artifacts or overlap. ^dReference restraint for the ^{15}N -filtered-filtered 2D NOESY. ^eReference restraint for the ^{13}C -filtered-filtered 2D NOESY.

chemical shifts of $\text{Fuca}(1\text{--}3)\text{GlcNAc}\beta\text{-OME}$ and $\text{Gal}\beta(1\text{--}4)\text{GlcNAc}\beta\text{-O}-(\text{CH}_2)_3\text{NH}_2$,²⁸ and compared them to those of methyl Le^x (2) (see Table 2 and Table S5 (Supporting Information)). The expected hydrophobic interactions between H6 of L-Fuc and H2 of D-Gal are not reflected in the shifts with deviations of only 0.01 and 0.05 ppm, respectively. However, another proton at the stacking interface exhibits a dramatic chemical shift change: H5 of L-Fuc resonates at 4.33 ppm in $\text{Fuca}(1\text{--}3)\text{GlcNAc}\beta\text{-OME}$ and at 4.83 ppm in methyl Le^x (2), resulting in a difference in the chemical shift of 0.50 ppm (at 293 K). Furthermore, the NMR shifts calculated for the optimized stacked geometry (Table 2; Table S5 (Supporting Information)) are in excellent agreement with the experimental data, thus supporting the experimental NMR conformation.

A close inspection of the structure ensemble of $\text{Le}^x\text{-FimH}$ (4) reveals that the C5–H5 bond of L-Fuc points toward O5 of D-Gal. The H5–O5 distance in the ensemble is 2.50 ± 0.01 Å. The sum of the corresponding van der Waals radii is 2.61 Å,³² indicating the presence of a C–H...O hydrogen bond. C5–O5 distances of 3.55–3.58 Å in the ensemble are also slightly shorter than the distance expected for the corresponding van der Waals separation (3.71 Å). The large H5 chemical shift deviation is a strong indication for such a nonconventional

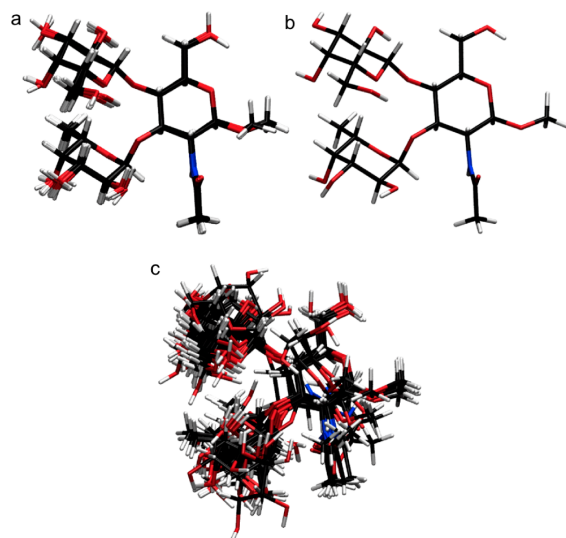


Figure 4. Calculated and refined structural ensembles of Le^x at 293 K using NOE restraints: (a) Le^x -FimH (4); (b) a representative structure thereof; (c) methyl Le^x (2).

Table 2. Experimental and Calculated^a Chemical Shifts (ppm) of Selected Le^x Protons at the Interface between Fucose and Galactose and Deviation from the Shifts of $\text{Fuc}\alpha(1-3)\text{GlcNAc}$ and $\text{Gal}\beta(1-4)\text{GlcNAc}$

proton	exptl			calcd		
	Le^x	$\Delta\delta(\text{Le}^x\text{-Fuc}\alpha(1-3)\text{GlcNAc})$	$\Delta\delta(\text{Le}^x\text{-Gal}\beta(1-4)\text{GlcNAc})$	Le^x	$\Delta\delta(\text{Le}^x\text{-Fuc}\alpha(1-3)\text{GlcNAc})$	$\Delta\delta(\text{Le}^x\text{-Gal}\beta(1-4)\text{GlcNAc})$
H3 of L-Fuc	3.90	0.08		3.90	-0.02	
H4 of L-Fuc	3.79	-0.01		3.66	0.05	
H5 of L-Fuc	4.84	0.51		4.90	0.83	
CH_2 of L-Fuc	1.18	0.01		1.13	-0.05	
H2 of D-Gal	3.50		0.04	3.64		0.04

^aCalculated using B3LYP/6-31G(d,p) in water with the polarizable conductor calculation model (CPCM).²⁹

hydrogen bond.³³ The C–H \cdots O hydrogen bond locks the Le^x conformation, resulting in a narrow cluster of ϕ/ψ torsion angles. Although C–H \cdots O hydrogen bonds are only about half as strong as “classical” O–H \cdots O hydrogen bonds, they are widespread and presumably represent 20–25% of the total number of hydrogen bonds in protein structures.³⁴ To our knowledge, intramolecular C–H \cdots O hydrogen bonds have not been explicitly reported in carbohydrate solution structures so far, in particular not in the context of stabilizing a certain conformation. However, during his studies of the solution conformation of Le^b in 1989, Lemieux proposed that hydrogen atoms in van der Waals contact with oxygen atoms of different sugar units are the reason for the conformational preferences.³⁵

In the crystal structures of methyl Le^x (2)¹² and most of the glycoproteins containing Le^x as part of their glycosylation or as ligand, a nonconventional C–H \cdots O hydrogen bond can be identified (Table 3) but have remained unnoticed so far. The

Table 3. Distance between C5 of L-Fuc and O5 of D-Gal in Le^x Crystal Structures with a Resolution <3.0 Å

glycan	PDB ^a or CSD ^b code	resolution (Å)	refinement method	Fuc C5–Gal O5 distance (Å)	ref
Me Le^x (2)	ABUCEF ^b		direct	3.269 3.304	12
Me Le^x (2)	1UZ8 ^a	1.8	Refmac 5.2	3.465 3.312	25a
$\text{Le}^x\text{-}\beta(1-3)\text{Gal}$	1SLS ^a	1.7	CNS 1.1	3.741	25b
$\text{Le}^x\text{-}\beta(1-3)\text{Gal}\beta(1-4)\text{Glc}$	3AP9 ^a	1.33	Refmac 5.5	3.778	25d
	2OX9 ^a	1.95	CNS 1.1	3.210 3.329 3.289 3.352	25c
$\text{Sia}\alpha(2-3)\text{Le}^x\text{-OMe}$	1G1T ^a	1.5	CNS	3.434	30
$\text{Sia}\alpha(2-3)\text{Le}^x\text{-OMe}$	2KMB ^a	2.0	X-PLOR 3.54	3.350 3.256 3.343	31
$\text{Sia}\alpha(2-3)\text{Le}^x\text{-}\beta(1-3\text{Gal})\text{-}\beta(1-4)\text{GalNAc-Thr}$	1G1S ^a	1.9	CNS	3.380	30
				3.374	

^aPDB (Protein Data Bank). ^bCSD (Cambridge Structural Database).

distance between C5 of L-Fuc and O5 of D-Gal varies between 3.21 and 3.78 Å in crystal structures in comparison to 3.56 ± 0.01 Å in our solution structure. We assume that the observed deviations result from using different force fields for structure refinement that prevent too close contacts by applying a van der Waals repulsion term. It is therefore not surprising that the smallest distances between C5 of L-Fuc and O5 of D-Gal , namely 3.269 and 3.304 Å, were obtained from the crystal structure of methyl Le^x (2), where direct assignment methods instead of force field calculations were applied.¹² We therefore consider this structure as the most reliable with regard to the C–H \cdots O hydrogen bond stabilizing Le^x . Other structures reported so far were determined by MD and NMR methods, and thus their ϕ/ψ glycosidic torsion angles are biased by the van der Waals repulsion terms in the applied force fields.

Computational Analysis. The structure based on distance restraints determined by NMR relied on the force field GLYCAM, which does not include any specialized terms for C–H \cdots O hydrogen bonding interactions but instead uses the Lennard–Jones potential function to keep atoms at ideal distances given by the sum of their van der Waals radii. Therefore, geometry optimization in the solvent phase using the density functional theory (DFT)³⁶ and ONIOM-(MP2:HF)³⁷ quantum chemical methods was used to refine the geometry. The *ab initio* optimization led to a shortening of the distance between H5 of L-Fuc and O5 of D-Gal typical for a C–H \cdots O hydrogen bond (Table S6, Supporting Information). The resulting interatomic distances were in good agreement with structural parameters observed in the crystal structure of Le^x .¹²

The DFT³⁶ optimized conformation served as starting point for a series of single point quantum mechanical calculations aimed at the quantification of the stacking interaction (Figure

5). The C–H \cdots O hydrogen bond between C5–H5 of L-Fuc and O5 of D-Gal seems to be the most prominent factor in

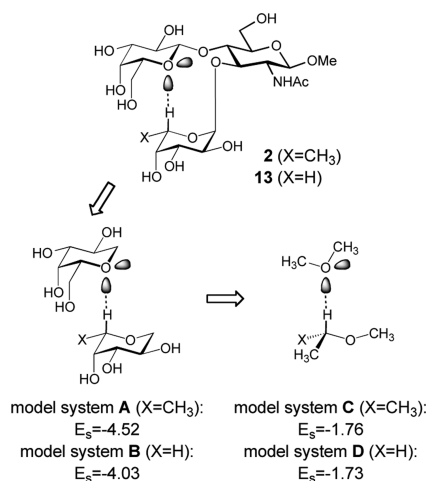


Figure 5. Model systems (A) 1-deoxy-galactose/1-deoxy-fucose, (B) 1-deoxy-galactose/1-deoxy-arabinose, (C) Me-OMe/Me-O-iPr, and (D) Me-OMe/Me-O-Et and the corresponding stabilization energies (E_s in kcal/mol).

stabilization, corresponding to almost 40% of the total stabilization energy as calculated at the MP2/aug-cc-pVQZ level including counterpoise correction (1.76 out of 4.52 kcal/mol, Table S7, Supporting Information). For comparison, the contribution of the C6 methyl group of L-Fuc toward stabilization is only 0.5 kcal/mol (compare model systems A and B, Figure 5). The calculated energy profile for a typical C–H \cdots O hydrogen bonding interaction in carbohydrates (Figure S8, Supporting Information) indicates an optimal H \cdots O distance between 2.35 and 2.45 Å, which is in excellent agreement with the distance observed in the Le^x crystal structure,¹² with the DFT optimized conformation (2.33 Å), and also with statistical averages derived from neutron diffraction crystal structures.³⁸ On the basis of calculations at the highest level (MP2/aug-cc-pVQZ), the C–H \cdots O hydrogen bond interaction energy at an optimal distance (2.4 Å) results in a value between 1.72 (value corrected for BSEE) and 1.98 kcal/mol (value without correction). The presence of the intramolecular C–H \cdots O hydrogen bond was confirmed also by localizing the bond critical point (Figure S9, Supporting Information) on the basis of the quantum theory of atoms in molecules.^{39,40}

DISCUSSION

In general, it is assumed that oligosaccharides are highly flexible and conformational restriction results predominantly from the exoanomeric effect⁴¹ and in special cases from steric effects and hydrophobic contacts.^{26,42} For the trisaccharide Le^x, for example, it could be demonstrated that two factors, the steric effect induced by the NHAc group adjacent to the linking position of L-fucose and the hydrophobic interaction of L-fucose with the β -face of D-galactose, are responsible for its low conformational flexibility.^{11,12} In the present communication, a nonconventional C–H \cdots O hydrogen bond³³ between H–C(5) of L-fucose and O(5) of D-galactose was identified as an additional factor. This nonconventional C–H \cdots O hydrogen

bond contributes to a reduction of the conformational flexibility and exhibits a novel dimension of the glycode.⁴³ We speculate that such interactions are widespread among glycoepitopes in mammals. A corresponding inspection of structures from the PDB is currently being performed.

Finally, the presented results uncovered a weakness of approaches based on molecular mechanics in being unable to produce an accurate geometry of C–H \cdots O hydrogen bonds due to van der Waals repulsion terms and to correctly evaluate its energetic contribution. In our model glycan Le^x such a bond contributes 40% to the stabilization of the 3D structure: i.e., a major contribution that should not be neglected. These results therefore potentially challenge the results of molecular modeling studies of carbohydrates. Our quantum mechanical calculations, the X-ray crystal structure of Le^x, and a statistical analysis of neutron diffraction studies of carbohydrates revealed the geometry of a typical C–H \cdots O hydrogen bond in carbohydrates.³⁸ This information will help to develop more accurate force fields for carbohydrates in the future.

EXPERIMENTAL METHODS

Commercially Available Carbohydrates. Methyl Le^x (Gal β 1,4-[Fuc β 1,3]GlcNAc β OMe) and the disaccharide Fuc α 1,3GlcNAc β OMe were purchased from Carbosynth/UK.

Protein Expression and Purification. All plasmids, bacterial strains, and DNA primers are given in Table S8 (Supporting Information). The plasmid pDsbA3 containing the carbohydrate recognition domain of the bacterial adhesin FimH linked to a 6His Tag (FimH-CRD-6His)⁴⁴ was used to generate the FimH-CRD-6His S78C mutant by site-directed mutagenesis.⁴⁵ The mutation was confirmed by double-strand DNA sequencing (Microsynth, Galbach, Switzerland). For uniform ¹³C/¹⁵N-labeling *E. coli* BL21(DE3) cells harboring the FimH-CRD-6His (S78C) encoding plasmid were cultivated in M9 minimal medium⁴⁶ supplemented with 1% (v/v) of Basal Medium Eagle vitamin mix solution (Sigma, Buchs, Switzerland) and 100 μ g/mL of ampicillin (Applchem, Baden-Dättwil, Switzerland) overnight at 37 °C and 300 rpm. Cells were centrifuged, washed with fresh M9 minimal medium, and further cultivated in M9 minimal medium without glucose and ammonium chloride for 2 h. The cells were centrifuged for 5 min at 5000 rpm, and the pellet was resuspended in 1 L M9 minimal medium containing ¹³C₆-glucose (2 g/L), ¹⁵N-ammonium chloride (1 g/L) and 1% (v/v) Basal Medium Eagle (BME) vitamin mix solution. The cells were allowed to grow to an OD₆₀₀ of 0.8 followed by the addition of IPTG at a final concentration of 1 mM and further cultivated at 30 °C and 160 rpm for 14 h. FimH-CRD-6His (S78C) was then extracted from the periplasmic space and purified on a Ni-NTA column as described previously.⁴⁴ The purity of the protein was verified by SDS-PAGE analysis, and the quantity (6.7 mg) was determined by HPLC⁴⁷ using BSA as standard.

FimH-Le^x Glycoprotein (4). A mixture of the ¹³C,¹⁵N-labeled FimH mutant (5.7 mg, 0.29 nmol) and maleimide 12 (13.4 mg, 17.1 μ mol) was dissolved in 2 mL of phosphate buffer (pH 7, 50 mM) and shaken (650 rpm) at 37 °C for 16 h. The mixture was lyophilized and purified by dialysis and ultrafiltration. The resulting mixture of protein and glycoprotein 4 was analyzed by SDS-PAGE and MALDI-TOF MS (Figure S2, Supporting Information).

NMR Spectroscopy of Carbohydrates, Proteins, and Glycoproteins. All NMR spectra were acquired on Bruker Avance or Avance III spectrometers equipped with triple-resonance cryogenic probes at a field strength of 500 or 900 MHz and a temperature of 293 K. The glycoprotein sample was dialyzed against water and enriched to a concentration of 0.5 mM using a centricon filter unit (Vivaspin, Sartorius stedim, Goettingen, Germany, 10 kDa cutoff). Samples were prepared in either 93% H₂O/7% D₂O or 100% D₂O using lyophilization for the preparation of the latter. 2D ¹³C F1-filtered F2-filtered NOESY²² and 2D ¹⁵N F1-filtered F2-filtered NOESY

spectra¹⁵ were recorded with a mixing time of 150 ms in D₂O and H₂O/D₂O, respectively, to obtain NOEs of the unlabeled glycan. At this mixing time the NOE build-up curves are assumed to be in the linear slope regions, where contributions from spin diffusion can be neglected.¹⁵ For the assignment of the free sugars, ¹H–¹³C HSQC, long-range ¹H–¹³C HSQC, ¹H–¹³C HMQC–COSY, and 2D TOCSY spectra were recorded.

Spectra were acquired and processed using Topspin 2.1 (Bruker) and analyzed with the software Sparky (T. D. Goddard and D. G. Kneller, SPARKY 3, University of California, San Francisco). Spectra were referenced to DSS by an external sample of 2 mM sucrose/0.5 mM DSS (Bruker standard).

Structure Calculation and Refinement. The structural template of methyl Le^x was generated using the Biomolecular Builder on the GLYCAM Web site (Woods Group. (2005–2012) GLYCAM Web. Complex Carbohydrate Research Center, University of Georgia, Athens, GA; <http://www.glycam.com>). Initial structures were calculated using CYANA 3.0.²³ Signal to noise (S/N) ratios of all NOE signals were extracted using the program Sparky (T. D. Goddard and D. G. Kneller, SPARKY 3, University of California, San Francisco) and converted to distances using the r^{-6} dependence and the GlcNAc H61–H62 cross-peaks (1.77 Å) as reference. S/N ratios of signals involving CH₂ and CH₃ groups were divided by a factor of 2 or 3, respectively. Upper limit restraints with an additional tolerance of 0.5 Å were applied. Out of 200 structures, the 30 structures with the lowest target function were further refined in AMBER 9^{24a} applying the Glycam06 force field.^{24b} A generalized Born model⁴⁸ was used to mimic solvent.

Calculation of ¹H Chemical Shifts. The geometry of the NMR-derived stacked conformation of Le^x was optimized using the density functional theory (DFT)³⁶ method at the B3LYP/6-31G(d,p) level of theory.⁴⁹ Solvent effects were accounted for using the CPCM model (implicit solvation model).²⁹ The NMR shielding tensors were calculated using the gauge-independent atomic orbital (GIAO)⁵⁰ method. Absolute values of shifts were adjusted relative to tetramethylsilane (its structure was fully optimized in solvent and an NMR spectrum was calculated using the same level of theory). The same settings were used for calculating shifts of the disaccharides. At the optimized geometry vibrational mode analysis was performed to confirm the stability of the obtained minimum. No imaginary frequencies were found. All ab initio geometry optimizations and spectral calculations were performed using Gaussian 09.⁵¹

Detailed Computational Analysis. As the DFT³⁶ methods are known to underestimate the fine dispersion interaction that might play an important role in stacking interaction of the two sugar units, a calculation including electron correlation (second-order Møller–Plesset perturbation theory, MP2)⁵² was performed for comparison. Methyl Le^x (2; C₂₁H₃₇NO₁₅) consists of 37 heavy atoms. Due to computational costs this hampers the use of large basis set and electron correlation for all atoms. Therefore, a geometry optimization run was set up for a two-layer ONIOM method. Parts of the Le^x molecule with the greatest impact on the stacking interaction (i.e., L-Fuc and D-Gal) were treated at the MP2/6-31G(d,p) level, while a smaller basis set along with a substantially less demanding level of theory—HF/6-31G(d)—was applied to the remaining part that has a small impact on the stacking interaction. For more details on the assignment of atoms to layers, see Figure S10 (Supporting Information).

Accurate Evaluation of Stacking Interaction. In order to better understand the nature and extent of stabilization between fucose and galactose in the stacked conformation, several high-level calculations were performed on simplified model systems. From the fully minimized structure obtained at the B3LYP/6-31G(d,p) CPCM²⁹ level, coordinates of fucose and galactose units were extracted and the OH groups at C1 of both units were replaced by hydrogen atoms. The interaction energy was then calculated as a difference of three single-point calculations (1-deoxyfucose, 1-deoxygalactose, and their stacked “complex”) in the gas phase at the MP2 level with 6-311++G(d,p), aug-cc-pVTZ, and aug-cc-pVQZ basis sets with counterpoise correction (CP)⁵³ in order to remove the error caused by the basis

set superposition (BSSE). For comparison a calculation using the OPLS-2005 force field was performed using the very same model system.

Accurate Evaluation of the C–H···O Hydrogen Bond. The model system for studying the C–H···O hydrogen bond between the two sugar units was constructed by extracting coordinates of C1, O5, C4, C5, C6 and attached hydrogen atoms of the fucose (isopropyl methyl ether) and C1, O5, C5 and attached hydrogen atoms of the galactose (dimethyl ether). The same methodology as that used for the evaluation of the stacking interaction was employed (Table S7, Supporting Information).

Energy Profile of the C–H···O Hydrogen Bond. The model system for the C–H···O hydrogen bond was used to obtain the energy profile of the C–H···O hydrogen bond as a function of the H···O distance. The interacting partners (iPro-O-Me and Me-O-Me) were aligned to an assumed ideal geometry (angle C–H···O set to full linear at 180°, H→O vector set at Me-O-Me angle axis with 35° deviation from the plane; see Figure S8 (Supporting Information)). The energy profile was obtained from a series of single-point calculations at the MP2 level with different correlation consistent basis sets and CP correction. The less computationally demanding level MP2/aug-cc-pVTZ was used to obtain a profile over a larger range of distances, while the best applicable level MP2/aug-cc-pVQZ was used to precisely localize the potential minimum in terms of preferred interatomic distance and interaction energy.

C–H···O Bond Critical Point. The molecular wave function of Le^x obtained at the B3LYP/6-31G(d,p) level was exported to the program AIMALL,⁵⁴ which was used to localize the bond critical points on the basis of the quantum theory of atoms in molecules.^{39,40}

■ ASSOCIATED CONTENT

📄 Supporting Information

Text giving synthetic procedures and characterization data for compounds 3, 7, 8, and 10–12, a diagram where rotational correlation time is plotted against the NOE/ROE enhancement correlation time for different field strengths (Figure S1), SDA-PAGE, MALDI-TOF MS (Figure S2) and ¹⁵N-HSQC (Figure S3) of 4, ¹³C-HSQC overlay of glycosylated ¹³C/¹⁵N-labeled glycoprotein 4 and Le^x 3 (Figure S4), ϕ/ψ plots of glycosidic linkages of solution structure of Le^x coupled to FimH (4) and Me Le^x (2) (Figures S5 and S6) and of reported structures (Figure S7 and Table S1), schematic representation of the model system to quantify the C–H···O bond and the corresponding potential curve (Figure S8), stereoview of the bond critical point of Me Le^x (2) (Figure S9), stereoview of Me Le^x (2) (Figure S10), glycosidic torsion angles of previously reported Le^x structures (Table S1), inter-residual and intra-residual NOEs and statistics of the NMR structure determination of 4 and 2 (Tables S2 and S3), ¹H–¹H distances of Le^x attached to FimH (4) used for the structure calculation (Table S4), experimental and calculated chemical shifts of the Le^x trisaccharide (Table S5), structural parameters and calculated stacking energy values (Tables S6 and S7), and bacterial strains, plasmids, and oligonucleotides used in this study (Table S8). This material is available free of charge via the Internet at <http://pubs.acs.org>.

■ AUTHOR INFORMATION

Corresponding Authors

*B.E.: beat.ernst@unibas.ch.

*M.S.: schubert@mol.biol.ethz.ch.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This research was supported by the Swiss National Science Foundation Sinergia Grant CRSII3_127333 to M.Z. and T.A. We thank Michael Wörle, Institute of Inorganic Chemistry, ETH, for his help in converting structures from the Cambridge Structural Database (CSD) format into PDB format, Markus Blatter for his advice for converting the different coordinate formats into Amber format, Robert J. Woods for providing the GLYCAM web site, and Brigitte Fiege, Institute of Molecular Pharmacy, University of Basel, for critical revision of the paper.

REFERENCES

- (1) (a) Bevilacqua, M. P.; Stengelin, S.; Gimbrone, M. A.; Seed, B. *Science* **1989**, *243*, 1160–1165. (b) Tedder, T. F.; Isaacs, C. M.; Ernst, T. J.; Demetri, G. D.; Adler, D. A.; Disteche, C. M. *J. Exp. Med.* **1989**, *170*, 123–133. (c) Johnston, G. I.; Cook, R. G.; McEver, R. P. *Cell* **1989**, *56*, 1033–1044.
- (2) (a) Barthel, S. R.; Gavino, J. D.; Descheny, L.; Dimitroff, C. J. *Expert Opin. Ther. Targets* **2007**, *11*, 1473–1491. (b) *Physiology and Pathophysiology of Leukocyte Adhesion*; Granger, D. N., Schmid-Schönbein, G. W., Eds.; Oxford University Press: Oxford, U.K., 1995.
- (3) Ernst, B.; Magnani, J. L. *Nat. Rev. Drug Discovery* **2009**, *8*, 661–677.
- (4) (a) Kansas, G. S. *Blood* **1996**, *88*, 3259–3287. (b) Polley, M. J.; Phillips, M. L.; Wayner, E.; Nudelman, E.; Singhal, A. K.; Hakomori, S.; Paulson, J. C. *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*, 6224–6228. (c) Phillips, M. L.; Nudelman, E.; Gaeta, F. C.; Perez, M.; Singhal, A. K.; Hakomori, S.; Paulson, J. C. *Science* **1990**, *250*, 1130–1132.
- (5) (a) Schwardt, O.; Kolb, H. C.; Ernst, B. In *The Organic Chemistry of Sugars*; Fügedi, P., Levy, D. E., Eds.; Marcel Dekker: New York, 2005. (b) Kaila, N.; Thomas, B. E. *Med. Res. Rev.* **2002**, *22*, 566–601. (c) Simanek, E. E.; McGarvey, G. J.; Jablonowski, J. A.; Wong, C.-H. *Chem. Rev.* **1998**, *98*, 833–862.
- (6) (a) Merritt, E. A.; Zhang, Z.; Pickens, J. C.; Ahn, M.; Hol, W. G.; Fan, E. *J. Am. Chem. Soc.* **2002**, *124*, 8818–8824. (b) Spiess, M. *Biochemistry* **1990**, *29*, 10009–10018.
- (7) (a) Scheffler, K.; Ernst, B.; Katopodis, A.; Magnani, J. L.; Wang, W. T.; Weisemann, R.; Peters, T. *Angew. Chem., Int. Ed.* **1995**, *34*, 1841–1844. (b) Harris, R.; Kiddle, G. R.; Field, R. A.; Milton, M. J.; Ernst, B.; Magnani, J. L.; Homans, S. W. *J. Am. Chem. Soc.* **1999**, *121*, 2546–2551. (c) Poppe, L.; Brown, G. S.; Philo, J. S.; Nikrad, P. V.; Shah, B. H. *J. Am. Chem. Soc.* **1997**, *119*, 1727–1736.
- (8) Somers, W. S.; Tang, J.; Shaw, G. D.; Camphausen, R. T. *Cell* **2000**, *103*, 467–479.
- (9) (a) Schweizer, D.; Patton, J. T.; Cutting, B.; Smieško, M.; Wagner, B.; Kato, A.; Weckerle, C.; Binder, F. P. C.; Rabbani, S.; Schwardt, O.; Magnani, J. L.; Ernst, B. *Chem. Eur. J.* **2012**, *18*, 1342–1351. (b) Kolb, H. C.; Ernst, B. *Chem. Eur. J.* **1997**, *3*, 1571–1578.
- (10) Titz, A.; Marra, A.; Cutting, B.; Smieško, M.; Papandreou, G.; Dondoni, A.; Ernst, B. *Eur. J. Org. Chem.* **2012**, 5534–5539.
- (11) (a) Ichikawa, Y.; Lin, C.-Y.; Dumas, D. P.; Shen, G.-J.; Garcia-Junceda, E.; Williams, M. A.; Bayer, R.; Ketcham, C.; Walker, L. E.; Paulson, J. C.; Wong, C.-H. *J. Am. Chem. Soc.* **1992**, *114*, 9283–9298. (b) Lin, C.-Y.; Hummel, C. W.; Huang, D.-H.; Ishikawa, Y.; Nicolaou, K. C.; Wong, C.-H. *J. Am. Chem. Soc.* **1992**, *114*, 5452–5454. (c) Ball, G. E.; O'Neill, R. A.; Schultz, J. E.; Lowe, J. B.; Weston, B. W.; Nagy, J. O.; Brown, E. G.; Hobbs, C. J.; Bednarski, M. D. *J. Am. Chem. Soc.* **1992**, *114*, 5449–5451. (d) Rutherford, T. J.; Spackman, D. G.; Simpson, P. J.; Homans, S. W. *Glycobiology* **1994**, *4*, 59–68. (e) Imberty, A.; Mikros, E.; Koca, J.; Mollicone, R.; Oriol, R.; Pérez, S. *Glycoconj. J.* **1995**, *12*, 331–349. (f) Veluraja, K.; Margulis, C. J. *J. Biomol. Struct. Dynamics* **2005**, *23*, 101–111. (g) Wormald, M. R.; Edge, C. J.; Dwek, R. A. *Biochem. Biophys. Res. Commun.* **1991**, *180*, 1214–1221. (h) Miller, K. E.; Mukhopadhyay, C.; Cagas, P.; Bush, C. A. *Biochem.* **1992**, *31*, 6703–6709. (i) Azurmendi, H. F.; Martin-Pastor, M.; Bush, C. A. *Biopolymers* **2002**, *63*, 89–98. (j) Homans, S. W.; Forster, M. *Glycobiology* **1992**, *2*, 143–151.
- (12) Pérez, S.; Mouhous-Riou, N.; Nifantev, N. E.; Tsvetkov, Y. E.; Bachet, B.; Imberty, A. *Glycobiology* **1996**, *6*, 537–542.
- (13) Bothnerby, A. A.; Stephens, R. L.; Lee, J. M.; Warren, C. D.; Jeanloz, R. W. *J. Am. Chem. Soc.* **1984**, *106*, 811–813.
- (14) Neuhaus, D.; Williamson, M. P. *The Nuclear Overhauser Effect in Structural and Conformational Analysis*; Wiley-VCH: New York, 2000.
- (15) Slynko, V.; Schubert, M.; Numao, S.; Kowarik, M.; Aebi, M.; Allain, F. H.-T. *J. Am. Chem. Soc.* **2009**, *131*, 1274–1281.
- (16) Choudhury, D.; Thompson, A.; Stojanoff, V.; Langermann, S.; Pinkner, J.; Hultgren, S. J.; Knight, S. D. *Science* **1999**, *285*, 1061–1066.
- (17) (a) Matsushita, T.; Nagashima, I.; Fumoto, M.; Ohta, T.; Yamada, K.; Shimizu, H.; Hinou, H.; Naruchi, K.; Ito, T.; Kondo, H.; Nishimura, S. I. *J. Am. Chem. Soc.* **2010**, *132*, 16651–16656. (b) Rao, Y.; Boons, G. J. *Angew. Chem., Int. Ed.* **2007**, *46*, 6148–6151. Teumelsan, N.; Huang, X. F. *J. Org. Chem.* **2007**, *72*, 8976–8979.
- (18) Rabbani, S.; Compostella, F.; Franchini, L.; Wagner, B.; Panza, L.; Ernst, B. *J. Carbohydr. Chem.* **2005**, *24*, 789–807.
- (19) Sato, S.; Ito, Y.; Nukada, T.; Nakahara, Y.; Ogawa, T. *Carbohydr. Res.* **1987**, *167*, 197–210.
- (20) Yashunsky, D. V.; Higson, A. P.; Ross, A. J.; Nikolaev, A. V. *Carbohydr. Res.* **2001**, *336*, 243–248.
- (21) 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.; Wynns, L.; Klemm, P.; Oscarson, S.; Knight, S. D.; Greve, H. D. *Mol. Microbiol.* **2005**, *55*, 441–455.
- (22) Peterson, R. D.; Theimer, C. A.; Wu, H. H.; Feigon, J. *J. Biomol. NMR* **2004**, *28*, 59–67.
- (23) (a) Herrmann, T.; Güntert, P.; Wüthrich, K. *J. Biomol. NMR* **2002**, *24*, 171–189. (b) Güntert, P. *Methods Mol. Biol. (N.Y.)* **2004**, *278*, 353–378.
- (24) (a) Case, D. A.; Cheatham, T. E., 3rd; Darden, T.; Gohlke, H.; Luo, R.; Merz, K. M., Jr.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R. J. *Comput. Chem.* **2005**, *26*, 1668–1688. (b) Kirschner, K. N.; Yongye, A. B.; Tschampel, S. M.; Gonzalez-Outeirino, J.; Daniels, C. R.; Foley, B. L.; Woods, R. J. *J. Comput. Chem.* **2008**, *29*, 622–655.
- (25) (a) van Roon, A. M. M.; Pannu, N. S.; de Vrind, J. P. M.; van der Marel, G. A.; van Boom, J. H.; Hokke, C. H.; Deelder, A. M.; Abrahams, J. P. *Structure* **2004**, *12*, 1227–1236. (b) Guo, Y.; Feinberg, H.; Conroy, E.; Mitchell, D. A.; Alvarez, R.; Blixt, O.; Taylor, M. E.; Weis, W. I.; Drickamer, K. *Nature Struct. Mol. Biol.* **2004**, *11*, 591–598. (c) Feinberg, H.; Taylor, M. E.; Weis, W. I. *J. Biol. Chem.* **2007**, *282*, 17250–17258. (d) Ideo, H.; Matsuzaka, T.; Nonaka, T.; Seko, A.; Yamashita, K. *J. Biol. Chem.* **2011**, *286*, 11346–11355.
- (26) Imberty, A.; Perez, S. *Chem. Rev.* **2000**, *100*, 4567–4588.
- (27) Su, Z.; Wagner, B.; Cocinero, E. J.; Ernst, B.; Simons, J. P. *Chem. Phys. Lett.* **2009**, *477*, 365–368.
- (28) Choudhury, M. I.; Minoura, N.; Uzawa, H. *Carbohydr. Res.* **2003**, *338*, 1265–1270.
- (29) Barone, V.; Cossi, M. *J. Phys. Chem. A* **1998**, *102*, 1995–2001.
- (30) Somers, W. S.; Tang, J.; Shaw, G. D.; Camphausen, R. T. *Cell* **2000**, *103*, 467–479.
- (31) Ng, K. K. S.; Weis, W. I. *Biochemistry* **1997**, *36*, 979–988.
- (32) www.ccdc.cam.ac.uk/products/csd/radii/table.php4.
- (33) The weak hydrogen bond. In *Structural Chemistry and Biology*; Desiraju, G. G. R.; Steiner, T., Eds.; Oxford University Press: Oxford, U.K., 1999.
- (34) Weiss, M. S.; Brandl, M.; Suhnel, J.; Pal, D.; Hilgenfeld, R. *Trends Biochem. Sci.* **2001**, *26*, 521–523.
- (35) Lemieux, R. U. *Chem. Soc. Rev.* **1989**, *18*, 347–374.
- (36) Hohenberg, P.; Kohn, W. *Phys. Rev.* **1964**, *136*, B864–B871.
- (37) Dapprich, S.; Komáromi, I.; Byun, K. S.; Morokuma, K.; Frisch, M. J. *J. Mol. Struct. (THEOCHEM)* **1999**, *462*, 1–21.
- (38) Steiner, T.; Saenger, W. *J. Am. Chem. Soc.* **1992**, *114*, 10146–10154.
- (39) Bader, R. F. W. *Atoms in Molecules: A Quantum Theory*; Oxford University Press: New York, 1994.
- (40) Bader, R. F. W. *Chem. Rev.* **1991**, *91*, 893–928.

- (41) (a) Lemieux, R. U. In *Molecular Rearrangements*; de Mayo, P., Ed.; Interscience: New York, 1964; pp 709–769. (b) Wolfe, S.; Whangbo, M.-H.; Radom, L. *Carbohydr. Res.* **1979**, *69*, 1–26.
- (42) Lemieux, R. U.; Bock, K.; Delbaere, L. T. J.; Koto, S.; Rao, V. S. *Can. J. Chem.* **1980**, *58*, 631–653.
- (43) (a) Gabius, H.-J. *Acta Anat.* **1998**, *161*, 110–129. (b) Ambrosi, M.; Cameron, N. R.; Davis, B. G. *Org. Biomol. Chem.* **2005**, *3*, 1593–1608. (c) Pilobello, K. T.; Mahal, L. K. *Curr. Opin. Chem. Biol.* **2007**, *11*, 300–305. (d) Gabius, H.-J.; André, S.; Jiménez-Barbero, J. J.; Romero, A.; Solis, D. *Trends Biochem. Sci.* **2011**, *36*, 298–313.
- (44) Rabbani, S.; Jiang, X. H.; Schwardt, O.; Ernst, B. *Anal. Biochem.* **2010**, *407*, 188–195.
- (45) Ling, M. M.; Robinson, B. H. *Anal. Biochem.* **1997**, *254*, 157–178.
- (46) Sambrook, J.; Russel, D. W. *Molecular Cloning*; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, NY, 2000.
- (47) Bitsch, F.; Aichholz, R.; Kallen, J.; Geisse, S.; Fournier, B.; Schlaeppli, J. M. *Anal. Biochem.* **2003**, *323*, 139–149.
- (48) Bashford, D.; Case, D. A. *Annu. Rev. Phys. Chem.* **2000**, *51*, 129–152.
- (49) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648–5652.
- (50) Wolinski, K.; Hilton, J. F.; Pulay, P. *J. Am. Chem. Soc.* **1990**, *112*, 8251–8260.
- (51) *Gaussian 09, Revision A.2*; Gaussian, Inc., Wallingford, CT, 2009.
- (52) (a) Möller, C.; Plesset, M. S. *Phys. Rev.* **1934**, *46*, 618–622. (b) Head-Gordon, M.; Pople, J. A.; Frisch, M. J. *Chem. Phys. Lett.* **1988**, *153*, 503–506.
- (53) Boys, S. F.; Bernardi, F. *Mol. Phys.* **1970**, *19*, 553–566.
- (54) Keith, T. A. *AIMAll (Version 13.05.06)*; TK Gristmill Software, Overland Park, KS, 2013; aim.tkgristmill.com.

3.1.1.1 Supporting Information

Supporting Information

Stabilization of branched oligosaccharides: Lewis^x benefits from a non-conventional C- H···O hydrogen bond

Mirko Zierke,[#] Martin Smieško,[#] Said Rabbani,[#] Thomas Aeschbacher,[¶] Brian Cutting,[#]
Frédéric H.-T. Allain,[¶] Mario Schubert,[¶] and Beat Ernst[#]

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

[¶] Institute of Molecular Biology and Biophysics, ETH Zürich, CH-8093 Zürich, Switzerland

Table of content:

Supplementary Figures	S2
Supplementary Tables	S11
Supplementary Methods	S18
Experimental Data	S24
References	S26

Supplementary Figures

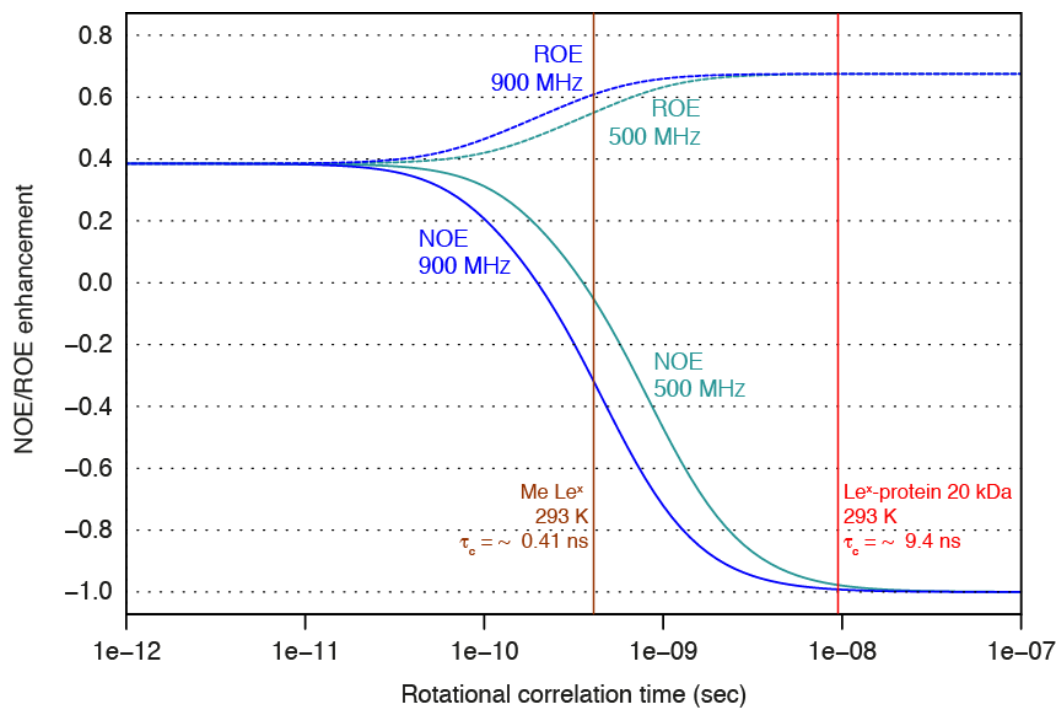


Figure S1. Maximal NOE and ROE enhancements calculated for a transient NOE experiment at two different field strengths in D₂O.^[S1]

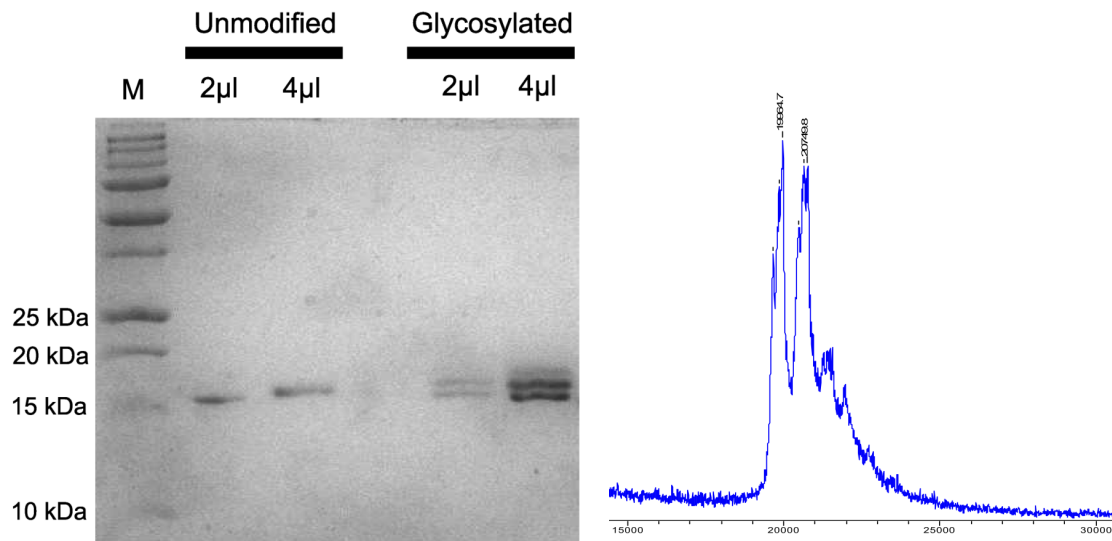


Figure S2. a.) SDS-PAGE analysis of unmodified and glycosylated uniformly ^{13}C , ^{15}N -labeled S78C-FimH. The glycosylated samples contained also unmodified protein, however this does not effect the measurements as the protein signals are suppressed in the NMR experiments. b.) MALDI-TOF MS analysis of the glycosylated uniformly ^{13}C , ^{15}N -labeled S78C-FimH.

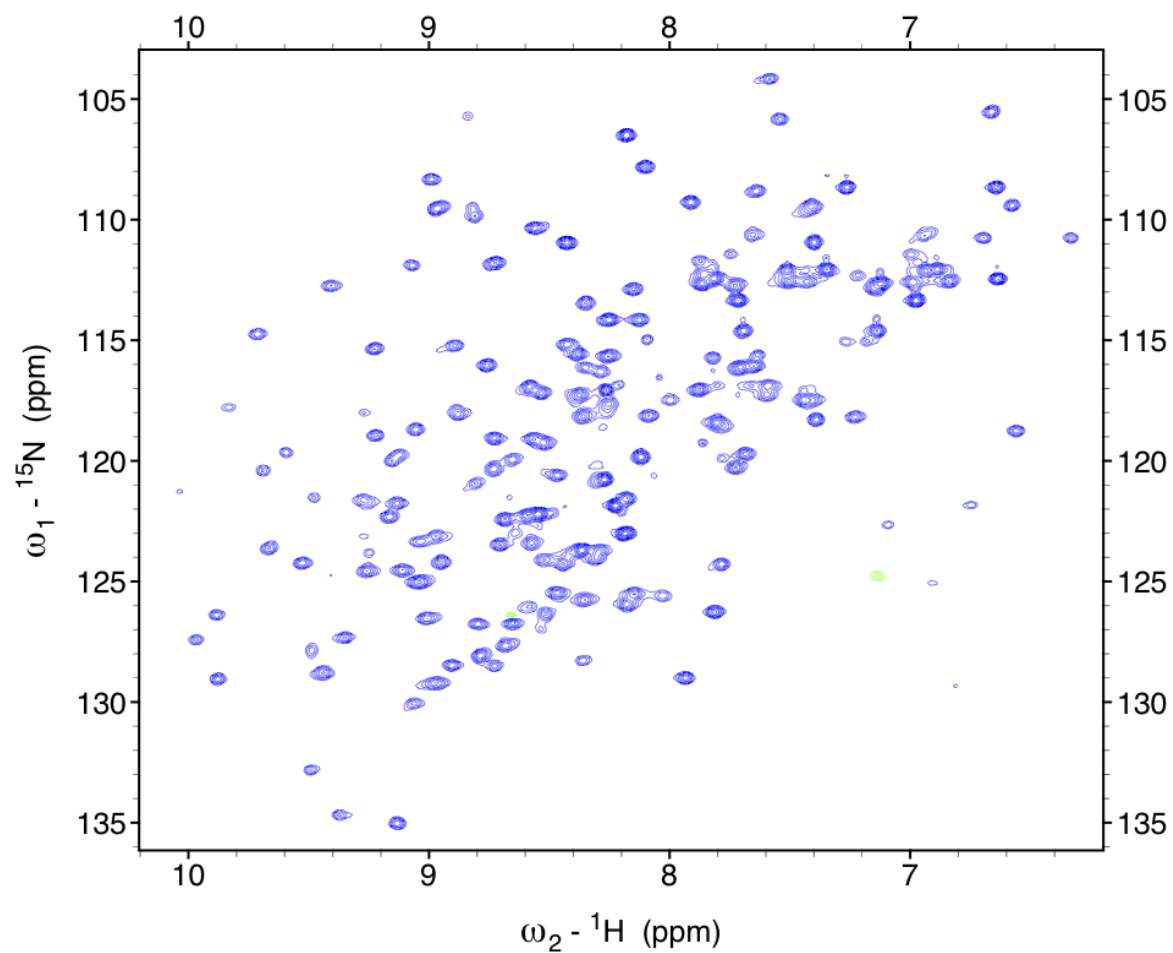


Figure S3. $^1\text{H}^{15}\text{N}$ -HSQC spectrum of the glycosylated ^{13}C , ^{15}N -labeled S78C-FimH (0.7 mM in 93% H_2O /7% D_2O) recorded at 900 MHz and 293K with 2 scans and 256 increments.

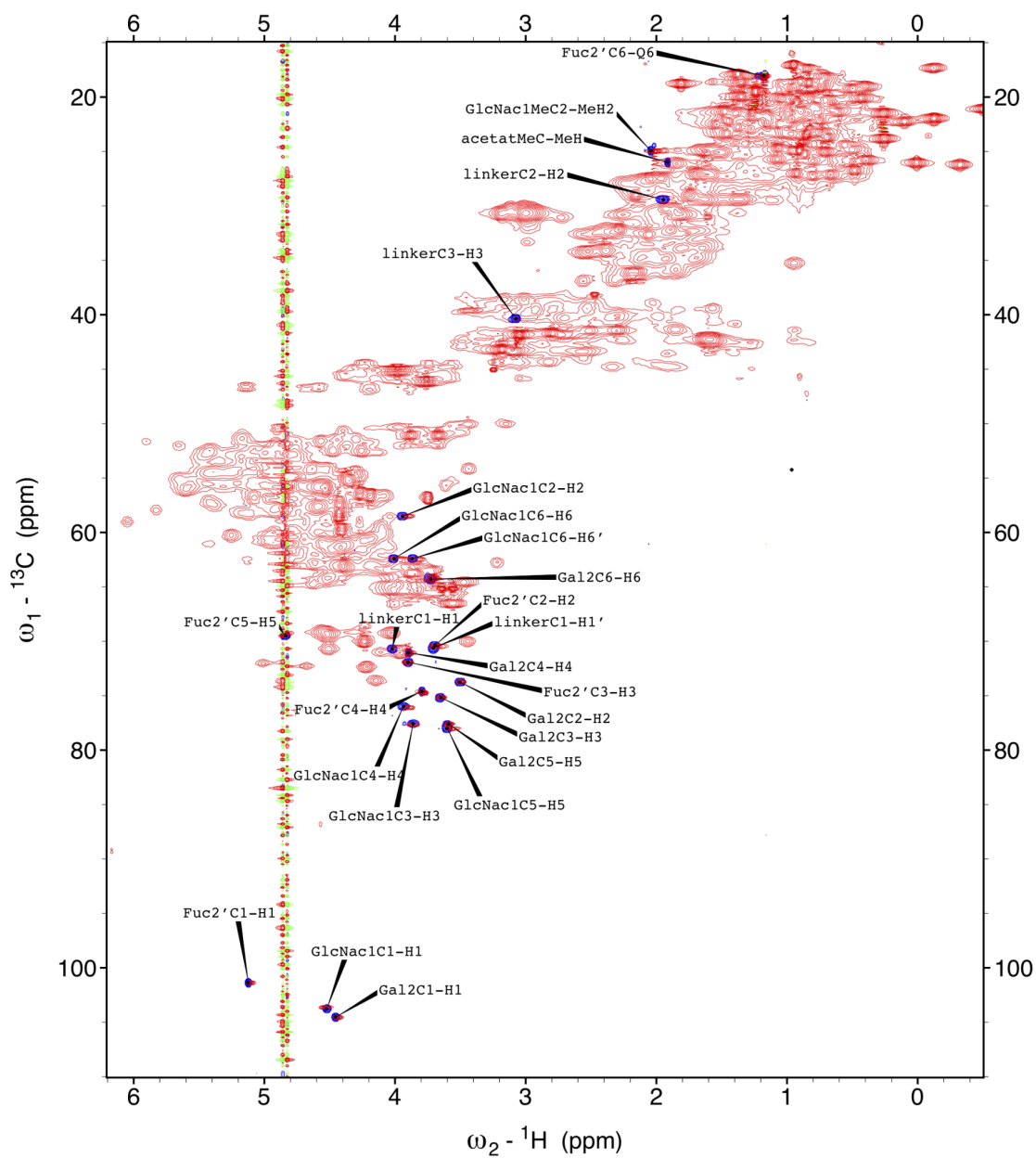


Figure S4. $^1\text{H}^{13}\text{C}$ -HSQC spectra of the glycosylated ^{13}C , ^{15}N -labeled S78C-FimH (red) and the free Le^x trisaccharide (blue). Spectra were recorded at 500 MHz and 293 K with 8 scans and 512 increments or 4 scans and 400 increments, respectively. The concentrations of the glycosylated protein and the trisaccharide were 0.7 mM and 12 mM, respectively (both in D_2O).

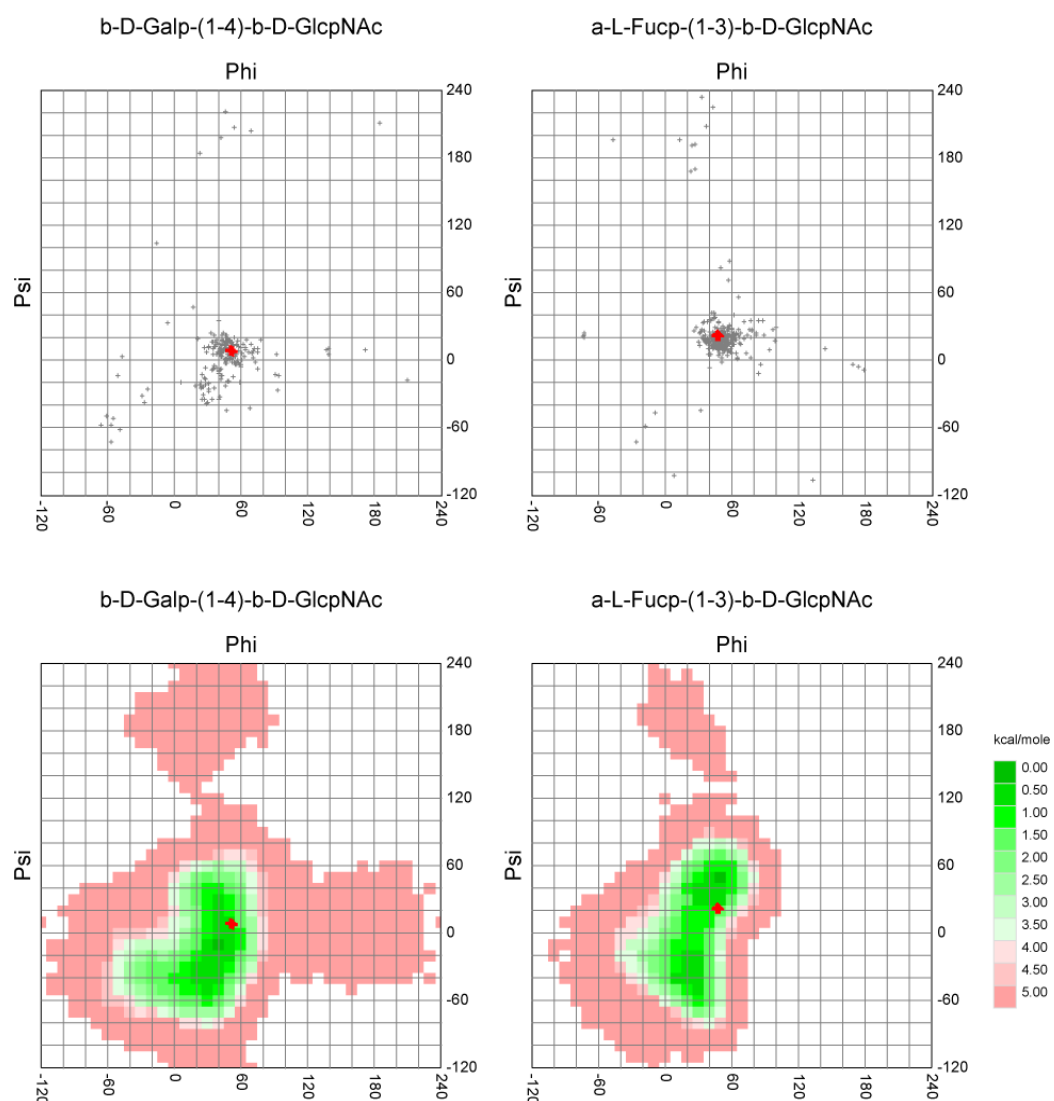


Figure S5. Phi-psi plots of the solution structure of Le^x (attached to FimH) superimposed with all entries in the PDB database containing the corresponding saccharide linkages are displayed in the upper panel whereas the bottom panel shows superimpositions on energy landscapes. Plots were generated with the software CARP.^[S2] Angles of the presented structural ensemble of Le^x consisting of 20 structures are designated by red crosses. The angles phi and psi are defined as H1-C1-O1-C'_x and C1-O1-C'_x-H'_x, respectively.

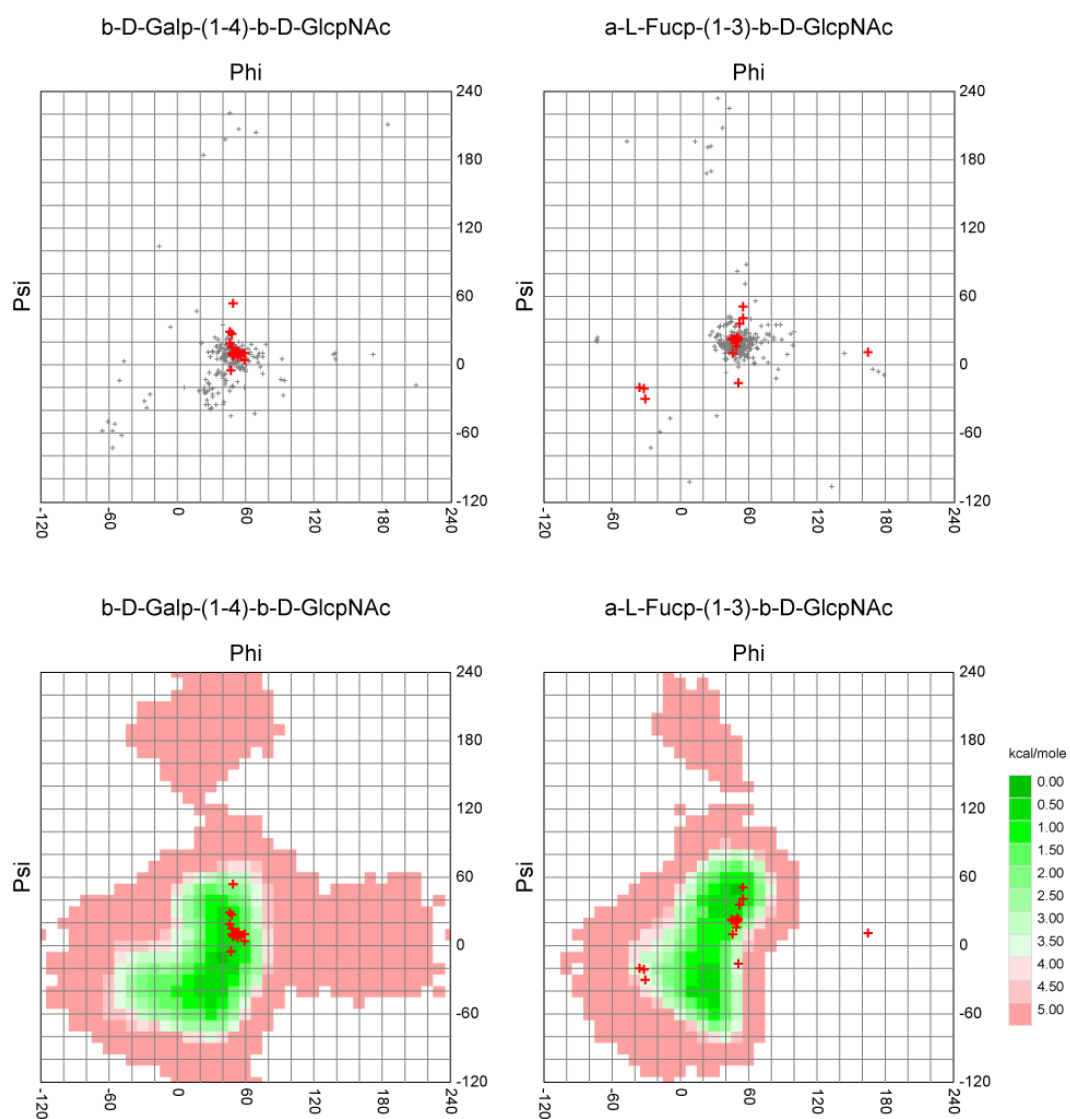


Figure S6. phi-psi plots of the solution structure of methyl Le^x (**2**) based on a 2D NOESY experiment. Note that these data were insufficient to derive a well-defined structure due to the unfavorable tumbling time. Superimpositions of phi-psi angles of all entries in the PDB database for a given linkage are shown in the upper panel whereas the bottom panel shows overlays on energy landscapes. Plots were generated with the software CARP.^[S2] Angles of 20 structures are shown by red crosses. The angles phi and psi are defined as H1-C1-O1-C'_x and C1-O1-C'_x-H'_x, respectively.

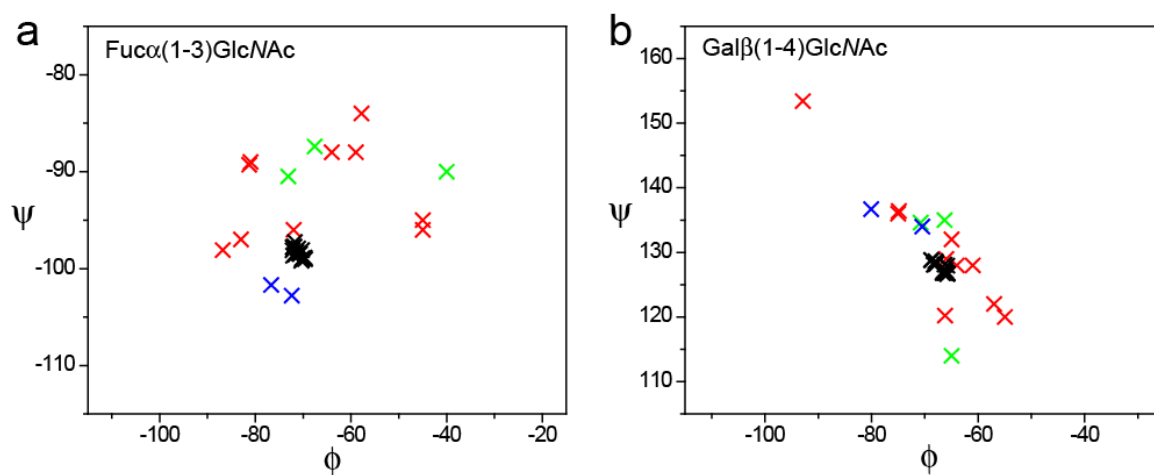


Figure S7. Comparison of glycosidic angles of the obtained NMR ensemble with previously reported Le^x structures/substructures. ϕ/ψ -angles of the Fuc α (1,3)GlcNAc linkage (a) and Gal β (1,4)GlcNAc linkage (b). Torsion angles of the here presented NMR ensemble are shown in black, angles of structures based on other NMR data and MD simulations are shown in red, of structures based on RDC data in green, and from the crystal structure of Le^x in blue (see Table S1). The torsion angles are defined as follows: ϕ O5–C1–O1–C'_x, ψ C1–O1–C'_x–C'_{x-1}.

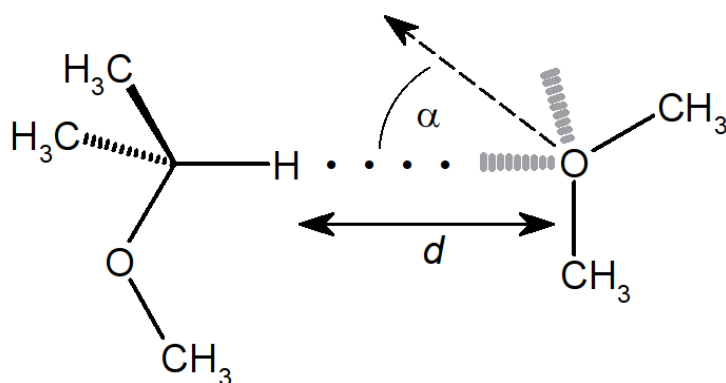
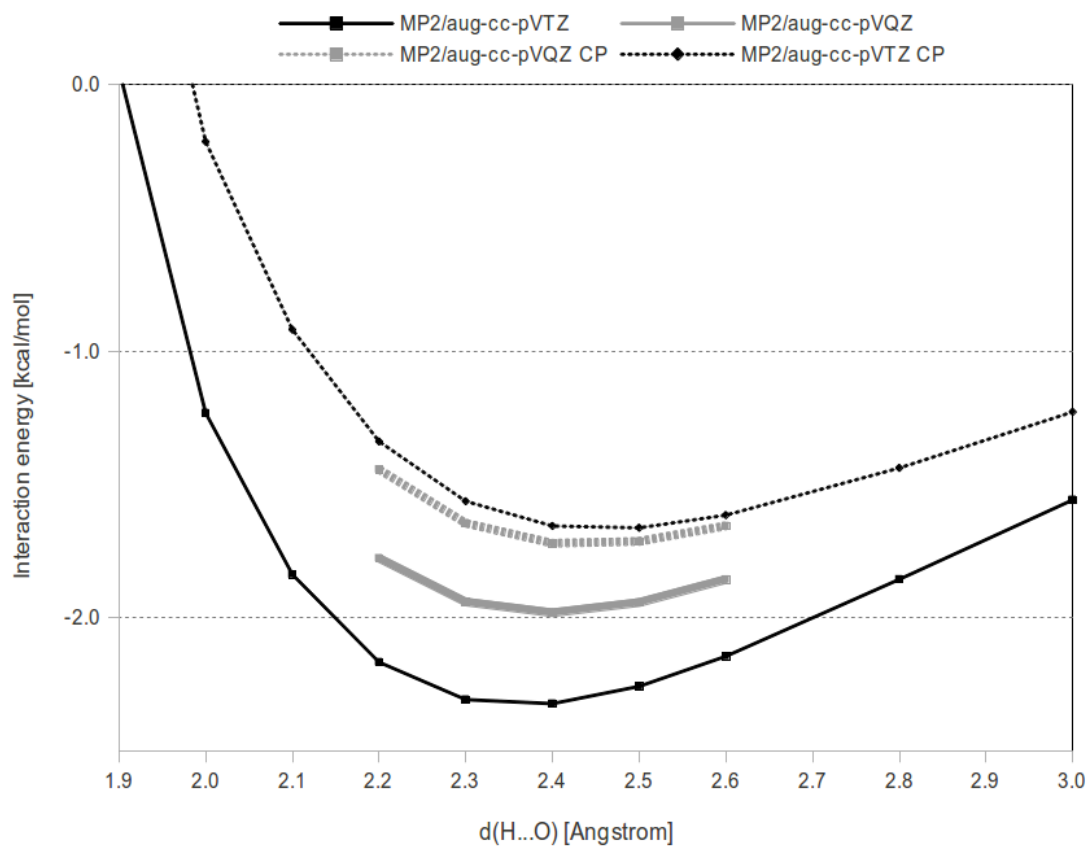


Figure S8. Schematic representation of the model system for quantification of the non-conventional C-H \cdots O hydrogen bond between H5 of L-fucose (resembling an isopropyl methyl ether) and O5 of D-galactose (resembling a dimethyl ether) of Le^x and corresponding energy potential curves as a function of the interatomic distance $d_{\text{H}\cdots\text{O}}$ ($\alpha = 35^\circ$).

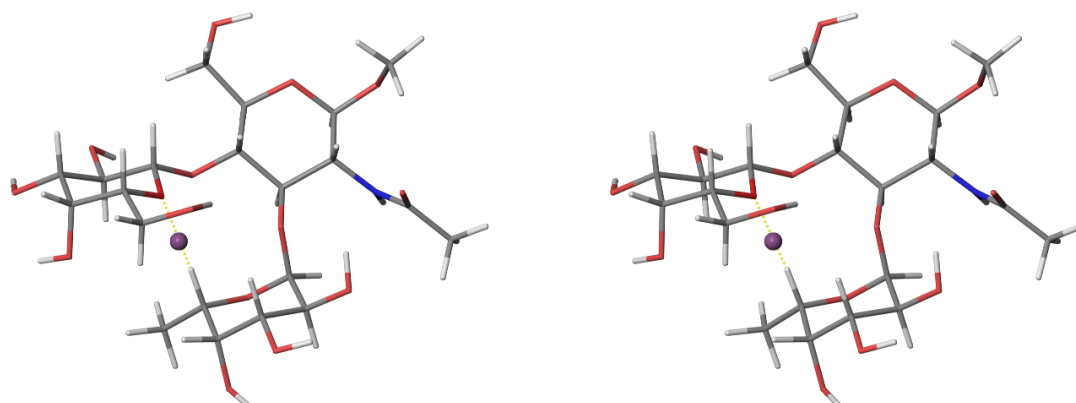


Figure S9. Stereo view of the bond critical point (signature: 3,-1; density $\rho_b = 0.01327$ au; displayed as a blue sphere) identified for the C-H...O hydrogen bond between C5-H5 of L-Fuc and O5 of D-Gal of Me Le^x (**2**) based on the B3LYP/6-31G(d,p) wavefunction.

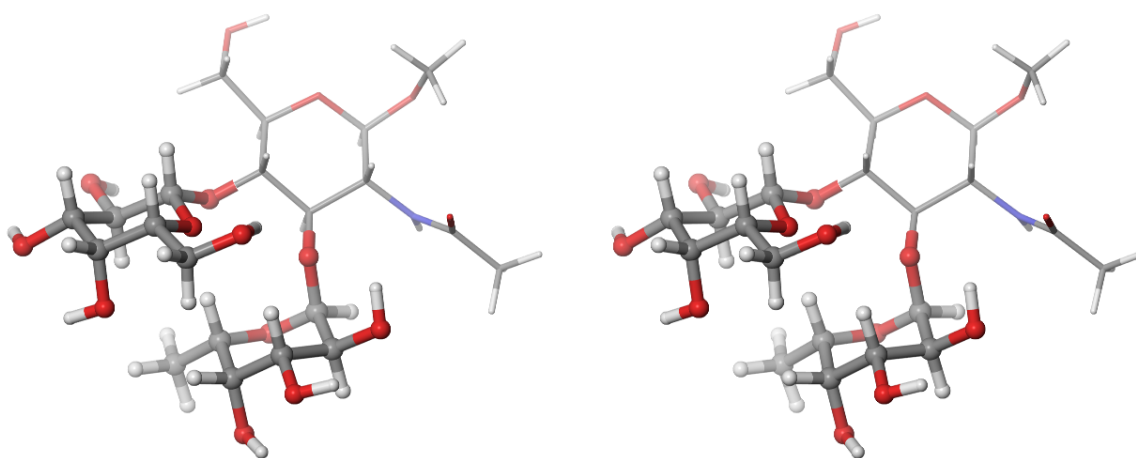


Figure S10. Stereo view of methyl Le^x (**2**) optimized using the ONIOM method.^[S3] Atoms displayed as balls and sticks were assigned to the high layer and atoms displayed as thin sticks were assigned in the low layer.

Supplementary Tables

Table S1. Glycosidic torsion angles of Le^x structures obtained by X-ray crystallography, NMR spectroscopy (NOE, ROE or RDC) and molecular modeling (MD).

Method	Fucα1,3GlcNAc torsion angles ^{a,b}				Galβ1,4GlcNAc torsion angles ^{a,b}				Reference
	phi IUPAC [°]	psi IUPAC [°]	phi NMR [°]	psi NMR [°]	phi IUPAC [°]	psi IUPAC [°]	phi NMR [°]	psi NMR [°]	
X-ray (ABUCEF)	-72.4	-102.8	40.4	20.9	-80.1	136.7	36.3	14.8	[S4]
	-76.7	-101.7	35.4	20.8	-70.5	134.0	44.3	15.5	[S4]
RDC (Le ^s)	-73.1	-90.5	(46.9)	(29.5)	-66.3	135.0	(53.7)	(15.0)	[S5]
RDC (LNF-3)	-67.6	-87.4	(52.4)	(32.6)	-70.8	134.6	(49.2)	(14.6)	[S5]
RDC+NOE (LNF-3)	-40.0	-90.0	(80.0)	(30.0)	-65.0	114.0	(55.0)	(-6.0)	[S5]
MD	-81.0	-89.0	(39.0)	(31.0)	-75.0	136.0	(45.0)	(16.0)	[S5]
SIMNOE	-45.0	-95.0	(75.0)	(25.0)	-55.0	120.0	(65.0)	(0.0)	[S6]
MD	-59.0	-88.0	(61.0)	(32.0)	-61.0	128.0	(59.0)	(8.0)	[S6]
MD	-64.0	-88.0	(56.0)	(32.0)	-64.0	128.0	(56.0)	(8.0)	[S6]
MD+NOE	-45.0	-96.0	(75.0)	(24.0)	-57.0	122.0	(63.0)	(2.0)	[S6]
MD+NOE	-83.0	-97.0	(37.0)	(23.0)	-65.0	132.0	(55.0)	(12.00)	[S6]
MD+ROE	-57.8	-84.0	(62.2)	(36.0)	-66.2	120.2	(53.8)	(0.2)	[S7]
MD+ROE	(-86.8)	(-98.1)	33.20	21.90	(-92.9)	(153.4)	27.10	33.40	[S8]
MD+ROE	(-72.0)	(-96.0)	48.00	24.00	(-66.0)	(129.0)	54.00	9.00	[S9]
MD	-81.3	(-89.3)	(38.7)	(30.7)	-74.9	(136.4)	(45.1)	(16.4)	[S10]
NOE (present work)	-71.1 ±0.9	-98.4 ±0.5	47.7±1.0	22.0±0.6	-67.0±1.1	127.6±0.8	52.5±1.3	8.1±0.8	

a) The glycosidic torsion angles phi and psi (IUPAC) are defined as O5-C1-O1-C'_x and C1-O1-C'_x-C'_{x-1}, respectively. The NMR definition for the glycosidic torsion angles phi and psi is H1-C1-O1-C'_x and C1-O1-C'_x-H'_x, respectively.

b) Values in brackets were interconverted between IUPAC nomenclature and NMR nomenclature by adding/subtracting 120°.

Table S2. Intra- and inter-residual NOEs of Le^x attached to FimH (**4**) and Le^x *O*-methylglycoside (**2**) at 293 K and 900 MHz and the corresponding distances.

proton pair	Le ^x attached to FimH (4)		Me Le ^x (2) (free)	
	average S/N of NOEs cross peaks	Corresponding ¹ H- ¹ H distance ^b [Å]	average S/N of NOEs cross peaks	Corresponding ¹ H- ¹ H distance ^b [Å]
intra				
Gal H1-H2	903	2.3 ^f		
Gal H1-H3	710	2.4	146	1.7
Gal H1-H5	1142	2.2	290 ^a	2.4
Gal H2-H3	2448 ^a	2.0 ^f		
Gal H3-H4	1581 ^a	2.1		
Gal H4-H5	1577 ^a	2.1		
GlcNAc H1-H3	207	3.0 ^f		
GlcNAc H1-H5	474	2.6	290 ^a	2.4
GlcNAc H2-Q8	75 ^a	4.2 ^e		
GlcNAc H5-H62	435	2.6	192 ^a	2.6
GlcNAc H61-H62 ^c	1004 ^a	1.77	3209 ^a	1.77
GlcNAc H61-H62 ^d	4577	1.77	1837 ^a	1.77
Fuc H1-H2	800	2.4	186	2.6
Fuc H1-Q6	106 ^a	4.0 ^e		
Fuc H3-H5	253 ^a	2.9		
Fuc H4-H5	590 ^a	2.5	178	2.6
Fuc H4-Q6	1357 ^a	2.6 ^e		
Fuc H5-Q6	844 ^a	2.8 ^e		
GlcNAc H1-HN2	60 ^a	3.6		
GlcNAc H3-HN2	72 ^a	3.5		
GlcNAc Q8-HN2	115 ^a	3.9 ^e		
inter				
Gal H1 - GlcNAc H4	910	2.3	206	2.6
Gal H1 - GlcNAc H62	438	2.6	206 ^a	2.6
Gal H1 - GlcNAc H61	795	2.4	226	2.5
Gal H2 - Fuc H5	209 ^a	3.0		
Gal H2 - Fuc Q6	939 ^a	2.8 ^e		
Gal Q6 - Fuc H3	718 ^a	2.7 ^e		
GlcNAc H3 - Fuc H1	286 ^a	2.8		
GlcNAc Q8 - Fuc H1	142	3.8 ^e		
GlcNAc HN2 - Fuc H1	129	3.2		

^{a)} Only one cross-peak was used because of artifacts or sever spectral overlap.

^{b)} The ¹H-¹H distances were calculated from experimentally obtained NOE intensities using the H61-H62 cross-peak of GlcNAc as a reference with a distance of 1.77 Å assuming a 1/*r*⁶ dependence of the NOE intensities. For the structure calculations the herein reported distances were increased by 0.5 Å tolerance and used as upper limit restraints.

^{c)} Reference restraints for the ¹⁵N-filtered-filtered NOESY.

^{d)} Reference restraints for the ¹³C-filtered-filtered NOESY.

^{e)} Based on signal to noise ratios from cross-peaks involving methyl or methylene protons that were divided by 3 and 2, respectively, due to their number of protons.

^{f)} The tolerance of these distances was increased by 1 Å.

Table S3. NMR structure determination statistics of Le^x attached to FimH (4), modeled as Le^x O-methylglycoside (2)

	Le ^x attached to FimH (4)	Me Le ^x (2), free (insufficient restraints)
NMR distance and dihedral restraints		
Total NOE restraints	28	9
Intra-residue	19	6
Inter-residue	9	3
Sequential ($ i - j = 1$)	6	3
Nonsequential ($ i - j > 1$)	3	0
Hydrogen bonds	0	0
Total dihedral angle restraints	0	0
HN-CO peptide bonds of acetamido	0	0
Sugar pucker	0	0
Structure statistics *		
Violations (mean and s.d.)		
Number of distance constraint violations > 0.1 Å	0±0	0.31±0.31
Max. distance constraint violation (Å)	0.04±0.00	0.08±0.08
Deviations from idealized geometry		
Bond lengths (Å)	0.0154±0.0001	0.0158±0.0004
Bond angles (°)	1.88±0.05	2.02±0.15
Heavy atom RMSD to mean (Å)	0.10±0.09	1.17±0.55
Glycosidic linkage phi / psi angles **		
Fucα(1,3)GlcNAc	-71.1±0.9/-98.4±0.5† (47.7±1.0 / 22.0±0.6)¶	-75.1±41.9/-103.9±20.7† (43.6±41.8 / 15.7±21.3)¶
Galβ(1,4)GlcNAc	-67.0±1.1/127.6±0.8† (52.5±1.3 / 8.1±0.8)¶	-68.3±3.5/ 132.5±11.1† (51.1±3.8 / 13.6±12.1)¶

* for an ensemble of 20 refined structures

** phi is defined as O₅-C₁-O_x-C'_x and psi as C₁-O_x-C'_x-C'_{x-1}

† extracted by XtalView [S11]

¶ NMR dihedral angles extracted automatically by CARP [S2], phi is defined as H₁-C₁-O₁-C'_x and psi as C₁-O₁-C'_x-H'_x

Table S4. ^1H - ^1H distances of Le^x attached to FimH (4) used for the structure calculation and measured in a representative structure of the obtained ensemble.

proton pair	Calculated ^1H - ^1H distances ^a [Å]	Applied upper distances restraint ^b [Å]	^1H - ^1H distances in representative model of the final ensemble [Å]
intra			
Gal H1-H2	2.3	3.3 ^e	3.06
Gal H1-H3	2.4	2.9	2.69
Gal H1-H5	2.2	2.7	2.56
Gal H2-H3	2.0	3.0 ^e	3.04
Gal H3-H4	2.1	2.6	2.46
Gal H4-H5	2.1	2.6	2.47
GlcNAc H1-H3	3.0	4.0 ^e	2.60
GlcNAc H1-H5	2.6	3.1	2.52
GlcNAc H2-Q8	4.2	4.7	3.80 ^f
GlcNAc H5-H62	2.6	3.1	3.06
GlcNAc H61-H62 ^c	1.77	–	1.77
GlcNAc H61-H62 ^d	1.77	–	1.77
Fuc H1-H2	2.4	2.9	2.40
Fuc H1-Q6	4.0	4.5	4.20 ^f
Fuc H3-H5	2.9	3.4	2.53
Fuc H4-H5	2.5	3.0	2.45
Fuc H4-Q6	2.6	3.1	2.46 ^f
Fuc H5-Q6	2.8	3.3	1.81 ^f
GlcNAc H1-HN2	3.6	4.1	2.64
GlcNAc H3-HN2	3.5	4.0	2.59
GlcNAc Q8-HN2	3.9	4.4	2.12 ^f
inter			
Gal H1 - GlcNAc H4	2.3	2.8	2.45
Gal H1 - GlcNAc H62	2.6	3.1	2.74
Gal H1 - GlcNAc H61	2.4	2.9	2.34
Gal H2 - Fuc H5	3.0	3.5	2.41
Gal H2 - Fuc Q6	2.8	3.3	2.36 ^f
Gal Q6 - Fuc H3	2.7	3.2	2.50 ^f
GlcNAc H3 - Fuc H1	2.8	3.3	2.52
GlcNAc Q8 - Fuc H1	3.8	4.3	3.01 ^f
GlcNAc HN2 - Fuc H1	3.2	3.7	2.38

a) See Table S2.

b) For the structure calculations upper limit restraints were generated from the calculated ^1H - ^1H distances by adding a 0.5 Å tolerance.

c) Reference restraints for the ^{15}N -filtered-filtered NOESY.

d) Reference restraints for the ^{13}C -filtered-filtered NOESY.

e) The tolerance of these distances was increased by 1 Å instead of a 0.5 Å tolerance.

f) For methyl groups the distance to the average position the three protons (pseudoatom) was extracted and 0.66 Å subtracted. This corresponds to the treatment of upper distance restraints to methyl groups in Amber calculations.

Table S5. Experimental and calculated chemical shifts of the Le^x trisaccharide (**3**), the Fuca(1-3)GlcNAc β methyl glycoside and Gal β (1-4)GlcNAc β methyl glycoside. All chemical shifts are given in ppm.

	Le ^x exp ^a	Le ^x theo ^b	Le ^x cryst1 ^c	Le ^x cryst2 ^c	Fuca(1-3) GlcNAc (exp) ^d	Gal β (1-4) GlcNAc (exp) ^d	Fuca(1-3) GlcNAc (theo)	Gal β (1-4) GlcNAc (theo)
GlcNAc H1	4.52	4.18	4.31	4.31	4.47	4.51	4.24	4.24
GlcNAc H2	3.94	3.92	3.98	3.99	3.82	3.76	3.94	3.83
GlcNAc H3	3.86	3.67	3.55	3.55	3.66	3.72	3.82	3.58
GlcNAc H4	3.94	3.36	3.42	3.43	3.51	3.72	3.33	3.10
GlcNAc H5	3.60	3.53	3.39	3.38	3.50	3.60	3.59	3.54
GlcNAc H61	3.87	3.80	3.69	3.69	3.77	3.84	3.80	3.80
GlcNAc H62	4.01	4.29	4.10	4.09	3.95	4.00	4.28	4.10
GlcNAc Q8	2.04	1.79	1.83	1.83	2.03	2.05	1.78	1.76
Fuc H1	5.12	5.29	5.01	5.00	4.99		5.70	
Fuc H2	3.69	3.88	3.55	3.55	3.71		3.88	
Fuc H3	3.90	3.90	3.69	3.68	3.82		3.92	
Fuc H4	3.79	3.66	3.55	3.54	3.80		3.61	
Fuc H5	4.84	4.90	4.73	4.78	4.32		4.06	
Fuc Q6	1.18	1.13	1.13	1.11	1.17		1.18	
Gal H1	4.45	4.18	4.06	4.18		4.47		4.16
Gal H2	3.50	3.64	3.49	3.66		3.54		3.60
Gal H3	3.65	3.59	3.46	3.59		3.67		3.58
Gal H4	3.90	3.80	4.12	4.11		3.93		3.84
Gal H5	3.59	3.58	3.39	3.34		3.73		3.58
Gal H61	3.73	3.57	3.59	3.71		3.76		3.58

^{a)} measured at 293 K of Le^x propanolamine aglycone (**3**) in D₂O.

^{b)} calculation based as presented in this paper.

^{c)} calculation based on the coordinates of the two models in the Le^x crystal structure ABUCEF^[S4]

^{d)} determined with Fuca1,3GlcNAc β methyl glycoside and Gal β (1-4)GlcNAc β propanolamine aglycone in D₂O at 293 K.

Table S6. Structural parameters of the C-H \cdots O bond in the stacked conformation of Me Le^x (**2**) as determined by experiment and calculated at different levels of theory.

Structural parameter	NMR	OPLS 2005	B3LYP/6-31G(d,p)	ONIOM(MP2/6-31G(d,p): HF/6-31G(d))	CSD
r(H \cdots O) [Å]	2.50 ± 0.01	2.464	2.333	2.220	2.290, 2.311
r(C \cdots O) [Å]	3.56 ± 0.01	3.483	3.424	3.277	3.269, 3.304
a(C-H \cdots O) [°]	165.0 ± 0.9	156.0	175.2	163.7	163.1, 176.4

Table S7. Stacking and C-H \cdots O hydrogen bonding energy calculated at different levels of theory in kcal/mol.

Model system	OPLS 2005	B3LYP/6-31G(d,p)	MP2/6-311++G(d,p) (CP*)	MP2/aug-cc-pVTZ (CP*)	MP2/aug-cc-pVQZ (CP*)
1-deoxy- Fuc \cdots 1-deoxy-Gal	-4.83	-2.62	-7.81 (-3.49)	-6.58 (-4.52)	n/a
Me-O- iPro \cdots Me-O-Me	-1.23	-1.34	-2.86 (-1.23)	-2.52 (-1.76)	-2.14 (-1.84)
1-deoxy- Ara \cdots 1-deoxy-Gal	-4.48	-2.75	-6.76 (-3.16)	-5.83 (-4.03)	n/a
Me-O- Et \cdots Me-O-Me	-1.32	-1.58	-2.51 (-1.30)	-2.34 (-1.73)	-2.05 (-1.80)

* Values in brackets denote interaction energies corrected for the BSSE error by the counterpoise correction.

Table S8. Bacterial strains, plasmids and oligonucleotides used in this study.

Strain, plasmid, or oligonucleotide	Relevant characteristics or sequence	Source or reference
<i>E. coli</i>		
BL21(DE3)	F ⁻ <i>ompT hsdS_B(r_B⁻ m_B⁻) gal dcm</i> (λDE3)	Novagen
pDsbA3	Cloning vector, P _{trc} , amp ^r	[S12]
pFimH1	FimH expression vector based on pDsbA3	[S13]
pFimH2	FimH-S78C expression vector based on pFimH1	This study
<u>Oligonucleotide Primers (5'→3')</u> ^a		
FimH-Extern Forward	CC <u>TCT AGA</u> ATG ATT GTA ATG AAA CGA GTT ATT ACC CTG	This study
FimH-Extern Revers	CC <u>AAG CTT</u> TCG GGC TTT GTT AGC AGC CGG ATC TCA GTG	This study
FimH-S78C Forward	ACC GTA AAA <i>TAT</i> TGT GGC AGT AGC TAT	This study
FimH-S78C Revers	ATA GCT ACT GCC <i>ACA</i> ATA TTT TAC GGT	This study

^{a)} Bold and underlined: cleavage sites for restriction enzymes XbaI and HindIII in forward and reverse primers, respectively. Bold and italic: codon of the mutated amino acid.

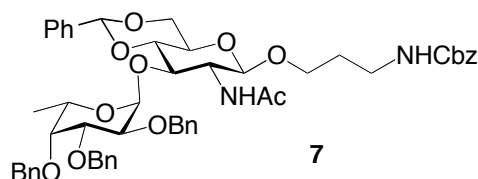
Supplementary Methods

General methods. Commercial materials (Sigma-Aldrich) were used without further purification, solvents were reagent grade (Acros). CH₂Cl₂ and MeOH were dried by passing through an Al₂O₃ (Fluka, type 5016 A basic) column. DMF extra dry (Acros) was used as is. All reactions were performed in oven dried glassware under an atmosphere of argon.

¹H and ¹³C NMR spectra were recorded on a Bruker Avance DMX-500 at room temperature. Chemical shifts are reported in ppm and referenced to TMS using residual solvent peaks.^[S14] For complex molecules the following prefixes were used: Fuc (fucose), Gal (galactose) and GlcNAc (*N*-acetyl glucosamine). The coupling constants (*J*) are reported in Hertz (Hz). Analytical TLC was performed on Merck silica gel 60 F₂₅₄ glass plates and visualized by UV light and charring with a molybdate solution (a 0.02 M solution of ammonium cerium sulfate dihydrate and ammonium molybdate tetrahydrate in aq. 10% H₂SO₄) by heating for 5 min at 140°C. Column chromatography was performed on a CombiFlash Companion (Teledyne-ISCO, Inc.) using RediSep® normal phase disposable flash columns (silica gel). Reversed phase chromatography was carried out with LiChroprep®RP-18 (Merck, 40-63 µm). Optical rotations were determined on a Perkin-Elmer Polarimeter 341. Low resolution mass spectra were measured on a Waters micromass ZQ. High resolution mass spectra (HRMS) were obtained on a micrOTOF spectrometer (Bruker Daltonics, Germany) equipped with a TOF hexapole detector. Purity of final compound was determined on an Agilent 1100 HPLC; detector: ELS, Waters 2420; column: Waters Atlantis dC18, 3 µm, 4.6 x 75 mm; eluents: A: water + 0.1% TFA; B: 90% acetonitrile + 10% water + 0.1% TFA; linear gradient: 0 - 1 min 5% B; 1 - 16 min 5 to 70% B; flow: 0.5 mL/min

MALDI-TOF and ESI-MS glycoprotein analyses were recorded by the Functional Genomic Center Zurich (FGCZ). VIVASPIN® 500 ultrafiltration tubes with 10000 MWCO PES membrane, and ZelluTrans/Roth dialysis membranes MWCO 8000-10000 were used for glycoprotein concentration and dialysis.

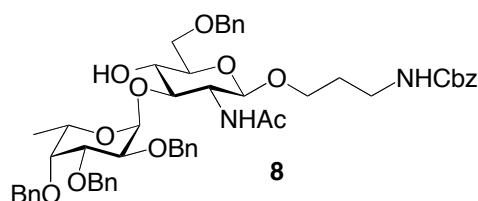
(3-*N*-Benzyloxycarbonylamino)propyl (2,3,4-tri-*O*-benzyl-6-deoxy- α -L-galactopyranosyl)-(1-3)-2-acetamido-4,6-*O*-benzylidene-2-deoxy- β -D-glucopyranoside (7)



Compound **5** (250 mg, 486 μmol), glycosyl donor **6** (465 mg, 972 μmol) and Bu_4NBr (392 mg, 1.22 mmol) were dried for 16 h at high vacuum. Powdered 4Å molecular sieves (600 mg) in DCM/DMF (5 ml, 4:1) was added and the suspension stirred for 4 h at r.t.. CuBr_2 (272 mg, 1.22 mmol) was dried for 20 h at 70°C and added to the suspension and the resulting mixture was stirred for 20 h at r.t.. The mixture was filtered over a short pad of celite and the filtrate was extracted (3x 20 ml) with NH_3 (25%)/ saturated NH_4Cl (1:9) and brine (20 ml). The aqueous layers were washed with dichloromethane (DCM, 3x 20 ml). The combined organic layers were dried (Na_2SO_4) and the solvent was removed in vacuo. The crude product was purified by flash chromatography (EE/toluene 30 – 60%) to give pure **7** (351 mg, 374 μmol , 77%).

R_f (PE/EE 2:3) 0.24; $[\alpha]_D^{22}$ - 68.4 (c 1.56, CHCl_3); $^1\text{H NMR}$ (500.1 MHz, CDCl_3) δ 7.50 – 7.23 (m, 25H, Ar-H), 5.96 (d, J = 6.8 Hz, 1H, *NHAc*), 5.49 (s, 1H, benzylidene *CH*), 5.15 – 5.03 (m, 3H, Fuc-H1, 2x CH_2 -Ph), 4.91 (d, J = 11.5 Hz, 1H, CH_2 -Ph), 4.86 (d, J = 11.3 Hz, 1H, CH_2 -Ph), 4.77 – 4.65 (m, 4H, *GlcNAc*-H1, 3x CH_2 -Ph), 4.58 (d, J = 11.4 Hz, 1H, CH_2 -Ph), 4.31 (dd, J = 10.4, 4.9 Hz, 1H, *GlcNAc*-H6), 4.17 – 4.03 (m, 3H, *GlcNAc*-H3, Fuc-H2, Fuc-H5), 3.93 (dd, J = 10.2, 2.7 Hz, 1H, Fuc-H3), 3.91 – 3.86 (m, 1H, O- CH_2), 3.74 (t, J = 10.3 Hz, 1H, *GlcNAc*-H6'), 3.64 – 3.54 (m, 3H, Fuc-H4, *GlcNAc*-H2, -H4), 3.52 – 3.39 (m, 2H, O- CH_2 , *GlcNAc*-H5), 3.39 – 3.25 (m, 1H, CH_2 -NH), 3.23 – 3.12 (m, 1H, CH_2 NH), 1.81 – 1.64 (m, 5H, 2x CH_2 - CH_2 - CH_2 , CO- CH_3), 0.86 (d, J = 6.4 Hz, 3H, Fuc-H6) ppm. $^{13}\text{C NMR}$ (125.8 MHz, CDCl_3) δ 171.2 (CO- CH_3), 156.7 (OCO-NH), 138.8 – 126.3 (25C, Ar-C), 101.6 (benzylidene *CH*), 101.58 (*GlcNAc*-C1), 98.51 (Fuc-C1), 80.9 (Fuc-C4), 79.9 (Fuc-C3), 77.7 (*GlcNAc*-C4), 77.1 (Fuc-C2), 75.7 (*GlcNAc*-C3), 75.1 (CH_2 -Ph), 74.2 (CH_2 -Ph), 72.8 (CH_2 -Ph), 68.9 (*GlcNAc*-C6), 67.5 (Fuc-C5), 67.1 (O- CH_2), 66.7 (*GlcNAc*-C5), 66.5 (CH_2 -Ph), 57.6 (*GlcNAc*-C2), 38.0 (CH_2 -NH), 29.6 (CH_2 - CH_2 - CH_2), 23.2 (CO- CH_3), 16.5 (Fuc-C6) ppm. ESI-MS Calcd for $\text{C}_{53}\text{H}_{60}\text{N}_2\text{O}_{12}$ [$\text{M}+\text{Na}$] $^+$: 939.40; Found: 939.40.

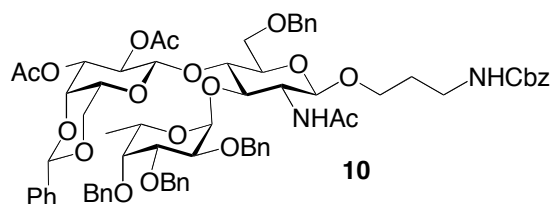
(3-*N*-Benzyloxycarbonylamino)propyl (2,3,4-tri-*O*-benzyl-6-deoxy- α -L-galactopyranosyl)-(1-3)-2-acetamido-6-*O*-benzyl-2-deoxy- β -D-glucopyranoside (8)



Compound **7** (166 mg, 181 μmol) and sodium cyanoborohydride (56.8 mg, 905 μmol) were suspended in THF (5ml) and treated with HCl in ether (1M). The completion of the reaction was monitored by TLC. The mixture was neutralized with sodium hydrogencarbonate and diluted with ethyl acetate (20ml). The mixture was washed with saturated aqueous NaHCO_3 (15 ml) and brine (15 ml). The organic layer was dried (Na_2SO_4) and concentrated. The crude product was purified by flash chromatography to yield **8** as a white solid (141 mg, 154 μmol , 85%).

R_f (PE/EE 2:3) 0.20; $[\alpha]_D^{22}$ -37.2 (c 1.84, CHCl_3); ^1H NMR (500.1 MHz, CDCl_3) δ 7.42 – 7.26 (m, 25H, Ar-H), 5.08 (s, 2H, 2x CH_2 -Ph), 4.96 (d, $J = 11.4$ Hz, 1H, CH_2 -Ph), 4.93 (d, $J = 2.5$ Hz, 1H, Fuc-H1), 4.84 – 4.78 (m, 2H, CH_2 -Ph), 4.75 (d, $J = 11.8$ Hz, 1H, CH_2 -Ph), 4.69 – 4.60 (m, 3H, GlcNAc-H1, 2x CH_2 -Ph), 4.60 – 4.52 (m, 2H, 2x CH_2 -Ph), 4.13 – 4.04 (m, 2H, Fuc-H5, Fuc-H2), 3.97 – 3.88 (m, 2H, Fuc-H3, O- CH_2), 3.83 – 3.76 (m, 1H, GlcNAc-H6), 3.70 – 3.53 (m, 4H, Fuc-H4, GlcNAc-H6', GlcNAc-H2, O- CH_2), 3.53 – 3.43 (m, 2H, GlcNAc-H3, GlcNAc-H5), 3.43 – 3.34 (m, 2H, GlcNAc-H4, CH_2 -NH), 3.24 – 3.15 (m, 1H, CH_2 -NH), 1.82 – 1.74 (m, 1H, CH_2 - CH_2 - CH_2), 1.73 – 1.67 (m, 1H, CH_2 - CH_2 - CH_2), 1.61 (s, 3H, CO- CH_3), 1.14 (d, $J = 6.4$ Hz, 3H, Fuc-H6). ^{13}C NMR (125.8 MHz, CDCl_3) δ 156.9 (OCO- NH_2), 138.7 – 127.7 (25C, Ar-C), 100.8 (GlcNAc-C1), 99.7 (Fuc-C1), 85.1 (GlcNAc-C3), 79.2 (Fuc-C3), 77.4 (Fuc-C4), 76.3 (Fuc-C2), 75.2 (CH_2 -Ph), 75.1 (GlcNAc-C5), 74.3 (CH_2 -Ph), 73.5 (CH_2 -Ph), 73.2 (CH_2 -Ph), 70.71 (GlcNAc-C4), 69.7 (GlcNAc-C6), 68.3 (Fuc-C5), 67.0 (O- CH_2), 66.7 (CH_2 -Ph), 55.7 (GlcNAc-C2), 37.8 (CH_2 -NH), 29.6 (CH_2 - CH_2 - CH_2), 23.0 (CO- CH_3), 16.8 (Fuc-C6) ppm. ESI-MS Calcd for $\text{C}_{53}\text{H}_{62}\text{NaN}_2\text{O}_{12}$ $[\text{M}+\text{Na}]^+$: 941.42; Found: 941.47.

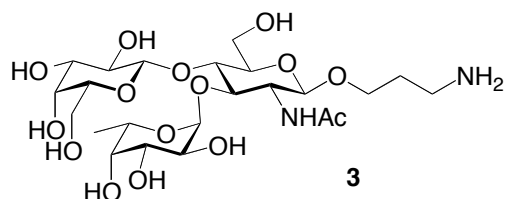
(3-*N*-Benzyloxycarbonylamino)propyl (2,3-di-*O*-acetyl-4,6-*O*-benzylidene- β -D-galactopyranosyl)-(1-4)-[(2,3,4-tri-*O*-benzyl-6-deoxy- α -L-galactopyranosyl)-(1-3)]-2-acetamido-6-*O*-benzyl-2-deoxy- β -D-glucopyranoside (**10**)



Compound **8** (781 mg, 0.850 mmol) and thioglycoside **9** (607 mg, 1.53 mmol) were dissolved in dry DCM (35 ml) and stirred together with powdered 4Å activated molecular sieves (10 g) for 4 h at r.t. DMTST (658 mg, 2.55 mmol) was dissolved in DCM (15 ml) and stirred together with powdered 4 Å activated molecular sieves (5g) for 4 h at r.t. as well. Both suspensions were combined and stirred for 16 h at r.t.. The mixture was filtered over a short pad of celite, diluted with DCM (100 ml), and washed with a saturated solution of NaHCO₃ (50 ml) and water (50 ml). The aqueous phases were extracted with DCM (3× 30 ml). The combined organic layers were dried (Na₂SO₄) and the solvent was removed in vacuo. The crude product was purified by flash chromatography (PE/EE 1:1) to yield **10** as a colorless oil (640 mg, 0.51 mmol, 60 %).

R_f (PE/EE 2:3) 0.25; [α]_D²² -24.6 (c 0.82, CHCl₃); [α]_D²² -24.6 (c 0.82, CHCl₃); ¹H NMR (500.1 MHz, CDCl₃) δ 7.58 – 7.13 (m, 30H, Ar-H), 5.74 (d, *J* = 6.4 Hz, 1H, NHAc), 5.54 (s, 1H, CH benzylidene), 5.29 (dd, *J* = 10.2, 8.3 Hz, 2H, Gal-H2), 5.11 – 5.00 (m, 3H, 2x Ph-CH₂, GlcNAc-H1), 4.94 (d, *J* = 2.7 Hz, 1H, Fuc-H1), 4.84 (d, *J* = 11.8 Hz, 1H, Ph-CH₂), 4.76 – 4.67 (m, 3H, Ph-CH₂, Gal-H3, Fuc-H5), 4.67 – 4.59 (m, 3H, Ph-CH₂, Gal-H6), 4.56 (d, *J* = 8.2 Hz, 1H, Gal-H1), 4.38 (d, *J* = 12.1 Hz, 1H, Ph-CH₂), 4.33 – 4.17 (m, 3H, 2x Ph-CH₂, Gal-H4), 4.12 (t, *J* = 9.4 Hz, 1H, GlcNAc-H3), 4.03 – 3.93 (m, 3H, Ph-CH₂, GlcNAc-H4, Fuc-H2), 3.92 – 3.80 (m, 3H, Fuc-H3, O-CH₂, GlcNAc-H6), 3.79 – 3.71 (m, 1H, GlcNAc-H6'), 3.56 (d, *J* = 11.2 Hz, 1H, Ph-CH₂), 3.53 – 3.47 (m, 1H, O-CH₂), 3.43 – 3.36 (m, 1H, GlcNAc-H5), 3.29 – 3.17 (m, 4H, 2x CH₂-NH₂, GlcNAc-H₂, Fuc-H4), 3.04 – 2.99 (m, 1H, Gal-H5), 2.12 (s, 3H, CO-CH₃), 2.02 (s, 3H, CO-CH₃), 1.76 – 1.67 (m, 2H, CH₂-CH₂-CH₂), 1.59 (s, 3H, CO-CH₃), 1.12 (d, *J* = 6.5 Hz, 3H, Fuc-H6). ¹³C NMR (125.8 MHz, CDCl₃) δ 170.8 (CO-CH₃), 169.9 (CO-CH₃), 168.9 (CO-CH₃), 156.7 (OCO-NH), 139.6-125.9 (36C, Ar-C), 99.9 (benzylidene-CH), 99.8 (Gal-C1), 99.6 (GlcNAc-C1), 98.2 (Fuc-C1), 79.6 (Fuc-C3), 78.9 (Fuc-C4), 76.0 (Fuc-C2), 75.1 (Ph-CH₂), 75.0 (GlcNAc-C5), 74.4 (GlcNAc-C4), 74.3 (Ph-CH₂), 73.8 (GlcNAc-C3), 73.5 (Ph-CH₂), 73.4 (Gal-C4), 72.2 (Gal-C3), 71.5 (Gal-C6), 69.2 (Ph-CH₂), 68.8 (Gal-C2), 68.0 (GlcNAc-C6), 67.0 (O-CH₂), 66.6 (Ph-CH₂), 66.5 (Fuc-C5), 66.4 (Gal-C5), 59.4 (GlcNAc-C2), 37.9 (CH₂-NH₂), 29.4 (CH₂-CH₂-CH₂), 23.3 (NH-CO-CH₃), 21.0 (CO-CH₃), 20.9 (CO-CH₃), 16.2 (Fuc-C6) ppm. ESI-MS Calcd for C₇₀H₈₀N₂NaO₁₉ [M+Na]⁺: 1275.53; Found: 1275.74.

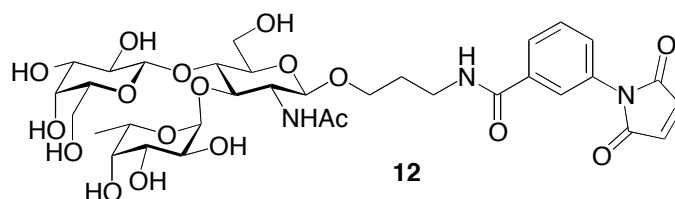
3-Aminopropyl (β-D-galactopyranosyl)-(1-4)-[(6-deoxy-α-L-galactopyranosyl)-(1-3)]-2-acetamido-2-deoxy-β-D-glucopyranoside (3)



Compound **10** (24.8 mg; 19.8 μmol) was dissolved in MeOH (1 ml), treated with NaOMe/MeOH (200 μl , 0.02 M) and stirred for 16 h at r.t.. The reaction was quenched with two drops of glacial acetic acid and concentrated in vacuo. The resulting alcohol **11** was dissolved in DCM/MeOH/AcOH/H₂O (1:1:2:2, 3 ml) and Pd(OH)₂/C (5 mg) as added. The suspension was stirred under an atmosphere of hydrogen for 18 h. The reaction mixture was filtered and purified by reversed phase flash chromatography and size exclusion chromatography to yield **3** as a white foam (8.9 mg, 15.2 μmol , 77% over 2 steps).

$[\alpha]_{\text{D}}^{22}$ -68.6 (*c* 0.96, D₂O); ¹H NMR (500.1 MHz, D₂O) δ 5.11 (d, *J* = 3.9 Hz, 1H, Fuc-H1), 4.85 – 4.79 (m, 2H, Fuc-H5), 4.51 (d, *J* = 8.3 Hz, 1H, GlcNAc-H1), 4.44 (d, *J* = 7.8 Hz, 1H, Gal-H1), 4.05 – 3.97 (m, 2H, O-CH₂, GlcNAc-H6), 3.96 – 3.81 (m, 7H, GlcNAc-H2, -H4, Fuc-H3, Gal-H4, GlcNAc-H3, -H6'), 3.80 – 3.77 (m, 1H, Fuc-H4), 3.76 – 3.66 (m, 4H, Gal-H6, -H6', O-CH₂, Fuc-H2), 3.64 (dd, *J* = 9.8, 3.2 Hz, 1H, Gal-H3), 3.62 – 3.56 (m, 2H, GlcNAc-H5, Gal-H5), 3.51 – 3.46 (m, 1H, Gal-H2), 3.10 – 3.04 (m, 2H, CH₂-NH₂), 2.03 (s, 3H, CO-CH₃), 1.98 – 1.90 (m, 2H, CH₂-CH₂-CH₂), 1.17 (d, *J* = 6.6 Hz, 3H, Fuc-H6) ppm; ¹³C NMR (125.8 MHz, CDCl₃) δ 177.1 (CO-Ac), 104.5 (Gal-C1), 103.7 (GlcNAc-C1), 101.3 (Fuc-C1), 78.0 (GlcNAc-C5), 77.6 (GlcNAc-C3), 77.5 (Gal-C5), 76.0 (GlcNAc-C4), 75.2 (Gal-C3), 74.6 (Fuc-C4), 73.7 (Gal-C2), 71.9 (Fuc-C3), 71.0 (Gal-C4), 70.7 (O-CH₂), 70.4 (Fuc-C2), 69.4 (Fuc-C5), 64.2 (Gal-C6), 62.4 (GlcNAc-C6), 58.5 (GlcNAc-C2), 40.3 (CH₂-NH₂), 29.4 (CH₂-CH₂-CH₂), 24.9 (CO-CH₃/Acetate), 18.0 (CO-CH₃), 18.0 (Fuc-C6) ppm; HR-MS (ESI) Calcd for C₂₃H₄₂N₂NaO₁₅ [M+Na]⁺: 609.2483; Found: 609.2484.

3-(3-(2,5-dioxo-2,5-dihydro-1H-pyrrol-1-yl)benzamido)propyl (β-D-galactopyranosyl)-(1-4)-[6-deoxy-α-L-galactopyranosyl]-2-acetamido-2-deoxy-β-D-glucopyranoside (12)

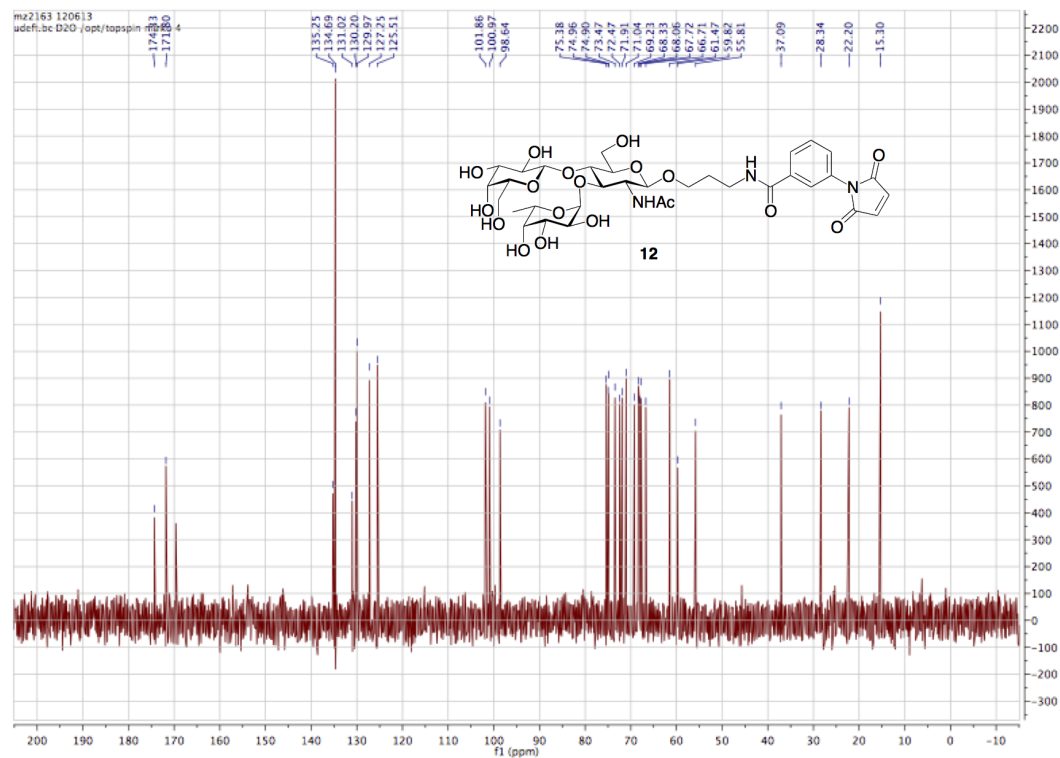
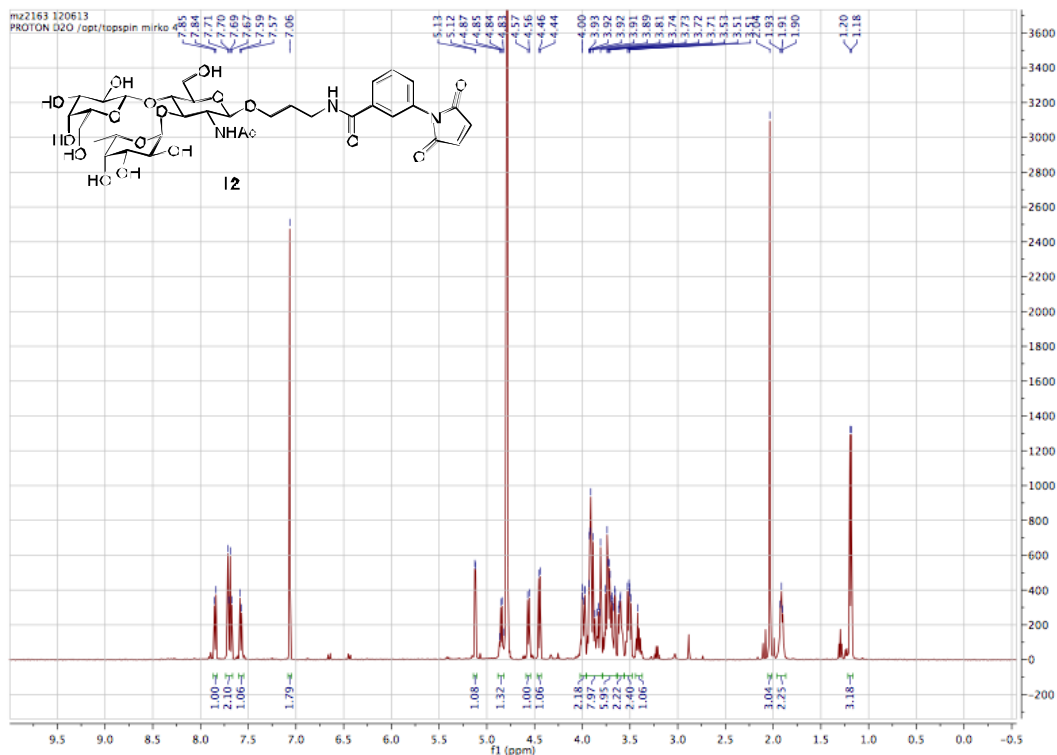


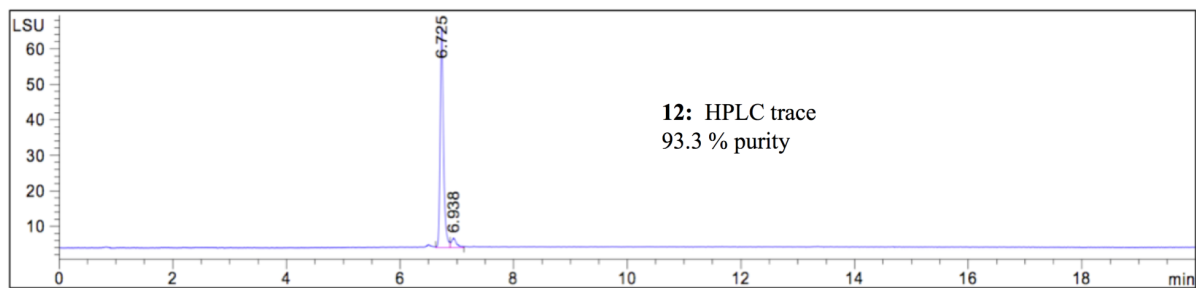
Compound **3** (11.0 mg, 18.8 μmol) was dissolved in H₂O (400 μl) and a solution of

3-maleimidobenzoic acid *N*-hydroxysuccinimide ester (11.8 mg, 37.5 μmol) in DMSO (1 ml) was added. The solution was stirred for 1.5 h at r.t. The water was removed in vacuo and the crude product purified by flash chromatography (MeCN/MeOH + 0.01% TFA, 1 to 3/4) and lyophilized to give **12** as a white foam (9.1 mg, 11.6 μmol , 62%).

R_f (DCM/MeOH 1.25:1) 0.20; $[\alpha]_D^{22}$ -53.7 (c 0.54, H_2O); ^1H NMR (500.1 MHz, D_2O) δ 7.87 – 7.82 (m, 1H, Ar-H), 7.73 – 7.67 (m, 2H, Ar-H), 7.60 – 7.56 (m, 1H, Ar-H), 7.06 (s, 2H, -CH=CH-), 5.12 (d, J = 4.0 Hz, 1H, Fuc-H1), 4.85 (q, J = 6.8 Hz, 1H, Fuc-H5), 4.56 (d, J = 8.0 Hz, 1H, GlcNAc-H1), 4.45 (d, J = 7.8 Hz, 1H, Gal-H1), 4.03 – 3.96 (m, 2H, O-CH₂, GlcNAc-H6), 3.96 – 3.64 (m, 14H, GlcNAc-H2, Gal-H4, Fuc-H3, GlcNAc-H4, -H3, -H6', Fuc-H4, Gal-H6, -H6', O-CH₂, Fuc-H2, Gal-H3), 3.64 – 3.55 (m, 2H, GlcNAc-H5, Gal-H5), 3.55 – 3.47 (m, 2H, Gal-H₂, CH₂-NH), 3.45 – 3.37 (m, 1H, CH₂-NH), 2.04 (s, 3H, CO-CH₃), 1.95 – 1.88 (m, 2H, CH₂-CH₂-CH₂), 1.19 (d, J = 6.6 Hz, 3H, Fuc-H6) ppm; ^{13}C NMR (126 MHz, D_2O) δ 174.3, 171.8, 169.7 (3C, CO), 134.7 (2C, CO-CH), 135.3, 131.0, 130.2, 130.0, 127.3, 125.5 (6C, Ar-C), 101.9 (Gal-C1), 101.0 (GlcNAc-C1), 98.6 (Fuc-C1), 75.4 (GlcNAc-C5), 75.0 (Gal-C5), 74.9 (GlcNAc-C3), 73.5 (GlcNAc-C4), 72.5 (Fuc-C2), 71.9 (Fuc-C4), 71.0 (Gal-C2), 69.2 (Fuc-C3), 68.3 (Gal-C4), 68.1 (O-CH₂), 67.7 (Fuc-C2), 66.7 (Fuc-C5), 61.5 (Gal-H6), 59.8 (GlcNAc-C6), 55.8 (GlcNAc-C2), 37.1 (CH₂-NH), 28.3 (CH₂-CH₂-CH₂), 22.2 (CO-CH₃), 15.30 (Fuc-C6) ppm; HR-MS (ESI) Calcd for $\text{C}_{34}\text{H}_{47}\text{N}_3\text{NaO}_{18}$ $[\text{M}+\text{Na}]^+$: 808.2752; Found: 808.2752.

Experimental Data





Supplementary References

- [S1] Neuhaus, D.; Williamson, M.P. *The Nuclear Overhauser Effect in Structural and Conformational Analysis*; Wiley-VCH: New York, 2000.
- [S2] Lütteke, T.; Frank, M.; von der Lieth, C.-W. *Nucleic Acids Res.* **2005**, *33*, 242-246.
- [S3] Dapprich, S.; Komáromi, I.; Byun, K. S.; Morokuma, K.; Frisch, M. J. *J. Mol. Struct. (Theochem)*, **1999**, *462*, 1-21.
- [S4] Perez, S.; Mouhous-Riou, N.; Nifant'ev, N. E.; Tsvetkov, Y. E.; Bachet, B.; Imberty, A. *Glycobiology* **1996**, *6*, 537-542.
- [S5] Azurmendi, H. F.; Martin-Pastor, M.; Bush, C. A. *Biopolymers* **2002**, *63*, 89-98.
- [S6] Miller, K. E.; Mukhopadhyay, C.; Cagas, P.; Bush, C. A. *Biochemistry* **1992**, *31*, 6703-6709.
- [S7] Homans, S. W.; Forster, M. *Glycobiology* **1992**, *2*, 143-151.
- [S8] Wormald, M. R.; Edge, C. J.; Dwek, R. A. *Biochem. Biophys. Res. Commun.* **1991**, *180*, 1214-1221.
- [S9] Ichikawa, Y.; Lin, Y. C.; Dumas, D. P.; Shen, G. J.; Garciajunceda, E.; Williams, M. A.; Bayer, R.; Ketcham, C.; Walker, L. E.; Paulson, J. C.; Wong, C. H. *J. Am. Chem. Soc.* **1992**, *114*, 9283-9298.
- [S10] Imberty, A.; Mikros, E.; Koca, J.; Mollicone, R.; Oriol, R.; Perez, S. *Glycoconj J.* **1995**, *12*, 331-349.
- [S11] D. E. McRee D. E. *J. Struct. Biol.* **1999**, *125*, 156-165.
- [S12] Hennecke, J.; Sebbel, P.; Glockshuber, R. *J. Mol. Biol.* **1999**, *286*, 1197-1215.
- [S13] Rabbani, S.; Jiang, X. H.; Schwardt, O.; Ernst, B. *Anal. Biochem.* **2010**, *407*, 188-195.
- [S14] Gottlieb, H. E.; Kotlyar, V.; Nudelman, A. *J. Org. Chem.* **1997**, *62*, 7512-7515.

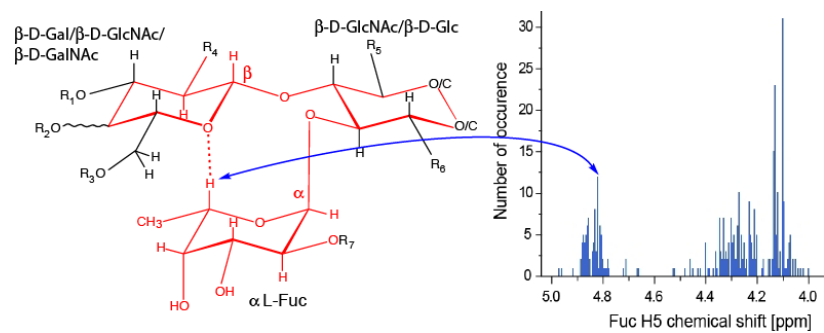
**3.1.2 A secondary structure in a wide range of fucosylated glyco-epitopes
(Manuscript)**

This manuscript is formatted according to guidelines of *PNAS (Proceedings of the National Academy of Sciences of the United States of America)*

Contributions:

- Synthesis of LDNF trisaccharide
- Discussion

Table of content



A secondary structure element in a wide range of fucosylated glyco-epitopes

Thomas Aeschbacher^a, Mirko Zierke^b, Martin Smieško^b, Mayeul Colloc^{c,1}, Jean-Maurice Mallet^c, Beat Ernst^b, Frédéric H.-T. Allain^{a,2}, Mario Schubert^{a,2}

^a Institute of Molecular Biology and Biophysics, ETH Zurich, CH-8093 Zurich, Switzerland; ^b Institute of Molecular Pharmacy, University of Basel, Klingelbergstr. 50, CH-4056 Basel, Switzerland; ^c Ecole Normale Supérieure, Département de Chimie, 24 rue Lhomond, 75005 Paris, France.

Submitted to Proceedings of the National Academy of Sciences of the United States of America

The increasing understanding of the essential role of carbohydrates in development, infection and in a wide range of diseases fuels a rapidly growing interest in the basic principles governing carbohydrate-protein interactions. An important yet much debated parameter for recognition is the flexibility or rigidity of oligosaccharides. Here we introduce a set of abundant trisaccharide motifs that adapt a single conformation stabilized by a non-conventional C-H...O hydrogen bond, a rigid structural motif in a broad class of oligosaccharides, found in bacteria to mammals that include Lewis blood group antigens. Combining NMR structure determination with density functional theory (DFT) and extensive database searches, we have identified the consensus motifs X-β1,4-[Fucα1,3]-Y and X-β1,3-[Fucα1,4]-Y that share the same three-dimensional architecture - a secondary structure we name [3,4]F-branch. The wide spectrum of possible modifications and extensions of these consensus motifs points towards a large variety of glyco-epitopes that nature generated using the same underlying architecture.

NMR spectroscopy | C-H...O hydrogen bond | secondary structure | solution conformation | glycan

Carbohydrates are found in great abundance on the cell surfaces of all organisms and are central for numerous biological processes, including cell-cell adhesion, cellular recognition and various signaling processes (1-3). By linking monosaccharide moieties, a wide and diverse range of highly specific and compact glyco-epitopes can be built. Part of this diversity is generated by branching a linear oligosaccharide with a fucose moiety through different types of linkages. Fucosylated glycans have important roles in tissue development, cell-adhesion (such as selectin-mediated leucocyte-endothelial adhesion), fertilization and host-microbe interactions (4-6). Fucose is an integral part of the human ABO and Lewis blood-group antigens (7, 8). Moreover, altered fucosylation has been observed in various cancer cells (9) and diseases, including type 1 diabetes mellitus, rheumatoid arthritis and cystic fibrosis (6).

Understanding the molecular basis of the specificity and affinity of carbohydrate-protein interactions is the key to understanding the glycode and its biological roles. In this context, information on the conformation of the carbohydrate ligands in the unbound and bound form is of particular interest. A ligand that is pre-formed in its bioactive conformation will be entropically favored over a flexible ligand, owing to an improved on-rate (k_{on}), which in turn leads to a higher affinity. In general, carbohydrates are considered to be flexible molecules (10), but there are a few reports on carbohydrates that adapt well-defined conformations (11-14). The ongoing debate whether a given carbohydrate adopts a single conformation in solution or whether it is flexible and adopts several conformations (15-19) remains unresolved, not least as a general explanation is missing of how a single conformation can be stabilized.

In the case of the blood-group antigen Lewis^x (Le^x, Galβ1,4[Fucα1,3]GlcNAcβ) it has been observed that there is a strong stacking interaction between Gal and Fuc (15, 20-24),

which was attributed to van-der-Waals contacts, steric hindrance especially with the adjacent GlcNAc moiety and the exo-anomeric effect (17, 25). We have recently determined the NMR structure of Le^x bound to a carrier protein to increase the tumbling time and found that a non-conventional C-H...O hydrogen bond stabilizes together with the exo-anomeric effect, steric hindrance and hydrophobic stacking interactions the conformation in this particular compound (26). The C-H...O hydrogen bond contributes ~1.8 kcal to the stabilization energy at the Gal-Fuc interface (total 4.5 kcal), locking the already quite restricted conformational space of that trisaccharide.

Here we introduce a general principle that explains how a large range of carbohydrates use such a C-H...O hydrogen bond to reach a stabilized conformation and thus a defined three-dimensional structure in solution. We have identified a conserved trisaccharide architecture that is common to many glyco-epitopes and can be compared to secondary structural elements in proteins, serving as an underlying scaffold for the presentation of various substitutions for different functions. Our findings are based on structural data obtained with an experimentally simple and robust approach for determining the 3D solution structure of small oligosaccharides, and on extensive searches through a carbohydrate database and through reported crystal structures to derive the minimal requirements for such a stabilization.

Significance

This study will change our understanding of functional entities in carbohydrates - the glycode. Whereas several thousand functional glycoepitopes are estimated in nature, knowledge of their functions is very limited. Understanding the specific recognition of those glycoepitopes on a molecular level is key to understand their function and in that context the glycoepitope conformation is central. We found with NMR spectroscopy, density functional theory calculations and crystal structure database analysis a common secondary structure element shared by many glycoepitopes that provides a scaffold for generating functionally different epitopes. Our results not only explain how natural carbohydrate ligands stabilize their bioactive conformation but will be very beneficial for designing potent mimetics for therapeutic applications.

Reserved for Publication Footnotes

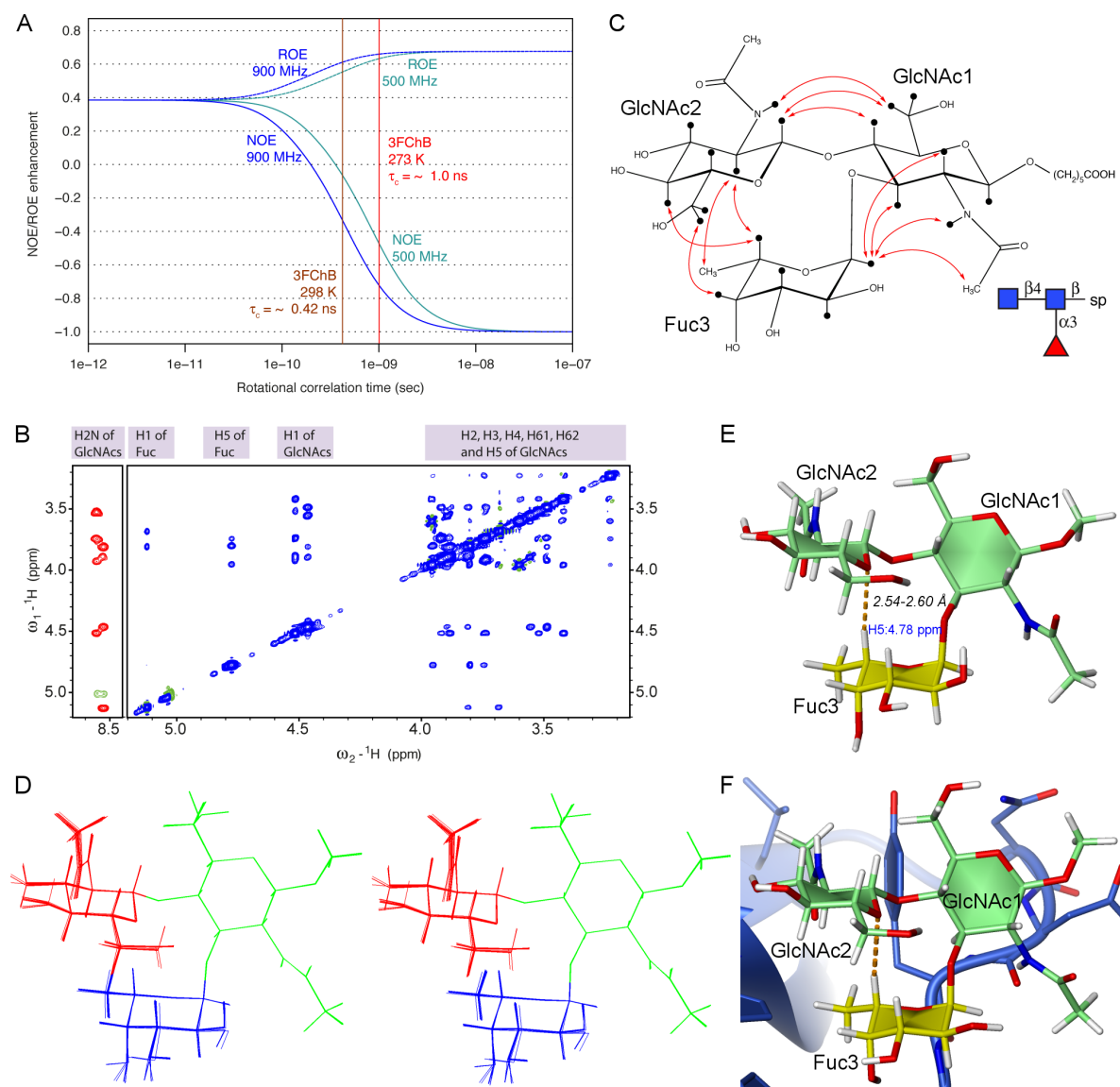


Fig. 1. Structure determination of α 1,3-fucosylated chitobiose (3FChB) in solution. (A) NOE enhancement factors of a transient NOE experiment as a function of the correlation time τ_c indicated for two magnetic fields corresponding to 500 and 900 MHz. Estimated correlation times of 3FChB (MW: 684 g/mol) in D_2O are indicated by vertical lines. (B) 2D NOESY spectra of 3FChB (GlcNAc β 1,4[Fuca1,3]GlcNAc-O(CH $_2$) $_5$ COOH, 2.8 mM) in either H_2O (red, left) or in D_2O (blue, right) recorded at 900 MHz and 277 K. Chemical shift assignments are indicated on the top for isolated resonances. (C) Schematic presentation of 3FChB showing the obtained inter-residue NOEs by red arrows. (D) Ensemble of the 20 best structures of 3FChB based on NOESY data recorded at 277 K. (E) Representative NMR structure of 3FChB. (F) Structure of 3FChB in complex with the lectin CCL2 (27).

Results

A general approach to the 3D-structure determination of carbohydrates by NMR spectroscopy

The most reliable approach to three-dimensional structure determination in solution is based on the nuclear Overhauser effect (NOE). Whereas nuclear Overhauser enhancement factors of an optimal value of -1 are observed for biopolymers, for small molecules with masses around 500 Da these factors are close to zero (Fig. 1A). A more favorable NOE enhancement factor can be achieved by measuring at low temperatures resulting in a larger viscosity and slower tumbling time. Thus, a decrease in

temperature from 25 °C to 0 °C increases the estimated tumbling time approximately by a factor of 2. In addition, the use of ultra-high magnetic fields further increases the absolute value of the NOE enhancement factor for small molecules (Fig. 1A). When applied to the free trisaccharide GlcNAc β 1,4[Fuca1,3]GlcNAc (α 1,3-fucosylated chitobiose; 3FChB), which is the target of the recently discovered *Coprinopsis cinerea* lectin 2 (CCL2) (27), excellent NOE cross-peaks in 2D NOESY spectra were obtained (Fig. 1B and C). To obtain pure absorptive NOE cross-peaks, a magnetic field of 900 MHz together with a temperature just above the freezing point were required. The exceptionally high

Table 1. Distances between H5 of Fuc and O5 of Gal/GlcNAc in the NMR structures, one Cambridge Structural Database (CSD) structure and Protein Databank (PDB) structures as evidence for C–H...O hydrogen bond. For more details see Table S5.

Method	Glycomotif	Type of structure or PDB/CSD entry identifier	Number of motifs	Distance H5(Fuc)-O5(Gal/GlcNAc)
NMR	3FChB	ensemble	20	2.58 ± 0.01 Å ^a
		DFT minimized	1	2.35 Å
	Le ^x	ensemble	20	2.50 ± 0.01 Å ^a
		DFT minimized	1	2.32 Å
	Le ^a	ensemble	20	2.49 ± 0.01 Å ^a
		DFT minimized	1	2.33 Å
	LDNF	ensemble	20	2.36 ± 0.01 Å ^a
		DFT minimized	1	2.40 Å
	Bv9	ensemble	20	2.54 ± 0.02 Å ^a
		DFT minimized	1	2.33 Å
LDFT	ensemble	20	2.53 ± 0.02 Å ^a	
	DFT minimized	1	2.47 Å	
X-ray (CSD)	Le ^x	ABUCEF	2	2.31 and 2.28 Å
X-ray (PDB)	Le ^x	1G1R, 1G1S, 1G1T, 1KMB, 1S1S, 1S1L, 1U28, 2KMB, 2OX9, 2R61, 2RDG, 2Z8L, 3AP9, 3KMB, 3PVD, 4DXG, 4KMB ^b	38	2.39 ± 0.25 Å ^c
		1CLY, 1CLZ, 1GSL, 1S3K, 2J1T, 3EYV, 3LEG, 3PA2, 3PUN, 4AH5, 4GW1 ^d	12	2.45 ± 0.17 Å ^c
	Le ^a	1FWV, 1W8H, 3ASR ^e	7	2.39 ± 0.10 Å ^c
	Le ^b	1LED, 3ASS, 3AST, 3LEK, 3SEJ, 4GWJ	7	2.25 ± 0.33 Å ^c
	3FChB (N-glycan cores)	1LK9, 1JU2, 1E4M, 1E6Q, 1E6S...	11 (only selection)	2.42 ± 0.15 Å ^c
	difucosylated N-glycan cores	1YM0, 2B9L, 2F9N, 2QQM, 3L9R, 3QW9, 4ARN, 4GWM, 4GWN, 4GZT ^f	17	2.40 ± 0.15 Å ^c
	Le ^x -like	1W8F	4	2.37 ± 0.03 Å ^c
	Le ^y -like	2O2L, 3EFX ^g	19	2.49 ± 0.15 Å ^c

^a average distances and standard deviation of the ensemble consisting of 20 structures

^b excluded 3ZV1 that displayed an elongated Le^x conformation

^c average distances and standard deviations of all motifs within one category; protons were added to the structures by Maestro (Schrödinger) because the crystal structures lacked protons (details in Table S5)

^d excluded 2WVG, 2WVK, 2 structures of 4AH5 that displayed an elongated Le^y conformation

^e excluded 3UET that displayed an elongated Le^a conformation

^f excluded 3UOP that displayed an elongated difucosylated N-glycan core conformation

^g excluded chain S of 2O2L that displayed an elongated Le^y-like conformation

number of 11 inter-residual NOEs (Fig. 1C) corresponding to 5.5 NOE distance restraints per glycosidic linkage was used for NMR structure determination and resulted in a well-defined ensemble (Fig. 1D, statistics in Table S1) that is virtually identical to the structure of the trisaccharide when bound to CCL2 (27) (Fig. 1F). The distal GlcNAc moiety stacks against Fuc (Fig. 1D and E). To our surprise the structure revealed the presence of a C–H...O hydrogen bond between H5 of Fuc and the ring oxygen O5 of GlcNAc2 (Fig. 1E), indicated by a C–H bond vector directly oriented towards the oxygen and a H...O distance of 2.6 Å, slightly shorter than expected from the sum of the van-der-Waals radii of 2.7 Å (28) (Table 1). However, the GLYCAM force field (29) used for the NMR structure determination does not allow close contact between non-bonded atoms implying that the H5...O5 distance might be even shorter. The C–H...O hydrogen bond in 3FChB results in a characteristic chemical shift of H5 of Fuc around 4.8 ppm as illustrated in a ¹³C HSQC (Fig. 2A). When the chemical shifts are compared to Fucα1,3GlcNAcβOME that lacks the stacking GlcNAc2, a large ¹H chemical shift difference is observed for the H5 of Fuc with $\Delta\delta = 0.45$ ppm, whereas the other Fuc chemical shifts hardly change (Fig. 2A, Table S2). Such a downfield shift is one of the characteristics of a hydrogen bond (30). The GlcNAc1 signals differ only slightly except for H4 and C4 to which the stacking GlcNAc2 is attached in 3FChB. A similar H5 Fuc chemical shift was observed previously for the trisaccharide Galβ1,4[Fucα1,3]GlcNAc, the Le^x blood group epitope, leading

to the initial identification of a nonconventional CH...O hydrogen bond as stabilizing element (26). To cross-validate the presented low temperature approach to NMR structure determination of small molecules in solution we also applied it to Le^x methyl glycoside (Table S1). A virtually identical structure as reported by Zierke *et al.* (26) was obtained (Fig. S1), but without the need for the chemical attachment to an isotopically labeled carrier protein.

A C–H...O hydrogen bond is indicated by a characteristic chemical shift

To further verify the presence of the C–H...O hydrogen bond we used density functional theory (DFT) (31) at the B3LYP/6-31G(d,p) level (32) using an implicit water solvation model to minimize the geometry of one representative of each structural ensemble and thus excluding the influence of a force field. The resulting structures of 3FChB and Le^x (Figs. S2 A and B) are very similar to the initial NMR structures but show shorter H5...O5 distances of 2.3–2.4 Å (Table 1), now significantly shorter than the sum of the van der Waals radii of 2.7 Å, further supporting the C–H...O hydrogen bond. Furthermore we used the wave functions obtained from the DFT calculations to localize and calculate the electron density at the bond critical points $\rho(r_c)$ within the C–H...O hydrogen bonds with Bader's atom in molecules (AIM) theory (33) (Table S3). The existence of bond critical points with a sufficiently high electron density $\rho(r_c)$ of 0.013 and 0.014 au is another characteristic of hydrogen bonds (30) and thus confirms the existence of the C–H...O hydrogen bond in 3FChB.

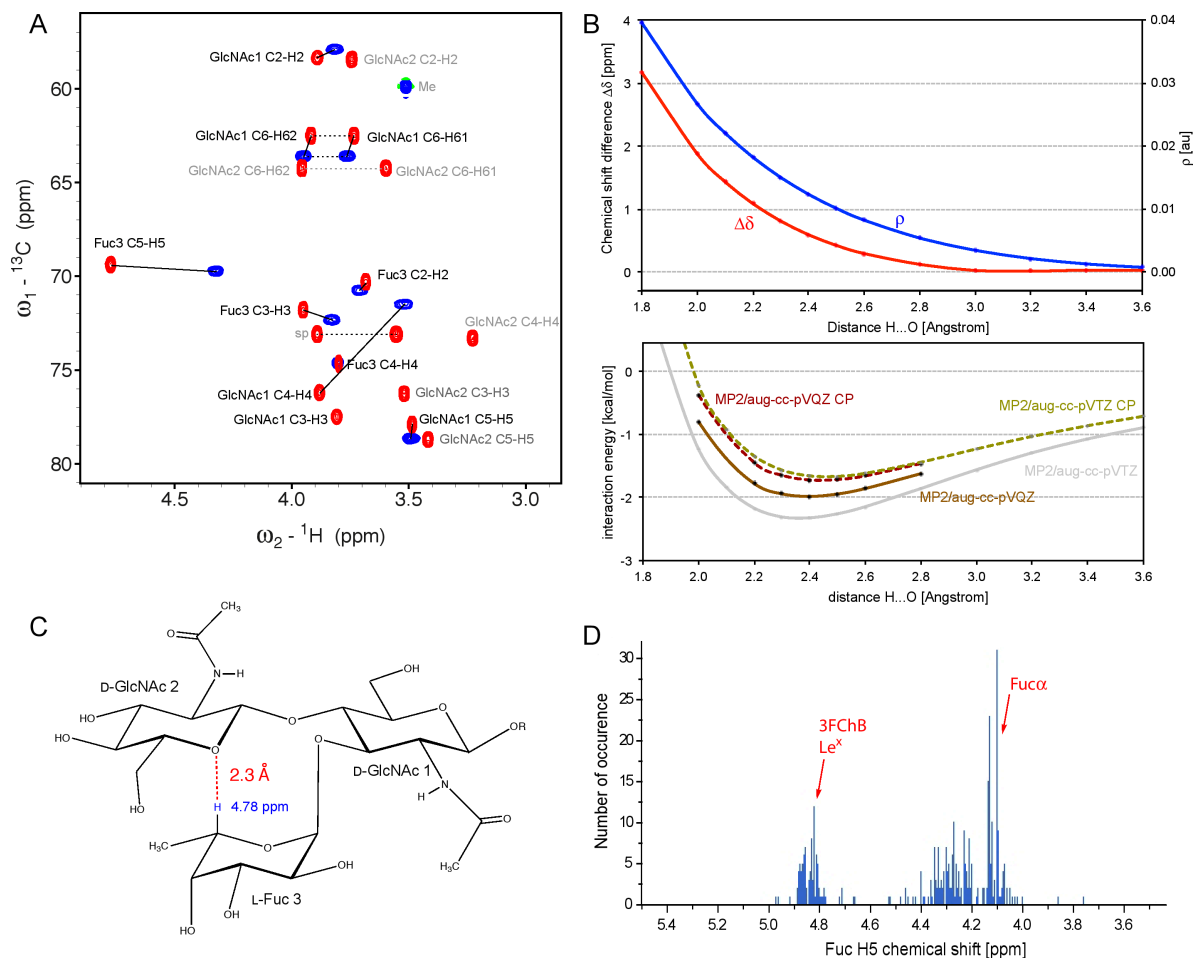
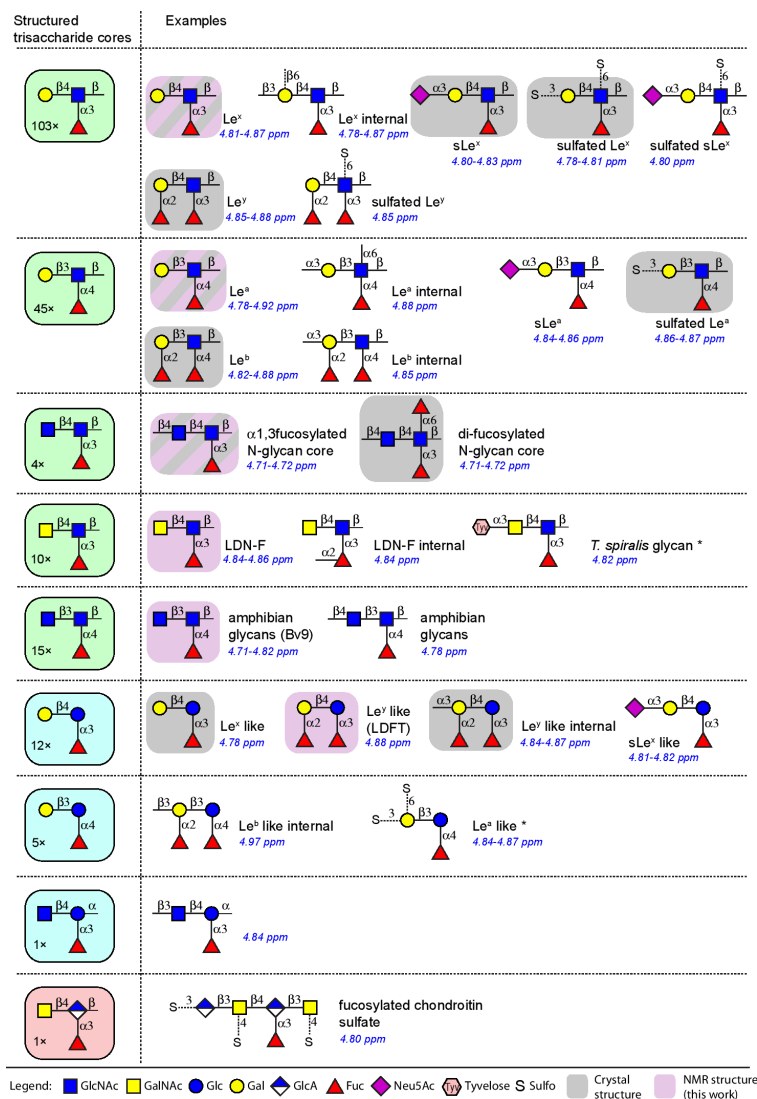


Fig. 2. Characteristic chemical shifts report a C-H...O hydrogen bond. (A) Influence of the stacking GlcNAc on the downfield shift of Fuc H5. Chemical shift differences between 3FChB (red) and Fu α 1,3GlcNAc that lacks the stacking GlcNAc (blue) illustrated by an overlay of their ^{13}C HSQC spectra. Signals of 3FChB are labeled and corresponding signals in the disaccharide are connected by lines to the 3FChB signals (black label). The stacking GlcNAc signals that are missing in the disaccharide are labeled grey. Signals of the methyl group and the spacer at the reducing end are indicated by 'Me' and 'sp'. (B) Dependence of the chemical shift δ in a C-H...O hydrogen bond and the electron density at the bond critical point $\rho(r_c)$ on the H-O distance as calculated with Bader's atom in molecules (AIM) theory using the model system (iPro-O-Me and Me-O-Me) introduced before (26). For comparison, the energetic minimum for the same model system as a function of the H-O distance (26) is shown at the bottom. (C) Schematic representation of α 1,3-fucosylated chitobiose displaying a C-H...O hydrogen bond. The characteristic chemical shift of H5 of Fuc is indicated in blue. (D) Histogram of all available chemical shifts of H5 of Fuc within the GLYCOSCIENCES.de database (37). The H5(Fuc) chemical shifts of α 1,3-fucosylated chitobiose and of α -l-fucose are indicated by a red arrow.

To further analyze the link between the characteristic downfield ^1H chemical shift of the H5 of Fuc and the C-H...O hydrogen bond, we used the Gauge-Independent Atomic Orbital (GIAO) method (34) within Gaussian 09 (35) to calculate the NMR shielding tensor of the DFT optimized structures of 3FChB and Le^x structures, respectively (Table S2). The agreement between the calculated and the experimental chemical shifts is good, especially considering that approximations are used to mimic the solvent. In particular the downfield chemical shifts of H5 of Fuc are excellently reproduced with predictions of 4.88 and 4.86 ppm for 3FChB and Le^x , respectively. The facts that the distances of 2.3-2.4 Å in the used DFT-optimized structures are typical for a C-H...O hydrogen bond (36) and that the chemical shift predictions agree very well with the experimental value, provides a direct support that the downfield chemical shift results from the hydrogen bond. Furthermore we calculated the dependence

of the chemical shift contribution of the hydrogen bond ($\Delta\delta$) on the H...O distance using the model system (iPro-O-Me and Me-O-Me) that we introduced before (26) (Fig. 2C). The shorter the hydrogen bond distance the higher is the downfield chemical shift contribution. This also goes in line with a higher electron density at the bond critical point at shorter distances. Previously the energetic minimum was identified at a distance of 2.3-2.5 Å for that model system using high level ab initio calculations (26). For this distance range the contribution of the C-H...O hydrogen bond to the chemical shift $\Delta\delta$ is calculated to be 0.4-0.8 ppm. This agrees well with the observed downfield shift of $\Delta\delta = \sim 0.45$ ppm observed for 3FChB and Le^x compared with Fu α 1,3GlcNAc. In conclusion, we demonstrated that the characteristic H5 Fuc chemical shift can serve as an indicator for the presence of the C-H...O hydrogen bond between the stacking saccharides of Fuc and either GlcNAc or Gal in the case of 3FChB or Le^x , respectively.



PDF

Fig. 3. Structural motifs within the results of the GLY-COSCIENCES.de database search (37) for characteristic chemical shifts of H5 of Fuc indicating a C-H...O hydrogen bonds. On the left, trisaccharide glycan cores that very likely adopt a defined 3D structure. The number of structures within each category is given as well. Typical examples are shown on the right. Structures indicated with an asterisk were absent in the database search but were added from two publications (42, 67). The associated chemical shifts of H5 of Fuc that are indicative of a C-H...O hydrogen bond are presented in blue. Glycans with a three-dimensional structure in the PDB databank are indicated by a light grey background. Glycans of which the solution structure was determined by NMR spectroscopy were highlighted in light magenta (this work).

Database search for characteristically downfield shifted resonance of H5 of fucose.

We explored whether such C-H...O hydrogen bonds are present in other glycan epitopes and we used the characteristic H5 Fuc chemical shift (Fig. 2 A and C) to perform an extensive search for C-H...O hydrogen bonds in glycans. We searched the NMR database of Glycosciences.de (37), which contains 628 entries for fucose H5 chemical shifts of which roughly 25% cluster around 4.84 ppm (Fig. 2D), the characteristic H5 chemical shifts of Le^x and 3FChB. 180 oligosaccharides displayed a chemical shift > 4.5 ppm that are summarized in Figure 3 (for details, see Table S4). In addition 21 values from publications were included (marked with + in Table S4). In general, only α1,3- and α1,4-fucosylated glycans were found, but no representatives with α1,2- and α1,6-fucosylation. Fucose in α1,3 position is always accompanied with a β1,4-linked saccharide, typically Gal, GlcNAc or GalNAc, whereas α1,4-linked fucose is always accompanied with

a β1,3-linked saccharide, typically Gal or GlcNAc, leading to the following consensus sequences: X-β1,3-[Fucα1,4]-Y and X-β1,4-[Fucα1,3]-Y in which X stands for the stacking saccharide and Y is a saccharide with equatorial hydroxy groups at positions 3 and 4. Nine different branched trisaccharide elements were observed as shown on the left of Fig. 3 (typical examples on the right). All classes display a chemical shift for H5 of Fuc between 4.7 and 5.0 ppm, indicative of a C-H...O hydrogen bond. This suggests that all of those trisaccharide core structures adopt a stabilized 3D structure similar to 3FChB and Le^x. Exceptions are four α1,2-fucosylated glycans that display a chemical shift close to the cut-off shift of 4.5 ppm and seem to be part of another cluster (Fig. 2D) probably reflecting a different kind of stabilization.

The first two categories contain the widely abundant Lewis-type blood group antigens Le^a, Le^b, Le^x and Le^y, as well as substituted versions thereof, including sialyl Le^x (sLe^x) and several sulfated Le^x, Le^y and sLe^x epitopes. Glycans containing

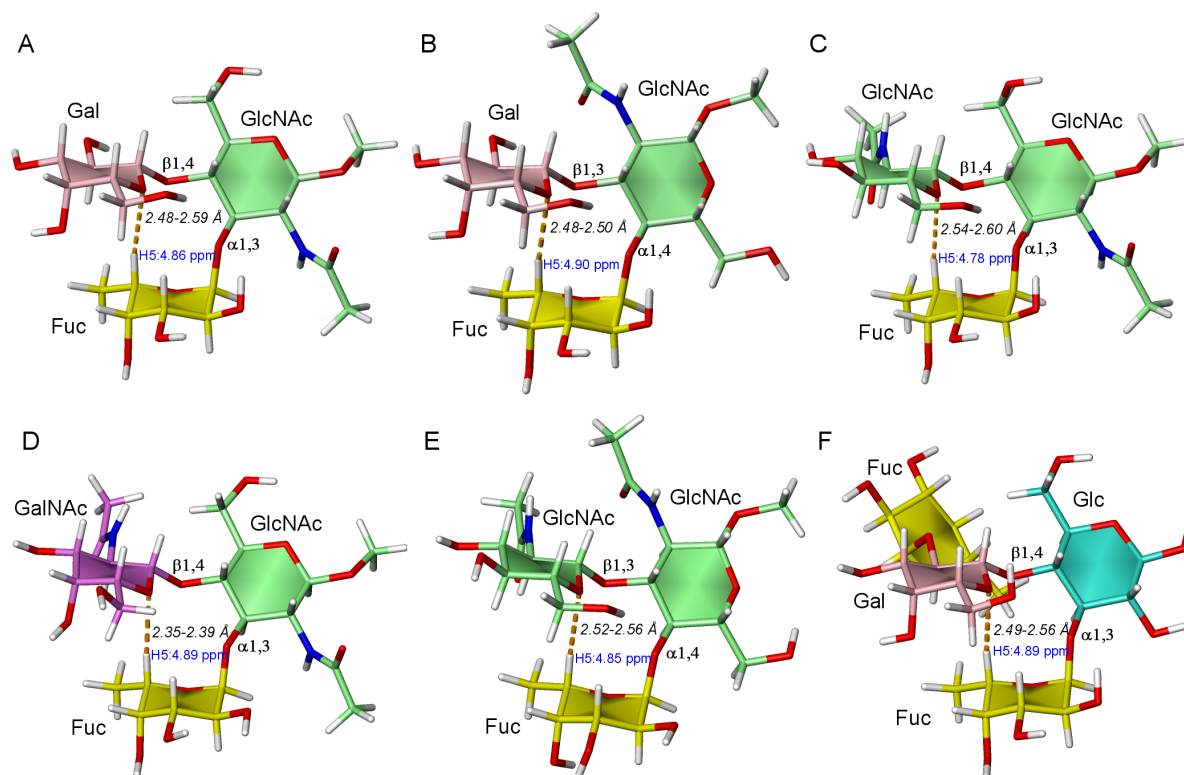


Fig. 4. Comparison of the three-dimensional structures of methyl Le^x , methyl Le^a , 3FChB, LDNF, Bv9 and LDFT. (A) Representative NMR structure of methyl Le^x . The C-H...O hydrogen bond is indicated by an orange dotted line and their C5-O5 distance is given. In addition the chemical shift of H5 of Fuc is indicated in blue. (B-F) Representative NMR structures of methyl Le^a , 3FChB, LDNF, Bv9 and LDFT, respectively.

3FChB are members of the third category that includes α 1,3-fucosylated N-glycan cores typically found in plants and invertebrates (38). Interestingly, several additional motifs are present. α 1,3-Fucosylated LactiNAc (LDNF) present in helminths (39-41) and sea squirt (42), species-specific amphibian egg jelly coats with an α 1,4-fucosylation (43) and fucosylated chondroitin sulfate (44) also display the downfield-shifted H5 resonance of their Fuc moiety. Moreover, Lewis-type-like antigens where the reducing end GlcNAc is substituted by Glc fulfill the search criterion as well. It is noteworthy that α 1,2-linked Fuc, as found for example in determinants for blood groups A, B and H, do not display the characteristic downfield chemical shift of H5 of Fuc, suggesting that those motifs are not stabilized by a comparable intramolecular C-H...O hydrogen bond.

NMR structure determination of carbohydrates with predicted stabilization

In a next step, we wanted to verify various structures of Figure 3 that were predicted to be stabilized by a C-H...O hydrogen bond, and we used our efficient approach by NMR spectroscopy to study the 3D solution structure of the blood group epitope Le^a , fucosylated LactiNAc (LDNF), the amphibian egg glycan Bv9 (43) and lactodifucotetraose (LDFT) each a representative of another trisaccharide category (see the Methods section and Table S1). We obtained well-defined structures that are markedly similar to 3FChB and Le^x and, most importantly, contained a comparable C-H...O hydrogen bond between H5 of Fuc and O5 of either Gal, GalNAc or GlcNAc (see Fig. 4). All structures showed comparable H5...O5 distances in the NMR ensembles and slightly shorter distances in the DFT-optimized structures (Table 1). The

bond critical points indicated the presence of a C-H...O hydrogen bond in each structure (Table S3) and the calculated chemical shifts of H5 of Fuc correspond well to the experimentally observed ones (Table S2) indicating the similarity of those four structures representing four additional categories in Figure 3 with Le^x and 3FChB.

To further analyze the similarity between the six different structures we used dihedral-angle plots of the glycosidic linkages (Fig. 5). The dihedral angles of the glycosidic linkages are basically identical for Le^x , Le^a , 3FChB, LDNF, Bv9 and LDFT indicating that they share the same underlying architecture of which the C-H...O hydrogen bond is an integral part.

Detecting structured, C-H...O hydrogen bond-stabilized carbohydrates in protein crystal structures

The well-defined NMR structures of Le^x , Le^a , 3FChB, LDNF, Bv9 and LDFT prompted us to analyze protein crystal structures containing other representatives of the trisaccharides displayed in Figure 3 either as ligand or within a glycan. Crystal structures for four out of the nine trisaccharide categories are available (marked with a grey background in Fig. 3; Tables 1 and S5). Due to its high abundance in N-glycan cores of plant proteins 3FChB is overrepresented. For comparison we included also a crystal structure of isolated Le^x from the Cambridge Structural Database (CSD). In almost all of the 132 extracted crystal structures from the PDB, the trisaccharide adopts the same conformation as we found for the six solution structures of Le^x , Le^a , 3FChB, LDNF, Bv9 and LDFT. Only in the case of 11 carbohydrate structures from 7 PDB entries the carbohydrates adopted a different extended conformation indicated by a large

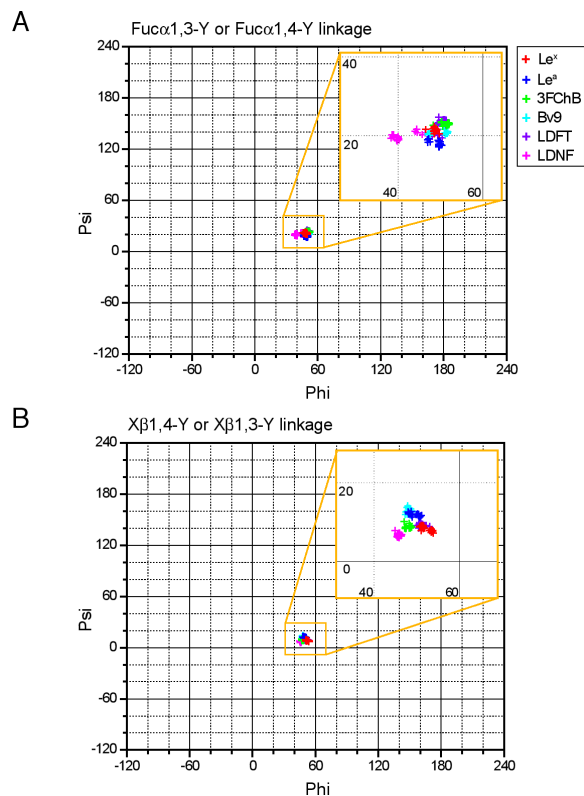


Fig. 5. Comparison of the glycosidic linkages of methyl Le^x, methyl Le^a, 3FChB, LDNF, Bv9 and LDFT. (A) Comparable phi-psi plot of the Fuc-Y linkages. We choose the NMR angle definition (phi: H₁-C₁-O₁-C_x; psi: C₁-O₁-C_x-H_x) rather than the X-ray dihedral angle definition (phi: O₅-C₁-O₁-C_x; psi: C₁-O₁-C_x-C_{x-1}), because the latter would lead to different angles between α 1,3- and α 1,4-fucosylated glycans, differing by $\sim 120^\circ$. Angles were extracted by CARP (71). (B) Comparable phi-psi plot of the X-Y linkages in which X stands for the saccharide that stacks with the Fuc and Y stands for saccharide to which Y and Fuc are connected.

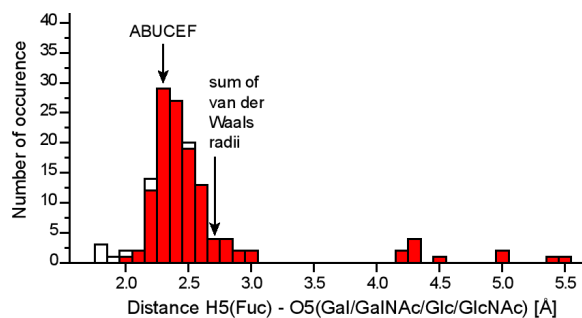


Fig. 6. Distribution of the H-O distances between Fuc H5 and its corresponding hydrogen bonded oxygen in glycan structures from the PDB. The sum of the van der Waals radii is indicated. The white bars represent the distances from the structure 3SEJ in which the carbohydrate coordinates are characterized by unusually high B-factors.

distance between the fucose and the stacking saccharide (Table S5).

In the remaining 123 glycan structures the distance between H5 of Fuc and O5 of Gal/GalNAc/GlcNAc clusters around ~ 2.3 Å

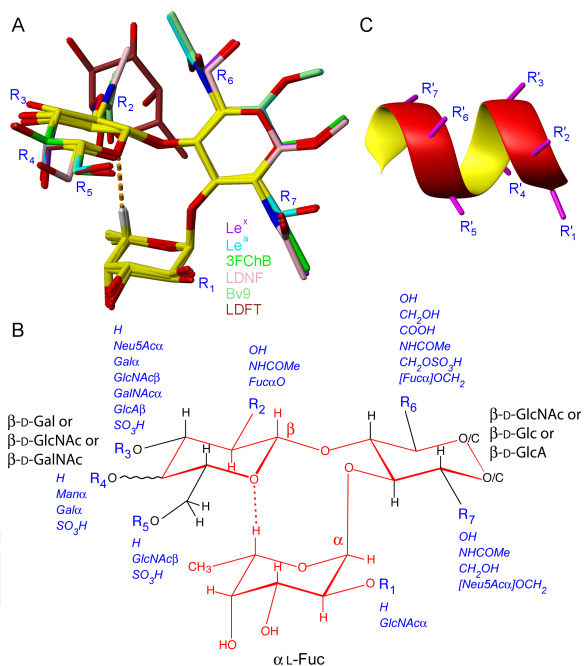


Fig. 7. Secondary structure in glycans. (A) Superposition of a representative NMR structure of Le^x (purple), Le^a (cyan), 3FChB (green), LDNF (pink), Bv9 (pale green) and LDFT (brown). The shared architecture is colored in yellow and the C-H \cdots O hydrogen bond is illustrated by a dotted, orange line. Protons are omitted for clarity except Fuc H5 that is colored in grey. (B) Consensus structure of branched trisaccharides stabilized by a C-H \cdots O hydrogen bond. Short modifications are given in blue. (C) Schematic presentation of a short α -helix. The amino acid site chains are indicated as functional groups R₁ to R₇.

(Fig. 6, Table 1 and S5). Even if some of these clustered distances are slightly biased due to repulsive terms between “non-bonded” atoms in the applied force-fields (26) they are clearly shorter than expected from the sum of the van-der-Waals radii of 2.7 Å (28). There is no difference in the H-O distance depending on the kind of trisaccharide (Table 1). In the 7 PDB entries with the extended conformations the distance is larger than 4 Å and the structures are significantly different. One of those exceptions is a glycan in complex with an inactivated glycoside hydrolase GH98 (45) that normally cleaves Gal β 1,4GlcNAc linkages – exactly the linkage that is distorted in the structure, so we speculate that this distortion is required for the enzymatic function.

In summary, 92% of crystal structures (Fig. 3, indicated by a grey background) confirm that the glyco-epitopes of Le^x, Le^y, Le^a, Le^b, 3FChB, difucosylated N-glycan cores, Le^a-like and a Le^y-like motifs (Glc instead of GlcNAc) share the same architecture and contain a C-H \cdots O hydrogen bond in their trisaccharide motif when bound or linked to protein.

Allowed substitutions

The present work suggests the same core architecture of all the glycan motifs in Figure 3 and Table S4 consisting of an α -L-Fuc that stacks to a Gal/GlcNAc/GalNAc and a GlcNAc/Glc/GlcA that connects both stacking saccharides with either α 1,3 and β 1,4 or α 1,4 and β 1,3 linkages (both in equatorial position). A superposition of the six NMR structures of 3FChB, Le^x, Le^a, LDNF, Bv9 and LDFT illustrates the shared architecture (Fig. 7A). The various substitutions and additional carbohydrate linkages are summarized in Figure 7B. Whereas the L-Fuc moiety is mostly not modified (except at O2 with R₁), modifications of

the Gal/GlcNAc/GalNAc moieties stacked upon Fuc are found on all positions: C2, O3, O4 and O6 (functional groups R₂, R₃, R₄, R₅). The GlcNAc/Glc/GlcA at the branch site can be connected in two orientations (swapping the equatorial O3 and O4 linkages) and can contain various equatorial substituents at C2 and C6 (R₆ and R₇). The α -L-Fuc moiety is the integral part of the conserved secondary structure since α -configuration is a prerequisite for the proper orientation of the C-H...O hydrogen bond. L-saccharides are rare among the common glycans and only α -L-rhamnose exhibiting deviating chirality from α -L-Fuc at C2 and C4 might fulfill the role of the α -L-saccharide. Interestingly, the recently assigned chemical shifts of the unusual O-antigen (-Rha α 1,2Rha α 1,4[GalNAc β 1,4]GlcNAc α 1,3-)_n shows an Rha H5 resonance at 4.61 ppm suggesting a similar stabilization (46).

Discussion

A simple, highly efficient method to obtain three-dimensional structures of small oligosaccharides with a molecular size > 500 Da is introduced using NOESY spectra measured at low temperature and ultra-high magnetic fields (≥ 900 MHz). The simplicity of this approach predestines it for a broader use and the application to other classes of molecules like natural products and small peptides including cyclic peptides.

The determined NMR structures together with existing protein/carbohydrate crystal structures of α 1,3- and α 1,4-fucose branched oligosaccharides (Fig. 3) show that all adopt a defined conformation and share the same spatial architecture. Common to those structures is a characteristic chemical shift of H5 of L-Fuc that is indicative of a C-H...O hydrogen bond between H5 of Fuc and the ring oxygen of a neighboring Gal/GlcNAc/GalNAc moiety. Six NMR solution structures, two crystal structures of an isolated carbohydrate and additionally 121 oligosaccharide structures found in the PDB database support a common architecture of X- β 1,4-[Fuc α 1,3]-Y or X- β 1,3-[Fuc α 1,4]-Y glycoepitopes (Fig. 3) that is stabilized by a C-H...O hydrogen bond. The consensus structure (Fig. 7B) provides a conceptual basis for understanding a large part of glyco-epitopes by their shared architecture that is similar to secondary structure elements in proteins. For example an α -helix as depicted in Fig. 7C the variety of amino acid side chains (R'₁-R'₇) provide a specific surface for interactions. In analogy, the recognition of the glyco-epitopes (Fig. 7B) is mediated by their exposed functional groups (R₁-R₇). Whereas in the displayed heptapeptide 12 backbone angles are mainly stabilized by four hydrogen bonds (~ 5 -6 kcal/mol each), the four glycosidic angles of the trisaccharides (Fig. 7B) are stabilized by one C-H...O hydrogen bond (~ 1.8 kcal/mol) together with other stacking interactions (~ 2.7 kcal/mol) (26), steric hindrance at the glycosidic linkages or sometimes between bulky substitutions, and further stabilization by the exo-anomeric effect. We name that structural element '[3,4]F-branch' standing for a glycan branch point with either α 1,3- or α 1,4-Fuc. Small modifications of the substitutions change the glyco-epitope that can lead to markedly different recognition events, for example a 6'-sulfo modification makes sLe^x invisible for L-selectin (47) but generates a glyco-epitope that can be recognized by Siglec-8 (48). Key to our conclusions were carbohydrate databases that included searchable chemical shift and structural data (37, 49).

Carbohydrates with a pre-defined, biologically active solution structure are perfect ligands since they do not require to change their conformation upon binding and have thus the advantage of a lower entropic penalty and a faster on-rate (k_{ON}) upon binding to their target lectin. An example is 3FChB, which is bound by the lectin CCL2 with a moderately high affinity (K_D = 1 μ M) and adopts the same conformation in solution and in the complex with the lectin (27). The advantage of a predefined conformation is best illustrated by sLe^x analogues that were developed as E-selectin antagonists (14, 21, 50). In the case of an replacement

of the GlcNAc moiety by either cyclohexane-1,2-diol or flexible ethylenglycol, a more than 100-fold higher inhibitory concentration IC₅₀ is observed for the flexible analogue compared to the cyclohexane derivative. Cyclohexane-1,2-diols mimic the GlcNAc and do not allow free rotation, which is also reflected by a H5 chemical shift of Fuc of 4.6–4.77 ppm indicative of the presence of a C-H...O hydrogen bond whereas it is only 4.12 ppm for the flexible version. This supports our hypothesis that the preformed secondary structure, stabilized by a C-H...O hydrogen bond, is a requirement for high affinity recognition.

Among the glycans adopting the presented secondary structure are highly important glyco-epitopes: sLe^x and 6-sulfo sLe^x, natural ligands of selectins that are involved in leucocyte rolling and homing (51, 52), sLe^x in the zona pellucida is crucial for human fertilization (53), 6'-sulfo sLe^x is specifically recognized by Siglec-8 (48), an immunosuppressive co-receptor present on eosinophiles, whereas 6-sulfo sLe^x is a ligand for Siglec-9 (54). In addition, LDNF – a central part of an allergenic sea squirt pentasaccharide (42) and present in glycans of parasitic helminths (39-41), and α 1,3-fucosylated N-glycan cores – major allergic determinants of pollen and insect venoms (38) are found. Surprisingly, among the stabilized glyco-epitopes is also fucosylated chondroitin sulfate found in crab (44) and sea cucumber (55) that was promoted as selectin antagonist (55).

Our work gives also insights into the molecular basis for better understanding cell differentiation, stem cell and cancer research. Among the here-presented structured glyco-epitopes (Fig. 3) are seven clusters of differentiation (CD), namely CD15 (Le^x, also known as stage specific embryonic antigen 1 or SSEA-1), CD15u (3S-Le^x), CD15s (sLe^x), CD15su (6S-Le^x), CD65 (internal Le^x), CD65s (internal 6S-Le^x) and CD174 (Le^x) according to the Glyco-CD database (56). For example, SSEA-1 (Le^x, CD15) was found as an enrichment marker for tumor-initiating cells (TICs) in human brain tumors(57) and the novel antibody clone 5750 that recognizes a certain internal Le^x motif, can select subpopulations of embryonic mouse brain cells displaying multipotency and increased neurosphere forming capacity (58). Being aware of the presence of a 'secondary structure' in all these glyco-epitopes might guide studies that focus on the molecular basis of their functions, in particular their roles in development, cell differentiation and cancer.

In summary, the discovery of a secondary structure element that includes a C-H...O hydrogen bond and thus stabilizes carbohydrate conformations adds a new dimension to the glyco-code and we speculate that more such stabilizing interactions among glyco-epitopes remain to be uncovered.

Materials and Methods

Estimation of the correlation time τ_c and the nuclear Overhauser enhancement. The correlation time was estimated according to the Stokes' law and an estimation of the hydrodynamic radius from the molecular weight according to Cavanagh et al. (59), assuming an average specific volume of the carbohydrate of 0.65 cm³/g and a hydration layer of 1.6 Å. Viscosity values reported for D₂O (60) were used. The nuclear Overhauser enhancement was calculated for a transient NOE experiment in dependence of the magnetic field according to Neuhaus and Williamson (61).

Carbohydrate samples. Le^x methyl glycoside, Le^a methyl glycoside, Fuc α 1,3GlcNAc methyl glycoside and LDFT were purchased from Carbosynth (UK). The synthesis of LDNF and of 3FChB is described in the SI Methods. Bv9 was a generous gift of the laboratory of Yann Guerardel (Lille). Carbohydrates were dissolved in either D₂O or 94% H₂O/6% D₂O with concentrations between 2.8 and 3.7 mM.

NMR Spectroscopy. All spectra were acquired on Bruker Avance III 500 MHz, 600 MHz and 900 MHz spectrometers equipped with cryogenetic triple-resonance probes. Spectra were processed in Topspin 2.1 (Bruker, Germany) and analyzed in

Sparky (T. D. Goddard and D. G. Kneller, SPARKY 3, University of California, San Francisco). Resonance assignment was achieved with 2D ^1H - ^{13}C HSQC, 2D ^1H - ^{13}C HMBC and 2D ^1H - ^1H TOCSY spectra with mixing times ranging from 13 ms to 120 ms. 2D ^1H - ^1H NOESY experiments were recorded with mixing times of 150 ms. All spectra are referenced to 2,2-dimethyl-2-silapentanesulfonic acid (DSS). ^{13}C chemical shifts were indirectly referenced using a scaling factor Ξ of 0.251449530 according to Markley et al. (62) Assigned ^1H - ^{13}C HSQC spectra of all investigated oligosaccharides are shown in the Supporting Information (Fig. S5).

NMR structure calculations. To verify our approach of obtaining NOEs at ultra-high fields together with low temperatures we recorded NOESY spectra at different temperatures. Excellent NOE cross-peaks in 2D NOESY spectra were obtained at low temperature illustrated for 3FChB in Fig. 1B whereas several key signals are missing in a corresponding spectrum at 298 K (Fig. S4A,B). That difference is even more drastic for Le^x (Fig. S4C,D). The conformation is not influenced by such a temperature change as judged from almost no chemical shift changes between natural abundance ^{13}C -HSQC spectra of 3FChB and Le^x recorded at the two different temperatures (Fig. S4E,F). In the case of 3FChB forty-two NOEs have been unambiguously assigned, including 11 inter-residue NOEs (Fig. 1C and Table S6). For the remaining glycans we could extract 29-42 distance restraints, including 9-21 inter-residue restraints (Figs. S1, S3 and Tables S7-S11). The NOE cross-peaks were quantified and converted into proton-proton distances that were used in structure calculations as upper-limit restraints (Tables S6-S11). More precisely, signal to noise (S/N) ratios of the NOE cross peaks were extracted using Sparky (T. D. Goddard and D. G. Kneller, SPARKY 3, University of California, San Francisco). The S/N ratios were correlated to distances assuming a r^{-6} dependence. The S/N ratios of cross peaks originating from degenerate CH_2 or CH_3 groups were divided by 2 or 3, respectively. Distances were calibrated with H3-H5 correlations, normally nicely isolated in the NOESY spectra, that correspond to a distance of 2.65 Å in a typical chair conformation. Initial calibration attempts with the fixed H61-H62 distance of 1.77 Å gave similar results, but due to the proximity of those signals to the diagonal, their S/N ratios could not be extracted reliably for all carbohydrates.

Initial coordinates were obtained by the Biomolecule Builder from the GLYCAM website (<http://glycam.ccruc.uga.edu/AMBER/index.html>) that were used to generate carbohydrate parameters in the CYANA library file. Preliminary structures were calculated by CYANA (63) starting from 300 structures with randomized conformation using upper limit constraints calculated from the S/N ratios and expanded by a 0.2 Å tolerance. The 30 best structures (lowest target function) were further refined with Amber 9.0 (64) using the Glycam06 force field (29) together with an implicit solvent model. 20 structures with the lowest distance violations were used for the final structural ensemble.

Geometry optimization and chemical shift calculation. The geometry of one representative of each NMR structure was optimized using density functional theory (DFT) (31) at the B3LYP/6-31G(d,p) level of theory (32). Solvent effects were accounted for applying the CPCM model (implicit solvent) (65). NMR shielding tensors were obtained using the gauge-independent atomic orbital (GIAO) method (34). Absolute chemical shift values were referenced to tetramethylsilane (fully optimized in solvent using the same level of theory). Vibrational mode analysis was performed at the optimized geometry to confirm the stability of the obtained minimum. No imaginary frequencies were found. All

ab initio geometry optimizations and chemical shift calculations were performed using Gaussian 09 (35).

Bond critical point of C-H...O hydrogen bonds. The molecular wave functions obtained at the B3LYP/6-31G(d,p) level were exported to the program AIMAll (66), which was used to localize and characterize the bond critical points on the basis of the quantum theory of atoms in molecules (AIM) (33).

Databank search for specific glycan chemical shifts. The chemical shift search of GLYCOSCIENCES.de (37) was used to search for fucose H5 chemical shifts between 4.5 and 5.5 ppm. The glycan structures and chemical shifts of the 180 results were extracted and transferred into a file. For glycans containing multiple fucose residues the chemical shifts were unambiguously assigned to each of them. The list contained one line per chemical shift with the associated glycomotif that was highlighted in case of a larger glycan. Finally the entries were ordered and categorized after their glycomotif. Data from five publications (43, 44, 67-69) were added to support categories with only few data. The full list of results is found in Table S4.

To estimate the percentage of glycans containing a C-H...O hydrogen bond, all entries containing a H5(Fuc) chemical shift were searched, revealing 628 entries. The 180 entries that displayed a H5(Fuc) chemical shift of 4.7–5.0 ppm present then ~25%. A similar search in structural data is described below.

Databank search for glycan coordinates. A substructure search within GLYCOSCIENCES.de (37) was used to search protein structures for the individual glycomotifs. A broader search for X- β 1,4-[Fuc α 1,3]-Y and X- β 1,3-[Fuc α 1,4]-Y did not reveal any additional results. However, a search using the GlycomeDB (49) revealed additional PDB entries. The obtained protein crystal structures containing either glycans or carbohydrate ligands were used to extract the distances between Fuc C5 and O5 of the stacking monosaccharide using XtalView (70). To all structures protons were added with Maestro (Schrödinger) and the distances measured between Fuc H5 and O5 of the stacking monosaccharide in order to detect C-H...O hydrogen bonds. The results are summarized in Table S5. The percentage of fucose-containing glycans that are stabilized by a C-H...O hydrogen bond was estimated from 4271 entries containing an α -L-fucose, 232 entries containing X- β 1,3-[Fuc α 1,4]-Y and 910 entries containing X- β 1,4-[Fuc α 1,3]-Y as found in GLYCOSCIENCES.de.

Accession codes. The atomic coordinates, chemical shifts, and restraints used for the structure calculations were deposited with the help of Protein Data Bank Japan (PDBj) and PDBj-BMRB in the Biological Magnetic Resonance Bank (BMRB) with the accession codes 21031, 21032, 21034, 21053 and 21054. The LDFT structure and restraints were deposited at the Protein Data Bank (PDB) with accession number 2MK1 (PDB accepts oligosaccharides with four or more residues) and the chemical shifts at the BMRB with accession number 19748.

Acknowledgement.

We are grateful for the generous gift of the Bv9 oligosaccharide from GlycoBase, Emmanuel Maes and Yann Guerardel (glycobase.univ-lille1.fr). We thank Gerhard Wider (ETH Zürich) for his suggestion to measure at low temperature, Markus Blatter (ETH Zürich) for his help with Amber, Oliver Schwardt for measuring high resolution mass spectra, Thomas Lütteke (University of Giessen) and Martin Frank (Biognos AB) for fruitful discussions and keeping the GLYCOSCIENCES.DE database running, Robert J. Woods (University of Georgia and NUI Galway) for providing the Glycam website, Markus Künzler for his comments on the manuscript, and Daniel Roe (University of Utah) and Garib Murshudov (MRC in Cambridge) for discussing the handling of C-H...O hydrogen bond in force fields. We like to acknowledge Markus Aebi, Robert Konrat, Yaakov Levi, Andreas Trabesinger and Vikram Pansé for helpful discussions and suggestions, María del Carmen Fernández-Alonso and Jesús Jiménez-Barbero for suggesting AIM, and Hans-Heinrich Limbach and José Elguero for sharing their enthusiasm for hydrogen bonds. This work was supported by the Swiss National Science Foundation (SNF sinergia grant CRSI3.127333).

1. Marth JD & Grewal PK (2008) Mammalian glycosylation in immunity. *Nature reviews*

- 8(11):874-887.
2. Taylor ME & Drickamer K (2007) Paradigms for glycan-binding receptors in cell adhesion. *Curr. Opin. Cell. Biol.* 19(5):572-577.
 3. Zhao YY, et al. (2008) Functional roles of N-glycans in cell signaling and cell adhesion in cancer. *Cancer science* 99(7):1304-1310.
 4. Staudacher E, Altmann F, Wilson IB, & Marz L (1999) Fucose in N-glycans: from plant to man. *Biochim Biophys Acta* 1473(1):216-236.
 5. Becker DJ & Lowe JB (2003) Fucose: biosynthesis and biological function in mammals. *Glycobiology* 13(7):41R-53R.
 6. Ma B, Simala-Grant JL, & Taylor DE (2006) Fucosylation in prokaryotes and eukaryotes. *Glycobiology* 16(12):158R-184R.
 7. Combs MR (2009) Lewis blood group system review. *Immunohematology* 25(3):112-118.
 8. Stanley P & Cummings RD (2009) Structures Common to Different Glycans. *Essentials of Glycobiology*, eds Varki A, Cummings RD, Esko JD, Freeze HH, Stanley P, Bertozzi CR, Hart GW, & Etzler ME (Cold Spring Harbor Laboratory Press, Cold Spring Harbor (NY)), 2nd Ed, pp 175-198.
 9. Heimburg-Molinero J, et al. (2011) Cancer vaccines and carbohydrate epitopes. *Vaccine* 29(48):8802-8826.
 10. DeMarco ML & Woods RJ (2008) Structural glycobiology: a game of snakes and ladders. *Glycobiology* 18(6):426-440.
 11. Slynko V, et al. (2009) NMR structure determination of a segmentally labeled glycoprotein using in vitro glycosylation. *J Am Chem Soc* 131(3):1274-1281.
 12. Yamamoto S, Yamaguchi T, Erdelyi M, Griesinger C, & Kato K (2011) Paramagnetic lanthanide tagging for NMR conformational analyses of N-linked oligosaccharides. *Chemistry* 17(34):9280-9282.
 13. Yu F, Wolff JJ, Amster IJ, & Prestegard JH (2007) Conformational preferences of chondroitin sulfate oligomers using partially oriented NMR spectroscopy of 13C-labeled acetyl groups. *J Am Chem Soc* 129(43):13288-13297.
 14. Kolb HC & Ernst B (1997) Development of tools for the design of selectin antagonists. *Chem-Eur J* 3(10):1571-1578.
 15. Imberty A & Perez S (2000) Structure, conformation, and dynamics of bioactive oligosaccharides: theoretical approaches and experimental validations. *Chem Rev* 100(12):4567-4588.
 16. Yuriev E, Farrugia W, Scott AM, & Ramslund PA (2005) Three-dimensional structures of carbohydrate determinants of Lewis system antigens: implications for effective antibody targeting of cancer. *Immunol Cell Biol* 83(6):709-717.
 17. Lemieux RU, Bock K, Delbaere LTJ, Koto S, & Rao VS (1980) The Conformations of Oligosaccharides Related to the Abh and Lewis Human-Blood Group Determinants. *Can J Chemistry* 58(6):631-653.
 18. Xu Q, Gitti R, & Bush CA (1996) Comparison of NMR and molecular modeling results for a rigid and a flexible oligosaccharide. *Glycobiology* 6(3):281-288.
 19. Meyer B (1990) Conformational Aspects of Oligosaccharides. *Top Curr Chem* 154:141-208.
 20. Miller KE, Mukhopadhyay C, Cagas P, & Bush CA (1992) Solution Structure of the Lewis X Oligosaccharide Determined by Nmr-Spectroscopy and Molecular-Dynamics Simulations. *Biochemistry* 31(29):6703-6709.
 21. Schwizer D, et al. (2012) Pre-organization of the core structure of E-selectin antagonists. *Chemistry* 18(5):1342-1351.
 22. Perez S, et al. (1996) Crystal and molecular structure of a histo-blood group antigen involved in cell adhesion: the Lewis x trisaccharide. *Glycobiology* 6(5):537-542.
 23. Azurmendi HF, Martin-Pastor M, & Bush CA (2002) Conformational studies of Lewis X and Lewis A trisaccharides using NMR residual dipolar couplings. *Biopolymers* 63(2):89-98.
 24. Ichikawa Y, et al. (1992) Chemical-Enzymatic Synthesis and Conformational-Analysis of Sialyl Lewis-X and Derivatives. *J Am Chem Soc* 114(24):9283-9298.
 25. Lemieux RU, Pavia AA, Martin JC, & Watanabe KA (1969) Solvation Effects on Conformational Equilibria. Studies Related to Conformational Properties of 2-M-ethoxytetrahydropyran and Related Methyl Glycopyranosides. *Can J Chemistry* 47(23):4427-4437.
 26. Zierke M, et al. (2013) Stabilization of branched oligosaccharides: Lewis benefits from a non-conventional C-H...O hydrogen bond. *J. Am. Chem. Soc.* 135:13464-13472.
 27. Schubert M, et al. (2012) Plasticity of the beta-Trefoil Protein Fold in the Recognition and Control of Invertebrate Predators and Parasites by a Fungal Defence System. *PLoS Pathog* 8(5):e1002706.
 28. Bondi A (1964) Van Der Waals Volumes and Radii. *J Phys Chem* 68(3):441-451.
 29. Kirschner KN, et al. (2008) GLYCAM06: a generalizable biomolecular force field. *Carbohydrates. J Comput Chem* 29(4):622-655.
 30. Arunan E, et al. (2011) Definition of the hydrogen bond (IUPAC Recommendations 2011). *Pure Appl Chem* 83(8):1637-1641.
 31. Hohenberg P & Kohn W (1964) Inhomogeneous Electron Gas. *Phys Rev B* 136(3B):B864-B871.
 32. Becke AD (1993) Density-Functional Thermochemistry. 3. The Role of Exact Exchange. *J Chem Phys* 98(7):5648-5652.
 33. Bader RFW (1991) A Quantum-Theory of Molecular-Structure and Its Applications. *Chem Rev* 91(5):893-928.
 34. Wolinski K, Hinton JF, & Pulay P (1990) Efficient Implementation of the Gauge-Independent Atomic Orbital Method for Nmr Chemical-Shift Calculations. *J Am Chem Soc* 112(23):8251-8260.
 35. Anonymous (2009) Gaussian 09, Revision A.2; (Gaussian, Inc., Wallingford, CT).
 36. Castellano RK (2004) Progress toward understanding the nature and function of C-H center dot center dot center dot O interactions. *Curr Org Chem* 8(10):845-865.
 37. Lutteke T, et al. (2006) GLYCOSCIENCES.de: an Internet portal to support glycomics and glycobiology research. *Glycobiology* 16(5):71R-81R.
 38. Hoffmann-Sommergruber K, Paschinger K, & Wilson IB (2011) Glycomarkers in parasitic infections and allergy. *Biochem Soc Trans* 39(1):360-364.
 39. Wuhrer M, et al. (2006) Gender-specific expression of complex-type N-glycans in schistosomes. *Glycobiology* 16(10):991-1006.
 40. van Die I, et al. (1999) Core alpha1-->3-fucose is a common modification of N-glycans in parasitic helminths and constitutes an important epitope for IgE from *Haemonchus contortus* infected sheep. *FEBS letters* 463(1-2):189-193.
 41. Aranzamendi C, et al. (2011) Glycan microarray profiling of parasite infection sera identifies the LDNF glycan as a potential antigen for serodiagnosis of trichinellosis. *Exp Parasitol* 129(3):221-226.
 42. Kato Y, et al. (2001) Determination of solution conformation of allergenically active pentasaccharitol obtained from sea squirt antigen. *Magn Reson Chem* 39(5):259-266.
 43. Coppin A, Florea D, Maes E, Cogalniceanu D, & Strecker G (2003) Comparative study of carbohydrate chains released from the oviductal mucins of the two very closely related amphibian species *Bombina orientalis* and *Bombina variegata*. *Biochimie* 85(1-2):53-64.
 44. Kitagawa H, et al. (1997) A novel pentasaccharide sequence GlcA(3-sulfate)(beta1-3)GalNAc(4-sulfate)(beta1-4)(Fuc alpha1-3)GlcA(beta1-3)GalNAc(4-sulfate) in the oligosaccharides isolated from king crab cartilage chondroitin sulfate K and its differential susceptibility to chondroitinases and hyaluronidase. *Biochemistry* 36(13):3998-4008.
 45. Higgins MA, et al. (2009) Differential recognition and hydrolysis of host carbohydrate antigens by *Streptococcus pneumoniae* family 98 glycoside hydrolases. *J Biol Chem* 284(38):26161-26173.
 46. Sawen E, et al. (2012) Structural studies of the O-antigenic polysaccharide from *Plesiomonas shigelloides* strain AM36565. *Carbohydr Res* 348:99-103.
 47. Lowe JB (2002) Glycosylation in the control of selectin counter-receptor structure and function. *Immunol Rev* 186:19-36.
 48. Bochner BS, et al. (2005) Glycan array screening reveals a candidate ligand for Siglec-8. *J Biol Chem* 280(6):4307-4312.
 49. Ranzinger R, Herget S, von der Lieth CW, & Frank M (2011) GlycoDB--a unified database for carbohydrate structures. *Nucleic Acids Res* 39:D373-376.
 50. Thoma G, Magnani JL, Patton JT, Ernst B, & Jahnke W (2001) Preorganization of the bioactive conformation of sialyl Lewis(X) analogues correlates with their affinity to E-selectin. *Angew Chem Int Ed* 40(10):1941-1945.
 51. Sperandio M (2006) Selectins and glycosyltransferases in leukocyte rolling in vivo. *Febs J* 273(19):4377-4389.
 52. McEver RP (2002) Selectins: lectins that initiate cell adhesion under flow. *Curr Opin Cell Biol* 14(5):581-586.
 53. Pang PC, et al. (2011) Human sperm binding is mediated by the sialyl-Lewis(x) oligosaccharide on the zona pellucida. *Science* 333(6050):1761-1764.
 54. Crocker PR, Paulson JC, & Varki A (2007) Siglecs and their roles in the immune system. *Nat Rev Immunol* 7(4):255-266.
 55. Borsig L, et al. (2007) Selectin blocking activity of a fucosylated chondroitin sulfate glycosaminoglycan from sea cucumber. Effect on tumor metastasis and neutrophil recruitment. *J Biol Chem* 282(20):14984-14991.
 56. Kumar S, Lutteke T, & Schwartz-Albiez R (2012) GlycoCD: a repository for carbohydrate-related CD antigens. *Bioinformatics* 28(19):2553-2555.
 57. Son MJ, Woolard K, Nam DH, Lee J, & Fine HA (2009) SSEA-1 is an enrichment marker for tumor-initiating cells in human glioblastoma. *Cell Stem Cell* 4(5):440-452.
 58. Hennen E, Czopka T, & Faisner A (2011) Structurally distinct LewisX glycans distinguish subpopulations of neural stem/progenitor cells. *J Biol Chem* 286(18):16321-16331.
 59. Cavanagh J, Fairbrother WJ, Palmer III AG, Rance M, & Skelton NJ (2006) *Protein NMR Spectroscopy: Principles and Practice* (Academic Press Inc., San Diego) 2nd Ed.
 60. Cho CH, Urquidí J, Singh S, & Robinson GW (1999) Thermal offset viscosities of liquid H₂O, D₂O, and T₂O. *J Phys Chem B* 103(11):1991-1994.
 61. Neuhaus D & Williamson MP (2000) *The Nuclear Overhauser Effect in Structural and Conformational Analysis* (Wiley-VCH, New York) 2nd Ed.
 62. Markley JL, et al. (1998) Recommendations for the presentation of NMR structures of proteins and nucleic acids. IUPAC-IUBMB-IUPAB Inter-Union Task Group on the Standardization of Data Bases of Protein and Nucleic Acid Structures Determined by NMR Spectroscopy. *J Biomol NMR* 12(1):1-23.
 63. Guntert P (2004) Automated NMR structure calculation with CYANA. *Methods Mol Biol* 278:353-378.
 64. Case DA, et al. (2005) The Amber biomolecular simulation programs. *J. Comput. Chem.* 26(16):1668-1688.
 65. Barone V & Cossi M (1998) Quantum calculation of molecular energies and energy gradients in solution by a conductor solvent model. *J Phys Chem A* 102(11):1995-2001.
 66. Keith TA (2013; aim.tkgristmill.com) AIMAll (Version 13.05.06) (Todd A. Keith, TK Christmill Software, Overland Park, KS).
 67. Kurutz JW & Kiessling LL (1997) Solution conformation of Lewis a--derived selectin ligands is unaffected by anionic substituents at the 3'- and 6'-positions. *Glycobiology* 7(3):337-347.
 68. Zhang J, Otter A, & Bundle DR (1996) Synthesis and conformational studies of the tyvelose capped, Lewis-x like tetrasaccharide epitope of *Trichinella spiralis*. *Bioorg Med Chem* 4(11):1989-2001.
 69. Ishizuka Y, Nemoto T, Fujiwara M, Fujita K, & Nakanishi H (1999) Three-dimensional structure of fucosylactoses in an aqueous solution. *J Carbohydr Chem* 18(5):523-533.
 70. McRee DE (1999) XtalView/Xfit--A versatile program for manipulating atomic coordinates and electron density. *J Struct Biol* 125(2-3):156-165.
 71. Lutteke T, Frank M, & von der Lieth CW (2005) Carbohydrate Structure Suite (CSS): analysis of carbohydrate 3D structures derived from the PDB. *Nucleic Acids Res* 33:D242-D246.

3.1.2.1 Supporting Information

Supporting Information

A secondary structure element in a wide range of fucosylated glyco-
epitopes

*Thomas Aeschbacher, Mirko Zierke, Martin Smieško, Mayeul Collot, Jean-Maurice Mallet,
Beat Ernst, Frédéric H.-T. Allain, Mario Schubert*

Supporting Figures S1-S5	S2
Supporting Tables S1-S11	S9
Supporting Methods	S30
Supporting References	S32

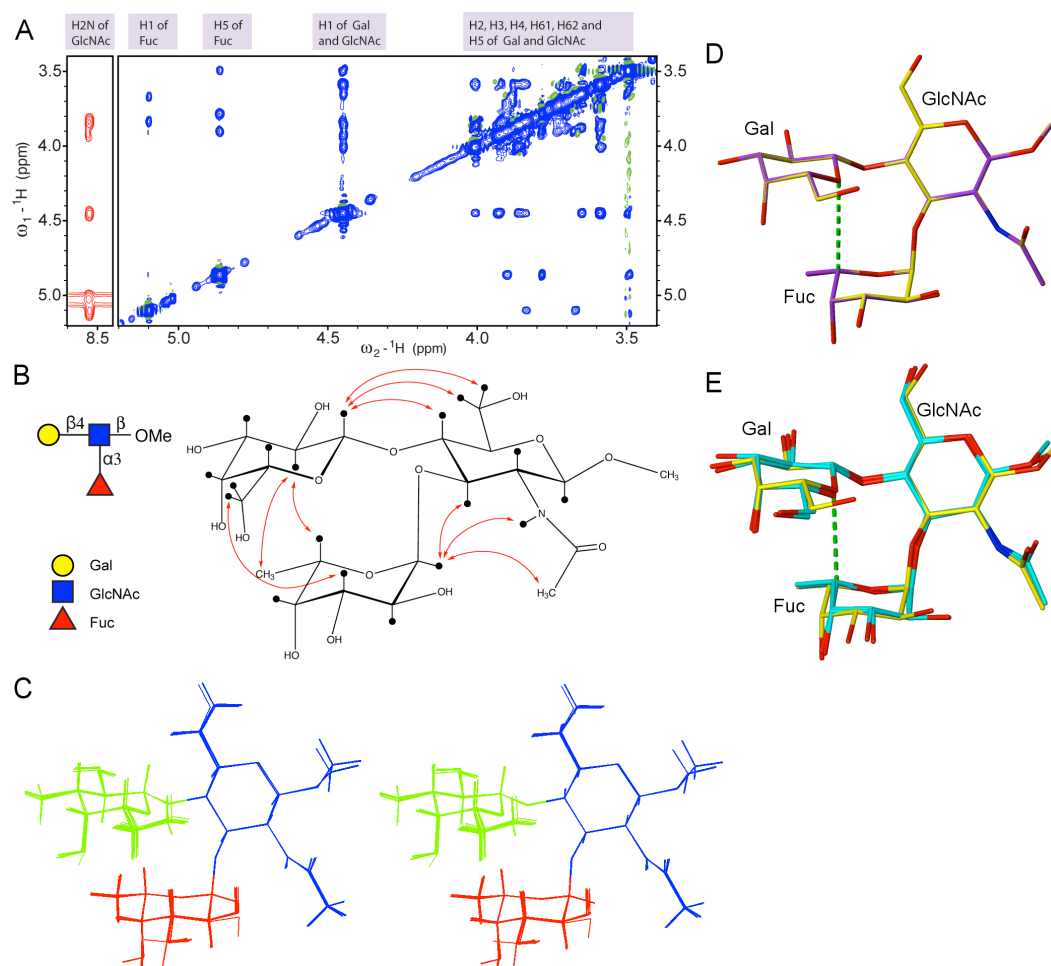


Figure S1. Three-dimensional structure determination of methyl Le^x. (A) 900 MHz 2D NOESY spectra of methyl Le^x (3.7 mM) in either H₂O (red, left) or in D₂O (blue, right) measured at 277 K. Chemical shift assignments are indicated on the top for isolated resonances. (B) Schematic presentation of methyl Le^x with the observed inter-residue NOEs indicated by red arrows. (C) The ensembles of 20 best structures based on NOESY data recorded at 277 K. (D) Most representative NMR structure (yellow) superimposed with a representative obtained from a Le^x-protein conjugate recorded at 293 K (purple)(1). For simplicity only heavy atoms are shown. The C-H...O hydrogen bond is shown as a dashed green line. (E) Comparison with the crystal structure ABUCEF of methyl Le^x that contained two molecules in the asymmetric unit (cyan) with the most representative NMR structure obtained from 277 K data (yellow).

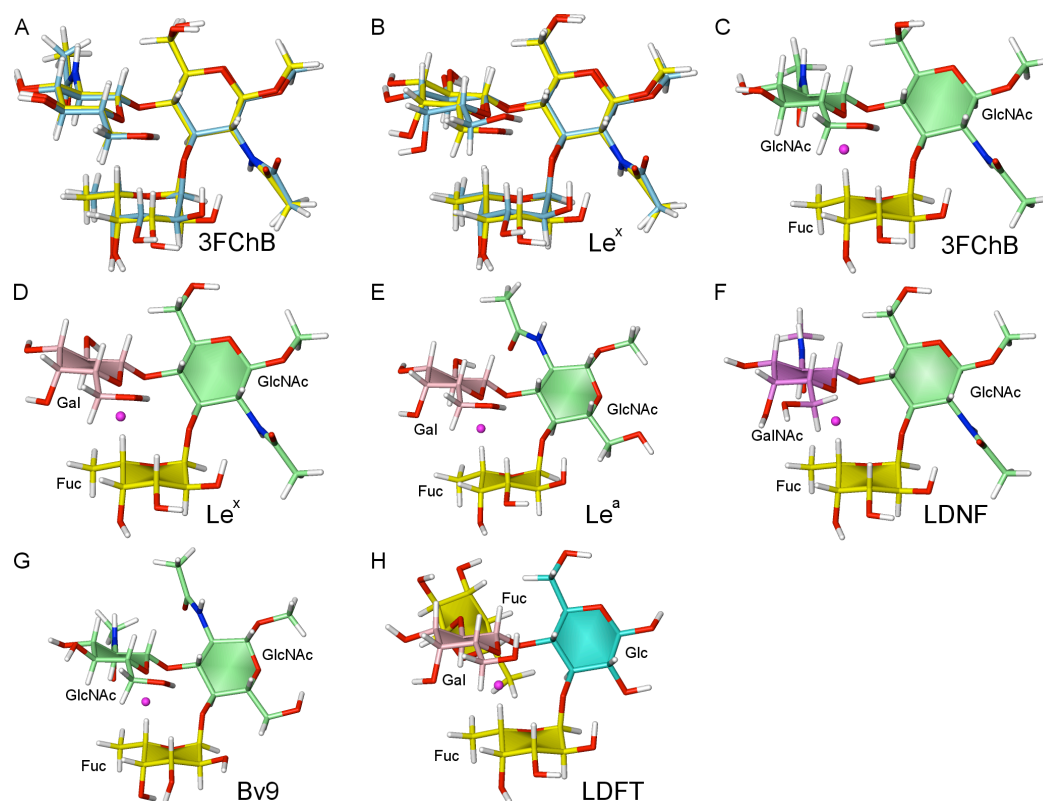


Figure S2. Structures of 3FChB, Le^x, Le^a, LDNF, Bv9 and LDFT optimized with density functional theory (DFT) and bond critical points of the C-H...O hydrogen bonds. (A) Overlay of the DFT-optimized structure (yellow) with the representative NMR structure (sky blue) of 3FChB. The latter was used as start structure for the DFT optimization. (B) Comparison of the DFT-optimized structure (yellow) with the representative NMR structure (sky blue) of Le^x. (C) Structure of 3FChB optimized with DFT with the bond critical point of the C-H...O hydrogen bond indicated as magenta sphere. (D) DFT-optimized structure of Le^x with the bond critical point. (E) DFT-optimized structure of Le^a with the bond critical point. (F) DFT-optimized structure of LDNF with the bond critical point. (G) DFT-optimized structure of Bv9 with the bond critical point. (H) DFT-optimized structure of LDFT with the bond critical point.

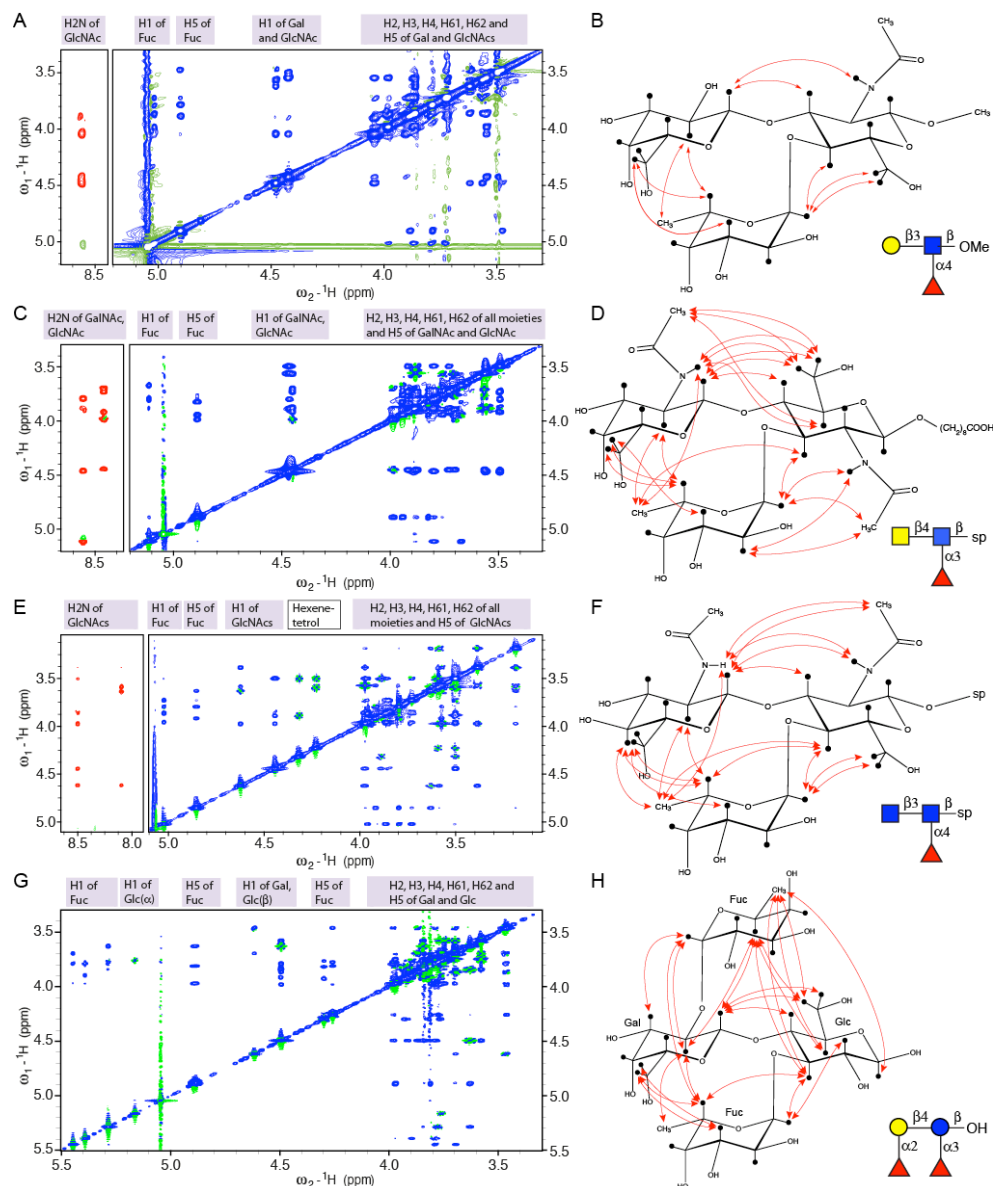


Figure S3. Structure determination of methyl Le^a, LDNF, Bv9 and LDFT. (A) 2D NOESY spectra of methyl Le^a (3.7 mM) in either H₂O (red, left) or in D₂O (blue, right) recorded at 900 MHz and 275 K. On the top, chemical shift assignments of isolated resonances are indicated. (B) Schematic presentation of methyl Le^a displaying the observed inter-residue NOEs by red arrows. (C) 2D NOESY spectra of LDNF (2.2 mM) in either H₂O (red, left) or in D₂O (blue, right) recorded at 900 MHz and 275 K. (D) Schematic presentation of LDNF displaying the observed inter-residue NOEs by red arrows. (E) 2D NOESY spectra of Bv9 (1 mM) in either H₂O (red, left) or in D₂O (blue, right) recorded at 900 MHz and 273 K. (F) Schematic presentation of Bv9 displaying the observed inter-residue NOEs by red arrows. The hexene-tetrol at the reducing end is indicated by 'sp'. (G) 2D NOESY spectrum of LDFT (3.2 mM) in D₂O recorded at 900 MHz and 275 K. (H) Schematic presentation of LDFT(β) displaying the observed inter-residue NOEs by red arrows. Only NOEs of the major β anomer are shown.

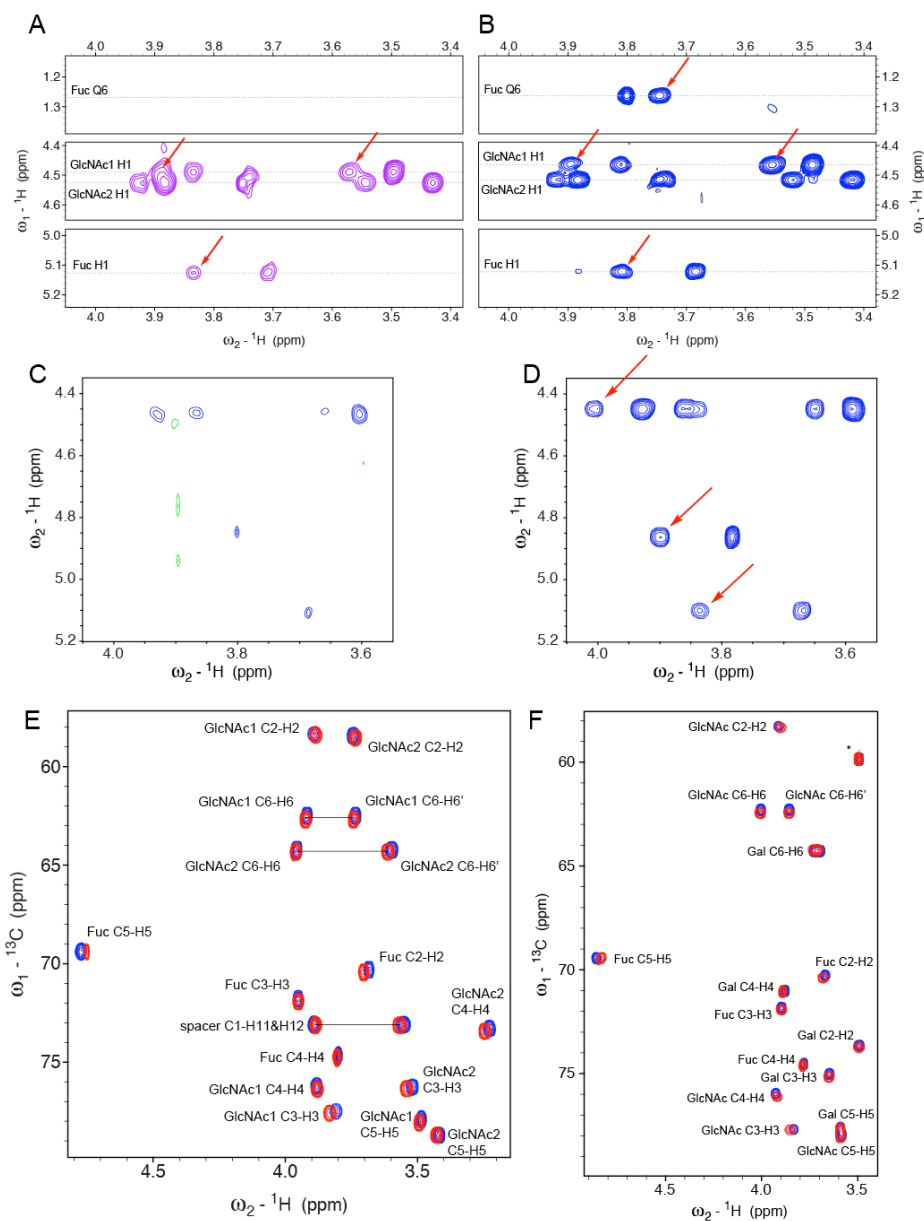


Figure S4. Influence of temperature on 2D NOESY spectra and chemical shifts of 3FChB and Le^x methyl glycoside. (A) Region of a 2D NOESY spectrum of 3FChB (2.8 mM) recorded at 298 K and 900 MHz. Intermolecular NOE cross-peaks are indicated by a red arrow. (B) Corresponding region of a 2D NOESY spectrum recorded at 277 K and 900 MHz. (C) Region of a NOESY spectrum of Le^x methyl glycoside (3.7 mM) recorded at 293 K and 900 MHz. (D) Corresponding region of 2D NOESY spectrum recorded at 277 K and 900 MHz. All spectra are referenced to DSS. (E) Overlay of two ¹H-¹³C HSQC spectra of 3FChB (2.8 mM) recorded at 900 MHz and either 277 K (blue) or 298 K (red). The spectra are referenced to DSS. (F) ¹H-¹³C HSQC spectra of Le^x methyl glycoside (3.7 mM) recorded at 900 MHz and either 277 K (blue) or 298 K (red). All spectra are referenced to DSS.

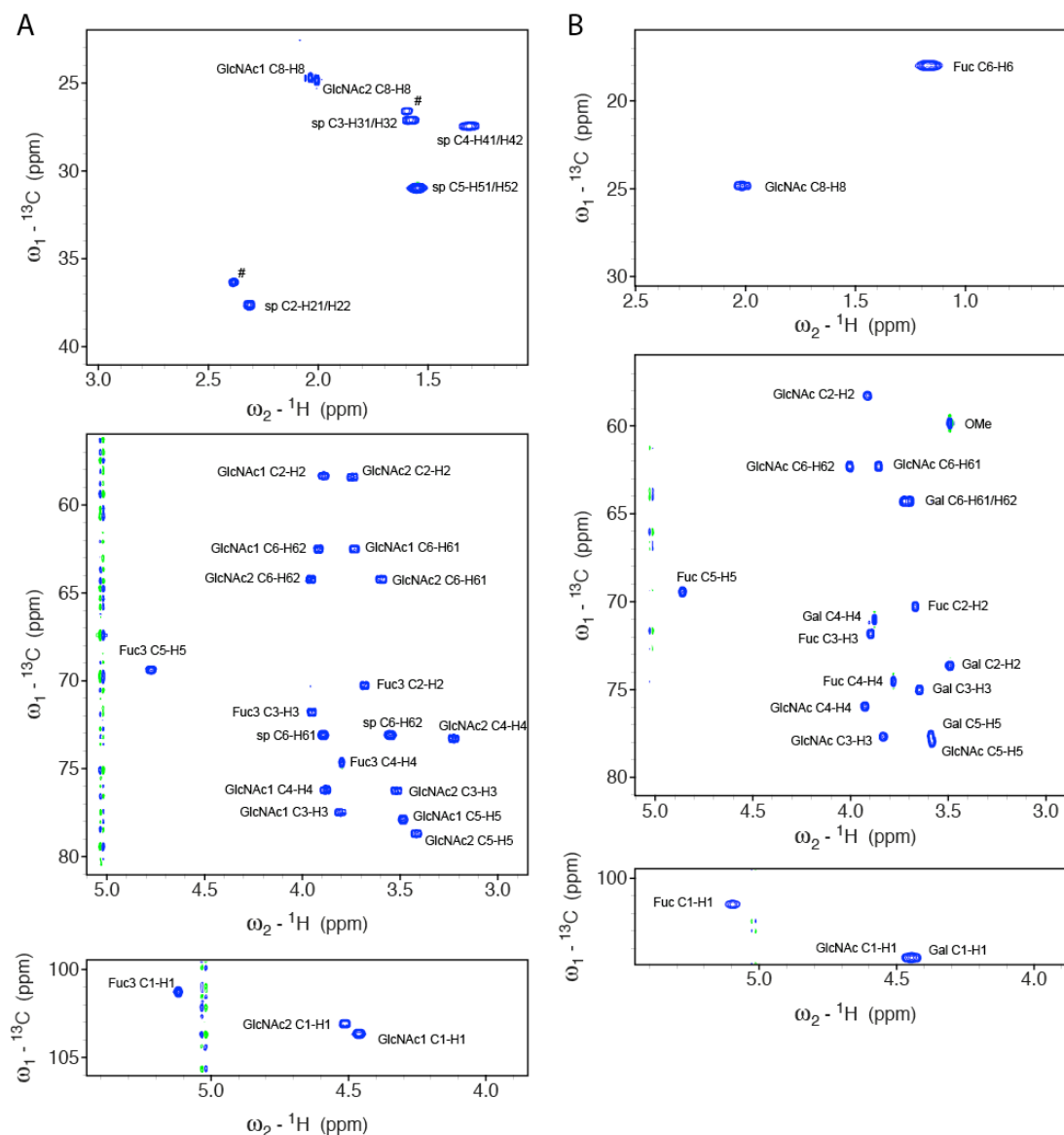


Figure S5. ^1H - ^{13}C HSQC spectra of all investigated oligosaccharides all referenced to DSS. (A) Three regions of the ^1H - ^{13}C HSQC spectrum of $\alpha,1,3$ -fucosylated chitobiose (2.8 mM) measured at 600 MHz and 277 K. The proximal N-acetylglucosamine that harbours the $\alpha,1,3$ branch is indicated as "GlcNAc1", the distal N-acetylglucosamine that stacks to the fucose is labeled "GlcNAc2". The signals of the linker 6-hydroxy hexanoate ($-\text{O}-(\text{CH}_2)_5\text{COOH}$) remaining from the chemical synthesis are labeled with "sp". The two correlations indicated with "#" originate from an impurity. (B) ^1H - ^{13}C HSQC spectrum of Le^x methylglycoside (3.7 mM) measured at 900 MHz and 277 K. (C) ^1H - ^{13}C HSQC spectrum of Le^a methylglycoside (3.7 mM) spectrum measured at 275 K. The two regions on the top are recorded at 900 MHz. The anomeric region (bottom) of a spectrum recorded at 500 MHz is shown at the bottom (better performance of ^{13}C decoupling compared to 900 MHz spectrum). (D) ^1H - ^{13}C HSQC spectrum of LDNF with a $\beta-(\text{CH}_2)_8\text{COOH}$ linker (2.2 mM) measured at 900 MHz and 275 K.

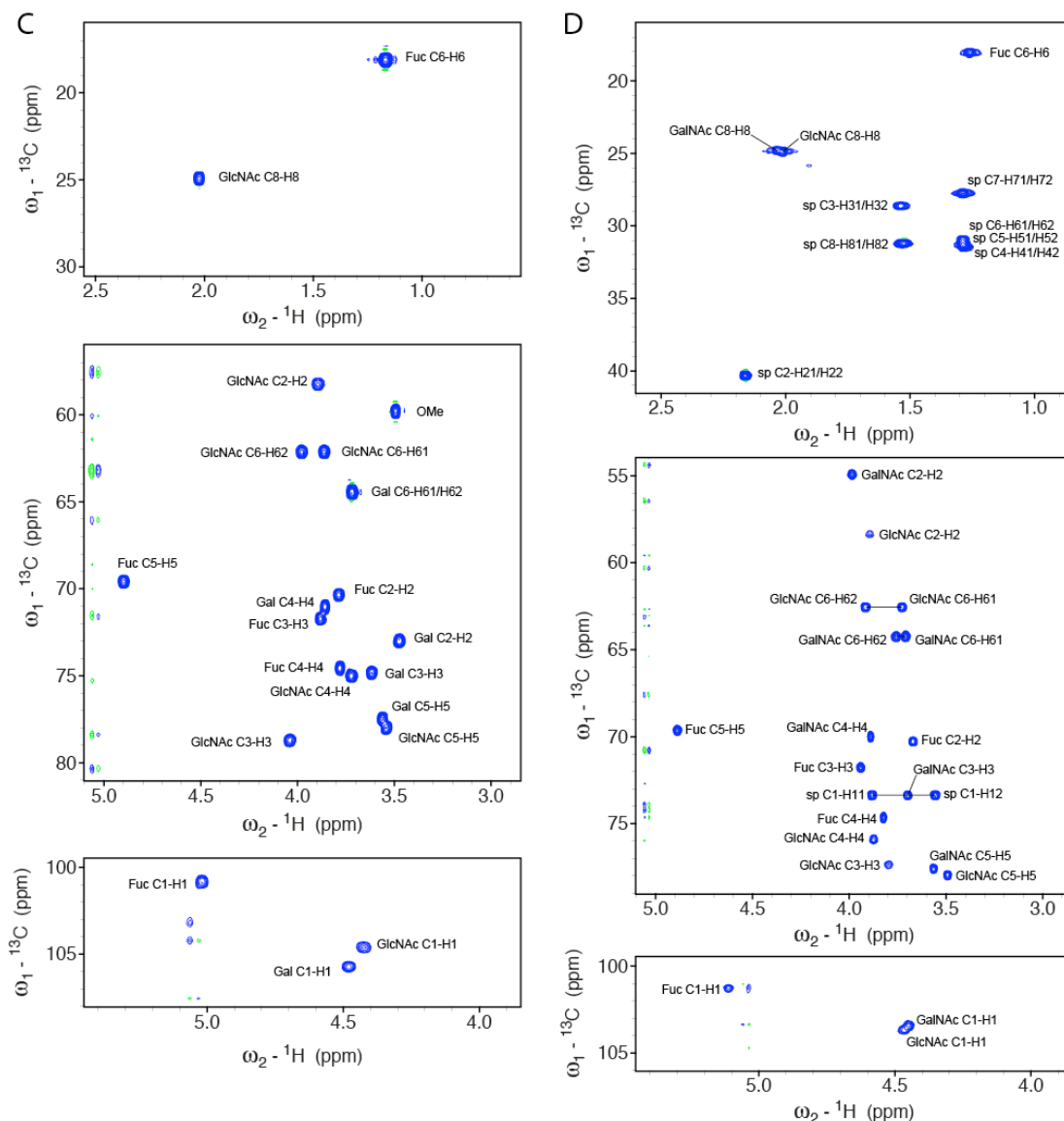


Figure S5 (continuation). (E) ^1H - ^{13}C HSQC spectrum of Bv9 with a β -hexene-tetrol moiety at the reducing end (1 mM) measured at 500 MHz and 273 K. Please note that 273 K was required so that the water signal was not overlapping with Fuc3 H1 but upfield of it. The proximal N-acetylglucosamine that harbours the α 1,3 branch is indicated as "GlcNAc1", the distal N-acetylglucosamine that stacks to the fucose is labeled "GlcNAc2". (F) ^1H - ^{13}C HSQC spectrum LDFT with free OH at the reducing end (3.2 mM) measured at 500 MHz and 275 K. The spectrum is referenced to DSS. Fuc3 corresponds to the α 1,2-linked fucose, Fuc4 to the α 1,3-linked fucose. Signals depending on the anomeric state of the glucose are labeled with (α) or (β).

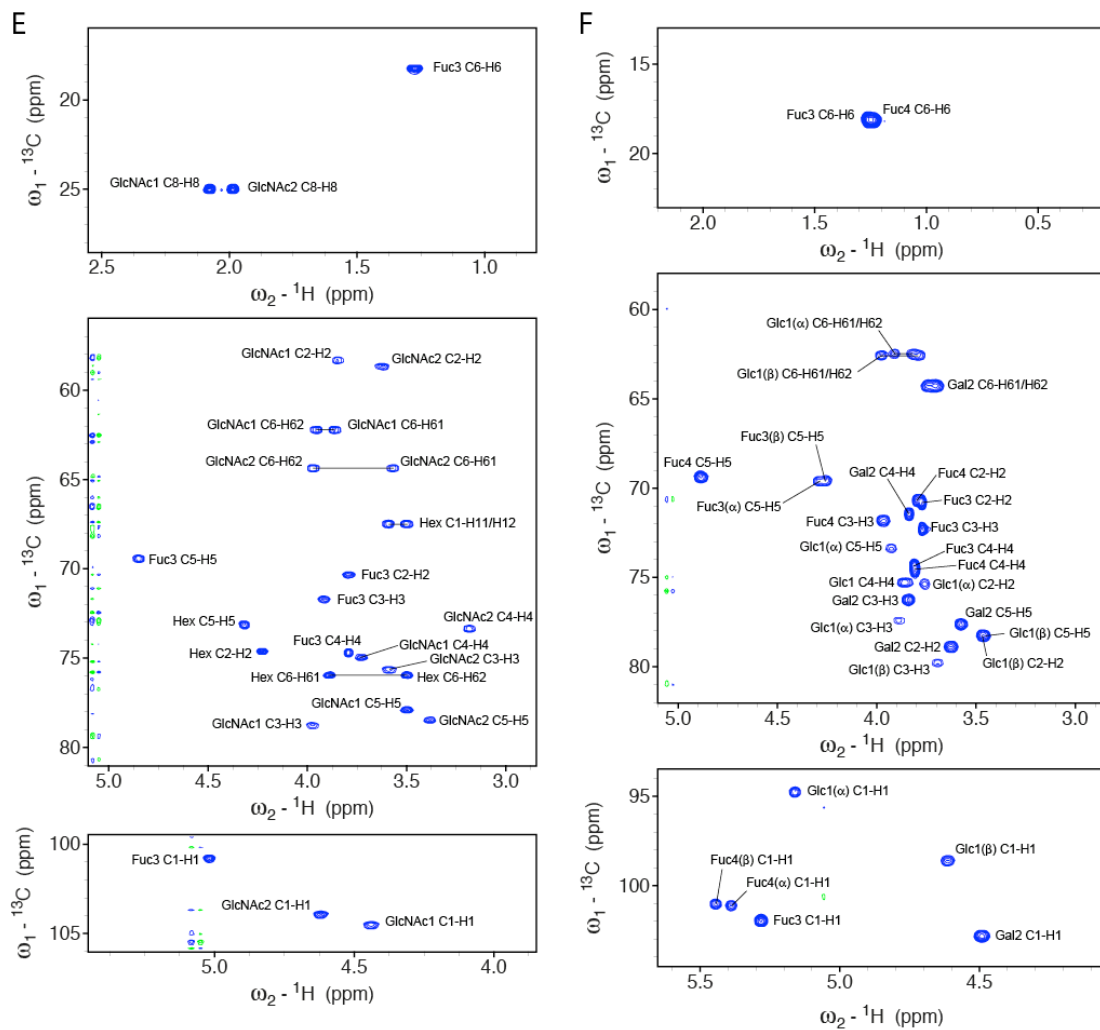


Figure S5 (continuation).

Table S1 NMR structure determination statistics of 3FChB (fucosylated chitobiose, GlcNAc β 1,4[Fuca1,3]GlcNAc β), methyl Le^x, methyl Le^a, LDNF (fucosylated LacdiNAc, GalNAc β 1,4[Fuca1,3]GlcNAc β) the amphibian glycan Bv9 (GlcNAc β 1,3[Fuca1,4]GlcNAc β) and LDFT (Lactodifucotetraose β anomer, Fuca1,2Gal β 1,4[Fuca1,3]GlcNAc β).

	3FChB	methyl Le ^x	methyl Le ^a	LDNF	Bv9	LDFT(β)
NMR distance and dihedral constraints						
Distance restraints						
Total NOE	42	29	29	42	32	42
Intra-residue	31	20	20	21	15	19
Inter-residue	11	9	9	21	17	23
Total dihedral angle restraints						
HN-CO peptide bonds of acetamido	2	0	0	0	0	0
Structure statistics						
Violations (mean and s.d.)						
Number of distance violations > 0.1 Å	0.0±0.0	0.0±0.0	0.0±0.0	0.0±0.0	0.0±0.0	0.0±0.0
Number of dihedral angle violations > 5°	0.0±0.0	0.0±0.0	0.0±0.0	0.0±0.0	0.0±0.0	0.0±0.0
Max. dihedral angle violation (°)	0.0±0.0	0.0±0.0	0.0±0.0	0.0±0.0	0.0±0.0	0.0±0.0
Max. distance constraint violation (Å)	0.04±0.00	0.07±0.00	0.00±0.00	0.09±0.00	0.01±0.01	0.09±0.01
Deviations from idealized geometry						
Bond lengths (Å)	0.01463± 0.0001	0.01537± 0.0002	0.0155± 0.0001	0.0143± 0.0001	0.0147± 0.0001	0.0166± 0.0002
Bond angles (°)	1.84±0.02	1.90±0.05	1.90±0.04	1.97±0.06	1.82±0.04	2.05±0.02
Average pairwise r.m.s. deviation* (Å)						
all heavy	0.05±0.02	0.15±0.15	0.14±0.13	0.26±0.16	0.14±0.13	0.07±0.03

* Pairwise r.m.s. deviation was calculated among 20 refined structures.

Table S2: Experimental chemical shift assignments and calculated chemical shifts using the GIAO method.

carbohydrates	Fuc α 1, 3GlcNAc	3FChB		Le ^x		Le ^a		LDNF		Bv9		LDFT (β)	
	exp.	exp.	GIAO calc.	exp.	GIAO calc.	exp.	GIAO calc.	exp.	GIAO calc.	exp.	GIAO calc.	exp.	GIAO calc.
GlcNAc/Glc (reducing end)													
H1	4.55	4.46	4.30	4.45	4.33	4.43	4.30	4.47	4.29	4.44	4.18	4.62	4.64
H2	3.83	3.89	3.92	3.91	3.95	3.90	3.98	3.89	3.94	3.85	3.99	3.47	3.38
H3	3.64	3.81	3.55	3.83	3.57	4.04	3.60	3.80	3.49	3.97	3.62	3.70	3.74
H4	3.49	3.88	3.91	3.93	3.43	3.73	3.41	3.88	3.47	3.73	3.33	3.87	3.65
H5	3.47	3.49	3.12	3.58	3.42	3.55	3.46	3.49	3.36	3.50	3.32	3.46	3.43
H61	3.72	3.92	3.93	3.86	4.06	3.98	4.29	3.91	3.94	3.96	4.18	3.98	3.87
H62	3.91	3.74	3.84	4.00	3.70	3.87	3.78	3.73	3.70	3.87	3.64	3.79	3.64
Q8	2.02	2.01	(1.83)	2.02	(1.84)	2.03	(1.84)	2.01	(1.83)	2.08	(1.88)	—	—
GlcNAc/Gal /GalNAc (stacking with Fuc)													
H1	—	4.52	4.80	4.44	4.21	4.48	4.44	4.45	4.14	4.62	4.64	4.49	4.28
H2	—	3.75	3.87	3.49	3.66	3.48	3.53	3.99	4.11	3.63	3.92	3.63	3.85
H3	—	3.52	3.44	3.65	3.62	3.62	3.55	3.70	3.66	3.59	3.51	3.84	3.77
H4	—	3.23	3.37	3.88	3.82	3.86	3.75	3.89	3.99	3.19	3.22	3.84	3.50
H5	—	3.42	3.52	3.59	3.59	3.57	3.57	3.66	3.64	3.39	3.42	3.58	3.35
H61	—	3.96	3.92	3.72	3.57	3.72	4.02	3.71	3.78	3.97	3.86	3.75	3.84
H62	—	3.60	3.68	3.70	4.05	3.72	3.57	3.76	4.13	3.57	3.61	3.70	3.95
Q8	—	2.04	(1.80)	—	—	—	—	2.03	(1.84)	1.99	(1.91)	—	—
Fuc (stacking)													
H1	4.98	5.12	5.02	5.10	5.02	5.02	4.67	5.11	4.97	5.02	4.70	5.45	5.30
H2	3.70	3.68	3.55	3.67	3.54	3.79	3.60	3.67	3.51	3.80	3.49	3.79	3.65
H3	3.83	3.95	3.62	3.90	3.60	3.88	3.67	3.94	3.60	3.92	3.44	3.97	3.78
H4	3.80	3.80	3.85	3.78	3.86	3.78	3.86	3.82	3.83	3.80	3.58	3.81	3.88
H5	4.33	4.78	4.88	4.86	4.92	4.90	4.92	4.89	4.85	4.85	5.10	4.89	4.89
Q6	1.16	1.26	(1.27)	1.17	(1.13)	1.17	(1.14)	1.26	(1.37)	1.27	(1.26)	1.24	(1.31)

^a The carbohydrates of which the experimental shifts were obtained were either β -methyl glycosides (Le^x, Le^a), contained a free OH (LDFT) or a spacer ($-\text{[CH}_2\text{]}_5\text{-NH}_2$ in case of Fuc α 1,3GlcNAc and LDNF; $-\text{[CH}_2\text{]}_5\text{-COONa}$ in 3FChB and hexenetetrol in Bv9). Calculations were obtained for the β -methyl glycosides except for LDFT for which the free OH as β -anomer was used.

^b Experimental values were measured at 273-278 K (see experimental section, the slightly different values were necessary to avoid overlap with the water signal) and referenced to DSS (0.00 ppm corresponds to 0.00 ppm referenced to TMS). Fuc H5 is shown in bold. The methylene groups H61 and H62 were not stereospecifically assigned.

^c Obtained with the GIAO chemical shift calculation in Gaussian 09 for a water environment (CPCM), values were referenced to TMS (separate calculation). These values are comparable to the experimental values because both TMS and DSS have identical ¹H chemical shifts (0.00 ppm). The predicted methyl group chemical shifts are the average of the three calculated chemical shifts, but since methyl groups rotate fast and sample a full cone, the average of three individual positions is not a good approximation. Therefore the values are in brackets.

Table S3: The hydrogen bond length (H...O), electron density $\rho(r_c)$, the Laplacian of the electron density $\nabla^2 \rho(r_c)$ and the ellipticity ϵ at the bond critical points of the DFT-minimized structures.

carbohydrate	Distance (H...O) [Å]	$\rho(r_c)$ at bond critical point [au]	$\nabla^2 \rho(r_c)$ at bond critical point [au]	ϵ at bond critical point [au]
3FChB	2.35	0.013	0.037	0.045
Lex	2.32	0.014	0.039	0.041
Lea	2.33	0.013	0.038	0.046
LDNF	2.41	0.012	0.034	0.019
Bv9	2.33	0.013	0.038	0.049
LDFT	2.47	0.010	0.030	0.031

Table S4: Carbohydrate structures with a characteristic Fuc H5 chemical shift between 4.5 and 5.0 ppm found by a GLYCOSCIENCES.DE database search, ordered in respect to the central glycomotifs, together with 21 additional values from the literature (marked by +).

# ^a	Carbohydrate structure ^b	Motif ^c	Chemical shift of Fuc H5 ^d	LINUCS ID or Reference ^e
	Le^x containing carbohydrates			
1	Galβ1,4[Fuα1,3]GlcNAcβ1,6[Galβ1,3GlcNAcβ1,3]Galβ1,4Glc	Le ^x	4.866	1670
2	Galβ1,4[Fuα1,3]GlcNAcβ1,4[Galβ1,4GlcNAcβ1,2]Manα1,3[Galβ1,4GlcNAcβ1,2 Manβ1,6]Manβ1,4GlcNAcβ1,4GlcNAcβ-Asn	Le ^x	4.835	274
3	Galβ1,4[Fuα1,3]GlcNAcβ1,6[Galβ1,3]GalNAc-ol	Le ^x	4.835	4286
4	Galβ1,4[Fuα1,3]GlcNAcβ1,3Galβ1,4GlcNAcβ1,6[Fuα1,2Galβ1,4GlcNAcβ1,3]GalNAc-ol	Le ^x	4.835	12503
5	Galβ1,4[Fuα1,3]GlcNAcβ1,3Galβ1,4GlcNAcβ1,6[Fuα1,2Galβ1,3GlcNAcβ1,3]GalNAc-ol	Le ^x	4.834	12501
6	Galβ1,4[Fuα1,3]GlcNAcβ1,3Galβ1,3GalNAc-ol	Le ^x	4.833	4285
7	Galβ1,4[Fuα1,3]GlcNAcβ1,3Galβ1,4Glc	Le ^x	4.832	212
8	Galβ1,4[Fuα1,3]GlcNAcβ1,4[Galβ1,4GlcNAcβ1,2]Manα1,3[Galβ1,4GlcNAcβ1,6[Galβ1,4GlcNAcβ1,2]Manβ1,6]Manβ1,4GlcNAcβ1,4GlcNAcβ-Asn	Le ^x	4.832	275
9	Galβ1,4[Fuα1,3]GlcNAcβ1,2[Galβ1,4GlcNAcβ1,6]Manα1,6[Galβ1,4GlcNAcβ1,4[Galβ1,4GlcNAcβ1,2]Manβ1,3]Manβ1,4GlcNAcβ1,4GlcNAcβ-Asn	Le ^x	4.832	276
10	Galβ1,4[Fuα1,3]GlcNAcβ1,4[Galβ1,4GlcNAcβ1,2]Manα1,6[Galβ1,4GlcNAcβ1,4[Galβ1,4GlcNAcβ1,2]Manβ1,3]Manβ1,4GlcNAcβ1,4GlcNAcβ-Asn	Le ^x	4.832	277
11	Neu5Acα2,3Galβ1,3[Neu5Acα2,6]GalNAcβ1,3[Galβ1,4[Fuα1,3]GlcNAcβ1,6]Galβ1,4Glc	Le ^x	4.832	1174
12	Galβ1,4[Fuα1,3]GlcNAcβ1,6[Neu5Acα2,6[Galβ1,3]GlcNAcβ1,3]Galβ1,4Glc	Le ^x	4.832	2835
13	Galβ1,4[Fuα1,3]GlcNAcβ1,3[Fuα1,2]Galβ1,4GlcNAcβ1,6[GlcNAcβ1,3]GalNAc-ol	Le ^x	4.831	12496
14	Galβ1,4[Fuα1,3]GlcNAcβ1,2Manα1,6Manβ1,4GlcNAcβ1,4[Fuα1,6]GlcNAcβ-Asn	Le ^x	4.83	273
15	Galβ1,4[Fuα1,3]GlcNAcβ1,6[GlcNAcβ1,3]GalNAc-ol	Le ^x	4.83	4288
16	Galβ1,4[Fuα1,3]GlcNAcβ1,4Manα1,3Manβ1,4GlcNAcβ1,4[Fuα1,6]GlcNAcβ-Asn	Le ^x	4.83	12533
17	Galβ1,4[Fuα1,3]GlcNAcβ1,2Manα1,3[Manα1,6]Manβ1,4GlcNAcβ1,4[Fuα1,6]GlcNAcβ-Asn	Le ^x	4.83	12538
18	Galβ1,4[Fuα1,3]GlcNAcβ1,3Galβ1,4GlcNAcβ1,2Manα1,6Manβ1,4GlcNAcβ1,4[Fuα1,6]GlcNAcβ-Asn	Le ^x	4.83	12548
19	Galβ1,4[Fuα1,3]GlcNAcβ1,2Manα1,6[Neu5Acα2,6Galβ1,4GlcNAcβ1,2Manβ1,3]Manβ1,4GlcNAcβ1,4[Fuα1,6]GlcNAcβ-Asn	Le ^x	4.829	278
20	Galβ1,4[Fuα1,3]GlcNAcβ1,6[Galβ1,4GlcNAcβ1,3]GalNAc-ol	Le ^x	4.829	4294
21	Galβ1,4[Fuα1,3]GlcNAcβ1,6[Neu5Acα2,3Galβ1,3]GalNAc-ol	Le ^x	4.829	4298
22	Neu5Acα2,6Galβ1,4GlcNAcβ1,3[Galβ1,4[Fuα1,3]GlcNAcβ1,6]Galβ1,4Glc	Le ^x	4.829	8313
23	Galβ1,4[Fuα1,3]GlcNAcβ1,6[Neu5Acα2,3Galβ1,3]Galβ1,4Glc	Le ^x	4.827	2838
24	Galβ1,4[Fuα1,3]GlcNAcβ1,6[Neu5Acα2,3Galβ1,3GlcNAcβ1,6]Galβ1,4Glc	Le ^x	4.825	2837
25	Galβ1,4[Fuα1,3]GlcNAcβ1,2Manα1,3Manβ1,4GlcNAcβ1,4[Fuα1,6]GlcNAcβ-Asn	Le ^x	4.822	272
26	Galβ1,4[Fuα1,3]GlcNAcβ1,6[Galβ1,4GlcNAcβ1,2]Manα1,3[Galβ1,4GlcNAcβ1,2Manβ1,6]Manβ1,4GlcNAcβ1,4GlcNAcβ-Asn	Le ^x	4.82	279
27	Galβ1,4[Fuα1,3]GlcNAcβ1,6[Fuα1,2Galβ1,3GlcNAcβ1,3]Galβ1,4Glc	Le ^x	4.819	1668
28	Galβ1,4[Fuα1,3]GlcNAcβ1,3[KDNα2,6]GalNAc-ol	Le ^x	4.818	3508
29	Galβ1,4[Fuα1,3]GlcNAcβ1,3GalNAc-o	Le ^x	4.813	4278
30	Galβ1,4[Fuα1,3]GlcNAcβ1,3[Galβ1,4[Fuα1,3]GlcNAcβ1,6]Galβ1,4Glc	Le ^x	4.829	6680
31	Galβ1,4[Fuα1,3]GlcNAcβ1,3[Galβ1,4[Fuα1,3]GlcNAcβ1,6]Galβ1,4Glc	Le ^x	4.821	6680
32	Galβ1,4[Fuα1,3]GlcNAcβ1,2[Galβ1,4[Fuα1,3]GlcNAcβ1,6]Manα1,6Manβ1,4GlcNAcβ1,4[Fuα1,6]GlcNAcβ-Asn	Le ^x	4.83	12544
33	Galβ1,4[Fuα1,3]GlcNAcβ1,2[Galβ1,4[Fuα1,3]GlcNAcβ1,6]Manα1,6Manβ1,4GlcNAcβ1,4[Fuα1,6]GlcNAcβ-Asn	Le ^x	4.83	12544
34	Galβ1,4[Fuα1,3]GlcNAcβ1,2 Manα1,3[Galβ1,4[Fuα1,3]GlcNAcβ1,2 Manα1,6]Manβ1,4GlcNAcβ1,4[Fuα1,6]GlcNAcβ-Asn	Le ^x	4.829	12545
35	Galβ1,4[Fuα1,3]GlcNAcβ1,2Manα1,3[Galβ1,4[Fuα1,3]GlcNAcβ1,2 Manα1,6]Manβ1,4GlcNAcβ1,4[Fuα1,6]GlcNAcβ-Asn	Le ^x	4.829	12546
36	Galβ1,4[Fuα1,3]GlcNAcβ1,2Manα1,3[Galβ1,4[Fuα1,3]GlcNAcβ1,2 Manα1,6]Manβ1,4GlcNAcβ1,4[Fuα1,6]GlcNAcβ-Asn	Le ^x	4.829	12546
37	Galβ1,4[Fuα1,3]GlcNAcβ1,2 Manα1,3[Galβ1,4[Fuα1,3]GlcNAcβ1,2	Le ^x	4.822	12545

	Man α 1,6]Man β 1,4GlcNAc β 1,4[Fuca α 1,6]GlcNAc β -Asn			
38	Fuca α 1,2Gal β 1,3[Fuca α 1,4]GlcNAc β 1,3[Gal β 1,4]Fuca α 1,3]GlcNAc β 1,6]Gal β 1,4GlcNAc β 1,6[Neu5Ac α 2,3Gal β 1,3]GalNAc-ol	Le ^x	4.81	12417
39	Fuca α 1,2Gal β 1,3[Fuca α 1,4]GlcNAc β 1,3[Gal β 1,4]Fuca α 1,3]GlcNAc β 1,6]Gal β 1,4Glc	Le ^x	4.823	8107
40	Gal β 1,4]Fuca α 1,3]GlcNAc β 1,3Gal β 1,4[Fuca α 1,3]GlcNAc β 1,3Gal β 1,3GalNAc-ol	Le ^x	4.851	12507
41	Neu5Ac α 2,3Gal β 1,3[Fuca α 1,4]GlcNAc β 1,3[Gal β 1,4]Fuca α 1,3]GlcNAc β 1,6]Gal β 1,4Glc	Le ^x	4.795	4124
42	Gal β 1,4]Fuca α 1,3]GlcNAc β 1,3Gal β 1,4GlcNAc β 1,6[Neu5Ac α 2,3Gal β 1,3[Fuca α 1,4]GalNAc β 1,3]Gal β 1,4Glc	Le ^x	4.789	4222
43	Gal β 1,4]Fuca α 1,3]GlcNAc β 1,6[Gal β 1,3[Fuca α 1,4]GlcNAc β 1,3]Gal β 1,4Glc	Le ^x	4.873	1671
44	Gal β 1,3GlcNAc β 1,3Gal β 1,4]Fuca α 1,3]GlcNAc β 1,6[Neu5Ac α 2,3Gal β 1,3]GalNAc-ol	Int Le ^x	4.806	12413
45	Gal β 1,4GlcNAc β 1,3Gal β 1,4]Fuca α 1,3]GlcNAc β 1,6[Neu5Ac α 2,3Gal β 1,3]GalNAc-ol	Int Le ^x	4.806	12414
46	Gal β 1,3GlcNAc β 1,3Gal β 1,4]Fuca α 1,3]GlcNAc β 1,6[Gal β 1,3]GalNAc-ol	Int Le ^x	4.806	12419
47	Gal β 1,3GlcNAc β 1,3Gal β 1,4]Fuca α 1,3]GlcNAc β 1,6[Fuca α 1,2Gal β 1,3GlcNAc β 1,3]Gal β 1,4Glc	Int Le ^x	4.801	14252
48	Fuca α 1,2Gal β 1,3[Fuca α 1,4]GlcNAc β 1,3[Gal β 1,4GlcNAc β 1,6]Gal β 1,4]Fuca α 1,3]GlcNAc β 1,6[Neu5Ac α 2,3Gal β 1,3]GalNAc-ol	Int Le ^x	4.81	12418
49	Gal β 1,3[Fuca α 1,4]GlcNAc β 1,3Gal β 1,4]Fuca α 1,3]GlcNAc β 1,6[Gal β 1,3GlcNAc β 1,3]Gal β 1,4Glc	Int Le ^x	4.801	2727
50	Gal β 1,3[Fuca α 1,4]GlcNAc β 1,3Gal β 1,4GlcNAc β 1,3Gal β 1,4]Fuca α 1,3]GlcNAc β 1,6[Neu5Ac α 2,3Gal β 1,3]GalNAc-ol	Int Le ^x	4.81	12416
51	Fuca α 1,2Gal β 1,3[Fuca α 1,4]GlcNAc β 1,3Gal β 1,4]Fuca α 1,3]GlcNAc β 1,3Gal β 1,4Glc	Int Le ^x	4.811	14239
52	GlcNAc β 1,3Gal β 1,4]Fuca α 1,3]GlcNAc β 1,9-9-hydroxy-Nonanoate-OMe	Int Le ^x	4.82	9296
53	Gal β 1,4GlcNAc β 1,3Gal β 1,4]Fuca α 1,3]GlcNAc β 1,9-9-hydroxy-Nonanoate-OMe	Int Le ^x	4.82	9297
54	Neu5Ac α 2,6Gal β 1,4GlcNAc β 1,3Gal β 1,4]Fuca α 1,3]GlcNAc β 1,9-9-hydroxy-Nonanoate-OMe	Int Le ^x	4.82	9299
55	Neu5Ac α 2,3Gal β 1,4GlcNAc β 1,3Gal β 1,4]Fuca α 1,3]GlcNAc β 1,9-9-hydroxy-Nonanoate-OMe	Int Le ^x	4.81	8860
56	Fuca α 1,2Gal β 1,3[Fuca α 1,4]GlcNAc β 1,3Gal β 1,4]Fuca α 1,3]GlcNAc β 1,3Gal β 1,4]Fuca α 1,3]GlcNAc β 1,3Gal β 1,4Glc	Int Le ^x	4.812	12761
57	Fuca α 1,2Gal β 1,3[Fuca α 1,4]GlcNAc β 1,3Gal β 1,4[Fuca α 1,3]GlcNAc β 1,3Gal β 1,4]Fuca α 1,3]GlcNAc β 1,3Gal β 1,4Glc	Int Le ^x	4.812	12761
58	Fuca α 1,2Gal β 1,3GlcNAc β 1,3Gal β 1,4]Fuca α 1,3]GlcNAc β 1,6[Fuca α 1,2Gal β 1,3[Fuca α 1,4]GlcNAc β 1,3]Gal β 1,4Glc	Int Le ^x	4.814	12762
59	Fuca α 1,2Gal β 1,3[Fuca α 1,4]GlcNAc β 1,3Gal β 1,4]Fuca α 1,3]GlcNAc β 1,6[Fuca α 1,2Gal β 1,3[Fuca α 1,4]GlcNAc β 1,3]Gal β 1,4Glc	Int Le ^x	4.8	12763
60	Gal β 1,4[Fuca α 1,3]GlcNAc β 1,3Gal β 1,4]Fuca α 1,3]GlcNAc β 1,3Gal β 1,3GalNAc-ol	Int Le ^x	4.865	12507
61	Neu5Ac α 2,3Gal β 1,4[Fuca α 1,3]GlcNAc β 1,3Gal β 1,4]Fuca α 1,3]GlcNAc β 1,9-9-hydroxy-Nonanoate-OMe	Int Le ^x	4.82	8861
62	Gal β 1,4GlcNAc β 1,3Gal β 1,4]Fuca α 1,3]GlcNAc β 1,6[Neu5Ac α 2,3Gal β 1,3[Fuca α 1,4]GalNAc β 1,3]Gal β 1,4Glc	Int Le ^x	4.776	4221
63	Neu5Ac α 2,3Gal β 1,4]Fuca α 1,3]GlcNAc β 1,6[Gal β 1,4GlcNAc β 1,3]GalNAc-ol	Sia Le ^x	4.83	12342
64	Neu5Ac α 2,3Gal β 1,4]Fuca α 1,3]GlcNAc β 1,6[Fuca α 1,2Gal β 1,3GlcNAc β 1,3]GalNAc-ol	Sia Le ^x	4.821	12697
65	Neu5Ac α 2,3Gal β 1,4]Fuca α 1,3]GlcNAc β 1,3Gal β 1,3GalNAc-ol	Sia Le ^x	4.82	12340
66	Neu5Ac α 2,3Gal β 1,4]Fuca α 1,3]GlcNAc β 1,6[GlcNAc β 1,3]GalNAc-ol	Sia Le ^x	4.82	12341
67	Neu5Ac α 2,3Gal β 1,4]Fuca α 1,3]GlcNAc β 1,6[Fuca α 1,2Gal β 1,3]GalNAc-ol	Sia Le ^x	4.82	12388
68	Neu5Ac α 2,3Gal β 1,4]Fuca α 1,3]GlcNAc β 1,3Gal β 1,3[Neu5Ac α 2,6]GalNAc-ol	Sia Le ^x	4.819	17378
69	Neu5Ac α 2,3Gal β 1,4]Fuca α 1,3]GlcNAc β 1,3[Neu5Ac α 2,6]GalNAc-ol	Sia Le ^x	4.804	12408
70	Neu5Ac α 2,3Gal β 1,4]Fuca α 1,3]GlcNAc β 1,6[Neu5Ac α 2,3Gal β 1,3]Gal β 1,4Glc	Sia Le ^x	4.798	6665
71	Neu5Ac α 2,3Gal β 1,4]Fuca α 1,3]GlcNAc β 1,3Gal β 1,4[Fuca α 1,3]GlcNAc β 1,9-9-hydroxy-Nonanoate-OMe	Sia Le ^x	4.82	8861
72	3S-Gal β 1,4]Fuca α 1,3]GlcNAc β 1,6[Fuca α 1,2Gal β 1,3]GalNAc-ol	3S-Le ^x	4.806	17382
73	3S-Gal β 1,4]Fuca α 1,3]GlcNAc β 1,6[Neu5Ac α 2,3Gal β 1,3]GalNAc-ol	3S-Le ^x	4.805	17388
74	3S-Gal β 1,4]Fuca α 1,3]GlcNAc β 1,3[Neu5Ac α 2,6]GalNAc-ol	3S-Le ^x	4.802	12406
75	Gal β 1,4[6S]]Fuca α 1,3]GlcNAc β 1,3[6S]Gal β 1,4[6S]GlcNAc-ol	6S-Le ^x	4.783	4412
76	Gal β 1,4[6S]]Fuca α 1,3]GlcNAc β 1,3Gal β 1,4[6S]GlcNAc-ol	6S-Le ^x	4.782	4410
77	Gal β 1,4[6S]]Fuca α 1,3]GlcNAc β 1,3[6S]Gal β 1,4[6S]GlcNAc β 1,3Gal-ol	6S-Le ^x	4.778	3526
78	Gal β 1,4[6S]]Fuca α 1,3]GlcNAc β 1,3Gal β 1,4[6S]]Fuca α 1,3]GlcNAc-ol	6S-Le ^x	4.782	4408
79	Gal β 1,3[Fuca α 1,4]GlcNAc β 1,3Gal β 1,4[6S]]Fuca α 1,3]GlcNAc β 1,3Gal β 1,4Glc	Int 6S-Le ^x	4.805	22552

80	Fuca1,2Galβ1,3GlcNAcβ1,3Galβ1,4[6S][Fuca1,3]GlcNAcβ1,3Galβ1,4Glcβ	Int 6S-Le ^x	4.823	22550
81	Fuca1,2Galβ1,3GlcNAcβ1,3Galβ1,4[6S][Fuca1,3]GlcNAcβ1,3Galβ1,4Glcα	Int 6S-Le ^x	4.823	22551
82	Galβ1,3[Fuca1,4]GlcNAcβ1,3Galβ1,4[6S][Fuca1,3]GlcNAcβ1,3Galβ1,4Glcβ	Int 6S-Le ^x	4.805	22553
83	Fuca1,2Galβ1,3[Fuca1,4]GlcNAcβ1,3Galβ1,4[6S][Fuca1,3]GlcNAcβ1,3Galβ1,4Glcα	Int 6S-Le ^x	4.819	22556
84	Fuca1,2Galβ1,3[Fuca1,4]GlcNAcβ1,3Galβ1,4[6S][Fuca1,3]GlcNAcβ1,3Galβ1,4Glcβ	Int 6S-Le ^x	4.819	22557
85	Neu5Acα2,3Galβ1,4[6S][Fuca1,3]GlcNAcβ1,6[Neu5Acα2,3Galβ1,3]GalNAc-ol	6S-Sia Le ^x	4.802	17385
	Le^y containing carbohydrates			
86	Fuca1,2Galβ1,4[Fuca1,3]GlcNAcβ1,6[Neu5Acα2,6Galβ1,4GlcNAcβ1,3]Galβ1,4Glc	Le ^y	4.88	8312
87	Fuca1,2Galβ1,4[Fuca1,3]GlcNAcβ1,3[Fuca1,2]Galβ1,4GlcNAcβ1,6[GlcNAcβ1,3]GalNAc-ol	Le ^y	4.88	12499
88	Fuca1,2Galβ1,4[Fuca1,3]GlcNAcβ1,3[Fuca1,2]Galβ1,4GlcNAcβ1,6[Fuca1,2GlcNAcβ1,3]GalNAc-ol	Le ^y	4.88	12504
89	Fuca1,2Galβ1,4[Fuca1,3]GlcNAcβ1,3[Neu5Acα2,3Galβ1,4GlcNAcβ1,6]GalNAc-ol	Le ^y	4.877	12695
90	Fuca1,2Galβ1,4[Fuca1,3]GlcNAcβ1,6[Fuca1,2Galβ1,3GlcNAcβ1,3]GalNAc-ol	Le ^y	4.874	8956
91	Fuca1,2Galβ1,4[Fuca1,3]GlcNAcβ1,3[Neu5Acα2,6]GalNAc-ol	Le ^y	4.874	12386
92	Fuca1,2Galβ1,4[Fuca1,3]GlcNAcβ1,3[Galβ1,4GlcNAcβ1,6]Galβ1,3GalNAc-ol	Le ^y	4.874	12508
93	Fuca1,2Galβ1,4[Fuca1,3]GlcNAcβ1,3[Fuca1,2Galβ1,4GlcNAcβ1,6]Galβ1,3GalNAc-ol	Le ^y	4.874	12511
94	Fuca1,2Galβ1,4[Fuca1,3]GlcNAcβ1,3Galβ1,3GalNAc-ol	Le ^y	4.873	6679
95	Fuca1,2Galβ1,4[Fuca1,3]GlcNAcβ1,6[GlcNAcβ1,3]GalNAc-ol	Le ^y	4.872	8949
96	Fuca1,2Galβ1,4[Fuca1,3]GlcNAcβ1,6[Fuca1,2Galβ1,3]GalNAc-ol	Le ^y	4.871	1667
97	Fuca1,2Galβ1,4[Fuca1,3]GlcNAcβ1,6[Neu5Acα2,3Galβ1,3]GalNAc-ol	Le ^y	4.87	12286
98	Fuca1,2Galβ1,4[Fuca1,3]GlcNAcβ1,3[Fuca1,2Galβ1,4GlcNAcβ1,6]GalNAc-ol	Le ^y	4.867	8953
99	Fuca1,2Galβ1,4[Fuca1,3]GlcNAcβ1,3GalNAc-ol	Le ^y	4.866	4284
100	Fuca1,2Galβ1,4[Fuca1,3]GlcNAcβ1,6[Galβ1,3]GalNAc-ol	Le ^y	4.85	4292
101	Fuca1,2Galβ1,4[Fuca1,3]GlcNAcβ1,3[Fuca1,2Galβ1,4[Fuca1,3]GlcNAcβ1,6]GalNAc-ol	Le ^y	4.873	12510
102	Fuca1,2Galβ1,4[Fuca1,3]GlcNAcβ1,3[Fuca1,2Galβ1,4[Fuca1,3]GlcNAcβ1,6]GalNAc-ol	Le ^y	4.873	12510
103	Fuca1,2Galβ1,4[6S][Fuca1,3]GlcNAcβ1,6[Neu5Acα2,3Galβ1,3]GalNAc-ol	6S-Le ^y	4.854	17381
	Le^a containing carbohydrates			
104	Galβ1,3[Fuca1,4]GlcNAc	Le ^a	4.881	1708
105	Galβ1,3[Fuca1,4]GlcNAcβ1,3[Neu5Acα2,6Galβ1,4GlcNAcβ1,6]Galβ1,4Glc	Le ^a	4.876	2836
106	Galβ1,3[Fuca1,4]GlcNAcβ1,3Galβ1,4Glc	Le ^a	4.871	1223
107	Galβ1,3[Fuca1,4]GlcNAcβ	Le ^a	4.78	668
108	Galβ1,3[Fuca1,4]GlcNAcβ1,3Galβ1,4[6S][Fuca1,3]GlcNAcβ1,3Galβ1,4Glcβ	Le ^a	4.883	22552
109	Galβ1,3[Fuca1,4]GlcNAcβ1,3Galβ1,4[6S][Fuca1,3]GlcNAcβ1,3Galβ1,4Glcβ	Le ^a	4.883	22553
110	Galβ1,4[Fuca1,3]GlcNAcβ1,6[Galβ1,3[Fuca1,4]GlcNAcβ1,3]Galβ1,4Glc	Le ^a	4.915	1671
111	Fuca1,2Galβ1,3[Fuca1,4]GlcNAcβ1,3[Galβ1,4[Fuca1,3]GlcNAcβ1,6]Galβ1,4GlcNAcβ1,6[Neu5Acα2,3Galβ1,3]GalNAc-ol	Le ^a	4.86	12417
112	Galβ1,3[Fuca1,4]GlcNAcβ1,3Galβ1,4[Fuca1,3]GlcNAcβ1,6[Galβ1,3GlcNAcβ1,3]Galβ1,4Glc	Le ^a	4.869	2727
113	Galβ1,3[Fuca1,4]GlcNAcβ1,3Galβ1,4GlcNAcβ1,3Galβ1,4[Fuca1,3]GlcNAcβ1,6[Neu5Acα2,3Galβ1,3]GalNAc-ol	Le ^a	4.86	12416
114	Galβ1,3[Fuca1,4]GlcNAcβ1,3Galβ1,4GlcNAcβ1,6[Neu5Acα2,3Galβ1,3[Fuca1,4]GalNAcβ1,3]Galβ1,4Glc	Le ^a	4.839	4223
115	Neu5Acα2,3Galβ1,4[Fuca1,3]GlcNAcβ1,6[Galβ1,3[Fuca1,4]GlcNAcβ1,3]Galβ1,4Glc	Le ^a	4.873	2826
116	Neu5Acα2,3Galβ1,3[Neu5Acα2,6][Fuca1,4]GlcNAcβ1,3Galβ1,4Glc	Int. Le ^a	4.878	2780
117	Neu5Acα2,3Galβ1,3[Fuca1,4]GlcNAcβ1,3Galβ1,4Glc	Sia Le ^a	4.864	1632
118	Galβ1,3GlcNAcβ1,3Galβ1,4GlcNAcβ1,6[Neu5Acα2,3Galβ1,3[Fuca1,4]GlcNAcβ1,3]Galβ1,4Glc	Sia Le ^a	4.838	4220
119	Neu5Acα2,3Galβ1,3[Fuca1,4]GlcNAcβ1,3[Neu5Acα2,6Galβ1,4GlcNAcβ1,3]Galβ1,4Glc	Sia Le ^a	4.837	4125

	,6]Galβ1,4Glc			
120	Neu5Acα2,3Galβ1,3[Fuα1,4]GlcNAcβ1,3[Galβ1,4GlcNAcβ1,6]Galβ1,4Glc	Sia Le ^a	4.836	4123
121	Galβ1,4[Fuα1,3]GlcNAcβ1,3Galβ1,4GlcNAcβ1,6[Neu5Acα2,3Galβ1,3[Fuα1,4]GlcNAcβ1,3]Galβ1,4Glc	Sia Le ^a	4.839	4222
122	Neu5Acα2,3Galβ1,3[Fuα1,4]GlcNAcβ1,3[Galβ1,4[Fuα1,3]GlcNAcβ1,6]Galβ1,4Glc	Sia Le ^a	4.836	4124
123	Galβ1,3[Fuα1,4]GlcNAcβ1,3Galβ1,4GlcNAcβ1,6[Neu5Acα2,3Galβ1,3[Fuα1,4]GlcNAcβ1,3]Galβ1,4Glc	Sia Le ^a	4.839	4223
124	Galβ1,4GlcNAcβ1,3Galβ1,4[Fuα1,3]GlcNAcβ1,6[Neu5Acα2,3Galβ1,3[Fuα1,4]GlcNAcβ1,3]Galβ1,4Glc	Sia Le ^a	4.838	4221
125	3S-Galβ1,3[Fuα1,4]GlcNAc	3S-Le ^a	4.874	13788
126	3S-Galβ1,3[Fuα1,4]GlcNAcβ	3S-Le ^a	4.861	13788
	Le^b containing carbohydrates			
127	Fuα1,2Galβ1,3[Fuα1,4]GlcNAcβ1,3Galβ1,3GalNAc-ol	Le ^b	4.88	5022
128	Fuα1,2Galβ1,3[Fuα1,4]GlcNAcβ1,3[Neu5Acα2,6Galβ1,4GlcNAcβ1,6]Galβ1,4Glc	Le ^b	4.866	2839
129	Fuα1,2Galβ1,3[Fuα1,4]GlcNAcβ1,3Galβ1,4GlcNAcβ1,6[GlcNAcβ1,3]GalNAc-ol	Le ^b	4.866	12497
130	Fuα1,2Galβ1,3[Fuα1,4]GlcNAcβ1,3Galβ1,4GlcNAcβ1,6[Galβ1,4GlcNAcβ1,3]GalNAc-ol	Le ^b	4.866	12506
131	Fuα1,2Galβ1,3[Fuα1,4]GlcNAcβ1,3Galβ1,4GlcNAcβ1,6[Fuα1,2Galβ1,3]GalNAc-ol	Le ^b	4.865	12502
132	Fuα1,2Galβ1,3[Fuα1,4]GlcNAcβ1,3Galβ1,3[Galβ1,4GlcNAcβ1,6]GalNAc-ol	Le ^b	4.865	12514
133	Fuα1,2Galβ1,3[Fuα1,4]GlcNAcβ1,3Galβ1,4Glc	Le ^b	4.86	1225
134	Fuα1,2Galβ1,3[Fuα1,4]GlcNAcβ1,3Galβ1,4GlcNAcβ1,3Galβ1,4GlcNAcβ1,6[Neu5Acα2,3Galβ1,3]GalNAc-ol	Le ^b	4.86	12415
135	Fuα1,2Galβ1,3[Fuα1,4]GlcNAcβ1,3[Fuα1,2]Galβ1,4GlcNAcβ1,6[GlcNAcβ1,3]GalNAc-ol	Le ^b	4.855	12505
136	Fuα1,2Galβ1,3[Fuα1,4]GlcNAcβ1,3[Galβ1,4GlcNAcβ1,3]GalNAc-ol	Le ^b	4.82	4291
137	Fuα1,2Galβ1,3[Fuα1,4]GlcNAcβ1,3Galβ1,4[Fuα1,3]GlcNAcβ1,6[Fuα1,2Galβ1,3[Fuα1,4]GlcNAcβ1,3]Galβ1,4Glc	Le ^b	4.853	12763
138	Fuα1,2Galβ1,3[Fuα1,4]GlcNAcβ1,3Galβ1,4[Fuα1,3]GlcNAcβ1,6[Fuα1,2Galβ1,3[Fuα1,4]GlcNAcβ1,3]Galβ1,4Glc	Le ^b	4.853	12763
139	Fuα1,2Galβ1,3[Fuα1,4]GlcNAcβ1,3[Galβ1,4[Fuα1,3]GlcNAcβ1,6]Galβ1,4Glc	Le ^b	4.861	8107
140	Fuα1,2Galβ1,3[Fuα1,4]GlcNAcβ1,3Galβ1,4[6S][Fuα1,3]GlcNAcβ1,3Galβ1,4Glcα	Le ^b	4.871	22556
141	Fuα1,2Galβ1,3[Fuα1,4]GlcNAcβ1,3Galβ1,4[6S][Fuα1,3]GlcNAcβ1,3Galβ1,4Glcβ	Le ^b	4.871	22557
142	Fuα1,2Galβ1,3[Fuα1,4]GlcNAcβ1,3Galβ1,4[Fuα1,3]GlcNAcβ1,3Galβ1,4Glc	Le ^b	4.86	14239
143	Fuα1,2Galβ1,3[Fuα1,4]GlcNAcβ1,3[Galβ1,4[Fuα1,3]GlcNAcβ1,6]Galβ1,4GlcNAcβ1,6[Neu5Acα2,3Galβ1,3]GalNAc-ol	Le ^b	4.86	12417
144	Fuα1,2Galβ1,3[Fuα1,4]GlcNAcβ1,3[Galβ1,4GlcNAcβ1,6]Galβ1,4[Fuα1,3]GlcNAcβ1,6[Neu5Acα2,3Galβ1,3]GalNAc-ol	Le ^b	4.86	12418
145	Fuα1,2Galβ1,3GlcNAcβ1,3Galβ1,4[Fuα1,3]GlcNAcβ1,6[Fuα1,2Galβ1,3[Fuα1,4]GlcNAcβ1,3]Galβ1,4Glc	Le ^b	4.868	12762
146	Fuα1,2Galβ1,3[Fuα1,4]GlcNAcβ1,3Galβ1,4[Fuα1,3]GlcNAcβ1,3Galβ1,4[Fuα1,3]GlcNAcβ1,3Galβ1,4Glc	Le ^b	4.853	12761
147	GalNAcα1,3[Fuα1,2]Galβ1,3[Fuα1,4]GlcNAcβ1,3Galβ1,4Glc	Int Le ^b	4.853	1271
148	GalNAcα1,3[Fuα1,2]Galβ1,3[Fuα1,4]GlcNAcβ1,3Gal	Int Le ^b	4.853	14257
	3-Fucosylated chitobiose containing carbohydrates			
149	Manα1,6[Xylβ1,2]Manβ1,4GlcNAcβ1,4[Fuα1,3]GlcNAcβ-Asn	3FChB	4.722	4566
150	Manα1,6[Manα1,3][Xylβ1,2]Manβ1,4GlcNAcβ1,4[Fuα1,3]GlcNAc	3FChB	4.721	1024
151	Manα1,6[Manβ1,3]Manβ1,4GlcNAcβ1,4[Fuα1,3][Fuα1,6]GlcNAcβ-Asn	3FChB	4.72	12525
152	Manα1,6[Manβ1,3]Manβ1,4GlcNAcβ1,4[Fuα1,3][Fuα1,6]GlcNAc	3FChB	4.71	2900
	4-Fucosylated GlcNAcβ1,3GlcNAc cont. carbohydrates			
153	GlcNAcβ1,3[Fuα1,4]GlcNAcβ1,4GalAα1,3Fuα		4.71	26716
+1	GlcNAcβ1,3[Fuα1,4]GlcNAcβ1,6-Hexene-tetrol		4.816	Coppin 2003
+2	GlcNAcβ1,3[Fuα1,4]GlcNAcβ1,6[Galβ1,3]GalNAc-ol		4.809	Coppin 2003
+3	GlcNAcβ1,3[Fuα1,4]GlcNAcβ1,6[GlcNAcα1,4Galβ1,3]GalNAc-ol		4.805	Coppin 2003

+4	GlcNAcβ1,3[Fucaα1,4]GlcNAcβ1,6[Fucaα1,2Galβ1,3]GalNAc-ol		4.809	Coppin 2003
+5	GlcNAcβ1,3[Fucaα1,4]GlcNAcβ1,6[Kdnα2,3Galβ1,3]GalNAc-ol		4.807	Coppin 2003
+6	GlcNAcβ1,3[Fucaα1,4]GlcNAcβ1,6[Kdnα2,3[Galβ1,4]Galβ1,3]GalNAc-ol		4.809	Coppin 2003
+7	GlcNAcβ1,3[Fucaα1,4]GlcNAcβ1,6[Kdnα2,3[GlcNAcα1,4Galβ1,4]Galβ1,3]GalNAc-ol		4.811	Coppin 2003
+8	GlcNAcβ1,3[Fucaα1,4]GlcNAcβ1,6[NeuAcα2,3[Galβ1,4]Galβ1,3]GalNAc-ol		4.809	Coppin 2003
+9	GlcNAcβ1,3[Fucaα1,4]GlcNAcβ1,6[NeuGcα2,3Galβ1,3]GalNAc-ol		4.809	Coppin 2003
+10	GlcNAcβ1,3[Fucaα1,4]GlcNAcβ1,6[NeuGcα2,3[Galβ1,4]Galβ1,3]GalNAc-ol		4.807	Coppin 2003
+11	GlcNAcβ1,3[Fucaα1,4]GlcNAcβ1,6[NeuGcα2,3[GlcNAcα1,4Galβ1,4]Galβ1,3]GalNAc-ol		4.812	Coppin 2003
+12	Galβ1,4GlcNAcβ1,3[Fucaα1,4]GlcNAcβ1,6[Kdnα2,3Galβ1,3]GalNAc-ol		4.79	Coppin 2003
+13	GlcNAcα1,4Galβ1,4GlcNAcβ1,3[Fucaα1,4]GlcNAcβ1,6[GlcNAcα1,4Galβ1,3]GalNAc-ol		4.775	Coppin 2003
+14	GlcNAcα1,4Galβ1,4GlcNAcβ1,3[Fucaα1,4]GlcNAcβ1,6[Kdnα2,3Galβ1,3]GalNAc-ol		4.78	Coppin 2003
	3-Fucosylated LAcDiNAc (LDNF) cont. carbohydrates			
154	GalNAcβ1,4[Fucaα1,3]GlcNAcβ1,2Manα1,3[GalNAcβ1,4GlcNAcβ1,2Manβ1,6]Manβ1,4GlcNAcβ1,4[Fucaα1,6]GlcNAc	LDNF	4.862	13355
155	GalNAcβ1,4[Fucaα1,3]GlcNAcβ1,2Manα1,6[GalNAcβ1,4GlcNAcβ1,2Manβ1,3]Manβ1,4GlcNAcβ1,4[Fucaα1,6]GlcNAc	LDNF	4.862	13356
156	GalNAcβ1,4[Fucaα1,3]GlcNAcβ1,2Manα1,3[Neu5Acα2,3Galβ1,4GlcNAcβ1,2Manβ1,6]Manβ1,4GlcNAcβ1,4[Fucaα1,6]GlcNAc	LDNF	4.861	13760
157	GalNAcβ1,4[Fucaα1,3]GlcNAcβ1,6[Galβ1,3]GalNAc-ol	LDNF	4.85	13039
158	GalNAcβ1,4[Fucaα1,3]GlcNAcβ1,3GalNAc-ol	LDNF	4.841	13884
159	GalNAcβ1,4[Fucaα1,3]GlcNAcβ1,2Manα1,3[GalNAcβ1,4[Fucaα1,3]GlcNAcβ1,2Manβ1,6]Manβ1,4GlcNAcβ1,4[Fucaα1,6]GlcNAc	LDNF	4.853	13757
160	GalNAcβ1,4[Fucaα1,3]GlcNAcβ1,2Manα1,3[GalNAcβ1,4[Fucaα1,3]GlcNAcβ1,2Manβ1,6]Manβ1,4GlcNAcβ1,4[Fucaα1,6]GlcNAc	LDNF	4.853	13757
161	GalNAcα1,2[Fucaα1,3]GalNAcβ1,4]GlcNAcβ1,3GalNAc-ol	iLDNF	4.837	2816
+15	Tyv α1,3GalNAcβ1,4[Fucaα1,3]GlcNAcβOMe	iLDNF	4.82	Zhang 1996
+16	Tyv β1,3GalNAcβ1,4[Fucaα1,3]GlcNAcβOMe	iLDNF	4.84	Zhang 1996
	Le^x-like containing carbohydrates			
162	Galβ1,4[Fucaα1,3]Glcβ	Le ^x like	4.78	27042
163	Neu5Acα2,3Galβ1,3[Neu5Acα2,6]GlcNAcβ1,3Galβ1,4[Fucaα1,3]Glcα	int Le ^x like	4.82	2781
164	Neu5Acα2,3Galβ1,4GlcNAcβ1,3Galβ1,4[Fucaα1,3]Glcα	int Le ^x like	4.819	2737
165	Neu5Acα2,3Galβ1,3[Neu5Acα2,6]GlcNAcβ1,3Galβ1,4[Fucaα1,3]Glcβ	int Le ^x like	4.81	2781
166	Neu5Acα2,3Galβ1,4GlcNAcβ1,3Galβ1,4[Fucaα1,3]Glcβ	int Le ^x like	4.806	2737
167	Neu5Acα2,3Galβ1,4[Fucaα1,3]Glcβ	Sia Le ^x like	4.81	1630
168	Neu5Acα2,3Galβ1,4[Fucaα1,3]Glcα	Sia Le ^x like	4.82	1630
	Le^y-like containing carbohydrates			
169	Galα1,3[Fucaα1,2]Galβ1,4[Fucaα1,3]Glcα	int Le ^y like	4.872	3993
170	Galα1,3[Fucaα1,2]Galβ1,4[Fucaα1,3]Glcβ	int Le ^y like	4.854	3993
171	GalNAcα1,3[Fucaα1,2]Galβ1,4[Fucaα1,3]Glcα	int Le ^y like	4.851	1273
172	GalNAcα1,3[Fucaα1,2]Galβ1,4[Fucaα1,3]Glcβ	int Le ^y like	4.835	1273
+17	Fucaα1,2Galβ1,4[Fucaα1,3]Glcβ	Le ^y like	4.88	Ishizuka 1999
	Le^a-like containing carbohydrates			
+18	3S-Galβ1,3[Fucaα1,4]Glcβ-sp	3S-Le ^a like	4.840	Kurutz 1997

+19	6S-Galβ1,3[Fuca1,4]Glcβ-sp	6S-Le ^a like	4.839	Kurutz 1997
+20	3S,6S-Galβ1,3[Fuca1,4]Glcβ-sp	3S,6S- Le ^a like	4.873	Kurutz 1997
	Le^b-like containing carbohydrates			
173	Galα1,3[Fuca1,2]Galβ1,3[Fuca1,4]Glc	int Le ^b like	4.97	14260
174	GalNAcα1,3[Fuca1,2]Galβ1,3[Fuca1,4]Glc	int Le ^b like	4.969	14259
	4-Fucosylated chitobiose-like containing carbohydrates			
175	Glcα1,2Glcβ1,3GlcNAcβ1,3[Fuca1,4]GlcαP	3FChB like	4.84	27088
	3-Fucosylated chondroitin sulfate			
+21	3S-GlcAβ1,3[4S]GalNAcβ1,4[Fuca1,3]GlcAβ1,3[4S]GalNAc		4.80	Kitagawa 1997
	Miscellaneous carbohydrates			
176	Galα1,3[Fuca1,2]Galβ1,3[Fuca1,4]Glc		4.665	14260
177	GalNAcα1,3[Fuca1,2]Galβ1,3[Fuca1,4]Glc		4.662	14259
178	Manα1,3[Fuca1,2]RhaαOMe		(4.96) ^f	27238
179	Xylα1,6Glcβ1,4[Galβ1,2Xylα1,6]Glcβ1,4[Fuca1,2Galβ1,2Xylα1,6]Glcβ1,4Glc-ol		4.523	3764
180	Xylα1,6Glcβ1,4[Xylα1,6]Glcβ1,4[Fuca1,2Galβ1,3Xylα1,6]Glcβ1,4Glc-ol		4.519	3762

^a Consecutive numbering of the results from the database search. Few additional data found in the literature are indicated by a + sign.

^b The motif with the characteristic Fuc H5 chemical shift is indicated in bold.

^c The following abbreviations are used to indicate the Glyco motif: Le^x – Lewis^x; Int Le^x – internal Lewis^x; Sia Le^x – sialyl Lewis^x; 3S-Le^x – 3-sulfo Lewis^x; 6S-Le^x – 6-sulfo Lewis^x; Int 6S-Le^x – internal 6-sulfo Lewis^x; 6S-Sia Le^x – 6-sulfo sialyl Lewis^x; Le^y – Lewis^y; 6S-Le^y – 6-sulfo Lewis^y; Le^a – Lewis^a; Int Le^a – internal Lewis^a; Sia Le^a – sialyl Lewis^a; 3S-Le^a – 3-sulfo Lewis^a; Le^b – Lewis^b; Int Le^b – internal Lewis^b; 3FChB – α1,3-fucosylated chitobiose; LDNF – fucosylated LacDiNAc; iLDNF – internal fucosylated LacDiNAc;

^d Chemical shift values are given in ppm.

^e Additional data are used, indicated with an +: from five publications (2-6).

^f Probably an artifact of the database, Kochetkov et al. reports an H5 chemical shift of 4.09 ppm for the same compound (number 13) (7).

Table S5: Three-dimensional coordinates of representatives for most structural motifs found in the Protein Databank (PDB) and the Cambridge Structural database (CSD).

PDB or CSD accession code	Chain and residue numbers	Carbohydrate structure	Distance H5–O5	Distance C5–O5	Resolution	Refinement program
		Le^x containing carbohydrates				
ABUCEF	1	Galβ1,4[Fucα1,3]GlcNAcβ-OMe (pure carbohydrate)	2.3	3.3		direct
ABUCEF	2	Galβ1,4[Fucα1,3]GlcNAcβ-OMe (pure carbohydrate)	2.3	3.3		direct
1SL6	G 1-3	Galβ1,4[Fucα1,3]GlcNAcα	2.4 ^a	3.4	2.25	CNS 1.1
1SL6	H 1-3	Galβ1,4[Fucα1,3]GlcNAcα	2.4 ^a	3.4	2.25	CNS 1.1
1SL6	I 1-3	Galβ1,4[Fucα1,3]GlcNAcα	2.9 ^a	3.9	2.25	CNS 1.1
1SL6	J 1-3	Galβ1,4[Fucα1,3]GlcNAcα	2.9 ^a	3.9	2.25	CNS 1.1
1SL6	K 1-3	Galβ1,4[Fucα1,3]GlcNAcα	3.0 ^a	4.00	2.25	CNS 1.1
1SL6	L 1-3	Galβ1,4[Fucα1,3]GlcNAcα	3.0 ^a	3.9	2.25	CNS 1.1
1UZ8	A 1213-1214	Galβ1,4[Fucα1,3]GlcNAcβ-OMe	2.4 ^a	3.5	1.8	REFMAC 5.2
1UZ8	H 1213-1214	Galβ1,4[Fucα1,3]GlcNAcβ-OMe	2.2 ^a	3.3	1.8	REFMAC 5.2
1SL5	C 1-4	Galβ1,4[Fucα1,3]GlcNAcβ1,3Gal	2.8 ^a	3.7	1.7	CNS 1.1
3ZW1	A 105-108	Galβ1,4[Fucα1,3]GlcNAcβ1,3Gal β	4.3 ^{a,b,c}	4.6 ^{b,c}	1.6	REFMAC 5.5
3ZW1	B 99-103	Galβ1,4[Fucα1,3]GlcNAcβ1,3Gal α/β	4.6 ^{a,b,c}	4.9 ^{b,c}	1.6	REFMAC 5.5
2OX9	E 1-3	Galβ1,4[Fucα1,3]GlcNAcβ	2.1 ^a	3.2	1.95	CNS 1.1
2OX9	F 1-3	Galβ1,4[Fucα1,3]GlcNAcβ	2.3 ^a	3.3	1.95	CNS 1.1
2OX9	G 1-3	Galβ1,4[Fucα1,3]GlcNAcβ	2.2 ^a	3.3	1.95	CNS 1.1
2OX9	H 1-3	Galβ1,4[Fucα1,3]GlcNAcβ	2.3 ^a	3.4	1.95	CNS 1.1
3AP9	A 155-160	Galβ1,4[Fucα1,3]GlcNAcβ1,3Gal β1,4Glc	2.8 ^a	3.8	1.33	REFMAC 5.5
3KMB	1 222-224	3S-Galβ1,4[Fucα1,3]GlcNAcβ-OMe	2.2 ^a	3.2	2.0	X-PLOR 3.54
3KMB	3 222-224	3S-Galβ1,4[Fucα1,3]GlcNAcβ-OMe	2.1 ^a	3.1	2.0	X-PLOR 3.54
4KMB	1 222-224	4S-Galβ1,4[Fucα1,3]GlcNAcβ-OMe	2.4 ^a	3.4	2.0	X-PLOR 3.54
4KMB	3 222-224	4S-Galβ1,4[Fucα1,3]GlcNAcβ-OMe	2.4 ^a	3.4	2.0	X-PLOR 3.54
2RDG	A 601-604	Neu5Acα2,3Galβ1,4[Fucα1,3]GlcNAcα	2.3 ^a	3.4	1.6	REFMAC 5.3
3PVD	A 1-4	Neu5Acα2,3Galβ1,4[Fucα1,3]GlcNAcα	2.2 ^a	3.3	1.9	PHENIX
3PVD	B 1-4	Neu5Acα2,3Galβ1,4[Fucα1,3]GlcNAcα	2.3 ^a	3.4	1.9	PHENIX
4DXG	A 301-304	Neu5Acα2,3Galβ1,4[Fucα1,3]GlcNAcα	2.4 ^a	3.4	2.5	REFMAC 5.5
2R61	A 601-604	Neu5Acα2,3Galβ1,4[Fucα1,3]GlcNAcβ	2.3 ^a	3.4	2.75	REFMAC 5.3
2Z8L	A 601-604	Neu5Acα2,3Galβ1,4[Fucα1,3]GlcNAcβ	2.3 ^a	3.4	1.65	REFMAC 5.3
1G1R	A 901-904	Neu5Acα2,3Galβ1,4[Fucα1,3]GlcNAcβ-OMe	2.6 ^a	3.6	3.4	CNS
1G1R	B 901-904	Neu5Acα2,3Galβ1,4[Fucα1,3]GlcNAcβ-OMe	2.0 ^a	3.1	3.4	CNS
1G1R	D 901-904	Neu5Acα2,3Galβ1,4[Fucα1,3]GlcNAcβ-OMe	2.4 ^a	3.5	3.4	CNS
1G1T	A 901-904	Neu5Acα2,3Galβ1,4[Fucα1,3]GlcNAcβ-OMe	2.4 ^a	3.4	1.5	CNS
1KMB	1 222-224	Neu5Acα2,3Galβ1,4[Fucα1,3]GlcNAcβ-OMe	2.3 ^a	3.4	2.1	X-PLOR 3.54
1KMB	2 222-224	Neu5Acα2,3Galβ1,4[Fucα1,3]GlcNAcβ-OMe	2.3 ^a	3.3	2.1	X-PLOR 3.54
1KMB	3 222-224	Neu5Acα2,3Galβ1,4[Fucα1,3]GlcNAcβ-OMe	2.3 ^a	3.3	2.1	X-PLOR 3.54
2KMB	1 222-224	Neu5Acα2,3Galβ1,4[Fucα1,3]GlcNAcβ-OMe	2.3 ^a	3.4	2.0	X-PLOR 3.54
2KMB	2 222-224	Neu5Acα2,3Galβ1,4[Fucα1,3]GlcNAcβ-OMe	2.2 ^a	3.3	2.0	X-PLOR 3.54
2KMB	3 222-224	Neu5Acα2,3Galβ1,4[Fucα1,3]GlcNAcβ-OMe	2.3 ^a	3.3	2.0	X-PLOR 3.54
1G1S	C 629-634	Neu5Acα2,3Galβ1,4[Fucα1,3]GlcNAcβ1,6[Galβ1,3]GalNacα-Thr616	2.4 ^a	3.4	1.9	CNS
1G1S	D 629-634	Neu5Acα2,3Galβ1,4[Fucα1,3]GlcNAcβ1,6[Galβ1,3]GalNacα-Thr616	2.3 ^a	3.4	1.9	CNS
		Le^y containing carbohydrates				

1GSL	A 252-255	Fuc α 1,2Gal β 1,4[Fuc α 1,3]GlcNAc β -OMe	2.5 ^a	3.5	2.00	X-PLOR
1S3K	H 223-226	Fuc α 1,2Gal β 1,4[Fuc α 1,3]GlcNAc α	2.4 ^a	3.5	1.9	CNS 1.0
2J1T	A 1150-1153	Fuc α 1,2Gal β 1,4[Fuc α 1,3]GlcNAc β	2.5 ^a	3.5	1.6	REFMAC 5.1
2WMG	A 1590-1593	Fuc α 1,2Gal β 1,4[Fuc α 1,3]GlcNAc β	5.4 ^{a,b,d}	6.2 ^{b,d}	2.3	REFMAC 5.2
2WMK	A 2006-2010	GalNAc α 1,3[Fuc α 1,2]Gal β 1,4[Fuc α 1,3]GlcNAc β	5.0 ^{a,b,d}	5.8 ^{b,d}	1.9	REFMAC 5.2
2WMK	B 2006-2010	GalNAc α 1,3[Fuc α 1,2]Gal β 1,4[Fuc α 1,3]GlcNAc β	5.0 ^{a,b,d}	5.8 ^{b,d}	1.9	REFMAC 5.2
3EYV	H 223-226	Fuc α 1,2Gal β 1,4[Fuc α 1,3]GlcNAc α	2.3 ^a	3.4	2.5	CNS 1.0
3LEG	A 1186	Fuc α 1,2Gal β 1,4[Fuc α 1,3]GlcNAc β	2.8 ^a	3.8	2.0	REFMAC 5.5
3PA2	A 1-4	Fuc α 1,2Gal β 1,4[Fuc α 1,3]GlcNAc β	2.7 ^a	3.7	1.48	PHENIX
3PUN	A 1-4	Fuc α 1,2Gal β 1,4[Fuc α 1,3]GlcNAc α	2.4 ^a	3.5	2.05	PHENIX
3PUN	B 1-4	Fuc α 1,2Gal β 1,4[Fuc α 1,3]GlcNAc α	2.2 ^a	3.3	2.05	PHENIX
4AH5	A 930-933	Fuc α 1,2Gal β 1,4[Fuc α 1,3]GlcNAc α	4.3 ^{a,b,e}	4.3 ^{b,e}	1.99	REFMAC5
4AH5	A 960-963	Fuc α 1,2Gal β 1,4[Fuc α 1,3]GlcNAc α	2.5 ^a	3.5	1.99	REFMAC5
4AH5	B 930-933	Fuc α 1,2Gal β 1,4[Fuc α 1,3]GlcNAc α	4.3 ^{a,b,e}	4.3 ^{b,e}	1.99	REFMAC5
4GWI	A 205	Fuc α 1,2Gal β 1,4[Fuc α 1,3]GlcNAc β	2.5 ^a	3.5	1.6	REFMAC5
1CLY	H 228-231	Fuc α 1,2Gal β 1,4[Fuc α 1,3]GlcNAc β -sp	2.3 ^a	3.4	2.51	X-PLOR
1CLZ	H 232-235	Fuc α 1,2Gal β 1,4[Fuc α 1,3]GlcNAc β -sp	2.4 ^a	3.5	2.78	X-PLOR
		Le^a containing carbohydrates				
1FWV	A 2001-2003	3S-Gal β 1,3[Fuc α 1,4]GlcNAc β	2.6 ^a	3.6	2.2	CNS
1W8H	A 331-333	Gal β 1,3[Fuc α 1,4]GlcNAc α / β	2.4 ^a	3.5	1.75	REFMAC 5.1
1W8H	B 331-333	Gal β 1,3[Fuc α 1,4]GlcNAc α / β	2.4 ^a	3.5	1.75	REFMAC 5.1
1W8H	C 331-333	Gal β 1,3[Fuc α 1,4]GlcNAc α / β	2.3 ^a	3.4	1.75	REFMAC 5.1
1W8H	D 331-333	Gal β 1,3[Fuc α 1,4]GlcNAc β	2.3 ^a	3.4	1.75	REFMAC 5.1
3ASR	A 1002-1004	Gal β 1,3[Fuc α 1,4]GlcNAc β -O-p-nitrophenol	2.3 ^a	3.3	1.6	COOT, REFMAC5
3ASR	B 1002-1004	Gal β 1,3[Fuc α 1,4]GlcNAc β -O-p-nitrophenol	2.4 ^a	3.4	1.6	COOT, REFMAC5
3UET	A 502-504	Gal β 1,3[Fuc α 1,4]GlcNAc β	4.4 ^{a,b,f}	4.4 ^{b,f}	2.1	COOT, REFMAC5
3UET	B 502-504	Gal β 1,3[Fuc α 1,4]GlcNAc β	4.2 ^{a,b,f}	4.2 ^{b,f}	2.1	COOT, REFMAC5
		Le^b containing carbohydrates				
1LED	A 252-255	Fuc α 1,2Gal β 1,3[Fuc α 1,4]GlcNAc β -OMe	2.6 ^a	3.6	2.0	X-PLOR, PROLSQ
3ASS	A 1001-1005	Fuc α 1,2Gal β 1,3[Fuc α 1,4]GlcNAc β -O-p-nitrophenol	2.5 ^a	3.5	1.6	COOT, REFMAC5
3ASS	B 1001-1005	Fuc α 1,2Gal β 1,3[Fuc α 1,4]GlcNAc β -O-p-nitrophenol	2.5 ^a	3.5	1.6	COOT, REFMAC5
3AST	A 1001-1005	Fuc α 1,2Gal β 1,3[Fuc α 1,4]GlcNAc β -O-p-nitrophenol	2.5 ^a	3.5	1.4	COOT, REFMAC5
3AST	B 1001-1005	Fuc α 1,2Gal β 1,3[Fuc α 1,4]GlcNAc β -O-p-nitrophenol	2.5 ^a	3.5	1.4	COOT, REFMAC5
3LEK	A1185	Fuc α 1,2Gal β 1,3[Fuc α 1,4]GlcNAc α	2.7 ^a	3.7	2.2	COOT, REFMAC5
3SEJ	B 1-6	Fuc α 1,2Gal β 1,3[Fuc α 1,4]GlcNAc β 1,3Gal β 1,4Glc	(1.9) ^{a,g}	2.9 ^g	3.04	REFMAC
3SEJ	C 1-6	Fuc α 1,2Gal β 1,3[Fuc α 1,4]GlcNAc β 1,3Gal β 1,4Glc	(2.5) ^{a,g}	3.4 ^g	3.04	REFMAC
3SEJ	C 532-537	Fuc α 1,2Gal β 1,3[Fuc α 1,4]GlcNAc β 1,3Gal β 1,4Glc	(1.8) ^{a,g}	2.9 ^g	3.04	REFMAC
3SEJ	D 1-6	Fuc α 1,2Gal β 1,3[Fuc α 1,4]GlcNAc β 1,3Gal β 1,4Glc	(1.8) ^{a,g}	2.8 ^g	3.04	REFMAC
3SEJ	E 1-6	Fuc α 1,2Gal β 1,3[Fuc α 1,4]GlcNAc β 1,3Gal β 1,4Glc	(2.2) ^{a,g}	3.2 ^g	3.04	REFMAC
3SEJ	F 1-6	Fuc α 1,2Gal β 1,3[Fuc α 1,4]GlcNAc β 1,3Gal β 1,4Glc	(2.2) ^{a,g}	3.1 ^g	3.04	REFMAC
3SEJ	G 1-6	Fuc α 1,2Gal β 1,3[Fuc α 1,4]GlcNAc β 1,3Gal β 1,4Glc	(1.8) ^{a,g}	2.7 ^g	3.04	REFMAC
3SEJ	J 1-6	Fuc α 1,2Gal β 1,3[Fuc α 1,4]GlcNAc β 1,3Gal β 1,4Glc	(2.0) ^{a,g}	2.9 ^g	3.04	REFMAC
4GWJ	A 205-208	Fuc α 1,2Gal β 1,3[Fuc α 1,4]GlcNAc β	2.5 ^a	3.6	1.6	REFMAC5

3-Fucosylated chitobiose containing carbohydrates						
1LK9	A 500-503	Man β 1,4GlcNAc β 1,4[Fuca1,3]GlcNAc β -Asn146	2.2 ^a	3.3	1.53	REFMAC 4.0.6
1JU2	B 544-547	Man β 1,4GlcNAc β 1,4[Fuca1,3]GlcNAc β -Asn392	2.3 ^a	3.4	1.45	CNS 1.1
1JU2	A 539-543	Man α 1,6Man β 1,4GlcNAc β 1,4[Fuca1,3]GlcNAc β -Asn135	2.2 ^a	3.3	1.45	CNS 1.1
1JU2	B 539-543	Man α 1,6Man β 1,4GlcNAc β 1,4[Fuca1,3]GlcNAc β -Asn135	2.3 ^a	3.4	1.45	CNS 1.1
1JU2	A 544-548	Man β 1,3Man β 1,4GlcNAc β 1,4[Fuca1,3]GlcNAc β -Asn392	2.4 ^a	3.4	1.45	CNS 1.1
1E4M	M 941-945	Xyl β 1,2Man β 1,4GlcNAc β 1,4[Fuca1,3]GlcNAc β -Asn265	2.5 ^a	3.5	1.2	REFMAC
1E4M	M 951-957	Man α 1,3[Man α 1,6][Xyl β 1,2]Man β 1,4GlcNAc β 1,4[Fuca1,3]GlcNAc β -Asn292	2.5 ^a	3.5	1.2	REFMAC
1E6Q	M 941-945	Xyl β 1,2Man β 1,4GlcNAc β 1,4[Fuca1,3]GlcNAc β -Asn265	2.5 ^a	3.5	1.35	REFMAC
1E6Q	M 951-957	Man α 1,3[Man α 1,6][Xyl β 1,2]Man β 1,4GlcNAc β 1,4[Fuca1,3]GlcNAc β -Asn292	2.7 ^a	3.7	1.35	REFMAC
1E6S	M 941-945	Xyl β 1,2Man β 1,4GlcNAc β 1,4[Fuca1,3]GlcNAc β -Asn265	2.5 ^a	3.5	1.35	REFMAC
1E6S	M 951-957	Man α 1,3[Man α 1,6][Xyl β 1,2]Man β 1,4GlcNAc β 1,4[Fuca1,3]GlcNAc β -Asn292	2.6 ^a	3.6	1.35	REFMAC
1YM0	A 1-5	Man β 1,3Man β 1,4GlcNAc β 1,4[Fuca1,3][Fuca1,6]GlcNAc β -Asn161	2.4 ^a	3.5	2.06	CNS
2B9L	A 416-419	GlcNAc β 1,4[Fuca1,3][Fuca1,6]GlcNAc β -Asn32	2.6 ^a	3.6	2.0	CNS 1.1
2F9N	B 1000-1003	GlcNAc β 1,4[Fuca1,3][Fuca1,6]GlcNAc β -Asn113	2.5 ^a	3.5	1.6	CNS 1.1
2F9N	D 1000-1003	GlcNAc β 1,4[Fuca1,3][Fuca1,6]GlcNAc β -Asn113	2.4 ^a	3.5	1.6	CNS 1.1
2QQM	D 1-4	GlcNAc β 1,4[Fuca1,3][Fuca1,6]GlcNAc β -Asn261	2.6 ^a	3.6	2.0	REFMAC 5
3L9R	A 501-507	Man β 1,3Man β 1,4GlcNAc β 1,4[Fuca1,3][Fuca1,6]GlcNAc β -Asn57	2.3 ^a	3.4	2.3	REFMAC 5.5
3L9R	C 501-507	Man β 1,4GlcNAc β 1,4[Fuca1,3][Fuca1,6]GlcNAc β -Asn57	2.5 ^a	3.6	2.3	REFMAC 5.5
3L9R	E 501-507	Man β 1,4GlcNAc β 1,4[Fuca1,3][Fuca1,6]GlcNAc β -Asn57	2.3 ^a	3.4	2.3	REFMAC 5.5
3L9R	G 501-507	Man β 1,4GlcNAc β 1,4[Fuca1,3][Fuca1,6]GlcNAc β -Asn57	2.4 ^a	3.5	2.3	REFMAC 5.5
3QW9	A 187-191	Man α 1,3[Man α 1,6]Man β 1,4GlcNAc β 1,4[Fuca1,3][Fuca1,6]GlcNAc β -Asn14(B)	2.3 ^a	3.3	1.84	REFMAC 5
3QW9	A 200-204	Man β 1,4GlcNAc β 1,4[Fuca1,3][Fuca1,6]GlcNAc β -Asn14	2.4 ^a	3.4	1.84	REFMAC 5
3U0P	A 301-304	GlcNAc β 1,4[Fuca1,3][Fuca1,6]GlcNAc β -Asn20	5.6 ^{a,b,h}	5.9 ^{b,h}	2.8	PHENIX
4ARN	C 501-507	Man α 1,3[Man α 1,6]Man β 1,4GlcNAc β 1,4[Fuca1,3][Fuca1,6]GlcNAc β -Asn140	2.4 ^a	3.5	2.41	BUSTER
4GWM	A 706-712	Man α 1,3[Man α 1,6]Man β 1,4GlcNAc β 1,4[Fuca1,3][Fuca1,6]GlcNAc β -Asn370	2.2 ^a	3.3	1.85	BUSTER
4GWM	B 705-710	Man α 1,3Man β 1,4GlcNAc β 1,4[Fuca1,3][Fuca1,6]GlcNAc β -Asn370	2.2 ^a	3.3	1.85	BUSTER
4GWM	B 711-714	GlcNAc β 1,4[Fuca1,3][Fuca1,6]GlcNAc β -Asn445	2.3 ^a	3.4	1.85	BUSTER
4GWN	A 705-711	Man α 1,3[Man α 1,6]Man β 1,4GlcNAc β 1,4[Fuca1,3][Fuca1,6]GlcNAc β -Asn370	2.6 ^a	3.6	3.0	BUSTER
4GZT	D 508-511	GlcNAc β 1,4[Fuca1,3][Fuca1,6]GlcNAc β -Asn146	2.2 ^a	3.3	2.19	REFMAC 5.5
Le ^x -like containing carbohydrates						
1W8F	A 1118-1121	GalNAc α 1,3Gal β 1,4[Fuca1,3]Glc β	2.3 ^a	3.4	1.05	REFMAC 5.1
1W8F	B 1118-1122	Gal β 1,4GalNAc α 1,3Gal β 1,4[Fuca1,3]Glc β	2.4 ^a	3.5	1.05	REFMAC 5.1
1W8F	C 1118-1120	Gal β 1,4[Fuca1,3]Glc β	2.4 ^a	3.5	1.05	REFMAC 5.1
1W8F	D 1118-1122	Gal β 1,4GalNAc α 1,3Gal β 1,4[Fuca1,3]Glc β	2.4 ^a	3.4	1.05	REFMAC 5.1
Le ^y -like containing carbohydrates						
2O2L	N 201-205	GalNAc α 1,3[Fuca1,2]Gal β 1,4[Fuca1,3]Glc β	2.3 ^a	3.4	2.53	REFMAC
2O2L	O 201-205	GalNAc α 1,3[Fuca1,2]Gal β 1,4[Fuca1,3]Glc β	2.4 ^a	3.5	2.53	REFMAC
2O2L	P 201-205	GalNAc α 1,3[Fuca1,2]Gal β 1,4[Fuca1,3]Glc β	2.6 ^a	3.6	2.53	REFMAC
2O2L	Q 201-205	GalNAc α 1,3[Fuca1,2]Gal β 1,4[Fuca1,3]Glc β	2.8 ^a	3.7	2.53	REFMAC

2O2L	R 201-205	GalNAc α 1,3[Fuc α 1,2]Gal β 1,4[Fuc α 1,3]Glc β	2.6 ^a	3.6	2.53	REFMAC
2O2L	S 201-205	GalNAc α 1,3[Fuc α 1,2]Gal β 1,4[Fuc α 1,3]Glc β	4.3 ^{a,b,i}	5.1 ^{b,i}	2.53	REFMAC
2O2L	T 201-205	GalNAc α 1,3[Fuc α 1,2]Gal β 1,4[Fuc α 1,3]Glc β	2.6 ^a	3.6	2.53	REFMAC
2O2L	U 201-205	GalNAc α 1,3[Fuc α 1,2]Gal β 1,4[Fuc α 1,3]Glc β	2.5 ^a	3.5	2.53	REFMAC
2O2L	V 201-205	GalNAc α 1,3[Fuc α 1,2]Gal β 1,4[Fuc α 1,3]Glc β	2.4 ^a	3.5	2.53	REFMAC
2O2L	W 201-205	GalNAc α 1,3[Fuc α 1,2]Gal β 1,4[Fuc α 1,3]Glc β	2.6 ^a	3.7	2.53	REFMAC
3EFX	D 201-205	GalNAc α 1,3[Fuc α 1,2]Gal β 1,4[Fuc α 1,3]Glc β	2.6 ^a	3.6	1.94	REFMAC
3EFX	E 201-205	GalNAc α 1,3[Fuc α 1,2]Gal β 1,4[Fuc α 1,3]Glc β	2.2 ^a	3.3	1.94	REFMAC
3EFX	F 201-205	GalNAc α 1,3[Fuc α 1,2]Gal β 1,4[Fuc α 1,3]Glc β	2.3 ^a	3.4	1.94	REFMAC
3EFX	G 201-205	GalNAc α 1,3[Fuc α 1,2]Gal β 1,4[Fuc α 1,3]Glc β	2.5 ^a	3.6	1.94	REFMAC
3EFX	H 201-205	GalNAc α 1,3[Fuc α 1,2]Gal β 1,4[Fuc α 1,3]Glc β	2.3 ^a	3.3	1.94	REFMAC
3EFX	I 201-205	GalNAc α 1,3[Fuc α 1,2]Gal β 1,4[Fuc α 1,3]Glc β	2.5 ^a	3.5	1.94	REFMAC
3EFX	J 201-205	GalNAc α 1,3[Fuc α 1,2]Gal β 1,4[Fuc α 1,3]Glc β	2.5 ^a	3.4	1.94	REFMAC
3EFX	K 201-205	GalNAc α 1,3[Fuc α 1,2]Gal β 1,4[Fuc α 1,3]Glc β	2.4 ^a	3.4	1.94	REFMAC
3EFX	L 201-205	GalNAc α 1,3[Fuc α 1,2]Gal β 1,4[Fuc α 1,3]Glc β	2.6 ^a	3.6	1.94	REFMAC
3EFX	M 201-205	GalNAc α 1,3[Fuc α 1,2]Gal β 1,4[Fuc α 1,3]Glc β	2.7 ^a	3.7	1.94	REFMAC

- ^a protons were added to the structures by Maestro (Schrödinger) because the crystal structures lacked protons; distances are in brackets when the heavy atoms are characterized by high B-factors
- ^b flipped away, value not used for average distance and histogram
- ^c the electron density is of good quality and the resolution is good; in this case the interactions with the protein seem to be so favorable that Le^x is distorted into an elongated conformation; in one of the structures the Fuc α 1,3GlcNAc linkage angles are significantly different and in the other structure the GlcNAc pucker is distorted
- ^d the two structures 2WMG and 2WMK of glycoside hydrolases (GH98, inactivated enzymes by a point mutant) show very unusual Gal β 1,4GlcNAc phi angles, however in that case this might be linked with the function of the WT enzymes – cleaving the very same Gal β 1,4GlcNAc linkage.
- ^e the electron density seems to be of good quality, the resolution is with 2 Å moderate; in this case the interactions with the protein seem to be so favorable that Le^y is distorted into an elongated conformation
- ^f the inactive double mutant D172A/E217A of the 1,3 -1,4- α -L-fucosynthase crystallized with Gal β 1,3[Fuc1,4]GlcNAc β 1,3Gal β 1,4Glc (LNFP II) showed only electron density of three moieties of which Fuc and Gal could be clearly assigned but GlcNAc was ‘partially ambiguous’ as stated in the corresponding publication (8); the fucosynthase is based on a natural fucosidase and in this context it might be plausible that the Fuc1,4GlcNAc linkage gets distorted by the enzyme to facilitate cleavage
- ^g high B-factors (74-122) compared to an average B-factor of 38 for the protein; values not used for the average distance
- ^h high B-factors (96-163) compared to an average B-factor of 68 for the protein; sugar puckers are very distorted including two boat conformations; axial O2,O3 and O4 in Fuc302; axial O3,O4 and acetamide at C2 in GlcNAc303; phi-psi angles in disallowed regions; value not used for average distance and histogram
- ⁱ sugar chain S in the binding site of protein chain I does show only very weak intensity at the fucose position; according to the associated publication (9) this site is characterized by weak electron density for Fuc α 1,3 and Glc β ; the occupancy in the pdb file is only 50%; all other nine binding sites that are equivalent in the two pentamers of the asymmetric unit show the expected stacked structure; value not used for average distance and histogram

Table S6. Intra- and inter-residual NOEs of 3FChB at 277 K and 900 MHz and their corresponding distances.

proton pair	average S/N of NOEs cross peaks	corresponding ^1H - ^1H distance ^b [Å]
intra		
GlcNAc2 H1-H2	489 ^a	2.79
GlcNAc2 H1-H3	553	2.73
GlcNAc2 H1-H5	1151.5	2.41
GlcNAc2 H2-H4	398.5	2.88
GlcNAc2 H2-Q8	100 ^a	4.36 ^e
GlcNAc2 H3-H5	499 ^a	2.78
GlcNAc2 H4-H62	409	2.87
GlcNAc2 H4-H61	255	3.11
GlcNAc2 H5-H62	462	2.81
GlcNAc2 H5-H61	684.5	2.63
(GlcNAc2 H61-H62, D ₂ O) ^c	7424.5	(1.77)
(GlcNAc2 H61-H62, H ₂ O) ^d	3303.5	(1.77)
GlcNAc2 HN2-H1	150.5	2.96
GlcNAc2 HN2-H2	219	2.78
GlcNAc2 HN2-H3	373	2.55
GlcNAc2 HN2-Q8	490	2.92 ^e
GlcNAc1 H1-H3	421.5	2.86
GlcNAc1 H1-H5	1256	2.38
GlcNAc1 H3-H5	496.5	2.78
GlcNAc1 H5-H61	750 ^a	2.59
GlcNAc1 H5-H62	762 ^a	2.59
GlcNAc1 HN2-H1	157	2.94
GlcNAc1 HN2-H2	108 ^a	3.13
GlcNAc1 HN2-H3	257.5	2.71
GlcNAc1 HN2-Q8	559.5	2.86 ^e
Fuc3 H1-H2	636	2.67
Fuc3 H1-H5	24 ^a	4.60
Fuc3 H1-Q6	117 ^a	4.25 ^e
Fuc3 H3-H4	374 ^a	2.91
Fuc3 H3-H5	655.5	2.65
(Fuc3 H3-H5, H ₂ O)	244 ^a	(2.73)
Fuc3 H4-H5	1354	2.35
Fuc3 H4-Q6	940 ^a	3.00 ^e
Fuc3 H5-Q6	3583.5	2.40 ^e
inter		
GlcNAc2 H1 - GlcNAc1 H4	1029 ^a	2.46
GlcNAc2 H1 - GlcNAc1 H62	444 ^a	2.83
GlcNAc2 HN2 - GlcNAc1 H62	140	3.00
GlcNAc2 H2 - Fuc3 H5	475 ^a	2.80
GlcNAc2 H2 - Fuc3 Q6	545 ^a	3.29 ^e
GlcNAc2 H62 - Fuc3 H4	204.5	3.22
GlcNAc2 H4 - Fuc3 H5	102.5	3.61
GlcNAc1 H3 - Fuc3 H1	406 ^a	2.87
GlcNAc1 H2 - Fuc3 H1	104 ^a	3.61 ^e
GlcNAc1 HN2 - Fuc3 H1	347 ^a	2.58
GlcNAc1 Q8 - Fuc3 H1	237 ^a	3.77 ^e

^a Only one cross-peak was used because of artifacts or severe spectral overlap.

^b The ^1H - ^1H distances were calculated from experimentally obtained NOE intensities using the H3-H5 cross-peak of Fuc3 as a reference with a distance of 2.65 Å assuming a $1/r^6$ dependence of the NOE intensities. For

the structure calculations the herein reported distances were increased by 0.2 Å tolerance and used as upper limit restraints.

^c Reference restraint for the NOESY measured in D₂O.

^d Reference restraint for the NOESY measured in H₂O.

^e Signal to noise ratios from cross-peaks involving methyl or methylene protons were divided by 3 and 2 for the calculation, respectively.

Table S7: Intra- and inter-residual NOEs of Le^x β-methyl glycoside at 277 K and 900 MHz and their corresponding distances.

proton pair	average S/N of NOEs cross peaks	corresponding ¹ H- ¹ H distance ^b [Å]
intra		
Gal2 H1-H2	345 ^a	2.76
Gal2 H1-H3	433 ^a	2.65
Gal2 H1-H5	824	2.38
Gal2 H3-H4	616 ^a	2.50
Gal2 H4-H5	992	2.31
Gal2 H4-Q6	375 ^a	3.05 ^e
GlcNAc1 H1-H3	240 ^a	2.93
GlcNAc1 H1-H5	837	2.38
GlcNAc1 H1-Q1	345	3.31
GlcNAc1 H5-H62	312 ^a	2.80
GlcNAc1 H5-H61	456 ^a	2.63
GlcNAc1 HN2-H1	87	2.84
GlcNAc1 HN2-H2	42 ^a	3.20
GlcNAc1 HN2-H3	114 ^a	2.71
GlcNAc1 HN2-Q8	203 ^a	2.96 ^e
Fuc3 H1-H2	403	2.69
Fuc3 H3-H5 ^c	437.5	2.65
(Fuc3 H3-H5, H2O) ^d	131 ^a	(2.65)
Fuc3 H4-H5	823	2.39
Fuc3 H4-Q6	472 ^a	3.14 ^e
Fuc3 H5-Q6	1580 ^a	2.57 ^{e,f}
inter		
Gal2 H1 - GlcNAc1 H4	744 ^a	2.43
Gal2 H1 - GlcNAc1 H61	416 ^a	2.67
Gal2 H1 - GlcNAc1 H62	312 ^a	2.80
Gal2 H2 - Fuc3 H5	323 ^a	2.79
Gal2 H2 - Fuc3 Q6	288 ^a	3.41 ^e
Gal2 Q6 - Fuc3 H3	381 ^a	3.04 ^e
GlcNAc1 H3 - Fuc3 H1	258 ^a	2.89
GlcNAc1 HN2 - Fuc3 H1	143 ^a	2.61
GlcNAc1 Q8 - Fuc3 H1	90 ^a	4.14 ^e

^a Only one cross-peak was used because of artifacts or severe spectral overlap.

^b The ¹H-¹H distances were calculated from experimentally obtained NOE intensities using the H3-H5 cross-peak of Fuc3 as a reference with a distance of 2.65 Å assuming a 1/r⁶ dependence of the NOE intensities. For the structure calculations the herein reported distances were increased by 0.2 Å tolerance and used as upper limit restraints.

^c Reference restraint for the NOESY measured in D₂O.

^d Reference restraint for the NOESY measured in H₂O.

^e Signal to noise ratios from cross-peaks involving methyl or methylene protons were divided by 3 and 2 for the calculation, respectively.

^f Because of overlap and close artifacts, the distance is not reliable and we used 6 Å as a conservative restraint.

Table S8: Intra- and inter-residual NOEs of Le^a β-methyl glycoside at 275 K and 900 MHz and their corresponding distances.

proton pair	average S/N of NOEs cross peaks	corresponding ¹ H- ¹ H distance ^b [Å]
intra		
Gal2 H1-H3	194.5	2.78
Gal2 H1-H5	394.5	2.47
Gal2 H3-H4	286 ^a	2.60
Gal2 H4-H5	369 ^a	2.49
Gal2 H4-Q6	556 ^a	2.62 ^e
Gal2 H5-Q6	1011 ^a	2.37 ^e
GlcNAc1 H1-H2	67.5	3.31
GlcNAc1 H1-H3	131	2.96
GlcNAc1 H1-H5	360.5	2.50
GlcNAc1 H2-H3	394	2.47 ^f
GlcNAc1 H3-H5	190	2.79
GlcNAc1 H5-H62	275 ^a	2.62 ^f
GlcNAc1 H5-H61	260 ^a	2.64
GlcNAc1 HN2-H1	283 ^a	3.04
GlcNAc1 HN2-H3	533.5	2.74
GlcNAc1 HN2-Q8	962 ^a	2.98 ^e
Fuc3 H1-H2	367.5	2.50
Fuc3 H3-H5 ^c	257	2.65
(Fuc3 H3-H5, H2O) ^d	645 ^a	(2.65)
Fuc3 H4-H5	419.5	2.44
Fuc3 H4-Q6	229 ^a	2.70
inter		
Gal2 H1 - GlcNAc1 H3	248.5	2.66
Gal2 H1 - GlcNAc1 HN2	440 ^a	3.50
Gal2 H2 - Fuc3 H5	212.5	2.74
Gal2 H2 - Fuc3 Q6	225 ^a	3.25 ^e
Gal2 Q6 - Fuc3 H3	427	2.73 ^e
Gal2 Q6 - Fuc3 H5	70	3.69 ^e
GlcNAc1 H4 - Fuc3 H1	335.5	2.53
GlcNAc1 H62 - Fuc3 H1	286 ^a	2.60
GlcNAc1 H61 - Fuc3 H1	134 ^a	2.95

^a Only one cross-peak was used because of artifacts or severe spectral overlap.

^b The ¹H-¹H distances were calculated from experimentally obtained NOE intensities using the H3-H5 cross-peak of Fuc3 as a reference with a distance of 2.65 Å assuming a 1/r⁶ dependence of the NOE intensities. For the structure calculations the herein reported distances were increased by 0.2 Å tolerance and used as upper limit restraints.

^c Reference restraint for the NOESY measured in D₂O.

^d Reference restraint for the NOESY measured in H₂O.

^e Signal to noise ratios from cross-peaks involving methyl or methylene protons were divided by 3 and 2 for the calculation, respectively.

^f Because of overlap and close artifacts, the distance is not reliable and we used 6 Å as a conservative restraint.

Table S9. Intra- and inter-residual NOEs of LDNF at 275 K and 900 MHz and their corresponding distances.

proton pair	average S/N of NOEs cross peaks	corresponding ^1H - ^1H distance ^b [Å]
intra		
GalNAc H1-H3	880	2.68
GalNAc H1-H5	1977	2.34
GalNAc HN2-H1	270	2.78
GalNAc HN2-H3	447	2.56
GalNAc HN2-H5	35	3.91
GalNAc HN2-Q8	530 ^a	2.49
GlcNAc H1-H3	416.5	3.04
GlcNAc H1-H5	1490.5	2.46
GlcNAc H3-H5	533	2.92
GlcNAc H5-H61	770 ^a	2.74
GlcNAc H5-H62	652 ^a	2.84
GlcNAc HN2-H1	192.5	2.94
GlcNAc HN2-H3	284.5	2.76
GlcNAc HN2-H5	28	4.06
GlcNAc HN2-Q8	618 ^a	2.91 ^e
Fuc H1-H2	965.5	2.64
Fuc H1-H5	150	3.60
Fuc H3-H4	1855	2.37
Fuc H3-H5 ^c	948	2.65
(Fuc H3-H5, H2O) ^d	361 ^a	(2.65)
Fuc H4-H5	1877	2.36
Fuc H4-Q6	1098 ^a	3.11 ^e
inter		
GalNAc H1 - GlcNAc H4	1579	2.43
GalNAc H1 - GlcNAc H62	1326 ^a	2.51
GalNAc H1 - GlcNAc H61	792 ^a	2.73
GalNAc Q8 - GlcNAc H5	69 ^a	4.93 ^e
GalNAc Q8 - GlcNAc H62	115 ^a	4.52 ^e
GalNAc Q8 - GlcNAc H61	113 ^a	4.54 ^e
GalNAc HN2 - GlcNAc H5	26.5	4.10
GalNAc HN2 - GlcNAc H62	213.5	2.89
GalNAc HN2 - GlcNAc H61	178.5	2.98
GalNAc H2 - Fuc H5	796.5	2.73
GalNAc H61 - Fuc H5	78.5	4.01
GalNAc H62 - Fuc H3	771	2.74
GalNAc H62 - Fuc H5	134.5	3.67
GalNAc H2 - Fuc Q6	734	3.32 ^e
GalNAc HN2 - Fuc Q6	18 ^a	5.25 ^e
GlcNAc H3 - Fuc H1	721	2.77
GlcNAc H3 - Fuc Q6	119 ^a	4.50 ^e
GlcNAc HN2 - Fuc H1	543 ^a	2.48
GlcNAc HN2 - Fuc H2	22	4.22
GlcNAc Q8 - Fuc H1	358.5	3.74 ^e
GlcNAc Q8 - Fuc H2	63 ^a	5.00 ^e

^a Only one cross-peak was used because of artifacts or severe spectral overlap.

^b The ^1H - ^1H distances were calculated from experimentally obtained NOE intensities using the H3-H5 cross-peak of Fuc as a reference with a distance of 2.65 Å assuming a $1/r^6$ dependence of the NOE intensities. For the structure calculations the herein reported distances were increased by 0.2 Å tolerance and used as upper limit restraints.

^c Reference restraint for the NOESY measured in D_2O .

^d Reference restraint for the NOESY measured in H_2O .

^e Signal to noise ratios from cross-peaks involving methyl or methylene protons were divided by 3 and 2 for the calculation, respectively.

Table S10. Intra- and inter-residual NOEs of Bv9 at 273 K and 900 MHz and their corresponding distances.

proton pair	average S/N of NOEs cross peaks	corresponding ^1H - ^1H distance ^b [Å]
intra		
GlcNAc2 H1-H3	174.5	2.83
GlcNAc2 H1-H5	393.5	2.47
GlcNAc2 HN2-H1	123 ^a	2.93
GlcNAc2 HN2-H2	34 ^a	3.63
GlcNAc2 HN2-H3	264 ^a	2.58
GlcNAc2 HN2-H5	15 ^a	4.16
GlcNAc1 H1-H3	166.5	2.86
GlcNAc1 H3-H5	250 ^a	2.67
GlcNAc1 H5-H62	256 ^a	2.66
GlcNAc1 HN2-H1	189 ^a	2.72
GlcNAc1 HN2-H3	159 ^a	2.80
Fuc3 H1-H2	525.5	2.36
Fuc3 H3-H4	661.5	2.27
Fuc3 H3-H5 ^c	260.5	2.65
(Fuc3 H3-H5, H2O) ^d	223 ^a	(2.65)
Fuc3 H4-Q6	401	2.96 ^e
inter		
GlcNAc2 H1 - GlcNAc1 H3	254	2.66
GlcNAc2 H1 - GlcNAc1 HN2	276 ^a	2.56
GlcNAc2 H1 - GlcNAc1 Q8	57 ^a	4.10 ^e
GlcNAc2 HN2 - GlcNAc1 HN2	15.5	4.13
GlcNAc2 HN2 - GlcNAc1 Q8	77 ^a	3.80 ^e
GlcNAc2 H2 - Fuc3 H5	190.5	2.79
GlcNAc2 H4 - Fuc3 H5	40	3.62
GlcNAc2 H62 - Fuc3 H3	161 ^a	2.87
GlcNAc2 H62 - Fuc3 H5	34.5	3.71
GlcNAc2 H2 - Fuc3 Q6	255	3.19 ^e
GlcNAc2 H4 - Fuc3 Q6	69 ^a	3.97 ^e
GlcNAc2 HN2 - Fuc3 Q6	17 ^a	4.89 ^e
GlcNAc1 H4 - Fuc3 H1	425	2.44
GlcNAc1 H61 - Fuc3 H1	377 ^a	2.49
GlcNAc1 H62 - Fuc3 H1	699 ^a	2.25
GlcNAc1 H4 - Fuc3 H5	61.5	3.37
GlcNAc1 H4 - Fuc3 Q6	47 ^a	4.23 ^e

^a Only one cross-peak was used because of artifacts or severe spectral overlap.

^b The ^1H - ^1H distances were calculated from experimentally obtained NOE intensities using the H3-H5 cross-peak of Fuc3 as a reference with a distance of 2.65 Å assuming a $1/r^6$ dependence of the NOE intensities. For the structure calculations the herein reported distances were increased by 0.2 Å tolerance and used as upper limit restraints.

^c Reference restraint for the NOESY measured in D₂O.

^d Reference restraint for the NOESY measured in H₂O.

^e Signal to noise ratios from cross-peaks involving methyl or methylene protons were divided by 3 and 2 for the calculation, respectively.

Table S11. Intra- and inter-residual NOEs of LDFT at 275 K and 900 MHz and their corresponding distances.

proton pair ^a	average S/N of NOEs cross peaks	corresponding ¹ H- ¹ H distance ^b [Å]
intra		
Fuc3 H1-H2 (β)	1857 ^c	2.12
Fuc3 H1-H5 (β)	112 ^c	3.39
(Fuc3 H3-H5 (β)) ^d	488 ^c	(2.65)
Fuc3 H4-H5 (β)	731 ^c	2.48
Fuc3 H4-Q6	670 ^c	3.74 ^f
Gal2 H1-H3	1403 ^c	2.75
(Gal2 H1-H5) ^e	1759	(2.65)
Gal2 H3-H5	2823.5	2.45
Glc1 H1-H3 (β)	332	2.83
Glc1 H1-H5 (β)	1137 ^c	2.30
Glc1 H2-H4 (β)	808 ^c	2.44
Glc1 H3-H5 (β)	777 ^c	2.45
Glc1 H5-H61 (β)	446.5	2.69
Fuc4 H1-H2 (β)	595 ^c	2.56
Fuc4 H1-H3 (β)	95 ^c	3.48
Fuc4 H1-H5 (β)	80	3.58
Fuc4 H1-Q6 (β)	72 ^c	4.38 ^f
Fuc4 H3-H5 (β)	703	2.49
Fuc4 H3-Q6 (β)	87 ^c	4.24 ^f
Fuc4 H4-H5 (β)	1033 ^c	2.34
Fuc4 H4-Q6	814 ^c	3.62 ^f
inter		
Fuc3 H1 - Gal2 H2	1218 ^c	2.82
Fuc3 H1 - Gal2 H3	236	3.70
Fuc3 H5 - Gal2 H1 (β)	105 ^c	3.42
Fuc3 H5 - Gal2 H2 (β)	66 ^c	3.70
Fuc3 H5 - Glc1 H3 (β)	49 ^c	3.89
Fuc3 H5 - Glc1 H5 (β)	483	2.65
Fuc3 H5 - Glc1 H61 (β)	177.5	3.14
Fuc3 Q6 - Glc1 H1 (β)	45 ^c	4.73 ^f
Fuc3 Q6 - Glc1 H3 (β)	187 ^c	3.73 ^f
Fuc3 Q6 - Glc1 H5 (β)	81 ^c	4.29 ^f
Fuc3 Q6 - Glc1 H61 (β)	161 ^c	3.83 ^f
Fuc3 H1 - Fuc4 Q6	223 ^c	4.49 ^f
Gal2 H1 - Glc1 H4 (β)	550 ^c	2.60
Gal2 H1 - Glc1 H61	465 ^c	3.31
Gal2 H1 - Glc1 H62 (β)	691 ^c	2.50
Gal2 H2 - Fuc4 H5 (β)	604	2.56
Gal2 H2 - Fuc4 Q6	786 ^c	3.64 ^f
Gal2 H61 - Fuc4 H3	597 ^c	3.17
Gal2 H62 - Fuc4 H3	194 ^c	3.83
Gal2 H61 - Fuc4 H5	65 ^c	4.59
Glc1 H2 - Fuc H1 (β)	72.5	3.64
Glc1 H3 - Fuc H1 (β)	640.5	2.53
Glc1 H3 - Fuc H5 (β)	106 ^c	3.42

^a The nomenclature of the residues is as follows, Glc1: reducing end Glc; Gal2: stacking Gal; Fuc3: α1,2-linked Fuc; Fuc4: α1,3-linked Fuc. If the cross-peaks can be distinguished between the anomers at the reducing end, only the β-anomer is used and indicated by "β".

^b The ¹H-¹H distances were calculated from experimentally obtained NOE intensities using the H3-H5 cross-peak of Fuc and Gal as a reference with a distance of 2.65 Å assuming a 1/*r*⁶ dependence of the NOE

intensities. For the structure calculations the herein reported distances were increased by 0.2 Å tolerance and used as upper limit restraints.

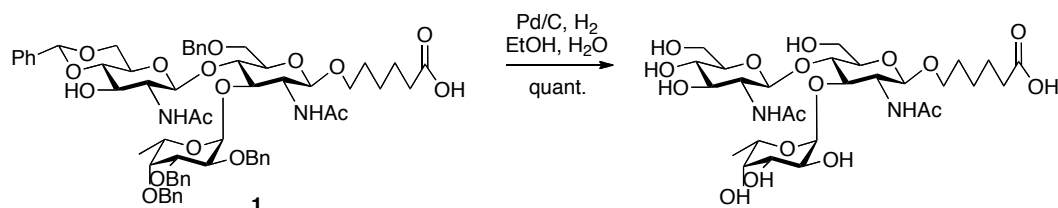
^c Only one cross-peak was used because of artifacts or severe spectral overlap.

^d Reference restraint for signals corresponding to the β-anomer at the reducing end (separated signals for α and β; NOESY measured in D₂O).

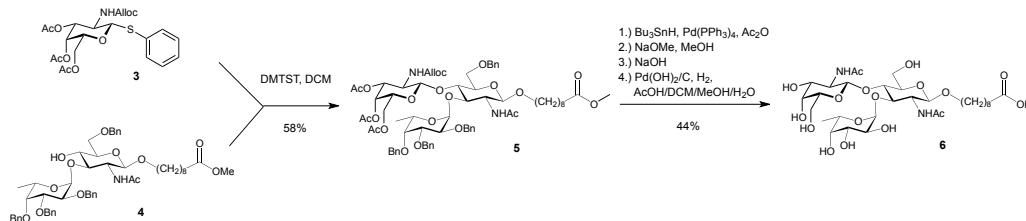
^e Reference restraint for signals indistinguishable for both anomers (no separate signals for α and β; NOESY measured in D₂O).

^f Signal to noise ratios from cross-peaks involving methyl or methylene protons were divided by 3 and 2 for the calculation, respectively.

SI Methods

Synthesis and characterisation of GlcNAc β 1,4[Fuca1,3]GlcNAc β 1-O-(CH₂)₅-COOH.

1 (10) (28 mg, 24 μ mol) was dissolved in a mixture of ethanol (2 mL) and water (1 mL). Pd/C 10% (12 mg) was added and the solution was stirred. Vacuum and H₂ were alternated and the mixture was allowed to stir under H₂ overnight. The mixture was filtered off through Celite and concentrated. The residue was dissolved in water and filtered through a 0.45 μ m syringe filter and concentrated. The residue was purified on a G15 column to give, after lyophilisation, 14 mg of the desired product (99%) as a white solid. $[\alpha]_{\text{D}}^{25}$ -80 (c 1.0, H₂O). MS FAB⁺-HRMS *m/z* [M+C₂H₅]⁺ calcd for C₃₀H₅₃O₁₇N₂ 713.3344, found 713.3357. A ¹H-¹³C HSQC spectrum with chemical shift assignment is shown in Figure S5A. Chemical shifts are deposited in the BMRB under the accession number 21032.

Synthesis and characterisation of GalNAc β 1,4[Fuca1,3]GlcNAc β 1-O-(CH₂)₈-COOH.

Compound **4** (11) (590 mg, 0.66 mmol) and thioglycoside **3** (791 mg, 1.64 mmol) were dissolved in dry DCM (25 mL) and stirred together with powdered 4 Å activated molecular sieves (4 g) for 4 h at rt. DMTST (509 mg, 1.97 mmol) was dissolved in DCM (15 mL) and stirred together with powdered 4 Å activated molecular sieves (2 g) for 4 h at room temperature as well. Both suspensions were combined and stirred for 16 h at room temperature. The mixture was filtered over a short pad of celite, diluted with DCM (50 mL), and washed with a saturated solution of NaHCO₃ (40 mL) and water (40 mL). The aqueous phases were extracted with DCM (3 × 30 mL). The combined organic layers were dried (Na₂SO₄) and the solvent was removed *in vacuo*. The crude product was purified by flash

chromatography (petroleum ether/EtOAc, 40% to 70%) to yield **5** as a colorless oil (268 mg, 211 μmol , 32%). Reactant **4** was recovered in 45% yield (263 mg, 293 μmol). $[\alpha]_{\text{D}}^{22}$ -44.2 (*c* 2.5, CHCl_3); ^1H NMR (500.1 MHz, CDCl_3): δ 7.42 – 7.22 (m, 20H, Ar-H), 5.94 – 5.83 (m, 1H, $\text{CH}_2=\text{CH}-\text{CH}_2$), 5.31 – 5.18 (m, 3H, 2x $\text{CH}_2=\text{CH}-\text{CH}_2$, GalNAc-H4), 5.16 – 5.09 (br s, 1H, Fuc-H1), 4.95 (d, *J* = 11.8 Hz, 1H, Ph- CH_2), 4.87 – 4.67 (m, 6H, 4x Ph- CH_2 , GalNAc-H3, GlcNAc-H1), 4.68 (d, *J* = 11.8 Hz, 1H, Ph- CH_2), 4.64 – 4.53 (m, 2H, Ph- CH_2 , $\text{CH}_2=\text{CH}-\text{CH}_2$), 4.49 (dd, *J* = 13.3, 5.4 Hz, 1H, $\text{CH}_2=\text{CH}-\text{CH}_2$), 4.46 – 4.35 (m, 2H, Ph- CH_2 , GalNAc-H1), 4.26 – 4.16 (m, 1H, Fuc-H5), 4.17 – 4.07 (m, 3H, GalNAc-H6, Fuc-H2, GlcNAc-H3), 4.00 (dd, *J* = 10.8, 5.9 Hz, 1H, GalNAc-H6'), 3.92 (t, *J* = 6.2 Hz, 1H, GlcNAc-H4), 3.88 – 3.83 (m, 1H, Fuc-H3), 3.78 – 3.67 (m, 5H, CH_2 linker, GlcNAc-H6, GlcNAc-H6', GalNAc-H2, GlcNAc-H2), 3.65 (s, 3H, COO-Me), 3.65 – 3.53 (m, 3H, Fuc-H4, GalNAc-H5, GlcNAc-H5), 3.40 – 3.32 (m, 1H, CH_2 linker), 2.28 (t, *J* = 7.5 Hz, 2H, CH_2 linker), 2.02 – 1.92 (m, 9H, 3x Ac), 1.83 – 1.75 (s, 3H, NH-CO- CH_3), 1.63 – 1.54 (m, 2H, CH_2 linker), 1.52 – 1.43 (m, 2H, CH_2 linker), 1.32 – 1.18 (m, 8H, CH_2 linker), 1.15 (d, *J* = 6.5 Hz, 3H, Fuc-H6); ^{13}C NMR (125.8 MHz, CDCl_3): δ 174.4, 170.5, 170.3, 170.3, 170.1 (4x Me-COO, COO-Me), 156.2 (NH-CO), 139.0, 138.7, 137.9, 128.8 – 127.3 (24x Ar-C), 132.7 ($\text{CH}_2=\text{CH}-\text{CH}_2$), 117.9 ($\text{CH}_2=\text{CH}-\text{CH}_2$), 100.2 (GalNAc-C1), 100.0 (GlcNAc-C1), 97.3 (Fuc-C1), 79.8 (Fuc-C3), 77.2 (Fuc-C4), 76.6 (Fuc-C2), 74.5 (Ph- CH_2), 74.4 (GlcNAc-C5), 74.3 (GlcNAc-C3), 73.9 (GlcNAc-C4), 73.6 (Ph- CH_2), 73.4 (Ph- CH_2), 72.7 (Ph- CH_2), 70.4 (GalNAc-C3), 70.4 (GalNAc-C5), 70.1 (GalNAc-C2), 69.6 (CH_2 linker), 69.3 (GlcNAc-C6), 66.7 (Fuc-C5), 66.3 (GalNAc-C4), 66.0 ($\text{CH}_2=\text{CH}-\text{CH}_2$), 60.7 (GalNAc-C6), 52.5 (GlcNAc-C2), 51.6 (COO-Me), 34.2, 29.5, 29.3, 29.3, 29.2, 25.9, 25.0 (7x CH_2 linker), 23.2 (NH-CO- CH_3), 20.7 (3x CH_3 -COO), 16.9 (Fuc-C6); ESI-MS: *m/z*: Calcd for $\text{C}_{68}\text{H}_{88}\text{N}_2\text{NaO}_{21}$ $[\text{M}+\text{Na}]^+$: 1291.58, found: 1291.65.

Compound **5** (87.0 mg, 68.5 μmol) was dissolved in DCM (2 mL) and $\text{Pd}(\text{PPh}_3)_4$ (1.6 mg, 1.37 μmol) and Bu_3SnH (21.9 mg, 75.4 μmol) were added. The solution was stirred for 5 min and Ac_2O (7.8 μL , 82.2 μmol) was added. The reaction mixture was stirred for 1 h at ambient temperature. The solvent was removed under reduced pressure and purified by flash chromatography (EtOAc/MeOH, 100% to 80%) to yield the intermediate (50.0 mg, 40.7 μmol , 60%) as a white solid. The intermediate (39.4 mg, 32.1 μmol) was suspended in MeOH (0.5 mL), treated with NaOMe/MeOH (320 μL , 0.02 M) and stirred for 16 h at room temperature. An aqueous solution of NaOH (64 μL , 10 M) was added and the reaction mixture stirred for 16 h at room temperature. The reaction was quenched with two drops of glacial acetic acid and concentrated *in vacuo*. The resulting alcohol was dissolved in DCM/MeOH/AcOH/ H_2O (1:1:2:2, 4 mL) and $\text{Pd}(\text{OH})_2/\text{C}$ (10.0 mg) was added. The suspension was stirred under an atmosphere of hydrogen for 18 h. The reaction mixture was filtered and purified by size exclusion chromatography to yield **6** as a white foam (17.5 mg, 23.6 μmol , 74%). $[\alpha]_{\text{D}}^{22}$ -49.0 (*c* 5.6, H_2O); HRMS: *m/z*: Calcd for $\text{C}_{31}\text{H}_{54}\text{N}_2\text{NaO}_{17}$ $[\text{M}+\text{Na}]^+$: 749.3320, found: 749.3323; A $^1\text{H}-^{13}\text{C}$ HSQC spectrum with chemical shift assignment is given in Fig. S5D. Chemical shifts are deposited in the BMRB under the accession number 21053.

SI References:

1. Zierke M, *et al.* (2013) Stabilization of branched oligosaccharides: Lewisx benefits from a non-conventional C-H...O hydrogen bond. *J. Am. Chem. Soc.* 135:13464-13472.
2. Kurutz JW & Kiessling LL (1997) Solution conformation of Lewis a--derived selectin ligands is unaffected by anionic substituents at the 3'- and 6'-positions. *Glycobiology* 7(3):337-347.
3. Coppin A, Florea D, Maes E, Cogalniceanu D, & Strecker G (2003) Comparative study of carbohydrate chains released from the oviducal mucins of the two very closely related amphibian species *Bombina bombina* and *Bombina variegata*. *Biochimie* 85(1-2):53-64.
4. Zhang J, Otter A, & Bundle DR (1996) Synthesis and conformational studies of the tyvelose capped, Lewis-x like tetrasaccharide epitope of *Trichinella spiralis*. *Bioorg Med Chem* 4(11):1989-2001.
5. Kitagawa H, *et al.* (1997) A novel pentasaccharide sequence GlcA(3-sulfate)(beta1-3)GalNAc(4-sulfate)(beta1-4)(Fuc alpha1-3)GlcA(beta1-3)GalNAc(4-sulfate) in the oligosaccharides isolated from king crab cartilage chondroitin sulfate K and its differential susceptibility to chondroitinases and hyaluronidase. *Biochemistry* 36(13):3998-4008.
6. Ishizuka Y, Nemoto T, Fujiwara M, Fujita K, & Nakanishi H (1999) Three-dimensional structure of fucosyllactoses in an aqueous solution. *J Carbohydr Chem* 18(5):523-533.
7. Kochetkov NK, Lipkind GM, Shashkov AS, & Nifant'ev NE (1991) N.m.r. and conformational analysis of some 2,3-disubstituted methyl alpha-L-rhamnopyranosides. *Carbohydr Res* 221:145-168.
8. Sakurama H, *et al.* (2012) 1,3-1,4-alpha-L-fucosynthase that specifically introduces Lewis a/x antigens into type-1/2 chains. *J Biol Chem* 287(20):16709-16719.
9. Holmner A, Askarieh G, Okvist M, & Krenzel U (2007) Blood group antigen recognition by *Escherichia coli* heat-labile enterotoxin. *J Mol Biol* 371(3):754-764.
10. Collot M, Wilson IB, Bublin M, Hoffmann-Sommergruber K, & Mallet JM (2011) Synthesis of cross-reactive carbohydrate determinants fragments as tools for in vitro allergy diagnosis. *Bioorg Med Chem* 19(3):1306-1320.
11. Srivastava OP, *et al.* (1999) US5874411A.

3.2 Acid pharmacophore orientation in sialyl Lewis^x mimics and synthesis of antagonists pre-organized in their bioactive conformation

3.2.1 “Bridging the gap” – Adjusting the acid pharmacophore in sLe^x mimics by ring closing metathesis (Manuscript)

Contributions

- Manuscript preparation
- Synthesis of compounds **4a-c**, **7-15**, **19-27**, **30-33** and **S7-S13**
- NMR spectra assignment and structure calculation of **3**

“Bridging the Gap” - Adjusting the Acid Pharmacophore in sLe^x Mimics by Ring Closing Metathesis

Mirko Zierke,[#] Martin Smiesko,[#] Norbert Varga,[#] Florian P. C. Binder,[#] Mario Schubert,[¶]
Roland Preston,[#] John T. Patton,[&] John L. Magnani,[&] Pascal Zihlmann,[#] Tobias
Mühlethaler,[#] and Beat Ernst^{#*}

[#] University of Basel, Klingelbergstr. 50, CH-4056 Basel, Switzerland

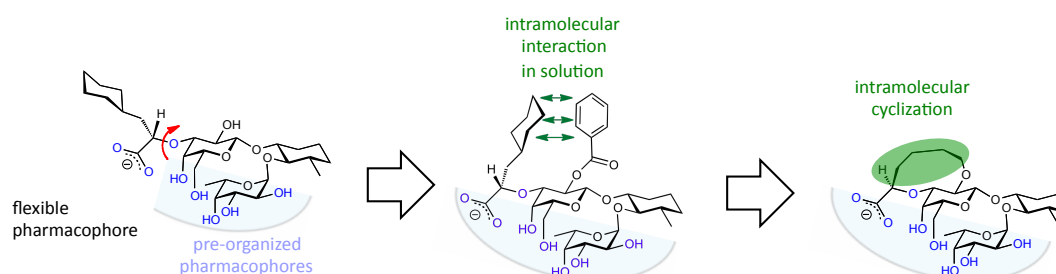
[¶] Institute of Molecular Biology and Biophysics, ETH Zürich, CH-8093 Zürich, Switzerland

[&] GlycoMimetics Inc., Gaithersburg, MD20878, USA

* Corresponding Author; E-mail: beat.ernst@unibas.ch

Key words: structural elucidation, sialyl Lewis^x (sLe^x) mimics, NOESY, NMR, molecular dynamics, solution conformation, pharmacophore pre-organization

Table of content



Abstract

Selectins play a crucial role in the body's defense mechanism against inflammation. They initiate, induced by an inflammatory stimulus, tethering and rolling of leukocytes on the endothelial surface of blood vessels. This is an essential step to extravasate from the blood to the inflamed tissue. However, an excessive infiltration of leukocytes is observed in several acute and chronic diseases, and therefore, selectins represent an interesting therapeutic target. Sialyl Lewis^x (sLe^x) is the natural ligand of selectins, however, the binding affinity is rather low and the high polarity and complexity of the molecule are challenges in the development of sLe^x mimics. In numerous studies the binding affinity and physicochemical parameters could be improved. Hydrophilic parts, which are not involved in binding, were replaced by hydrophobic moieties. Furthermore, the rigid core structure Lewis^x could be mimicked by non-carbohydrate moieties and hence, 5 of the 6 pharmacophores are pre-organized in the bioactive conformation in solution. Introducing an additional benzoate in 2'-position gave a further improvement in binding affinity (**2**→**3**). We could solve the conformation of **3** in solution by NMR spectroscopy and disclose an intramolecular σ - π interaction, that helps to pre-organized the acid pharmacophore in the bioactive conformation. A similar effect was obtained by introducing aromatic substituents, which can form interresidual π - π interactions. Taking advantage of this observation we synthesized a series of selectin antagonists that 'bridge the gap' by replacing the intramolecular hydrophobic interaction via a covalent bond using ring closing metathesis, and pre-organize all pharmacophores in the bioactive conformation.

Introduction

Selectins are an extensively studied class of carbohydrate binding proteins. They mediate the first contact and rolling of leukocytes on endothelia cells, followed by tighter integrin binding and leukocyte extravasation from blood circulation into the diseased or infected tissue.¹ This mechanism is crucial in both, the adaptive and innate immune response.^{2,3} Many acute and chronic diseases like asthma,⁴ psoriasis,⁵ reperfusion injury⁶ and rheumatoid arthritis⁷ are characterized by an excessive extravasations of leukocytes to the inflamed tissue. Furthermore, several types of tumors harness and exploit the selectin dependent mechanisms, as it is used by migrating leukocytes, to metastasize.^{8,9} Thus, selectins are a promising target to treat these diseases.

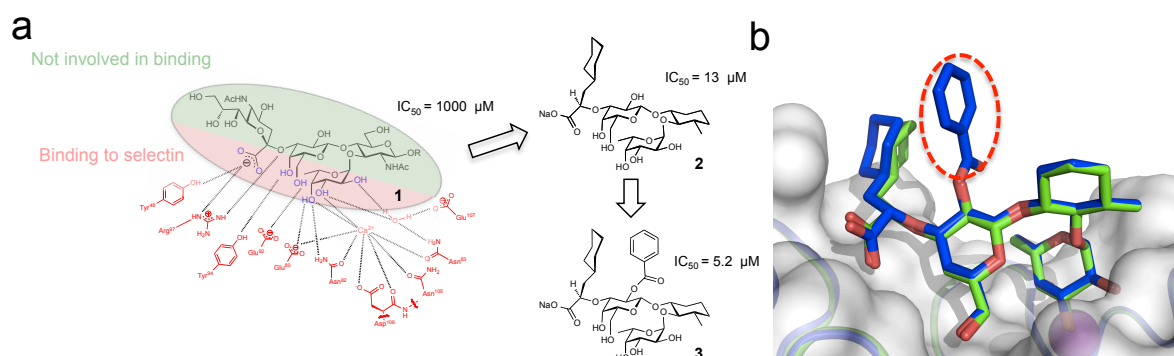


Figure 1. a) Schematic presentation of interactions between sLe^x (**1**) (pharmacophores highlighted in blue) and E-selectin. In sLe^x mimic **2**¹⁰, hydrophilic parts of **1** not involved in binding are replaced by hydrophobic moieties, leading to an 80 fold increase in binding affinity.¹¹ An additional benzoyl group (**2**→**3**) improved the affinity to 5.2 μM. b) Overlay of crystal structures¹² of compounds **2** (green) and **3** (blue) in complex with E-selectin (the benzoate group of **3** is highlighted by a red circle).

All native ligands of the selectins contain the common tetrasaccharide epitope sialyl Lewis^x [Neu5Acα(2-3)Gal[Fucα(1-3)]β(1-4)GlcNAc, sLe^x (**1**)]. Because the binding affinities of sLe^x (**1**) is rather weak, it is often improved by a multivalent presentation.^{13,14} For the development of efficient drug-like selectin antagonists mimicking the sLe^x epitope, it is necessary to overcome the problems of low affinity, the high polarity of the sLe^x lead structure and to reduce the structural complexity.¹⁵

Structural information on sLe^x (**1**) in solution and also bound to E-selectin are of major importance for the design of drug-like mimetics. The conformation of **1** bound to E-selectin was elucidated by NMR¹⁶⁻¹⁸ and crystallography,^{19,20} and disclosed the bioactive conformation and the pharmacophoric groups (Figure 1a). The solution conformation was investigated by several groups,^{17,18,21} and can be roughly divided in two structural elements, a rigid Galβ(1-4)[Fucα(1-3)]GlcNAc (Le^x-core) and a Neu5NAc moiety linked by a rather flexible glycosidic bond to the D-galactose unit.²¹ The trisaccharide core is stabilized by an exo-anomeric effect,^{22,23} hydrophobic interactions,²⁴ steric effects^{10,25,26} and a nonconventional hydrogen bond,²⁷ and is almost identical to the bound conformation. The Neu5NAc is rather flexible and adopts mainly two conformations in solution, one of them similar to the bound conformation.²¹

For the development of sLe^x mimics, the GlcNAc residue which task is the adjustment of the D-Gal and L-Fuc moiety, and the D-Neu5NAc which has just weak contacts to the protein were replaced by hydrophobic mimics (compound **2**, Figure 1a).^{10,28} The stacked L-fucose and D-galactose moieties are crucial, as they bear 5 of the 6 pharmacophores. The

unique orientation leads to a tight stacking and pre-organization of the pharmacophores in the bioactive conformation. Overall the affinity could be improved from sLe^x up to 80 fold (**1**→**2**), measured in a cell free competitive binding assay.¹⁰ A benzoate in 2'-position of the D-galactose improves the affinity by an additional factor of 2.5 (**2**→**3**).¹⁰ Thoma *et al.* presumed that the 2' benzylation has a stabilizing effect on the trisaccharide mimic core, and further improves its pre-organization.²⁹

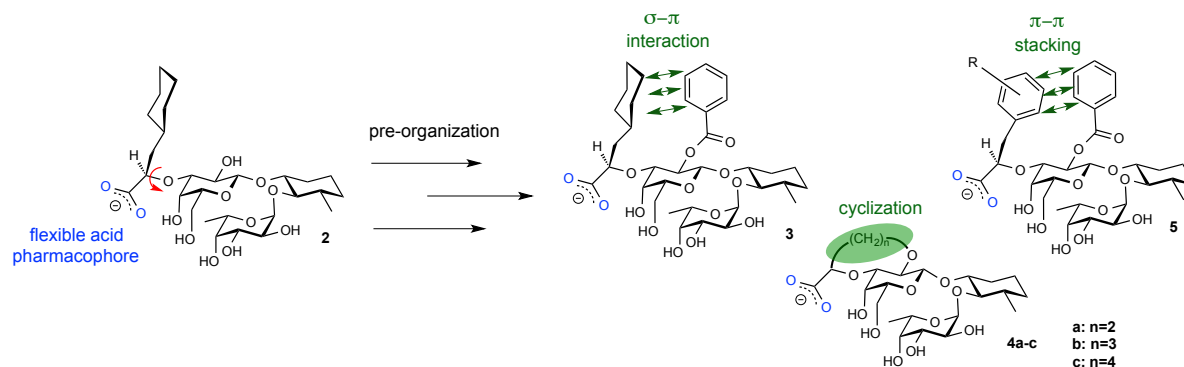


Figure 2. Compound **2** with a flexible acid pharmacophore and strategies to pre-organization the acid functionality in the bioactive conformation.

In recently solved crystal structures, **2** and **3** (cocrystallized with E-selectin) showed an almost identical binding mode (Figure 1b).¹² The benzoate in antagonist **3** is pointing to the solvent, clearly disproving speculations on an increase in binding affinity due to benzoate-protein interactions. However, the cyclohexane moiety of the cyclohexyl lactic acid in **3** is in spatial proximity to the benzoate enabling an intramolecular σ - π interactions.¹² An interaction of the two residues in solution could contribute to a preferable orientation of the acid pharmacophore in the bioactive conformation. The pre-organized ligand **3** should show an entropic gain over its flexible counterpart **2** due to the reduced conformational entropy that is associated with restricting rotors.^{30,31}

To confirm this hypothesis, the aim of this communication was to elucidation the conformation of **3** in solution. Furthermore, compounds that are proposed to pre-organize the acid pharmacophore in solution by intramolecular aromatic π - π interactions, or by intramolecular cyclisation were developed (Figure 2).

Results

Compound **3** is pre-organized in the bioactive conformation in solution.

To prove our hypothesis, that in compound **3** the acid pharmacophore is pre-organized in the bioactive conformation by an interaction between the benzoate and the cyclohexyl lactic acid, we solved the conformation in solution. Recently, we demonstrated that NOESY NMR spectroscopy is an effective tool to analyze the conformation of oligosaccharides at ~275 K at 900 MHz high field in solution.³² We adopted the method to evaluate the solution conformation of **3**.

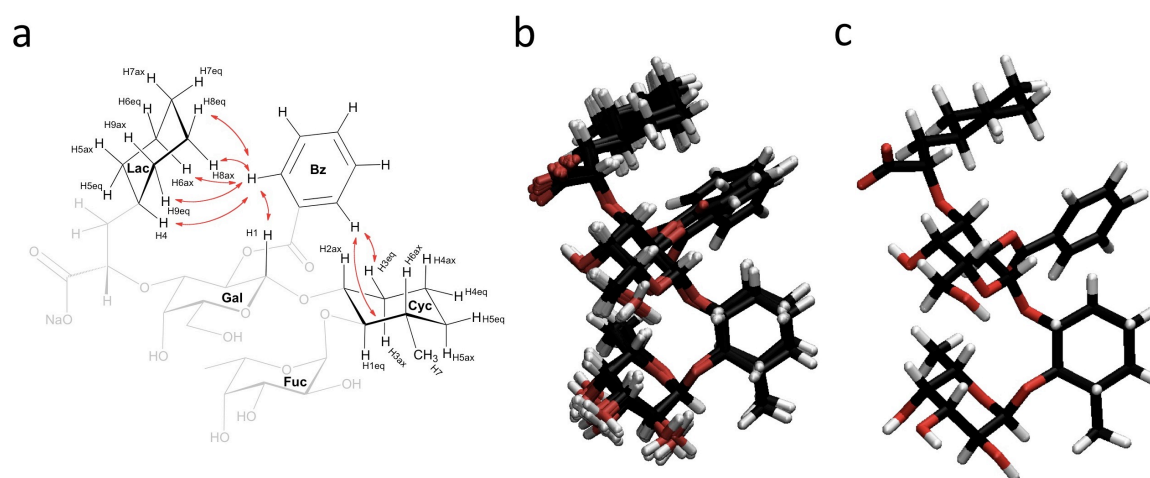


Figure 3. a) Schematic overview of interresidual NOEs between cyclohexan moieties and H2-benzoate (red arrows) protons; b) Ensemble of 20 lowest energy conformations of **3** at 277 K in solution, calculated and refined using experimental NOE restraints; c) Crystal structure of **3**.¹²

The chemical shift assignment of the resonances of compound **3** was achieved by homo- and heteronuclear 2D spectra recorded in D₂O (SFigure 1, SFigure2). Using a 2D NOESY spectrum, we could extract 56 unambiguous NOE cross-peaks (36 intraresidual and 20 interresidual) between various non-exchangeable protons (STable 1 and statistics in STable 2). Specifically 10 interresidual NOEs between the benzoate and the two cyclohexane residues could be extracted (Figure 3a and STable1). With 57 NOE derived distance restraints we were able to calculate a well-defined structural ensemble of compound **3** using Cyana³³ and subsequent refinement with AMBER⁹³⁴ using the GAFF force field (Figure 3b). A comparison of this ensemble of solution conformations and the crystal structure of **3** displays an obvious similarity (Figure 3b and 3c). Furthermore, the dihedral angles of conformations of **3** observed in solution, crystal and by molecular dynamic calculations showed values in a narrow range in the core conformation/acid orientation plot (SFigure 3;

statistics in STable 2).³⁵ Another confirmation for the interactions between the cyclohexyl lactic acid and the benzoate are the pronounced upfield shifts of cyclohexane protons with spatial proximity to the aromatic ring in **3** compared to **2** (SFigure 4).

The results propose that the stabilization of the cyclohexyl lactic acid by the benzoate in **3** is established already in the solution, that guarantees a pre-organization of the acid pharmacophore before the binding event.

Benzoylation in 2'-position of the D-galactose residue effects affinity in lactic acid derivatives.

To confirm the assumption that intermolecular hydrophobic interactions have an effect on affinity, we synthesized a series of sLe^x mimics bearing 2 aromatic moieties in spatial proximity. They are uniformly benzoylated in 2' position of the D-galactose residue and in contrast to **3**, the (*S*)-cyclohexyl lactic acid residue is replaced by a (*S*)-phenyl lactic acid moiety (\rightarrow **5a**), a phenyl group modified with electron donating (\rightarrow **5b**), or electron withdrawing (\rightarrow **5c**) substituents. We expected a strong π - π interaction between the aromatic residues, which would stabilize the acid orientation in a similar way to **3** and could lead to improved binding affinities. All three antagonists (**5a**, **5b** and **5c**) showed indeed favorable binding affinities, just slightly worse than the (*S*)-cyclohexyl lactic acid **3**. Substitutions on the phenyl rings had only a small effect on the potency of the ligands.

To quantify the effects of 2'-benzoylation, 2'-hydroxy derivatives **6a-c** were synthesized. Debenzoylation (**5a-c** \rightarrow **6a-c**) reduced affinities by a factor of 5 to 6, whereas debenzoylation of **3** just led to a 3.2 time reduction. Possible reasons are the reduced desolvation penalties, and different interactions between the phenyl/cyclohexane groups and protein side chains for **2** vs. **6a-c**.

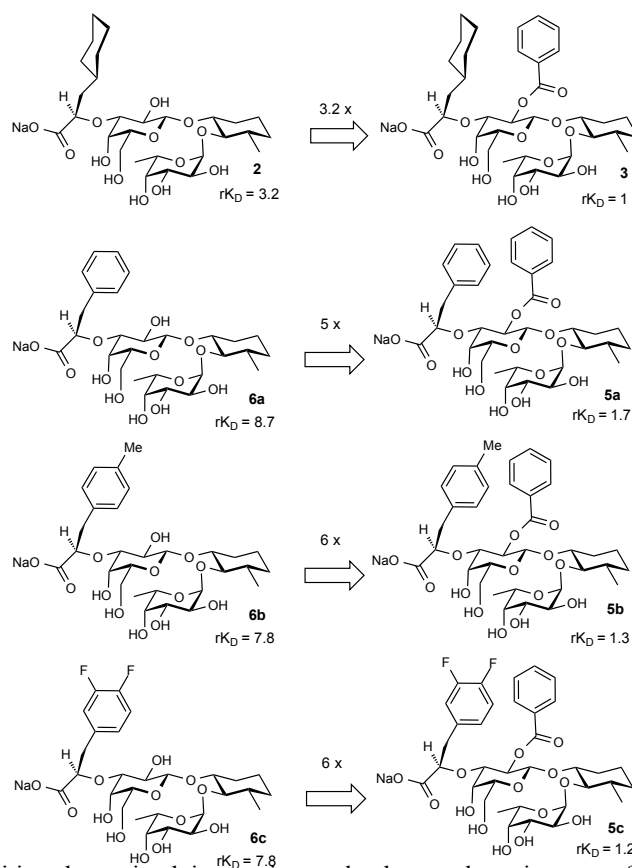
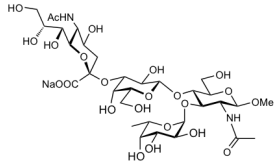
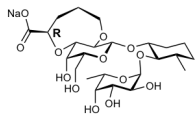
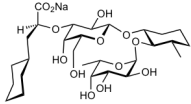
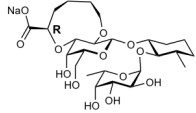
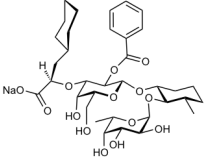
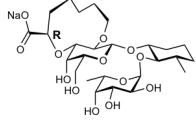
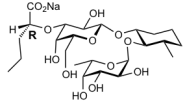


Figure 2. Binding affinities determined in a microscale thermophoresis assay for benzoates **3**, **5a-c** and compounds **2** and **6a-c** that lack the benzoate substitution in 2'-position of the D-galactose moiety. The rK_D s values are reported relative to lead compound **3**.

Acid pre-organization by intramolecular cyclisation.

In a second attempt we replaced the benzoate and cyclohexane residues by an intramolecular alkyl bridge, to explore whether we can detect a similar effect to the intramolecular hydrophobic interaction. By modifying the chain length, we could optimize the balance between rigidity and retaining a certain degree of flexibility for the optimal adjustment of the acid pharmacophore upon binding in bigger cycles. The acid pharmacophore pre-organization in different ring sizes was evaluated by molecular dynamic calculations and displayed in acid orientation / core conformation plots (SFigure 5). We decided to synthesize cyclic compounds **4a-c** as they show the best pharmacophore orientation.

Table 1. Affinities of selectin antagonists **1**, **2**, **3**, **4a-c**, **7**. K_D values in μM were determined by a microscale thermophoresis assay.^[16] $rK_{D\text{MST}}$ are in comparison to **3**.

Compound	No	$K_{D\text{MST}}$ [μM]	$rK_{D\text{MST}}$	Compound	No	$K_{D\text{MST}}$ [μM]	$rK_{D\text{MST}}$
	1	649.3 ± 66.7	147.9 ± 2.4		4c	23.3 ± 2.3	5.3 ± 0.1
	2	14.5 ± 1.5	3.2 ± 0.0		4a	11.4 ± 1.2	2.6 ± 0.0
	3	4.5 ± 0.4	1		4b	130.1 ± 13.2	29.6 ± 0.6
	7	4692.6 ± 433.0	1068.7 ± 56.1				

The compounds were synthesized and to evaluate their binding properties the $K_{D\text{S}}$ were measured in a microscale thermophoresis assay.

The best binding affinity was observed in the 9-membered ring **4a** with a K_D of 11.4 μM . The binding affinity was improved by a factor of 2 to the smaller 8-membered cycle **4c** and by a factor of 11 to the 10-membered cycle **4b**. The strong differences in binding affinity of the medium-size rings are probably due to changes in transannular interactions, conformational strains and a varying degree of pre-organization due to restricted rotations.^{36,37} **4a** showed an improved affinity to the flexible compound **2** by a factor of 1.3, but a drop in affinity by a factor of 2.5 compared to compound **3**, which is stabilized by intramolecular hydrophobic interactions.

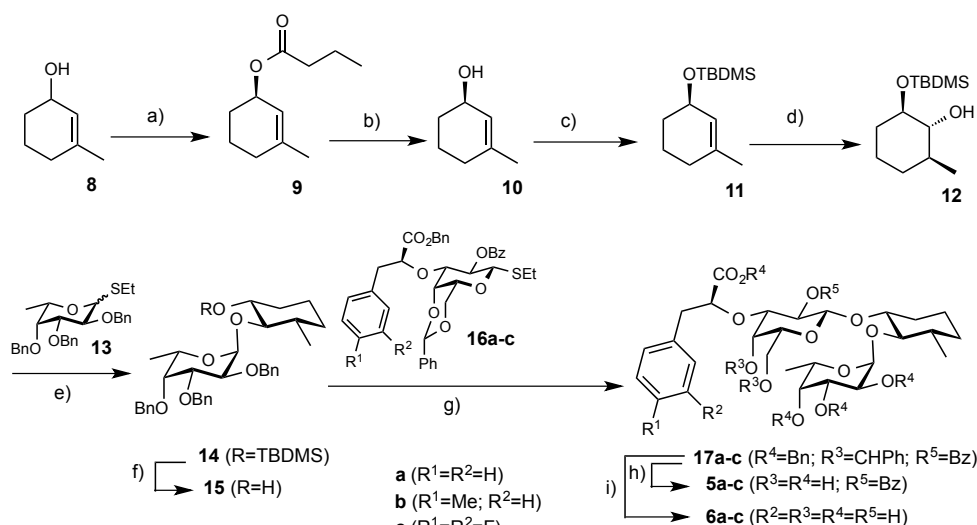
As control, we synthesized compound **7** with a flexible alkyl chain. The binding affinity for **7** is 4.7 mM, what is in agreement with the negative result in the acid orientation / core conformation plot obtained by MD simulations (SFigure 6). It predicted a 120° deviation of the acid orientation compared to the crystal structure, probably due to an intramolecular interaction of the acid and the hydroxy group in 2'-position of the D-galactose residue.

Synthesis of aromatic ligands **5a-c** and **6a-c**.

We recently reported the synthesis of the Fuc α (1-3)GlcNAc mimic **15**.¹⁰ In this communication we describe a faster and more efficient approach to synthesize **15** from commercially available racemic seudenol (3-methyl-2-cyclohexenol, **8**).³⁸

Using optimized conditions for the butanoylation of racemic seudenol with immobilized *Candida antarctica* lipase C (Novozym 435),³⁹ we could isolate (*R*)-seudenolester **9**.⁴⁰ Subsequent saponification with NaOH afforded (*R*)-seudenol **10** in 84% yield and 97.5% enantiomeric excess (*ee*). Since the protection group of the hydroxy group in **10** has to be stable under strongly basic and acidic conditions, should not hamper fucosylation by steric bulk, and finally allow cleavage under mild conditions orthogonal to benzyl protecting groups, a *tert*-butyldimethylsilyl (TBS) ether **11** was synthesized. Hydroboration followed by oxidation yielded all-*trans* **12** in 81% over two steps. Fucosylation of **12** under *in situ* anomerisation conditions⁹ gave **14**, which was smoothly deprotected with tetrabutylammonium fluoride, affording pseudodisaccharide **15**¹⁰ in excellent yield over two steps. Starting from racemic seudenol, this short sequence allowed the gram scale synthesis of **15** in 27% overall yield, requiring only four chromatographic purifications.

In the following step disaccharide mimic **15** was glycosylated with the corresponding galactose donors **16a-c** functionalized with different lactic acid derivatives (see scheme S3) using dimethyl(methylthio)sulfonium triflate (DMTST) as promoter, which gave products **17a-c** in good yields. In order to remove the benzyl protecting groups as well as the benzyliden acetal function, compounds **17a-c** were hydrogenated under heterogeneous conditions using Pd(OH)₂/C as catalyst resulting in final ligands **5a-c** with a benzoyl moiety in position 2 of the galactose unit. To obtain the fully deprotected antagonists **6a-c**, compounds **17a-c** were additionally treated with LiOH after the hydrogenation step.



Scheme 1. Synthesis of sLe^x mimics, a) Novozyme 435, vinylbutyrate, heptane; b) aqueous NaOH, MeOH, 84%; c) TBSCl, imidazol, DMAP, CH_2Cl_2 ; d) (i) BH_3 , THF, THF; (ii) H_2O_2 , aqueous NaOH, 81% from **16**; e) $CuBr_2$, DTBMP, TBAB, CH_2Cl_2 , DMF, mol. sieves (4\AA), 87%; f) TBAF, THF, quant.; g) DMTST, DCM, mol. sieves (4\AA) (**a**: 79%, **b**: 80%, **c**: 83%); h) H_2 , $Pd(OH)_2/C$, MeOH (**a**: 42%, **b**: 37%, **c**: 63%) i) (i) H_2 , $Pd(OH)_2/C$, MeOH; (ii) LiOH, MeOH/ H_2O (**a**: 12%, **b**: 41%, **c**: 59%).

Synthesis of cyclic ligands **4a-c**.

The key step of the synthesis was ring closure metathesis (RCM) reaction of a (*R*)-1-methoxy-1-oxo-pent-4-en-2-yl group in 3' position of the D-galactose and a terminal alkene with variable chain length in its 2'-position (Scheme 2).

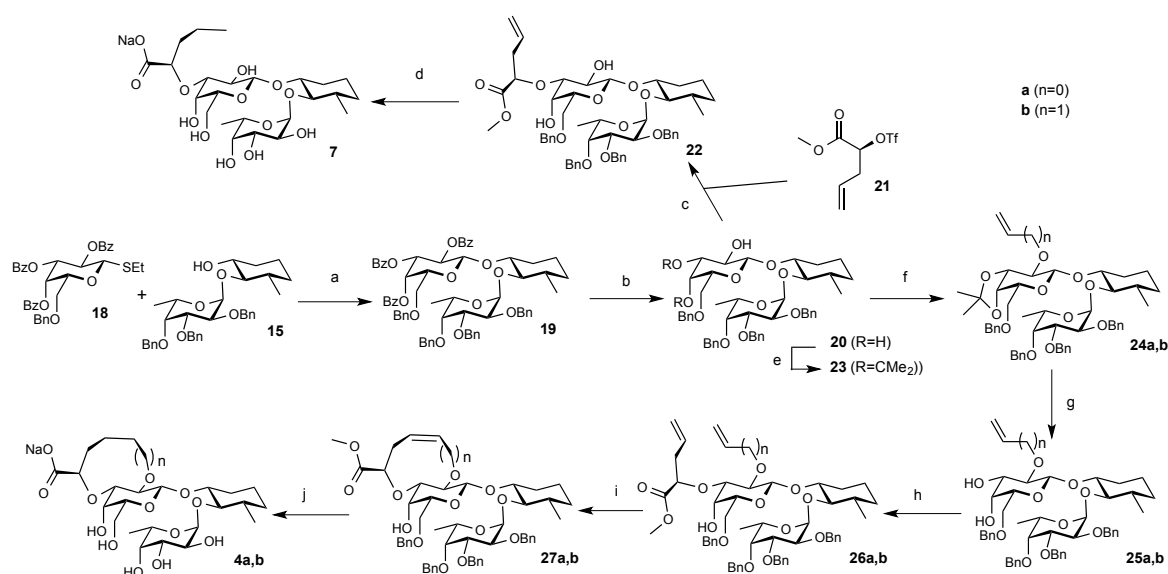
Galactosylation of **15** with galactosyl donor **18**⁴¹ promoted by DMTST afforded **19** β -selectively.¹⁰ The acetyl protection groups were hydrolysed under Zemplén conditions and gave **20**, the precursor for macrocycles **4a,b** and also for the non-cyclic control compound **7**.

By tin mediated alkylation of 3'-position of the D-galactose residue, alkene **22** was obtained. Debenzylation by hydrogenolysis, subsequent saponification with lithium hydroxide followed by ion exchange chromatography gave **7**.

For the synthesis of macrocycles **4a,b**, we selective introduced an acetonide protection group to **20** to get alcohol **23**. Alkylation of **23** with allylbromid, or 3-butenyl triflate lead to **24a** and **24b**, respectively.

The acetal protection group was removed with AcOH to give diols **25a,b**. Triflate **21** (Scheme S3) was introduced in 3'-position by selective tin-mediated alkylation of the D-galactose moiety and subsequently underwent RCM with Grubbs 2nd generation catalyst.

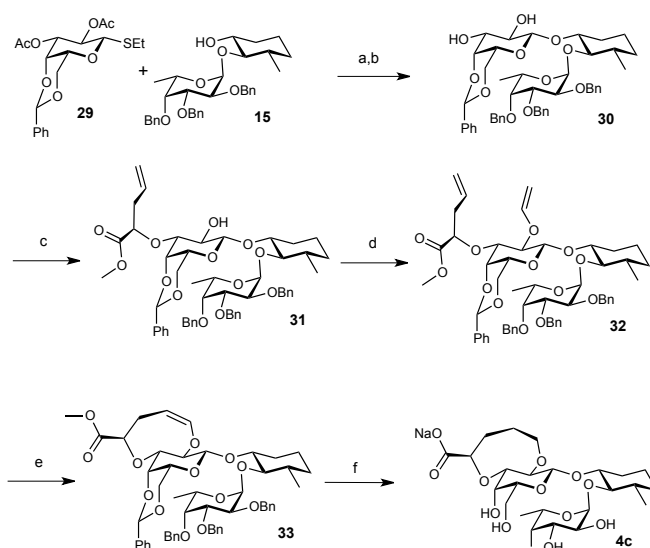
Debenzylation and saponification of the methylester with LiOH followed by ion exchange chromatography gave the 9- and 10-member macrocycles **4a** and **4b**.



Scheme 2. a) DMTST, DCM, 4Å mol.sieves, 79%; b) NaOMe, MeOH, 81%; c) (i) Bu₂SnO, MeOH; (ii) CsF, **21**, DCM 57%; d) Pd(OH)₂/C, dioxane/H₂O (4:1), H₂, 52%; (ii) LiOH, H₂O, 52%; e) Me₂C(OMe)₂, CuSO₄, PPTS, acetone, 83%; f) NaH, Br(CH)_nCH=CH₂, DMF [**a**: 85%; **b**: 34% (42% of reactant recovered)]; g) AcOH (80%), 40°C [**a**: 78%; **b**: 91%]; h) (i) Bu₂SnO, MeOH; (ii) CsF, triflate **21**, DCM [**a**: 57% **b**: 93%]; i) Grubbs 2nd gen. cat., DCM [**a**: 86%; **b**: 93%]; j) (i) Pd(OH)₂/C, dioxane/H₂O (4:1), H₂; (ii) LiOH, H₂O [**a**: 57%; **b**: 25% over three steps).

Synthesis of the smaller 8-membered macrocycle was not successful using a similar approach, as the vinyl group introduced in 2'-position of **23** turned out to be unstable under acetone deprotection conditions. We therefore modified the synthetic route and introduced the vinyl functionality in the last step before the RCM (Scheme 3).

Thus, galactosylation of alcohol **15** with thiogalactoside **29** and DMTST promotion followed by hydrolysis of the benzoyl groups yielded **30**. Alkylation via the tin acetal method in 3' position and subsequent transvinylation catalysed by an iridium complex catalyst following the procedure of Chayajarus gave **32** in good yields.^{42,43} The 8-membered ring of **33** was established by RCM of dialkene **32**. Deprotection and sodium exchange column chromatography gave **4c** in 35 % overall yield from **30**.



Scheme 3. a) DMTST, DCM, 4Å mol.sieves, 55%; b) NaOMe, MeOH, 94%; c) (i) Bu₂SnO, MeOH; (ii) CsF, triflate **21**, DCM, 85% over 2 steps; d) Na₂CO₃, vinyl acetate, Ir Cat., toluene, 70%; e) Grubbs 2nd Gen. Cat., DCM 86%; f) Pd(OH)₂/C, dioxane/H₂O (4:1), H₂; (ii) LiOH, H₂O, 68%.

Conclusions

We were able to determine the solution structure of the potent E-selectin antagonist **3**, which mimics the conformation of the rigid Le^x core of the natural E-selectin ligand sLe^x (**1**). In addition, the acid pharmacophore is pre-organized by hydrophobic interaction of a benzoate residue in 2'-position and a (*S*)-cyclohexyl lactic acid moiety in 3'-position of the D-galactose moiety. We assumed that the pre-organization of the acid functionality is a major factor for the improved binding affinity over the more flexible antagonist **2**. Therefore, we designed and synthesized a series of cyclic E-selectin antagonists in which an intramolecular aliphatic bridge adjusts the acid functionality in the bioactive conformation already before binding.

The ring size of the compounds had a strong influence on the binding properties. The best binding affinity was observed in medium-size cycle **4a**, where a 9-membered ring adjusts the acid pharmacophore. It shows a 1.5 fold increase in binding affinity over the flexible mimic **2** but it is not as potential than compound **3**, in which the acid is pre-organized by intramolecular hydrophobic interactions. The increase in affinity could be due to interactions of the cyclohexyl lactic acid in **3** with residues Glu98 and Lys99 of the protein and a stabilizing water cluster around the molecule.⁴⁴

In the future the binding mode and thermodynamic profile for E-selectin/ligand binding will be determined for the cyclic selectin ligands **4a-c**.

Experimental methods

Synthetic procedures and characterization of compounds are described in the supporting information.

Molecular modeling. The 3D structure of the carbohydrate mimic molecules was built in the Maestro⁴⁵ molecular modeling environment and a conformational analysis was performed using Macromodel⁴⁶ by performing 5000 steps of the mixed torsional/low-mode sampling in combination with the OPLS-2005 force-field in implicit solvent conditions (water).

Global minimum identified in the conformational search was then placed in a periodic boundary system (cube, side length 40.0 Å) filled with explicit water molecules (TIP3P parametrization). The system was minimized and a simulated annealing simulation (gradual heating of the system to the temperature of 400 K in 330 ps followed by slow gradual cooling to 300 K in 500 ps) was performed in Desmond^{47,48} in order to bring the system into equilibrium. The production phase molecular dynamics simulation of the total duration of 48 ns was performed. Two critical structural descriptors – the core conformation and the acid orientation – were monitored every 4.8 ps (10 000 data points were saved).

The pre-organization of the mimic molecule was evaluated by plotting the frequency of the two critical parameters in a two-dimensional color-coded plot using an in-house computer program.

NMR spectroscopy for three-dimensional structure determination. Unless indicated otherwise spectra were measured at a 750 MHz Bruker Avance III spectrometer equipped with a TXI triple-resonance probe at 275 K. For complete chemical shift assignment of compounds **2** and **3** they were dissolved in D₂O at concentrations of 6.6 mM **3** and 7.6 mM **2**, respectively. Assignment was achieved using the following 2D spectra: ¹³C-¹H HSQC, ¹³C-¹H HMBC, ¹³C-¹H HMQC-COSY⁴⁹ and TOCSY ¹H-¹H TOCSY (mixing times 13 ms and 120 ms). The ¹³C-¹H HMQC-COSY spectrum was especially useful for assigning the cyclohexane resonances, see Supplementary Fig. 2. Complete chemical shift assignments for compounds **3** and **4** can be found in Supplementary Figs. 1 and 6. A 2D NOESY spectrum of compound **3** was recorded at 900 MHz with a mixing time of 150 ms, 96 transients and 2,048×530 points. Spectra were processed using Topspin 2.1 and analyzed by Sparky⁵⁰. All spectra were referenced to DSS according to Markley et al.⁵¹

Structure Calculation and Refinement. Initial structures were calculated using CYANA 3.0.⁵² Signal to noise (S/N) ratios of all NOE signals were extracted using the program Sparky⁵⁰ and converted to distances using the r^{-6} dependence and the Cyc H3a–H3b cross-peaks (1.77 Å) as reference. S/N ratios of signals involving CH₂ and CH₃ groups with coinciding ¹H resonances were divided by a factor of 2 or 3, respectively. Upper limit restraints with an additional tolerance of 0.5 Å were applied. Out of 200 structures, the 30 structures with the lowest target function were further refined in AMBER 9³⁴ applying the general AMBER force field (GAFF).⁵³ A generalized Born model was used to mimic solvent.⁵⁴ Initial charge parameters were generated by Antechamber for the GAFF force field.

Microscale Thermophoresis. The measurements were performed using a Monolith NT.115 (NanoTemper Technologies GmbH, Munich, Germany). The experiments were carried out at 25°C with 50% LED power, 50% laser power, laser on time of 30 sec, and laser off time of 5 sec with standard treated capillaries. The ligands were diluted 1:1 with a starting concentration of 1 – 4 mM with assay buffer (20 mM HEPES, 150 mM NaCl, 1 mM CaCl₂, pH 7.4). In addition 0.05% Tween 20 and 10% DMSO were added. The protein E-Selectin_{LEC2} was purified and labeled according to Zihlmann *et al.*⁵⁵ The protein concentration was determined by HPLC-UV against a BSA standard.^{56,57} Throughout all experiments the E-Selectin_{LEC2} concentration was 0.2 µM. Every measurement was carried out three times. Datapoints were normalized using the bound and unbound borders achieved by NanoTemper Analysis 1.2.205 software (NanoTemper Technologies GmbH, Munich, Germany) and analyzed with GraphPad Prism 5.0 (GraphPad Software, La Jolla, CA, USA).

References

- (1) Butcher, E. C.; Picker, L. J. *Science* **1996**, *272*, 60.
- (2) Tosi, M. F. *J. Allerg. Clin. Immunol.* **2005**, *116*, 241.
- (3) Schön, M. P.; Ludwig, R. J. *Expert Opin. Therap. Targets* **2005**, *9*, 225.
- (4) Romano, S. *Treatm. Respirat. Med.* **2005**, *4*, 85.
- (5) Bock, D.; Philipp, S.; Wolff, G. *Expert Opin. Invest. Drugs* **2006**, *15*, 963.
- (6) Lefer, D. J.; Flynn, D. M.; Phillips, M. L.; Ratcliffe, M.; Buda, A. J. *Circulation* **1994**, *90*, 2390.
- (7) Sfikakis, P. P.; Mavrikakis, M. *Clin. Rheum.* **1999**, *18*, 317.

- (8) Witz, I. In *The Link Between Inflammation and Cancer*; Dalgleish, A., Haefner, B., Eds.; Springer US: **2006**; Vol. 130, p 125.
- (9) Witz, I. P. *Immun. Lett.* **2006**, *104*, 89.
- (10) Schwizer, D.; Patton, J. T.; Cutting, B.; Smieško, M.; Wagner, B.; Kato, A.; Weckerle, C.; Binder, F. P. C.; Rabbani, S.; Schwardt, O.; Magnani, J. L.; Ernst, B. *Chem. – Eur. J.* **2012**, *18*, 1342.
- (11) Binder, F. P. C.; Lemme, K.; Preston, R. C.; Ernst, B. *Angew. Chem. Int. Ed.* **2012**, *51*, 7327.
- (12) Preston, R. C., PhD Thesis, University of Basel, **2014**.
- (13) Spiess, M. *Biochemistry* **1990**, *29*, 10009.
- (14) Mammen, M.; Choi, S. K.; Whitesides, G. M. *Angew. Chem. Int. Ed.* **1998**, *37*, 2755.
- (15) Ernst, B.; Magnani, J. L. *Nat. Rev. Drug. Discov.* **2009**, *8*, 661.
- (16) Scheffler, K.; Ernst, B.; Katopodis, A.; Magnani, J. L.; Wang, W. T.; Weisemann, R.; Peters, T. *Angew. Chem. Int. Ed.* **1995**, *34*, 1841.
- (17) Poppe, L.; Brown, G. S.; Philo, J. S.; Nikrad, P. V.; Shah, B. H. *J. Am. Chem. Soc.* **1997**, *119*, 1727.
- (18) Harris, R.; Kiddle, G. R.; Field, R. A.; Milton, M. J.; Ernst, B.; Magnani, J. L.; Homans, S. W. *J. Am. Chem. Soc.* **1999**, *121*, 2546.
- (19) Somers, W. S.; Tang, J.; Shaw, G. D.; Camphausen, R. T. *Cell* **2000**, *103*, 467.
- (20) Lin, Y. C.; Hummel, C. W.; Huang, D. H.; Ichikawa, Y.; Nicolaou, K. C.; Wong, C. H. *J. Am. Chem. Soc.* **1992**, *114*, 5452.
- (21) Lemieux, R. U. *Molecular rearrangements* Interscience, New York, **1964**.
- (22) Wolfe, S.; Whangbo, M. H.; Mitchell, D. J. *Carbohydr. Res.* **1979**, *69*, 1.
- (23) Titz, A.; Marra, A.; Cutting, B.; Smiesko, M.; Papandreou, G.; Dondoni, A.; Ernst, B. *Eur. J. Org. Chem.* **2012**, 5534.
- (24) Simanek, E. E.; McGarvey, G. J.; Jablonowski, J. A.; Wong, C. H. *Chem. Rev.* **1998**, *98*, 833.
- (25) Kaila, N.; Thomas, B. E. *Med. Res. Rev.* **2002**, *22*, 566.
- (26) Zierke, M.; Smieško, M.; Rabbani, S.; Aeschbacher, T.; Cutting, B.; Allain, F. H. T.; Schubert, M.; Ernst, B. *Journal of the American Chemical Society* **2013**, *135*, 13464.
- (27) Ishida, T. *J. Phys. Chem. B* **2010**, *114*, 3950.
- (28) Thoma, G.; Banteli, R.; Jahnke, W.; Magnani, J. L.; Patton, J. T. *Angewandte Chemie* **2001**, *113*, 3756.

- (29) Khan, A. R.; Parrish, J. C.; Fraser, M. E.; Smith, W. W.; Bartlett, P. A.; James, M. N. *G. Biochemistry* **1998**, *37*, 16839.
- (30) DeLorbe, J. E.; Clements, J. H.; Whiddon, B. B.; Martin, S. F. *ACS Med. Chem. Lett.* **2010**, *1*, 448.
- (31) Zierke, M. PhD Thesis, Chapter 3.1.2, **2015**.
- (32) Herrmann, T.; Güntert, P.; Wüthrich, K. *Journal of Biomolecular NMR* **2002**, *24*, 171.
- (33) Case, D. A.; Cheatham, T. E.; Darden, T.; Gohlke, H.; Luo, R.; Merz, K. M.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R. J. *Journal of Computational Chemistry* **2005**, *26*, 1668.
- (34) Kolb, H. C.; Ernst, B. *Chemi. – Eur. J.* **1997**, *3*, 1571.
- (35) Mallinson, J.; Collins, I. *Future Med. Chem.* **2012**, *4*, 1409.
- (36) Smith, B.; March, J. *March's Advanced Organic Chemistry: Reactions, Mechanisms and Structure (5th Edition)* **2001**, 184.
- (37) Binder, F. P. C., PhD Thesis, Basel, **2011**.
- (38) ter Halle, R.; Bernet, Y.; Billard, S.; Bufferne, C.; Carlier, P.; Delaitre, C.; Flouzat, C.; Humblot, G.; Laigle, J. C.; Lombard, F.; Wilmouth, S. *Org. Process Res. Dev.* **2004**, *8*, 283.
- (39) Using column chromatography on silica with CH₂Cl₂ as eluent, 6 elutes first.
- (40) Ernst, B.; Wagner, B.; Baisch, G.; Katopodis, A.; Winkler, T.; Ohrlein, R. *Can. J. Chem.-Rev.* **2000**, *78*, 892.
- (41) Grindley, T. *Adv Carbohydr Chem Biochem* **1998**, *53*, 17.
- (42) Chayajarus, K.; Chambers, D. J.; Chughtai, M. J.; Fairbanks, A. J. *Org. Lett.* **2004**, *6*, 3797.
- (43) Maestro, version 9.3, Schrödinger, LLC, New York, NY, **2012**.
- (44) MacroModel, version 9.9, Schrödinger, LLC, New York, NY, **2012**.
- (45) Desmond Molecular Dynamics System, version 3.1, D. E. Shaw Research, New York, NY, **2012**.
- (46) Bowers, K. J.; Chow, E.; Xu, H.; Dror, R. O.; Eastwood, M. P.; Gregersen, B. A.; Klepeis, J. L.; Kolossvary, I.; Moraes, M. A.; Sacerdoti, F. D.; Salmon, J. K.; Shan, Y.; Shaw, D. E. Proceedings of the 2006 ACM/IEEE Conference on Supercomputing (SC06), Tampa, FL, 11 to 17 November **2006** (ACM Press, New York, **2006**).
- (47) Fesik, S. W.; Gampe, R. T.; Zuiderweg, E. R. P. *Journal of the American Chemical Society* **1989**, *111*, 770.

- (48) Goddard, T. D.; Kneller, D. G. *SPARKY 3*, University of California, San Francisco.
- (49) Markley, J.; Bax, A.; Arata, Y.; Hilbers, C. W.; Kaptein, R.; Sykes, B.; Wright, P.; Wüthrich, K. *Journal of Biomolecular NMR* **1998**, *12*, 1.
- (50) Güntert, P.; Mumenthaler, C.; Wüthrich, K. *Journal of Molecular Biology* **1997**, *273*, 283.
- (51) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. *Journal of Computational Chemistry* **2004**, *25*, 1157.
- (52) Bashford, D.; Case, D. A. *Annual Review of Physical Chemistry* **2000**, *51*, 129.
- (53) Zihlmann, P.; Sager, C.; Ernst, B. *Manuscript in preparation*.
- (54) Mesch, S.; Lemme, K.; Koliwer-Brandl, H.; Strasser, D. S.; Schwardt, O.; Kelm, S.; Ernst, B. *Carbohydrate Research* **2010**, *345*, 1348.
- (55) Bitsch, F.; Aichholz, R.; Kallen, J.; Geisse, S.; Fournier, B.; Schlaeppli, J. M. *Analytical Biochemistry* **2003**, *323*, 139.

3.2.1.1 Supporting Information

“Bridging the Gap” - Adjusting the Acid Pharmacophore in sLe^x Mimics by Ring Closing Metathesis

*Mirko Zierke,[#] Martin Smiesko,[#] Norbert Varga,[#] Florian P. C. Binder,[#] Mario Schubert,[¶]
Roland Preston,[#] John T. Patton,[&] John L. Magnani,[&] Pascal Zihlmann,[#] Tobias
Mühlethaler,[#] and Beat Ernst^{#*}*

[#] University of Basel, Klingelbergstr. 50, CH-4056 Basel, Switzerland

[¶] Institute of Molecular Biology and Biophysics, ETH Zürich, CH-8093 Zürich, Switzerland

[&] GlycoMimetics Inc., Gaithersburg, MD20878, USA

^{*} Corresponding Author; E-mail: beat.ernst@unibas.ch

Table of content:

Supplementary Methods	130
Supplementary Figures	179
Supplementary Tables	185
Supplementary References	186

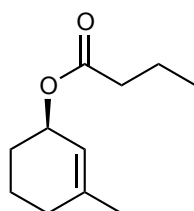
Supplementary Methods

Experimental

General methods. Commercial materials (Sigma-Aldrich) were used without further purification, solvents were reagent grade (Acros). CH₂Cl₂ and MeOH were dried by passing through an Al₂O₃ (Fluka, type 5016 A basic) column. DMF extra dry (Acros) was used as is. All reactions were performed in oven dried glassware under an atmosphere of argon.

¹H and ¹³C NMR spectra were recorded on a Bruker Avance DMX-500 at room temperature. Chemical shifts are reported in ppm and referenced to TMS using residual solvent peaks.^[1] For complex molecules the following prefixes were used: Fuc (fucose), Gal (galactose) and GlcNAc (*N*-acetyl glucosamine). The coupling constants (*J*) are reported in Hertz (Hz). Analytical TLC was performed on Merck silica gel 60 F₂₅₄ glass plates and visualized by UV light and charring with a molybdate solution (a 0.02 M solution of ammonium cerium sulfate dihydrate and ammonium molybdate tetrahydrate in aq. 10% H₂SO₄) by heating for 5 min at 140°C. Column chromatography was performed on a CombiFlash Companion (Teledyne-ISCO, Inc.) using RediSep® normal phase disposable flash columns (silica gel). Reversed phase chromatography was carried out with LiChroprep®RP-18 (Merck, 40-63 μm). Optical rotations were determined on a Perkin-Elmer Polarimeter 341. Low resolution mass spectra were measured on a Waters micromass ZQ. High resolution mass spectra (HRMS) were obtained on a micrOTOF spectrometer (Bruker Daltonics, Germany) equipped with a TOF hexapole detector. MALDI-TOF and ESI-MS glycoprotein analyses were recorded by the Functional Genomic Center Zurich (FGCZ). VIVASPIN® 500 ultrafiltration tubes with 10000 MWCO, 10 kDa molecular weight cutoff PES membrane, and ZelluTrans/Roth dialysis membranes MWCO 8000-10000 were used for glycoprotein concentration and dialysis

(*R*)-3-Methylcyclohex-2-en-1-yl butyrate (9)

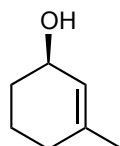


Immobilized Novozyme 435 (222 mg, 444 U, EC 232-619-9) was added to a solution of **8** (10.0 g, 89 mmol) and vinyl butyrate (22.6 mL, 20.3 g, 178 mmol) in heptane (90 mL). The

mixture was stirred at 23°C and 200 rpm. After 2 h 25 min the mixture was filtered and volatiles were evaporated at 60°C and 10 mbar to give 12 g of a clear oil. Column chromatography on silica (CH₂Cl₂) yielded pure **9** (7.50 g, 41 mmol, 46%).

$[\alpha]_D^{22} +168.7$ (*c* 9.28, CHCl₃); ¹H NMR (500.1 MHz, CDCl₃): δ = 5.44 (m, 1H, H-2), 5.23 (m, 1H, H-1), 2.24 (t, *J* = 7.4 Hz, 2H, COCH₂CH₂CH₃), 2.02-1.84 (m, 2H, H-4, H-4'), 1.81-1.56 (m, 9H, H-5, H-5', H-6, H-6', -CH₃, COCH₂CH₂CH₃), 0.92 (t, *J* = 7.4 Hz, 3H, COCH₂CH₂CH₃); ¹³C NMR (125.8 MHz, CDCl₃): δ 173.5 (COCH₂CH₂CH₃), 141.0 (C-3), 120.2 (C-2), 68.6 (C-1), 36.7 (COCH₂CH₂CH₃), 30.0 (C-4), 28.1 (C-6), 23.8 (-CH₃), 19.1, 18.7 (2C, C-5, COCH₂CH₂CH₃), 13.7 (COCH₂CH₂CH₃); MS (ESI) *m/z*: Calcd for C₁₁H₁₈NaO₂⁺ [M+Na]⁺: 205.12; found: 204.83; elemental analysis calcd (%) for C₁₁H₁₈O₂ (182.26): C 72.49, H 9.95; found: C 72.87, H 9.65.

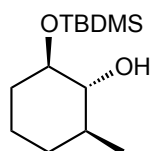
(*R*)-seudenol (**10**)



A solution of NaOH in H₂O (10.3 mL, 4N) was slowly added to a solution of seudenol butyrate **9** (3.50 g, 19 mmol) in MeOH (30 mL) at 0°C and stirred at 0°C for 5 h. The mixture was diluted with H₂O (25 mL) and extracted with CH₂Cl₂ (3 x 20 mL). The combined organic layers were washed with brine (25 mL) and dried over Na₂SO₄. Filtration and evaporation of volatiles (200 mbar, 40°C) gave spectroscopically pure (*R*)-seudenol (**10**) (1.81 g, 16.0 mmol, 84%) as a clear oil, which was directly used in the next step.

$[\alpha]_D +91.7$ (*c* 0.74, CHCl₃); HPLC: 97.5% ee, 98% purity; NMR data were in accordance with literature.⁷

(1*R*,2*R*,3*S*)-1-[(*tert*-butyldimethylsilyl)oxy]-3-methylcyclohexan-2-ol (**12**)

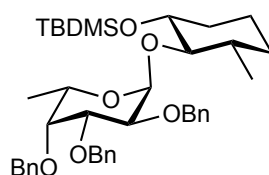


Imidazol (4.40 g, 65 mmol) was added to a solution of (*R*)-seudenol **10** (3.50 g, 31 mmol), DMAP (cat.), and TBSCl (7.31 g, 48 mmol) in anhydrous CH₂Cl₂ (65 mL) at r.t. under argon. After stirring for 15 h, the reaction mixture was quenched with satd. aq. NaHCO₃ (50 mL) and extracted with CH₂Cl₂ (20 mL). The organic layer was washed with aq. HCl (20

mL, 0.01 N), satd. aq. NaHCO₃ (20 mL), and brine (20 mL) and dried over Na₂SO₄. Filtration and evaporation of volatiles (200 mbar, 40°C) gave the TBS ether **11** as clear oil. A solution of BH₃·THF (60 mL, 1M in THF) was slowly added to a solution of the crude TBS ether (**11**) in anhydrous THF (60 mL) under argon at 0°C. After stirring for 2 h at rt, the reaction mixture was cooled to 0°C again and aq. NaOH (180 mL, 3N) followed by aq. H₂O₂ (180 mL, 30%) were slowly added *via* dropping funnel (**CAUTION**: strong gas development). The mixture was stirred at 0°C for 1 h, subsequently acidified to pH 3 by slow addition of 10% aq. HCl *via* dropping funnel (**CAUTION**: strong gas development) and extracted with DCM (2 x 300 mL). The extracts were dried over Na₂SO₄, filtered, concentrated (100 mbar, 40°C) and purified by column chromatography (petroleum ether/Et₂O 98.5/1.5) to yield pure **12** (6.20 g, 25 mmol, 81%) as clear oil.

$[\alpha]_D^{22}$ - 13.7 (*c* 3.14, CHCl₃); ¹H NMR (500.1 MHz, CDCl₃): δ = 3.34 (m, 1H, H-1), 2.92 (dd, *J* = 8.5, 10.0 Hz, 1H, H-2), 2.47 (s, 1H, OH), 1.81 (m, 1H, H-6), 1.63-1.56 (m, 2H, H-4, H-5), 1.41 (m, 1H, H-3), 1.34-1.99 (m, 2H, H-5', H-6'), 1.04-0.92 (m, 4H, H-4', Me), 0.91-0.83 (m, 9H, SiC(CH₃)₃), 0.07 (s, 3H, SiCH₃), 0.05 (s, 3H, SiCH₃); ¹³C NMR (125.8 MHz, CDCl₃): δ = 81.0 (C-2), 77.0 (C-1), 37.0 (C-3), 33.4, 33.9 (2C, C-4, C-6), 25.9 (3C, SiC(CH₃)₃), 23.6 (C-5), 18.5 (-CH₃), 18.1 (SiC(CH₃)₃), -3.9, -4.6 (SiCH₃); HR-MS (ESI) *m/z*: calcd for C₁₃H₂₈NaO₂Si⁺ [M+Na]⁺: 267.1751; found: 267.1752.

[(1R,2R,3S)-1-(*tert*-Butyldimethylsilyloxy)-3-methyl-cyclohex-2-yl] 2,3,4-tri-O-benzyl-6-deoxy- α -L-galactopyranoside (14**)**

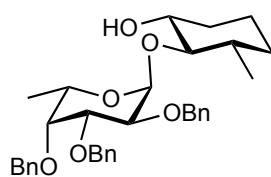


Ethylthio fucoside **13**¹² (3.90 g, 8.15 mmol) and TBAB (4.00 g, 12.4 mmol) were dried at high vacuum overnight. Powdered activated molecular sieves 4 Å (5.0 g), compound **12** (1.00 g, 4.09 mmol), 2,6-di-*tert*-butyl-4-methylpyridine (2.50 g, 12.2 mmol), anhydrous DCM (35 ml) and DMF (5 ml) were added and the mixture was stirred for 4 h at rt under argon. CuBr₂ (2.70 g, 12.1 mmol), dried under high vacuum overnight at 70°C, was added and the resulting dark mixture was stirred at rt under argon. After completion of the reaction (17 h), the solution was filtered through a pad of celite and the filtrate was washed with a solution of satd. aq. NH₄Cl and aqueous NH₃ (9/1 (v/v), 2 x 200 mL) and brine (100 mL).

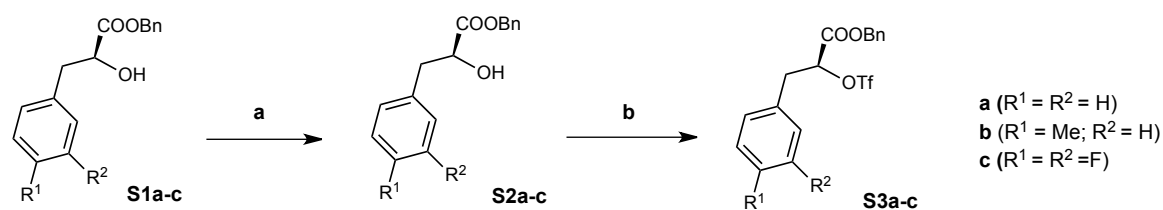
The aqueous layers were extracted with DCM (2 x 200 mL) and the combined organic layers were dried (Na_2SO_4) and concentrated. Column chromatography on silica (petroleum ether/EtOAc 98/2 to 97/3) gave the pseudodisaccharide **14** as clear oil (2.34 g, 3.54 mmol, 87%).

$[\alpha]_{\text{D}} - 53.7$ (*c* 2.1, CHCl_3); $^1\text{H NMR}$ (500.1 MHz, CDCl_3): $\delta = 7.47\text{--}7.27$ (m, 15H, 3 C_6H_5), 5.16 (d, $^3J = 3.4$ Hz, 1H, Fuc H-1), 5.03 (A of AB, $^2J = 11.6$ Hz, 1H, CH_2Ph), 4.89, 4.85, 4.78, 4.76 (4d, $^2J = 11.8$ Hz, 4H, CH_2Ph), 4.70 (B of AB, $^2J = 11.6$ Hz, 1H, CH_2Ph), 4.26 (q, $^3J = 6.4$ Hz, 1H, Fuc H-5), 4.10 (dd, $^3J = 3.4$, 10.2 Hz, 1H, Fuc H-2), 4.05 (dd, $^3J = 2.6$, 10.2 Hz, 1H, Fuc H-3), 3.75 (m, 1H, H-1), 3.70 (m, 1H, Fuc H-4), 3.36 (t, $^3J = 6.4$ Hz, H-2), 1.88–1.77 (m, 2H, H-3, H-6a), 1.76–1.68 (m, 2H, H-4a, H-5a), 1.43 (m, 1H, H-6b), 1.34–1.11 (m, 8H, Fuc-H6, $-\text{CH}_3$, H-4b, H-5b), 0.93 (s, 9H, $\text{SiC}(\text{CH}_3)_3$), 0.09 (s, 3H, SiCH_3), 0.06 (s, 3H, SiCH_3); $^{13}\text{C NMR}$ (125.8 MHz, CDCl_3): δ 139.1, 138.9, 138.8, 128.4, 128.3, 128.2, 127.6, 127.5 (18C, 3 C_6H_5), 96.8 (Fuc C-1), 81.5 (C-2), 79.3 (Fuc C-3), 78.2 (Fuc C-4), 76.7 (Fuc C-2), 74.9 (CH_2Ph), 73.6 (CH_2Ph), 73.3 (C-1), 73.0 (CH_2Ph), 66.4 (Fuc C-5), 35.6 (C-3), 33.2 (C-6), 31.1 (C-5), 26.1 (3C, $\text{SiC}(\text{CH}_3)_3$), 19.8 (C-4), 18.9 (Fuc C-6), 18.2 ($\text{SiC}(\text{CH}_3)_3$), 17.0 (CH_3), -3.9, -5.0 (2C, SiCH_3); HR-MS (ESI) *m/z*: Calcd for $\text{C}_{40}\text{H}_{56}\text{NaO}_6\text{Si}^+$ [$\text{M}+\text{Na}$] $^+$: 683.3738; found: 683.3740.

[(1R,2R,3S)-1-Hydroxy-3-methyl-cyclohex-2-yl] 2,3,4-tri-O-benzyl-6-deoxy- α -L-galactopyranoside (15)

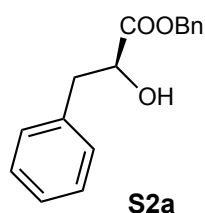


Compound **14** (2.10 g, 3.18 mmol) was dissolved in a solution of TBAF in THF (20 mL, 1M) and stirred for 24 h at rt. The solution was diluted with DCM (50 mL) and washed with H_2O (100 mL). The aq. layer was extracted with DCM (2 x 50 mL) and the combined organic layers were dried (Na_2SO_4) and concentrated. Column chromatography on silica (petroleum ether/EtOAc 80/20) gave **15** as white solid (1.74 g, 3.18 mmol, quant.); $[\alpha]_{\text{D}} - 42.0$ (*c* 0.45, CHCl_3); NMR data were in accordance with literature.⁶

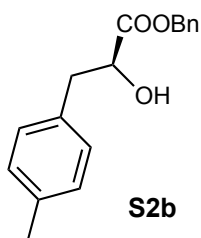


Scheme S1. Synthesis of triflate derivatives **S3a-c** a) BnOH, TsOH, benzene (**a**: 88%); or (i) NaNO₂, H₂SO₄, H₂O (ii) BnOH, TsOH, benzene (**b**: 57%, **c**: 50%); b) Tf₂O, lutidine, DCM (**a**: 98%, **b**: 97%, **c**: 95%).

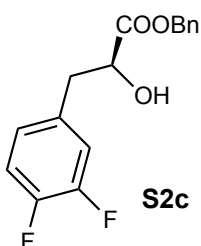
(*R*)-Benzyl 2-hydroxy-3-phenylpropanoate (**S2a**)



To a solution of 2-hydroxy-3-phenyl-2-(*R*)-propanoic acid **S1a** (1.94 g, 11.6 mmol) in benzene (25 mL) benzyl alcohol (1.62 mL, 15.6 mmol) and *p*-toluenesulfonic acid (200 mg, 1.2 mmol) were added. The solution was refluxed at 80°C until completion of the reaction (3 h, TLC: DCM/MeOH, 19:1 and petroleum ether/EtOAc, 4:1). The reaction mixture was allowed to cool down to rt, diluted with EtOAc (50 mL) and washed with satd. aq. NaHCO₃ (20 mL), water (20 mL) and brine (10 mL). The organic layer was dried over Na₂SO₄, filtered and solvent was removed *in vacuo*. The crude product was purified by flash chromatography (petroleum ether/EtOAc, gradient 0-25%) to afford **S2a** as colorless wax (2.62 g, 10.2 mmol, 88%). R_f = (petroleum ether/EtOAc, 4:1) 0.50; $[\alpha]_D^{22} +53.5$ (c 0.69, CHCl₃); ¹H NMR (500.1 MHz, CDCl₃): δ = 7.54 - 7.06 (m, 10H, Ar-H), 5.18 (s, 2H, Ar-CH₂), 4.50 (br s, 1H, H-2), 3.22 - 3.06 (m, 1H, H-3), 3.09 - 2.90 (m, 1H, H-3'); ¹³C NMR (125.8 MHz, CDCl₃): δ = 174.0 (C-1), 136.1, 135.0, 129.5, 128.7, 128.7, 128.6, 128.4 (8C, Ar-C); 71.3 (C-2); 67.5 (Ar-CH₂); 40.5 (C-3); ESI-MS: m/z : Calcd for C₁₆H₁₆NaO₃ [M+Na]⁺: 279.3, found: 278.9.

(R)-Benzyl 2-hydroxy-3-(4-methylphenyl)-propanoate (S2b)

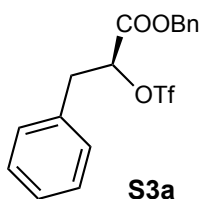
A solution of sodium nitrite (11.7 g, 0.17 mol) in water (40 mL) was added dropwise to a stirred and ice-cooled solution of *p*-methyl-D-phenylalanine **S1b** (5.10 g, 28.4 mmol) in 1 M H₂SO₄ (50 mL) over 1 h, and the resulting mixture was stirred for 24 h at rt. The reaction mixture was extracted with Et₂O (3 x 150 mL), the combined organic phases were washed with brine (30 mL), dried over Na₂SO₄, filtered and concentrated *in vacuo*. The crude product (4.85 g) was dissolved in benzene (30 mL) and the resulting solution was treated with benzyl alcohol (3.63 mL, 35.0 mmol, 1.3 eq) and *p*-toluenesulfonic acid (463 mg, 2.70 mmol, 0.1 eq). The solution was refluxed until completion of the reaction (5 h, TLC: DCM/MeOH, 19:1 and petroleum ether/EtOAc, 4:1). The reaction mixture was allowed to cool down to rt, diluted with EtOAc (150 mL) and washed with satd. aq. NaHCO₃ (30 mL), water (30 mL) and brine (20 mL). The organic layer was dried over Na₂SO₄, filtered and concentrated *in vacuo*. The crude product was purified by flash chromatography (petroleum ether/EtOAc, gradient 0-25%) to afford **S2b** as colorless wax (4.2 g, 15.16 mmol, 53% over two steps). R_f = (petroleum ether/EtOAc, 4:1) 0.52; [α]_D²² -46.5 (*c* 0.63, CHCl₃); ¹H NMR (500.1 MHz, CDCl₃): δ = 7.40 - 7.30 (m, 5H, Ar-H), 7.05 (dd, *J* = 8.1 Hz, 4H, Ar-H); 5.15 - 5.02 (m, 2H, Ar-CH₂), 4.47 (dd, *J* = 4.7, 6.4 Hz, 1H, H-2), 3.09 (dd, *J* = 4.7, 13.9 Hz, 1H, H-3), 2.95 (dd, *J* = 6.4, 13.9 Hz, 1H, H-3'), 2.31 (s, 3H, Ar-Me); ¹³C NMR (125.8 MHz, CDCl₃): δ = 174.2 (C-1), 136.6, 135.2, 133.0, 129.5, 129.3, 128.8, 128.7 (8C, Ar-C), 71.5 (C-2), 67.5 (Ar-CH₂), 40.2 (C-3), 21.2 (Ar-Me); ESI-MS: *m/z*: Calcd for C₁₇H₁₈NaO₃ [M+Na]⁺: 293.3, found: 292.9.

(R)-Benzyl 2-hydroxy-3-(3,4-difluorophenyl)-propanoate (S2c)

A solution of sodium nitrite (2.05 g, 39.8 mmol) in water (10 mL) was added dropwise to a

stirred and ice-cooled solution of 3,4-difluoro-D-phenylalanine **S1c** (1.00 g, 4.97 mmol) in 1 M H₂SO₄ (10 mL) over 30 min, and the resulting mixture was stirred for 2 d at rt. The reaction mixture was extracted with Et₂O (3 x 50 mL), the combined organic phases were washed with brine (20 mL), dried over Na₂SO₄, filtered and concentrated *in vacuo*. The crude product (790 mg) was dissolved in benzene (10 mL) and the resulting solution was treated with benzyl alcohol (0.55 mL, 5.08 mmol) and *p*-toluenesulfonic acid (62 mg, 0.39 mmol). The solution was refluxed at 80°C until completion of the reaction (4 h, TLC: DCM/MeOH, 19:1 and petroleum ether/EtOAc, 4:1). The reaction mixture was allowed to cool down to rt, diluted with EtOAc (50 mL) and washed with satd. aq. NaHCO₃ solution (20 mL), water (20 mL) and brine (10 mL). The organic layer was dried over Na₂SO₄, filtered and concentrated *in vacuo*. The crude product was purified by flash chromatography (Tol/EtOAc, gradient 0-10%) to afford **S2c** as colorless wax (720 mg, 2.46 mmol, 50% over two steps). R_f = (petroleum ether/EtOAc, 4:1) 0.38; [α]_D²² +61.9 (*c* 0.76, CHCl₃); ¹H NMR (500.1 MHz, CDCl₃): δ = 7.43 - 7.21 (m, 6H, Ar-H), 7.03 - 6.91 (m, 2H, Ar-H), 6.87 - 6.77 (m, 1H, Ar-H), 5.26 - 5.11 (m, 2H, Ar-CH₂), 4.45 (q, *J* = 5.3 Hz, 1H, H-2), 3.07 (dd, *J* = 4.0, 14.1 Hz, 1H, H-3), 2.91 (dd, *J* = 6.2, 14.1 Hz, 1H, H-3'); ¹³C NMR (125.8 MHz, CDCl₃): δ = 173.7 (C-1), 151.3 - 148.2 (m, 2C, ArF-C), 134.8 (Ar-C) 133.1 (m, Ar-C), 128.9, 128.8 (2C, Ar-C), 125.5 (m, Ar-C), 118.4 (d, *J* = 17.1 Hz, Ar-C), 116.9 (d, *J* = 17.1 Hz, Ar-C), 70.8 (C-2); 67.7 (Ar-CH₂); 39.3 (C-3). ESI-MS: *m/z*: Calcd for C₁₆H₁₄F₂NaO₃ [M+Na]⁺: 315.3, found: 314.6.

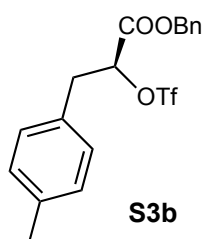
(R)-Benzyl 3-phenyl-2-(trifluoromethyl)sulfonyloxy-propanoate (S3a)



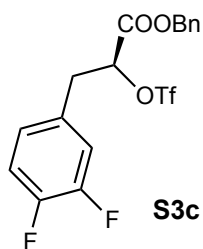
To a stirred solution of **S2a** (2.44 g, 8.35 mmol) and lutidine (1.16 mL, 10.8 mmol) in dry DCM (30 mL) Tf₂O (1.68 mL, 10.0 mmol) was added dropwise at -60°C under argon. The reaction was stirred at -60°C for 20 min, then at -5°C for 40 min. TLC (petroleum ether/EtOAc, 9:1) indicated completion of the reaction and the mixture was quenched by addition of water (10 mL) at 0°C. The organic phase was separated and the water phase was extracted with DCM (50 mL). The combined organic phases were washed with water (20 mL), dried over Na₂SO₄, filtered and concentrated *in vacuo*. The crude product was purified

by flash chromatography using a short silica pad (DCM/petroleum ether, 1:1) to afford (**S3a**) as pale orange oil. (3.17 g, 8.17 mmol, 98%). R_f = (petroleum ether/EtOAc, 9:1) 0.50; $[\alpha]_D^{22} +29.5$ (c 1.33, CHCl_3); $^1\text{H NMR}$ (500.1 MHz, CDCl_3): δ = 7.41 - 7.12 (m, 10H, Ar-H), 5.30 - 5.20 (m, 1H, H-2), 5.24 (s, 2H, Ar- CH_2), 3.34 (dd, J = 3.6, 14.6 Hz, 1H, H-3), 3.21 (dd, J = 8.5, 14.6 Hz, 1H, H-3'); $^{13}\text{C NMR}$ (126 MHz, CDCl_3): δ = 166.5 (C-1), 134.2, 133.2, 129.5, 128.9, 128.8, 128.7, 128.9, 127.8 (8C, Ar-C), 83.8 (C-2), 68.4 (Ar- CH_2), 38.2 (C-3).

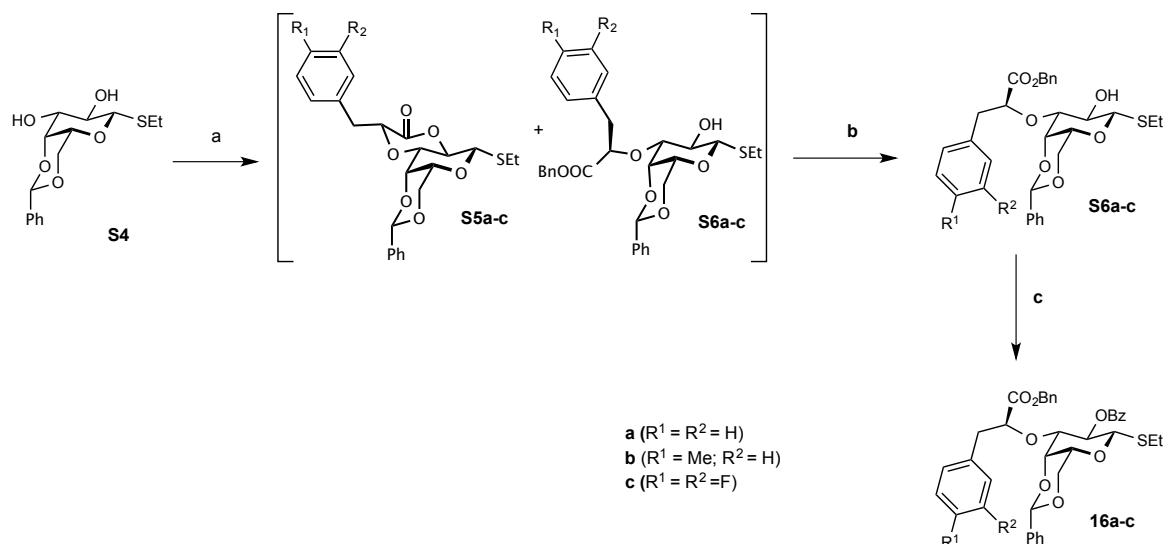
(R)-Benzyl 3-(4-methylphenyl)-2-(trifluoromethyl)sulfonyloxy-propanoate (S3b)



To a stirred solution of **S2b** (4.10 g, 15.2 mmol) and lutidine (2.29 mL, 19.7 mmol) in dry DCM (30 mL) a solution of TF_2O in dry DCM (1 M, 18.2 mL, 18.2 mmol) was added at -80°C under argon. The reaction was allowed to warm up to -20°C (2 h) and stirred at that temperature for an additional 1 h. TLC (petroleum ether/EtOAc, 9:1) indicated completion of the reaction and the mixture was quenched by addition of water (10 mL) at 0°C . The organic phase was separated and the water phase was extracted with DCM (2 x 50 mL). The combined organic phases were washed with water (20 mL), dried over Na_2SO_4 , filtered and concentrated *in vacuo*. The crude product was purified by flash chromatography using a short silica pad (DCM/petroleum ether, 1:1) to afford **S3b** as pale grey oil (5.90 g, 14.7 mmol, 97%). R_f = (petroleum ether/EtOAc, 9:1) 0.45; $[\alpha]_D^{22} -26.1$ (c 1.22, CHCl_3); $^1\text{H NMR}$ (500 MHz, CDCl_3): δ = 7.39 - 7.27 (m, 5H, Ar-H), 7.09 (d, J = 7.9 Hz, 2H, Ar-H), 7.04 (d, J = 7.9 Hz, 2H, Ar-H), 5.26 (dd, J = 4.5, 8.1 Hz, 1H, H-2), 5.23 (s, 2H, Ar- CH_2), 3.29 (dd, J = 4.5, 14.6 Hz, 1H, H-3), 3.18 (dd, J = 8.1, 14.6 Hz, 1H, H-3'), 2.32 (s, 3H, Ar-Me); $^{13}\text{C NMR}$ (126 MHz, CDCl_3): δ = 166.6 (C-1), 137.7, 134.4, 130.1, 129.6, 129.5, 129.0, 128.8, 128.8 (8C, Ar-C), 84.0 (C-2), 68.5 (Ar- CH_2), 38.0 (C-3), 21.2 (Ar-Me).

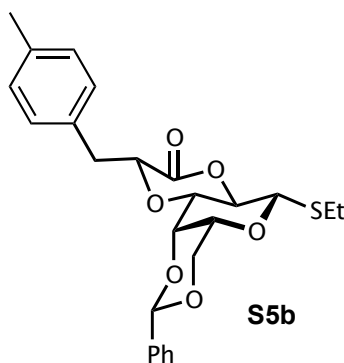
(R)-Benzyl 3-(3,4-difluorophenyl)-2-(trifluoromethyl)sulfonyloxy-propanoate (S3c)

To a stirred solution of **S2c** (670 mg, 2.29 mmol) and lutidine (0.35 mL, 2.97 mmol) in dry DCM (10 mL) a solution of Tf₂O in dry DCM (1 M, 2.75 mL, 2.75 mmol) was added at -80°C under argon. The reaction was allowed to warm up to -20°C (1 h) and stirred at this temperature for an additional 1 h. TLC (petroleum ether/EtOAc, 9:1) indicated completion of the reaction and the mixture was quenched by addition of water (10 mL) at 0°C. The organic phase was separated and the water phase was extracted with DCM (2 x 30 mL). The combined organic phases were washed with water (10 mL), dried over Na₂SO₄, filtered and concentrated *in vacuo*. The crude product was purified by flash chromatography using a short silica pad (DCM/petroleum ether, 1:1) to afford **S3c** as colorless oil (900 mg, 2.12 mmol, 93%). R_f = (petroleum ether/EtOAc, 9:1) 0.40; $[\alpha]_D^{22}$ +31.1 (c 0.62, CHCl₃); ¹H NMR (500 MHz, CDCl₃): δ = 7.41 - 7.22 (m, 5H, Ar-H), 7.09 - 7.00 (m, 1H, Ar-H), 7.00 - 6.92 (m, 1H, Ar-H), 6.89 - 6.83 (m, 1H, Ar-H), 5.32 - 5.14 (m, 3H, H-2, Ar-CH₂), 3.28 (dd, J = 3.6, 14.8 Hz, 1H, H-3), 3.18 (dd, J = 7.9, 14.8 Hz, 1H, H-3'); ¹³C NMR (126 MHz, CDCl₃): δ = 166.0 (C-1); 151.5 - 148.8 (m, 2C, ArF-C), 134.1 (Ar-C) 130.0 (m, Ar-C), 129.1, 128.8 (2C, Ar-C), 125.7 (m, Ar-C), 118.6 (d, J = 17.6 Hz, Ar-C), 117.6 (d, J = 17.4 Hz, Ar-C), 82.9 (C-2), 68.6 (Ar-CH₂), 37.3 (C-3).



Scheme S2. Synthesis of galactose derivatives **22a-c**: a) (i) Bu_2SnO , MeOH; (ii) CsF, DME, **S3a-c** ((lactone) **S5a**: 16%, **S5b**: 25%, **S5c**: 25%; (open) **S6a**: 25%, **S6b**: 0%, **S6c**: 5%); b) BnOH, DMAP (cat.) (**S6a**: 73%, **S6b**: 80%, **S6c**: 65%); c) BzCl, py, DMAP (cat.) (**16a**: 87%, **16b**: 72%, **16c**: 64%).

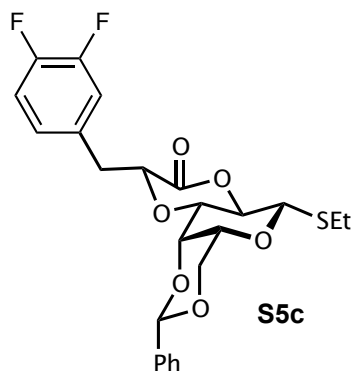
Ethyl 4,6-*O*-benzylidene-2,3-*O*-((*R*)-3-(4-methylphenyl)-1-oxopropan-1,2-diyl)-1-thio- β -D-galactopyranoside (S5b**)**



Thiogalactoside **S4** (50 mg, 0.16 mmol) and Bu_2SnO (60 mg, 0.24 mmol) were dissolved in dry MeOH (3 mL) and refluxed at $65^\circ C$ under argon for 4 h. The solvent was removed under reduced pressure and the crude product was dried in high vacuum for 16 h. The residue was dissolved in dry DME (3 mL) under argon and a solution of triflate **S3b** (193 mg, 0.48 mmol) in dry DME (2 mL) was added followed by the addition of CsF (73 mg, 0.48 mmol, dried in high vacuum at $100^\circ C$ for 16 h). The turbid solution was stirred for 3 h after which TLC (petroleum ether/EtOAc, 9:1, DCM/MeOH, 19:1) indicated still presence of starting material **S4** but no triflate **S3b** so another portion of **S3b** (128 mg, 0.32 mmol) was added and the mixture was stirred for additional 3 h. A solution of KF (10%) in aq. KH_2PO_4 (1 N, 10 mL) was added, and after stirring for 1 h, the reaction was extracted with

DCM (3 x 20 mL). The combined organic phases were washed with brine (10 mL) dried over Na₂SO₄ and the solvents were removed *in vacuo*. The crude product was purified by flash chromatography (petroleum ether/EtOAc, gradient 0-40%) to afford **S5b** as colorless wax (22.6 mg, 0.04 mmol, 25%). R_f = (petroleum ether/EtOAc, 7:3) 0.28; [α]_D²² -44.4 (c 1.25, CHCl₃); ¹H NMR (500 MHz, CDCl₃): δ = 7.48 - 7.42 (m, 2H, Ar-H), 7.38 - 7.31 (m, 3H, Ar-H), 7.19 (d, *J* = 7.6 Hz, 2H, Ar-H), 7.11 (d, *J* = 7.6 Hz, 2H, Ar-H), 5.52 (s, 1H, Ph-CH), 4.85 (t, *J* = 5.2 Hz, 1H, Lac-H2), 4.62 (t, *J* = 9.6 Hz, 1H, Gal-H2), 4.35 - 4.28 (m, 2H, Gal-H1, Gal-H6), 4.23 (s, 1H, Gal-H4), 4.02 (d, *J* = 12.4 Hz, 1H, Gal-H6'), 3.35 (s, 1H, Gal-H5), 3.30 - 3.23 (m, 2H, Gal-H3, Lac-H3), 3.16 (dd, *J* = 4.5, 14.5 Hz, Lac-H3'), 2.86 - 2.70 (m, 2H, SCH₂CH₃), 2.33 (s, 3H, Ar-Me), 1.32 (t, *J* = 7.5 Hz, 3H, SCH₂CH₃); ¹³C NMR (126 MHz, CDCl₃): δ = 168.5 (Lac-C1), 137.3, 136.7, 133.3, 129.6, 129.3, 129.2, 128.3, 126.4 (8C, Ar-C), 101.3 (Ph-CH), 81.9 (Gal-C1), 75.4 (Lac-C2), 73.6 (Gal-C4), 73.4 (Gal-C2, Gal-C3), 69.9 (Gal-C5), 69.2 (Gal-C6), 37.7 (Lac-C3), 23.5 (SCH₂CH₃), 21.1 (Ar-Me), 15.0 (SCH₂CH₃); ESI-MS: *m/z*: Calcd for C₂₅H₂₈NaO₆S [M+Na]⁺: 479.5, found: 479.1.

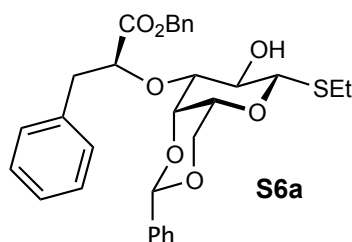
Ethyl 4,6-*O*-benzylidene-2,3-*O*-((*R*)-3-(3,4-difluorophenyl)-1-oxopropan-1,2-diyl)-1-thio-β-D-galactopyranoside (5c)



Thiogalactoside derivative **4** (208 mg, 0.66 mmol) and Bu₂SnO (247 mg, 1.00 mmol) were dissolved in dry MeOH (10 mL) and refluxed at 65°C under argon for 4 h. The solvent was removed under reduced pressure and the crude product was dried in high vacuum for 16 h. This residue was dissolved in dry DME (4 mL) under argon and a solution of triflate **3c** (533 mg, 1.33 mmol) in dry DME (6 mL) was added followed by the addition of CsF (302 mg, 2.00 mmol), dried in high vacuum at 100°C for 16 h. The turbid solution was stirred for 3 h after which TLC (petroleum ether/EtOAc, 9:1, DCM/MeOH, 19:1) indicated still presence of starting material **4** but no triflate **3c** so another portion of **3c** (266 mg, 0.66 mmol) was added and the mixture was stirred for additional 16 h. A solution of KF (10%) in

aq. KH_2PO_4 (1 N, 10 mL) was added, and after stirring for 1 h, the reaction was extracted with DCM (3 x 50 mL). The combined organic phases were washed with brine (10 mL), dried over Na_2SO_4 and the solvents were removed *in vacuo*. The crude product was purified by flash chromatography (petroleum ether/EtOAc, gradient 0-50%) to afford **S5c+S6c** as colorless wax (94 mg, 0.196 mmol, 30%) containing 17 mol% of **S5c**. R_f = (petroleum ether/EtOAc, 3:2) 0.30; ^1H NMR (500.1 MHz, CDCl_3): δ = 7.49 - 7.42 (m, 2H, Ar-H), 7.41 - 7.32 (m, 3H, Ar-H), 7.20 - 6.98 (m, 3H, Ar-H), 5.55 (s, 1H, Ph-CH), 4.80 (br s, 1H, Lac-H2), 4.70 (t, J = 9.6 Hz, 1H, Gal-H2), 4.44 (d, J = 9.5 Hz, 1H, Gal-H1), 4.35 (d, J = 12.6 Hz, 1H, Gal-H6), 4.30 (s, 1H, Gal-H4), 4.05 (d, J = 12.5 Hz, 1H, Gal-H6'), 3.50 - 3.40 (m, 2H, Gal-H5, Gal-H3), 3.25 - 3.12 (m, 2H, Lac-H3), 2.88 - 2.71 (m, 2H, SCH_2CH_3), 1.33 (t, J = 7.4 Hz, 3H, SCH_2CH_3); ^{13}C NMR (125.8 MHz, CDCl_3): δ = 168.2 (Lac-C1), 153.0 - 148.8 (m, 2C, ArF-C), 137.3 (Ar-C) 133.3 (m, Ar-C), 129.3 (Ar-C), 128.6 (m, Ar-C), 128.3, 126.4 (Ar-C), 125.8 (m, Ar-C), 118.7 (d, J = 17.2 Hz, Ar-C), 117.1 (d, J = 16.8 Hz, Ar-C), 101.3 (Ph-CH), 81.9 (Gal-C1), 74.5 (Lac-C2), 74.1 (Gal-C3), 73.7 (Gal-C4), 73.2 (Gal-C2), 69.9 (Gal-C5), 69.2 (Gal-C6), 36.7 (Lac-C3), 23.7 (SCH_2CH_3), 15.0 (SCH_2CH_3); ESI-MS: m/z : Calcd for $\text{C}_{24}\text{H}_{24}\text{F}_2\text{NaO}_6\text{S}$ $[\text{M}+\text{Na}]^+$: 501.5, found: 501.1.

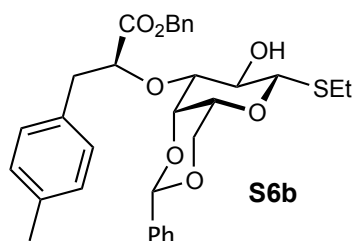
Ethyl 4,6-O-benzylidene-3-O-[(R)-1-(benzyloxy)-1-oxo-3-phenylpropan-2-yl]-1-thio- β -D-galactopyranoside (S6a)



Thiogalactoside **S4** (482 mg, 1.544 mmol, 1 eq) and Bu_2SnO (575 mg, 2.31 mmol) were dissolved in dry MeOH (15 mL) and refluxed at 65°C under argon for 4 h. The solvent was removed under reduced pressure and the crude product was dried in high vacuum for 16 h. The residue was dissolved in dry DME (10 mL) under argon and a solution of triflate **3a** (2 g, 5.14 mmol) in dry DME (10 mL) was added followed by the addition of CsF (700 mg, 4.62 mmol, dried in high vacuum at 100°C for 16 h). The turbid solution was stirred for 5 h after which TLC (petroleum ether/EtOAc, 9:1, DCM/MeOH, 19:1) indicated still presence of starting material **S4** but no triflate **S3a** so another portion of **S3a** (1 g, 2.57 mmol) was added and the mixture was stirred for additional 16 h. A solution of KF (10%) in aq. KH_2PO_4 (1 N, 10 mL) was added, and after stirring for 1 h, the reaction was extracted with

DCM (3 x 70 mL). The combined organic phases were washed with brine (20 mL) dried over Na_2SO_4 and the solvents were removed *in vacuo*. The crude product was purified by flash chromatography (petroleum ether/EtOAc, gradient 0-50%) to afford a 1:1.6 mixture of **S5a** (lactone) and **S6a** (open) (330 mg, 41%), which was used in the following reaction without further purification. ESI-MS: m/z : Calcd for **S5a** (lactone) $\text{C}_{24}\text{H}_{26}\text{NaO}_6\text{S}$ $[\text{M}+\text{Na}]^+$: 465.5, found: 465.2; Calcd for **S6a** (open) $\text{C}_{31}\text{H}_{34}\text{NaO}_7\text{S}$ $[\text{M}+\text{Na}]^+$: 573.7, found: 573.1. To a solution of the mixture (330 mg) obtained in the previous reaction in toluene (2 mL) benzyl alcohol (1 mL) and DMAP (cat.) were added. The reaction was heated up to 60°C for 4 h, then the solvents were removed *in vacuo* and the residue was purified by flash chromatography (petroleum ether/EtOAc, gradient 0-50%) to afford pure **S6a** as white foam (260 mg, 0.47 mmol, 30% starting from **S4**). R_f (petroleum ether/EtOAc, 7:3) 0.33; $[\alpha]_D^{22}$ -44.9 (c 0.31, CHCl_3); ^1H NMR (500.1 MHz, CDCl_3): δ = 7.52 - 7.47 (m, 2H, Ar-H), 7.41 - 7.31 (m, 5H, Ar-H), 7.28 - 7.20 (m, 8H, Ar-H), 5.34 (s, 1H, Ph-CH), 5.10 (d, J = 12.1 Hz, 1H, Ar- CH_2), 5.04 (d, J = 12.1 Hz, 1H, Ar- CH_2), 4.62 (dd, J = 4.8, 8.4 Hz, 1H, Lac-H2), 4.26 - 4.16 (m, 3H, Gal-H1, Gal-H4, Gal-H6), 3.88 - 3.82 (m, 2H, Gal-H2, Gal-H6'), 3.37 (dd, J = 3.4, 9.2 Hz, 1H, Gal-H3), 3.29 (br s, 1H, Gal-H5), 3.14 - 3.03 (m, 2H, Lac-H3), 2.80 - 2.62 (m, 2H, SCH_2CH_3), 2.09 (d, J = 1.9 Hz, 1H, OH), 1.28 (t, J = 7.5 Hz, 3H, SCH_2CH_3); ^{13}C NMR (125.8 MHz, CDCl_3): δ = 172.3 (Lac-C1), 138.0, 137.0, 135.4, 129.6, 128.9, 128.6, 128.5, 128.5, 128.5, 128.3, 128.1, 126.9, 126.4 (12C, Ar-C), 100.9 (Ph-CH), 85.2 (Gal-C1), 81.8 (Gal-C3), 80.4 (Lac-C2), 75.2 (Gal-C4), 70.2 (Gal-C5), 69.2 (Gal-C6), 68.7 (Gal-C2), 66.8 (Ph- CH_2), 39.2 (Lac-C3), 23.0 (SCH_2CH_3), 21.1 (Ar-Me), 15.2 (SCH_2CH_3); ESI-MS: m/z : Calcd for $\text{C}_{31}\text{H}_{34}\text{NaO}_7\text{S}$ $[\text{M}+\text{Na}]^+$: 573.7, found: 573.1.

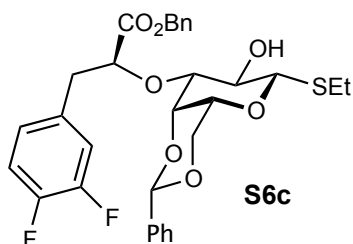
Ethyl 4,6-*O*-benzylidene-3-*O*-[(*R*)-1-(benzyloxy)-1-oxo-3-(4-methylphenyl)propan-2-yl]-1-thio- β -D-galactopyranoside (S6b**)**



To a solution of **S5b** (38 mg, 0.083 mmol) in benzyl alcohol (0.5 mL) DMAP (cat.) was added. The reaction was heated to 60°C for 4 h, then the reaction mixture was purified by flash chromatography (petroleum ether/EtOAc, gradient 0-50%) to afford **S6b** as colorless

wax (30 mg, 0.053 mmol, 65%). R_f = (petroleum ether/EtOAc, 7:3) 0.30; $[\alpha]_D^{22}$ -79.3 (c 0.25, CHCl_3); $^1\text{H NMR}$ (500 MHz, CDCl_3): δ = 7.53 - 7.45 (m, 2H, Ar-H), 7.41 - 7.30 (m, 6H, Ar-H), 7.12 (d, J = 7.4 Hz, 2H, Ar-H), 7.04 (d, J = 7.4 Hz, 2H, Ar-H), 5.34 (s, 1H, Ph-CH), 5.09 (d, J = 12.1 Hz, 1H, Ar- CH_2), 5.03 (d, J = 12.1 Hz, 1H, Ar- CH_2), 4.66 - 4.54 (m, 1H, Lac-H2), 4.26 - 4.20 (m, 2H, Gal-H1, Gal-H6), 4.18 (s, 1H, Gal-H4), 3.91 - 3.82 (m, 2H, Gal-H2, Gal-H6'), 3.32 (d, J = 9.2 Hz, 1H, Gal-H3), 3.29 (s 1H, Gal-H5), 3.10 - 2.98 (m, 2H, Lac-H3), 2.81 - 2.71 (m, 2H, SCH_2CH_3), 2.30 (s, 3H, Ar-Me), 1.28 (t, J = 7.5 Hz, 3H, SCH_2CH_3); $^{13}\text{C NMR}$ (125.8 MHz, CDCl_3): δ = 172.4 (Lac-C1), 138.3, 136.3, 135.4, 133.8, 129.5, 129.1, 128.9, 128.6, 128.5, 128.5, 128.1, 126.4 (12C, Ar-C), 100.3 (Ph-CH), 85.1 (Gal-C1), 81.9 (Gal-C3), 80.6 (Lac-C2), 75.1 (Gal-C4), 70.2 (Gal-C5), 69.2 (Gal-C6), 68.6 (Gal-C2), 66.8 (Ph- CH_2), 38.8 (Lac-C3), 23.0 (SCH_2CH_3), 21.1 (Ar-Me), 15.2 (SCH_2CH_3); ESI-MS: m/z : Calcd for $\text{C}_{32}\text{H}_{36}\text{NaO}_7\text{S}$ $[\text{M}+\text{Na}]^+$: 587.7, found: 587.2.

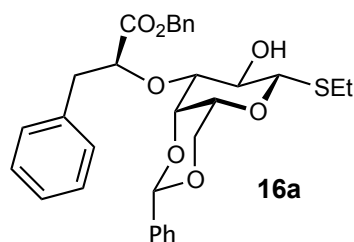
Ethyl 4,6-*O*-benzylidene-3-*O*-[(*R*)-1-(benzyloxy)-1-oxo-3-(3,4-difluorophenyl)propan-2-yl]-1-thio- β -D-galactopyranoside (S6c**)**



To a solution of **S5c** (90.0 mg, 0.188 mmol) in benzyl alcohol (1 mL) DMAP (cat.) was added. The reaction was heated to 60°C for 3 h, then the reaction mixture was purified by flash chromatography (petroleum ether/EtOAc, gradient 0-50%) to afford **S6c** as colorless wax (88 mg, 0.15 mmol, 80%). R_f = (petroleum ether/EtOAc, 3:2) 0.32; $[\alpha]_D^{22}$ -35.8 (c 0.35, CHCl_3); $^1\text{H NMR}$ (500.1 MHz, CDCl_3): δ = 7.52 - 7.47 (m, 2H, Ar-H), 7.40 - 7.32 (m, 6H, Ar-H), 7.31 - 7.24 (m, 2H, Ar-H), 7.12 - 7.05 (m, 1H, Ar-H), 6.92 - 6.82 (m, 2H, Ar-H), 5.42 (s, 1H, Ph-CH), 5.11 (d, J = 12.0 Hz, 1H, Ar- CH_2), 5.05 (d, J = 12.0 Hz, 1H, Ar- CH_2), 4.75 (t, J = 5.7 Hz, 1H, Lac-H2), 4.29 - 4.18 (m, 3H, Gal-H1, Gal-H4, Gal-H6a), 3.93 - 3.87 (m, 2H, Gal-H2, Gal-H6b), 3.47 (d, J = 9.3 Hz, 1H, Gal-H3), 3.34 (s 1H, Gal-H5), 3.10 - 2.95 (m, 2H, Lac-H3), 2.82 - 2.64 (m, 2H, SCH_2CH_3), 2.31 (br s, 1H, OH), 1.30 (t, J = 7.4 Hz, 3H, SCH_2CH_3); $^{13}\text{C NMR}$ (125.8 MHz, CDCl_3): δ = 172.0 (Lac-C1), 153.5 - 148.1 (m, 2C, ArF-C), 130.0, 135.3, 128.9, 128.7, 128.6, 128.1 (6C, Ar-C), 126.3 (Ar-C), 125.5 (m, Ar-C), 118.8 (d, J = 17.7 Hz, Ar-C), 116.6 (d, J = 17.0 Hz, Ar-C), 100.9 (Ph-CH), 85.7

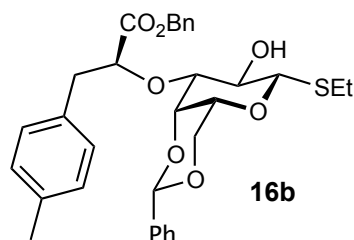
(Gal-C1), 80.5 (Gal-C3), 79.3 (Lac-C2), 75.6 (Gal-C4), 70.4 (Gal-C5), 69.2 (Gal-C6), 69.1 (Gal-C2), 66.8 (Ph-CH₂), 38.2 (Lac-C3), 23.0 (SCH₂CH₃), 15.4 (SCH₂CH₃); ESI-MS: *m/z*: Calcd for C₃₁H₃₂F₂NaO₇S [M+Na]⁺: 573.7, found: 573.1.

Ethyl 2-*O*-benzoyl-4,6-*O*-benzylidene-3-*O*-[(*R*)-1-(benzyloxy)-1-oxo-3-phenylpropan-2-yl]-1-thio-β-D-galactopyranoside (16a)



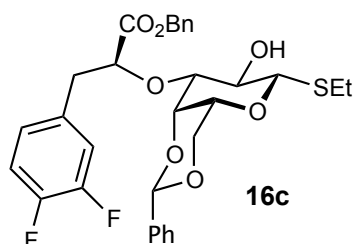
To a solution of **S6a** (250 mg, 0.45 mmol) in pyridine (2 mL) DMAP (6 mg, 0.05 mmol) and benzoyl chloride (0.16 mL, 1.36 mmol) were added at 0°C. The reaction was stirred at rt for 3h after which TLC (petroleum ether/EtOAc, 3:2) indicated completion of the reaction. The solvent was removed *in vacuo*, the crude residue was taken up in DCM (100 mL), washed with 0.5 M aq. HCl (20 mL), satd aq. NaHCO₃ (20 mL) and brine (20 mL), dried over Na₂SO₄ and concentrated *in vacuo*. The crude product was purified by flash chromatography (petroleum ether/EtOAc, gradient 0-50%) to afford **16a** as white solid (260 mg, 0.40 mmol, 87%). *R*_f = (petroleum ether/EtOAc, 7:3) 0.45; [α]_D²² -22.5 (*c* 2.0, CHCl₃); ¹H NMR (500.1 MHz, CDCl₃): δ = 7.95 (d, *J* = 7.3 Hz, 2H, Ar-H), 7.62 - 7.51 (m, 3H, Ar-H), 7.46 - 7.28 (m, 10H, Ar-H), 7.02 (t, *J* = 7.3 Hz, 1H, Ar-H), 6.98 - 6.90 (m, 2H, Ar-H), 7.02 (d, *J* = 7.2 Hz, 2H, Ar-H), 5.67 (t, *J* = 9.7 Hz, 1H, Gal-H2), 5.39 (s, 1H, Ph-CH), 5.05 - 4.95 (m, 2H, Ar-CH₂), 4.44 (d, *J* = 9.7 Hz, 1H, Gal-H1), 4.34 (t, *J* = 6.2 Hz, 1H, Lac-H2), 4.30 - 4.24 (m, 2H, Gal-H6, Gal-H4), 3.88 (d, *J* = 12.2 Hz, Gal-H6'), 3.68 (d, *J* = 7.6 Hz, 1H, Gal-H3), 3.34 (s, 1H, Gal-H5), 2.92 - 2.82 (m, 3H, Lac-H3, SCH₂CH₃), 2.77 - 2.66 (m, 1H, SCH₂CH₃), 1.22 (t, *J* = 7.5 Hz, 3H, SCH₂CH₃); ¹³C NMR (126 MHz, CDCl₃): δ = 172.0 (Lac-C1), 164.8 (Bz-CO), 137.9, 136.0, 135.3, 133.1, 129.9, 129.8, 129.4, 128.9, 128.7, 128.6, 128.4, 128.2, 128.1, 126.5, 126.4 (16C, Ar-C), 100.9 (Ph-CH), 82.9 (Gal-C1), 81.3 (Lac-C2), 79.9 (Gal-C3), 75.3 (Gal-C4), 70.3 (Gal-C5), 69.6 (Gal-C2), 69.2 (Gal-C6), 66.7 (Ph-CH₂), 39.2 (Lac-C3), 22.7 (SCH₂CH₃), 14.8 (SCH₂CH₃); ESI-MS: *m/z*: Calcd for C₃₈H₃₈O₈S [M+Na]⁺: 577.8, found: 577.2.

Ethyl **2-O-benzoyl-4,6-O-benzylidene-3-O-[(R)-1-(benzyloxy)-1-oxo-3-(4-methylphenyl)propan-2-yl]-1-thio-β-D-galactopyranoside (16b)**



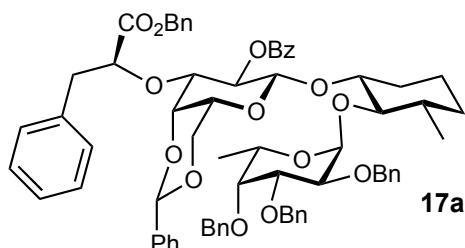
To a solution of **S6b** (195 mg, 0.34 mmol) in pyridine (2 mL) DMAP (4 mg, 0.03 mmol) and benzoyl chloride (0.12 mL, 1.04 mmol) were added at 0°C. The reaction was stirred at rt for 16 h after which TLC (petroleum ether/EtOAc, 3:2) indicated completion of the reaction. The solvent was removed *in vacuo*, the crude residue was taken up in DCM (100 mL), washed with 0.5 M aq. HCl (20 mL), satd aq. NaHCO₃ (20 mL) and brine (20 mL), dried over Na₂SO₄ and concentrated *in vacuo*. The crude product was purified by flash chromatography (petroleum ether/EtOAc, gradient 0-50%) to afford **16b** as white solid (166 mg, 0.25 mmol, 72%). R_f = (petroleum ether/EtOAc, 7:3) 0.48; $[\alpha]_D^{22}$ -26.3 (*c* 0.45, CHCl₃); ¹H NMR (500.1 MHz, CDCl₃): δ = 7.94 (d, *J* = 7.7 Hz, 2H, Ar-H), 7.64 - 7.50 (m, 3H, Ar-H), 7.49 - 7.29 (m, 8H, Ar-H), 7.25 - 7.17 (m, 2H, Ar-H), 6.81 - 6.68 (m, 4H, Ar-H), 5.67 (t, *J* = 9.7 Hz, 1H, Gal-H2), 5.39 (s, 1H, Ph-CH), 5.01 (s, 2H, Ar-CH₂), 4.43 (d, *J* = 9.6 Hz, 1H, Gal-H1), 4.35 - 4.23 (m, 3H, Gal-H6, Lac-H2, Gal-H4), 3.88 (d, *J* = 12.2 Hz, Gal-H6'), 3.67 (d, *J* = 9.6 Hz, 1H, Gal-H3), 3.35 (s 1H, Gal-H5), 2.93 - 2.66 (m, 4H, Lac-H3, SCH₂CH₃), 2.18 (s, 3H, Ar-Me), 1.23 (t, *J* = 7.7 Hz, 3H, SCH₂CH₃); ¹³C NMR (125.8 MHz, CDCl₃): δ = 172.1 (Lac-C1), 164.8 (Bz-CO), 137.9, 135.8, 135.3, 133.0, 132.9, 129.9, 129.8, 129.3, 128.9, 128.8, 128.6, 128.6, 128.5, 128.3, 128.2, 126.5 (16C, Ar-C), 101.0 (Ph-CH), 82.9 (Gal-C1), 81.7 (Lac-C2), 79.9 (Gal-C3), 75.3 (Gal-C4), 70.3 (Gal-C5), 69.5 (Gal-C2), 69.2 (Gal-C6), 66.7 (Ph-CH₂), 38.8 (Lac-C3), 22.7 (SCH₂CH₃), 21.0 (Ar-Me), 14.7 (SCH₂CH₃); ESI-MS: *m/z*: Calcd for C₃₉H₄₀NaO₈S [M+Na]⁺: 691.8, found: 691.2.

Ethyl **2-O-benzoyl-4,6-O-benzylidene-3-O-[(R)-1-(benzyloxy)-1-oxo-3-(3,4-difluorophenyl)propan-2-yl]-1-thio-β-D-galactopyranoside (16c)**



To a solution of **S6c** (73 mg, 0.2 mmol) in pyridine (2 mL) DMAP (1.5 mg, 0.01 mmol) and benzoyl chloride (0.045 mL, 0.37 mmol) were added at 0°C. The reaction was stirred at rt for 16 h after which TLC (petroleum ether/EtOAc, 3:2) indicated completion of the reaction. The solvent was removed *in vacuo*, the crude residue was taken up in DCM (100 mL), washed with 0.5 M aq. HCl (20 mL), sat. NaHCO₃ (20 mL) and brine (20 mL), dried over Na₂SO₄ and concentrated *in vacuo*. The crude product was purified by flash chromatography (petroleum ether/EtOAc, gradient 0-50%) to afford **16c** as white solid (55 mg, 0.08 mmol, 64%). R_f = (petroleum ether/EtOAc, 7:3) 0.40; $[\alpha]_D^{22}$ -35.3 (*c* 0.28, CHCl₃); ¹H NMR (500.1 MHz, CDCl₃): δ = 7.92 (d, *J* = 7.7 Hz, 2H, Ar-H), 7.65 - 7.51 (m, 3H, Ar-H), 7.50 - 7.32 (m, 8H, Ar-H), 7.32 - 7.22 (m, 2H, Ar-H), 6.77 - 6.53 (m, 3H, Ar-H), 5.65 (t, *J* = 9.6 Hz, 1H, Gal-H2), 5.42 (s, 1H, Ph-CH), 5.13 (d, *J* = 11.9 Hz, 1H, Ar-CH₂), 5.02 (d, *J* = 11.9 Hz, 1H, Ar-CH₂), 4.43 (d, *J* = 9.7 Hz, 1H, Gal-H1), 4.38 - 4.23 (m, 3H, Lac-H2, Gal-H6a, Gal-H4), 3.89 (d, *J* = 12.3 Hz, Gal-H6b), 3.69 (d, *J* = 9.5 Hz, 1H, Gal-H3), 3.36 (s, 1H, Gal-H5), 2.93 - 2.66 (m, 4H, Lac-H3, SCH₂CH₃), 1.22 (t, *J* = 7.5 Hz, 3H, SCH₂CH₃); ¹³C NMR (125.8 MHz, CDCl₃): δ = 171.6 (Lac-C1), 164.7 (Bz-CO), 160.8 - 140.9 (m, 2C, Ar-C), 137.9, 135.1, 133.7, 133.3 (4C, Ar-C), 133.0 (m, Ar-C), 130.2, 129.7, 129.0, 128.8, 128.8, 128.8, 128.5, 128.5, 128.2, 126.4 (10C, Ar-C), 125.7 - 125.4 (m, Ar-C), 118.4 (d, *J* = 17.2 Hz, Ar-C), 116.6 (d, *J* = 17.0 Hz, Ar-C), 101.0 (Ph-CH), 82.9 (Gal-C1), 80.8 (Lac-C2), 79.8 (Gal-C3), 75.5 (Gal-C4), 70.3 (Gal-C5), 69.6 (Gal-C2), 69.1 (Gal-C6), 67.0 (Ph-CH₂), 38.2 (Lac-C3), 22.7 (SCH₂CH₃), 14.7 (SCH₂CH₃); ESI-MS: *m/z*: Calcd for C₃₈H₃₆F₂NaO₈S [M+Na]⁺: 713.7, found: 713.1.

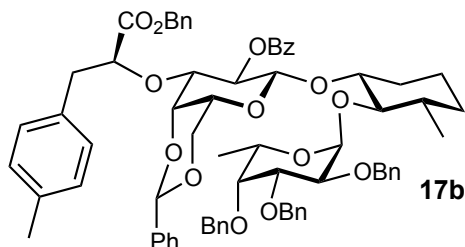
(1*R*,2*R*,3*S*)-2-[(2,3,4-Tri-*O*-benzyl-6-deoxy- α -L-galactopyranosyl)oxy]-3-methyl-cyclohex-1-yl 2-*O*-benzoyl-4,6-*O*-benzylidene-3-*O*-[(*R*)-1-(benzyloxy)-1-oxo-3-phenylpropan-2-yl]- β -D-galactopyranoside (17a**)**



A solution of donor **16a** (101 mg, 0.159 mmol) and acceptor **15²** (101 mg, 0.185 mmol) in dry DCM (5 mL) was added *via* syringe to activated 4Å molecular sieves (0.5 g) under argon. A suspension of dimethyl(methylthio)sulfonium triflate (DMTST, 119 mg, 0.462

mmol) and activated 4Å molecular sieves (0.5 g) in DCM (5 mL) was prepared in a second flask under argon. Both suspensions were stirred at rt for 4 h, then the DMTST suspension was added *via* syringe to the other suspension. The reaction was stopped after 18 h, filtered through celite and the celite was washed with DCM (20 mL). The filtrate was successively washed with satd. aq. NaHCO₃ (25 mL) and water (25 mL), dried over Na₂SO₄, filtered and concentrated. The crude product was purified by flash chromatography (petroleum ether/EtOAc, gradient 0-50%) to afford **17a** as white solid (138 mg, 0.12 mmol, 79%). R_f = (petroleum ether/EtOAc, 3:2) 0.35; $[\alpha]_D^{22}$ -78.6 (*c* 0.39, CHCl₃); ¹H NMR (500.1 MHz, CDCl₃): δ = 7.99 (d, *J* = 7.7 Hz, 2H, Ar-H), 7.63 (d, *J* = 7.5 Hz, 2H, Ar-H), 7.59 (t, *J* = 7.2 Hz, 1H, Ar-H), 7.45 (t, *J* = 7.4 Hz, 2H, Ar-H), 7.35 - 7.14 (m, 23H, Ar-H), 7.07 - 6.93 (m, 5H, Ar-H), 5.57 (t, *J* = 8.8 Hz, 1H, Gal-H2), 5.42 (s, 1H, Ph-CH), 5.02 - 4.85 (m, 5H, 2 Ph-CH₂, Fuc-H1), 4.80 (d, *J* = 11.7 Hz, 1H, Ph-CH₂), 4.69 (d, *J* = 11.7 Hz, 1H, Ph-CH₂), 4.63 (d, *J* = 11.5 Hz, 1H, Ph-CH₂), 4.58 (d, *J* = 11.3 Hz, 1H, Ph-CH₂), 4.51 (d, *J* = 7.9 Hz, 1H, Gal-H1), 4.36 - 4.16 (m, 4H, Lac-H2, Gal-H6, Gal-H4, Fuc-H2), 4.00 - 3.87 (m, 3H, Gal-H6', Fuc-H3, Fuc-H5), 3.71 - 3.57 (m, 2H, Gal-H3, Fuc-H4), 3.51 (t, *J* = 9.5 Hz, 1H, MeCy-H2), 3.33 - 3.22 (m, 2H, Gal-H5, Fuc-H4), 3.13 (t, *J* = 9.5 Hz, 1H, MeCy-H1), 2.98 - 2.85 (m, 2H, Lac-H3), 1.85 - 1.77 (m, 1H, MeCy-H3), 1.58 - 1.41 (m, 3H, MeCy-H6, MeCy-H4), 1.25 (d, *J* = 6.3 Hz, 3H, Fuc-H6), 1.01 (d, *J* = 6.0 Hz, 3H, MeCy-Me), 1.16 - 0.81 (m, 3H, MeCy-H4', MeCy-H5); ¹³C NMR (125.8 MHz, CDCl₃): δ = 171.8 (Lac-C1), 164.5 (Bz-CO), 139.8, 139.6, 138.8, 138.2, 136.0, 135.2, 133.0, 130.1, 129.7, 129.5, 128.7, 128.6, 128.6, 128.5, 128.5, 128.4, 128.1, 128.0, 128.0, 127.8, 127.5, 127.4, 127.3, 127.0, 126.8, 126.5, 126.0 (28C, Ar-C), 99.7 (Gal-C1), 99.5 (Ph-CH), 98.3 (Fuc-C1), 81.3 (MeCy-C1), 81.1 (Lac-C2), 80.7 (MeCy-C2), 79.7 (Fuc-C3), 79.1 (Fuc-C4), 89.9 (Gal-C3), 75.7 (Fuc-C5), 74.9 (2C, Gal-C4, Fuc-C4), 74.5 (Ph-CH₂), 71.6 (Gal-C2), 71.2 (Ph-CH₂), 69.2 (Gal-C6), 66.7 (Ph-CH₂), 66.4 (Ph-CH₂), 66.3 (Gal-C5), 39.5 (MeCy-C6), 39.2 (Lac-C3), 33.5 (MeCy-C4), 31.1 (MeCy-C3), 23.3 (MeCy-C5), 18.7 (MeCy-Me), 16.2 (Fuc-C6); ESI-MS: *m/z*: Calcd for C₇₀H₇₄NaO₁₄[M+Na]⁺: 1162.3, found: 1161.4.

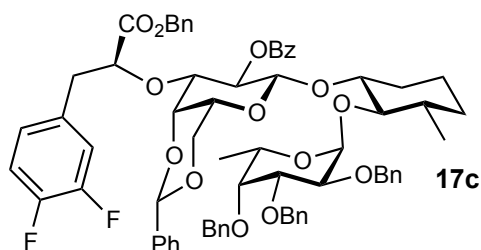
(1*R*,2*R*,3*S*)-2-[(2,3,4-Tri-*O*-benzyl-6-deoxy- α -L-galactopyranosyl)oxy]-3-methyl- cyclohex-1-yl 2-*O*-benzoyl-4,6-*O*-benzylidene-3-*O*-[(*R*)-1-(benzyloxy)-1-oxo-3-(4-methylphenyl)-propan-2-yl]- β -D-galactopyranoside (17b**)**



A solution of donor **16b** (30 mg, 0.044 mmol, 1 eq) and acceptor **15²** (30 mg, 0.058 mmol) in dry DCM (1.5 mL) was added *via* syringe to activated 4Å molecular sieves (0.2 g) under argon. A suspension of dimethyl(methylthio)sulfonium triflate (DMTST, 35 mg, 0.134 mmol) and activated 4Å molecular sieves (0.2 g) in DCM (1.5 mL) was prepared in a second flask under argon. Both suspensions were stirred at rt for 4 h, then the DMTST suspension was added *via* syringe to the other suspension. The reaction was stopped after 18 h, filtered through celite and the celite was washed with DCM (10 mL). The filtrate was successively washed with satd. aq. NaHCO₃ (15 mL) and water (15 mL), dried over Na₂SO₄, filtered and concentrated. The crude product was purified by flash chromatography (petroleum ether/EtOAc, gradient 0-50%) to afford **17b** as white solid (41 mg, 0.035 mmol, 80%). R_f = (petroleum ether/EtOAc, 3:2) 0.33; $[\alpha]_D^{22}$ -79.3 (*c* 0.25, CHCl₃); ¹H NMR (500.1 MHz, CDCl₃): δ = 7.98 (d, *J* = 7.7 Hz, 2H, Ar-H), 7.64 (d, *J* = 7.4 Hz, 2H, Ar-H), 7.59 (t, *J* = 7.2 Hz, 1H, Ar-H), 7.39 - 7.12 (m, 27H, Ar-H), 6.87 (d, *J* = 7.5 Hz, 2H, Ar-H), 6.26 (d, *J* = 7.5 Hz, 2H, Ar-H), 5.56 (t, *J* = 9.6 Hz, 1H, Gal-H2), 5.43 (s, 1H, Ph-CH), 5.04 - 4.90 (m, 5H, 2 Ph-CH₂, Fuc-H1), 4.80 (d, *J* = 11.7 Hz, 1H, Ph-CH₂), 4.68 (d, *J* = 11.7 Hz, 1H, Ph-CH₂), 4.63 (d, *J* = 11.3 Hz, 1H, Ph-CH₂), 4.58 (d, *J* = 11.3 Hz, 1H, Ph-CH₂), 4.51 (d, *J* = 9.7 Hz, 1H, Gal-H1), 4.30 - 4.15 (m, 4H, Lac-H2, Gal-H6, Gal-H4, Fuc-H2), 4.00 - 3.86 (m, 3H, Gal-H6', Fuc-H3, Fuc-H5), 3.71 - 3.57 (m, 2H, Gal-H3, Fuc-H4), 3.55 (t, *J* = 9.5 Hz, 1H, MeCy-H2), 3.30 - 3.21 (m, 2H, Gal-H5, Fuc-H4), 3.13 (t, *J* = 9.5 Hz, 1H, MeCy-H1), 2.93 - 2.85 (m, 2H, Lac-H3), 2.18 (s, 3H, Ar-Me), 1.86 - 1.74 (m, 1H, MeCy-H3), 1.55 - 1.40 (m, 3H, MeCy-H6, MeCy-H4), 1.25 (d, *J* = 6.0 Hz, 3H, Fuc-H6), 1.01 (d, *J* = 6.1 Hz, 3H, MeCy-Me), 1.16 - 0.81 (m, 3H, MeCy-H4', MeCy-H5); ¹³C NMR (125.8 MHz, CDCl₃): δ = 172.0 (Lac-C1), 164.5 (Bz-CO), 139.8, 139.6, 138.8, 138.2, 135.9, 135.3, 132.9, 132.9, 130.1, 129.7, 129.4, 128.8, 128.7, 128.6, 128.6, 128.5, 128.5, 128.4, 128.1, 128.0, 128.0, 127.8, 127.6, 127.5, 127.3, 127.0, 126.8, 126.0 (28C, Ar-C), 99.7 (Gal-

C1), 99.5 (Ph-CH), 98.3 (Fuc-C1), 81.4 (Lac-C2, MeCy-C1), 80.7 (MeCy-C2), 79.7 (Fuc-C3), 79.1 (Fuc-C4), 79.0 (Gal-C3), 75.7 (Fuc-C5), 74.9 (2C, Gal-C4, Fuc-C4), 74.5 (Ph-CH₂), 71.6 (Gal-C2), 71.2 (Ph-CH₂), 69.2 (Gal-C6), 66.7 (Ph-CH₂), 66.4 (Ph-CH₂), 66.3 (Gal-C5), 39.5 (MeCy-C6), 38.8 (Lac-C3), 33.5 (MeCy-C4), 31.1 (MeCy-C3), 23.3 (MeCy-C5), 18.7 (MeCy-Me), 16.2 (Fuc-C6); ESI-MS: *m/z*: Calcd for C₇₁H₇₆NaO₁₄ [M+Na]⁺: 1176.3, found: 1175.6.

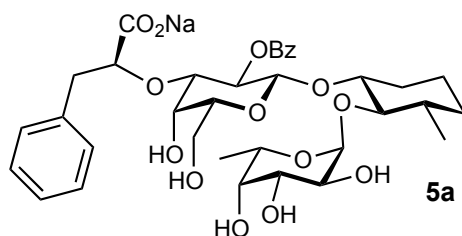
(1*R*,2*R*,3*S*)-2-[(2,3,4-Tri-*O*-benzyl-6-deoxy- α -L-galactopyranosyl)oxy]-3-methylcyclohex-1-yl 2-*O*-benzoyl-4,6-*O*-benzylidene-3-*O*-[(*R*)-1-(benzyloxy)-1-oxo-3-(3,4-difluorophenyl)propan-2-yl]- β -D-galactopyranoside (17c**)**



A solution of donor **16c** (53 mg, 0.076 mmol) and acceptor **15**² (50 mg, 0.092 mmol) in dry DCM (1.5 mL) was added *via* syringe to activated 4Å molecular sieves (0.25 g) under argon. A suspension of dimethyl(methylthio)sulfonium triflate (DMTST, 60 mg, 0.230 mmol) and activated 4Å molecular sieves (0.25 g) in DCM (1.5 mL) was prepared in a second flask under argon. Both suspensions were stirred at rt for 4 h, then the DMTST suspension was added *via* syringe to the other suspension. The reaction was stopped after 18 h, filtered through celite and the celite was washed with DCM (10 mL). The filtrate was successively washed with satd. aq. NaHCO₃ (20 mL) and water (20 mL), dried over Na₂SO₄, filtered and concentrated. The crude product was purified by flash chromatography (petroleum ether/EtOAc, gradient 0-50%) to afford **17c** as colorless wax (41 mg, 0.035 mmol, 80%). *R*_f = (petroleum ether/EtOAc, 3:2) 0.28; [α]_D²² -84.6 (*c* 0.60, CHCl₃); ¹H NMR (500.1 MHz, CDCl₃): δ = 7.96 (d, *J* = 7.5 Hz, 2H, Ar-H), 7.66 - 7.55 (m, 3H, Ar-H), 7.45 (t, *J* = 7.6 Hz, 2H, Ar-H), 7.40 - 7.14 (m, 23H, Ar-H), 6.89 - 6.81 (m, 1H, Ar-H), 6.74 - 6.60 (m, 2H, Ar-H), 5.56 (t, *J* = 9.0 Hz, 1H, Gal-H2), 5.47 (s, 1H, Ph-CH), 5.09 (d, *J* = 12.0 Hz, 1H, Ph-CH₂), 4.98 (d, *J* = 12.0 Hz, 1H, Ph-CH₂), 4.95 - 4.90 (m, 2H, Fuc-H1, Fuc-H2), 4.80 (d, *J* = 11.7 Hz, 1H, Ph-CH₂), 4.69 (d, *J* = 11.7 Hz, 1H, Ph-CH₂), 4.63 (d, *J* = 11.5 Hz, 1H, Ph-CH₂), 4.58 (d, *J* = 11.3 Hz, 1H, Ph-CH₂), 4.51 (d, *J* = 7.9 Hz, 1H, Gal-H1), 4.34 - 4.24 (m, 3H, Lac-H2, Gal-H6, Gal-H4), 4.20 (d, *J* = 11.3 Hz, 1H, Ph-CH₂), 4.00 - 3.88 (m,

3H, Gal-H6', Fuc-H2, Fuc-H3), 3.68 (d, $J = 8.2$ Hz, 1H, Gal-H3), 3.63 (d, $J = 11.3$ Hz, 1H, Ph-CH₂), 3.57 - 3.46 (m, 1H, MeCy-H2), 3.33 - 3.23 (m, 2H, Gal-H5, Fuc-H4), 3.13 (t, $J = 9.5$ Hz, 1H, MeCy-H1), 2.94 - 2.80 (m, 2H, Lac-H3), 1.85 - 1.77 (m, 1H, MeCy-H3), 1.56 - 1.40 (m, 3H, MeCy-H6, MeCy-H4), 1.25 (d, $J = 6.3$ Hz, 3H, Fuc-H6), 1.01 (d, $J = 6.4$ Hz, 3H, MeCy-Me), 1.16 - 0.81 (m, 3H, MeCy-H4', MeCy-H5); ¹³C NMR (126 MHz, CDCl₃): $\delta = 171.5$ (Lac-C1), 164.4 (Bz-CO), 150.9 - 145.1 (m, 2C, ArF-C), 139.8, 139.6, 138.8, 138.1, 135.0, 133.1, 132.9, 130.1, 129.8, 129.6, 129.0, 128.8, 128.7, 128.6, 128.5, 128.2, 128.1, 128.1, 128.0, 127.8, 127.6, 127.4, 127.3, 127.0, 126.8, 125.9, 125.5 (26C, Ar-C), 118.5 (d, $J = 17.3$ Hz, Ar-C), 116.9 (d, $J = 17.3$ Hz, Ar-C), 99.6, 99.6 (2C, Gal-C1, Ph-CH), 98.3 (Fuc-C1), 81.3 (MeCy-C1), 80.7 (Lac-C2), 80.5 (MeCy-C2), 79.7 (Fuc-C3), 79.1 (Fuc-C4), 79.9 (Gal-C3), 75.7 (Fuc-C2), 75.1 (Gal-C4), 74.9 (Ph-CH₂), 74.5 (Ph-CH₂), 71.8 (Gal-C2), 71.2 (Ph-CH₂), 69.2 (Gal-C6), 66.7 (Ph-CH₂), 66.4, 66.3 (2C, Gal-C5, Fuc-C5), 39.5 (MeCy-C6), 38.2 (Lac-C3), 33.5 (MeCy-C4), 31.1 (MeCy-C3), 23.3 (MeCy-C5), 18.7 (MeCy-Me), 16.2 (Fuc-C6); ESI-MS: m/z : Calcd for C₇₀H₇₂F₂NaO₁₄ [M+Na]⁺: 1198.3, found: 1198.9.

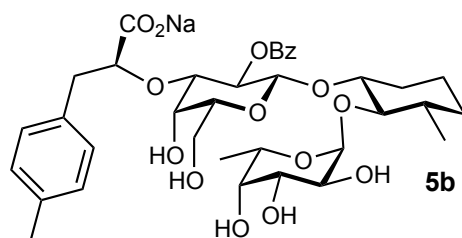
(1*R*,2*R*,3*S*)-2-[(α -L-Fucopyranosyl)oxy]-3-methyl-cyclohex-1-yl 2-*O*-benzoyl-3-*O*-[sodium (*S*)-1-carboxy-2-phenylethyl]- β -D-galactopyranoside (5a**)**



A suspension of **17a** (76 mg, 0.068 mmol) and Pd(OH)₂/C (50 mg, 10% Pd) in MeOH (5 mL) was hydrogenated (2 bar H₂) at rt. After 1 h TLC (silica: petroleum ether/EtOAc, 3:2; C-18: H₂O/MeOH, 1:1) indicated completion of the reaction. The reaction mixture was filtered through celite and concentrated *in vacuo*. The crude product was purified by reversed phase column chromatography (C-18, H₂O/MeOH, gradient 0-100%) followed by a Dowex 50 (Na⁺ form) ion exchange column, a Sephadex G15 column, and lyophilization from water to yield the product **5a** as white foam (20.3 mg, 0.028 mmol, 42%). R_f = (C-18, H₂O/MeOH, 1:1) 0.48; [α]_D²² -97.6 (*c* 0.35, MeOH); ¹H NMR (500 MHz, CDCl₃): $\delta = 7.92$ (d, $J = 8.1$ Hz, 2H, Ar-H), 7.76 (t, $J = 7.5$ Hz, 1H, Ar-H), 7.57 (t, $J = 7.9$ Hz, 2H, Ar-H), 7.04 - 7.96 (m, 2H, Ar-H), 6.87 - 6.75 (m, 2H, Ar-H), 5.19 (m, 1H, Gal-H2), 5.04 (d, $J = 4.0$ Hz, 1H, Fuc-H1), 4.90 (q, $J = 6.6$ Hz, 1H, Fuc-H5), 4.70 (d, $J = 8.1$ Hz, 1H, Gal-H1), 4.12

(dd, $J = 3.2, 9.7$ Hz, 1H, Lac-H2), 4.04 (d, $J = 2.5$ Hz, 1H, Gal-H4), 3.91 - 3.75 (m, 5H, Gal-H6, Fuc-H2, Fuc-H3, Fuc-H4), 3.69 - 3.62 (m, 2H, Gal-H3, Gal-H5), 3.57 (m, 1H, MeCy-H1), 3.06 (t, $J = 9.4$ Hz, 1H, MeCy-H2), 3.00 (dd, $J = 3.1, 14.3$ Hz, 1H, Lac-H3), 2.74 (dd, $J = 9.7, 14.3$ Hz, 1H, Lac-H3'), 1.95 (s, 1H, Ar-Me), 1.97 - 1.88 (m, 1H, MeCy-H6), 1.55 - 1.39 (m, 3H, MeCy-H3, MeCy-H4a, MeCy-H5), 1.30 (d, $J = 8.1$ Hz, 3H, Fuc-H6), 1.14 (m, 1H, MeCy-H5'), 1.01 (d, $J = 6.4$ Hz, 3H, MeCy-Me), 1.82 - 0.76 (m, 2H, MeCy-H4', MeCy-H6'); ^{13}C NMR (126 MHz, D_2O): $\delta = 181.1$ (Lac-C1), 167.8 (Bz-CO), 137.9, 134.0, 129.9, 129.0, 128.7, 128.6, 128.1, 126.2 (8C, Ar-C), 98.7, 98.7 (2C, Gal-C1, Fuc-C1), 83.5 (MeCy-C2), 82.4 (Lac-C2), 80.9 (MeCy-C2), 79.8 (Gal-C3), 74.2 (MeCy-C1), 72.1, 72.0 (2C, Gal-C5, Gal-C2), 69.3 (Fuc-C4), 68.2 (Fuc-C2), 66.5 (Fuc-C5), 66.4 (Gal-C4), 61.7 (Gal-C6), 39.6 (Lac-C3), 38.5 (MeCy-C3), 32.9 (MeCy-C4), 30.8 (MeCy-C6), 22.3 (MeCy-C5), 18.0 (MeCy-C6), 15.4 (Fuc-C6); HR-MS: m/z : Calcd for $\text{C}_{35}\text{H}_{45}\text{Na}_2\text{O}_{14}[\text{M}+\text{Na}]^+$: 735.2605, found: 735.2607.

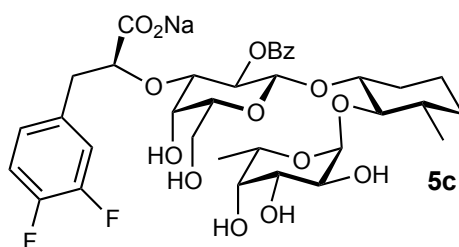
(1*R*,2*R*,3*S*)-2-[(α -L-Fucopyranosyl)oxy]-3-methyl-cyclohex-1-yl 2-*O*-benzoyl-3-*O*-[sodium (*S*)-1-carboxy-2-(4-methylphenyl)-ethyl]- β -D-galactopyranoside (5b**)**



A suspension of **17b** (40 mg, 0.035 mmol) and $\text{Pd}(\text{OH})_2/\text{C}$ (50 mg, 10% Pd) in MeOH (5 mL) was hydrogenated (2 bar H_2) at rt. After 3 h TLC (silica: petroleum ether/EtOAc, 3:2; C-18: $\text{H}_2\text{O}/\text{MeOH}$, 1:1) indicated completion of the reaction. The reaction mixture was filtered through celite and concentrated *in vacuo*. The crude product was purified by reversed phase column chromatography (C-18, $\text{H}_2\text{O}/\text{MeOH}$, gradient 0-100%) followed by a Dowex 50 (Na^+ form) ion exchange column, a Sephadex G15 column, and lyophilization from water to yield the product **5b** as colorless foam (9.4 mg, 0.013 mmol, 37%). $R_f =$ (C-18, $\text{H}_2\text{O}/\text{MeOH}$, 1:1) 0.42; $[\alpha]_{\text{D}}^{22} -64.5$ (c 0.23, MeOH); ^1H NMR (500.1 MHz, CDCl_3): $\delta = 7.92$ (d, $J = 7.9$ Hz, 2H, Ar-H), 7.77 (t, $J = 7.2$ Hz, 1H, Ar-H), 7.59 (t, $J = 7.7$ Hz, 2H, Ar-H), 6.92 (d, $J = 7.5$ Hz, 2H, Ar-H), 6.59 (d, $J = 7.5$ Hz, 2H, Ar-H), 5.19 (t, $J = 8.7$ Hz, 1H, Gal-H2), 5.03 (d, $J = 2.8$ Hz, 1H, Fuc-H1), 4.90 (dd, $J = 6.5, 13.0$ Hz, 1H, Fuc-H5), 4.71 (d, $J = 8.1$ Hz, 1H, Gal-H1), 4.08 (d, $J = 10.4$ Hz, 1H, Lac-H2), 4.03 (s, 1H, Gal-H4), 3.91 - 3.73 (m, 5H, Gal-H6, Fuc-H2, Fuc-H3, Fuc-H4), 3.69 - 3.60 (m, 2H, Gal-H3, Gal-H5), 3.57

(t, $J = 13.2$ Hz, 1H, MeCy-H1), 3.08 (t, $J = 9.3$ Hz, 1H, MeCy-H2), 2.96 (d, $J = 13.7$ Hz, 1H, Lac-H3a), 2.69 (m, 1H, Lac-H3b), 1.95 (s, 1H, Ar-Me), 1.93 (m, 1H, MeCy-H6), 1.58 - 1.40 (m, 3H, MeCy-H3, MeCy-H4, MeCy-H5), 1.30 (d, $J = 6.5$ Hz, 3H, Fuc-H6), 1.14 (m, 1H, MeCy-H5'), 1.01 (d, $J = 6.4$ Hz, 3H, MeCy-Me), 1.93 - 0.76 (m, 2H, MeCy-H4', MeCy-H6'); ^{13}C NMR (125.8 MHz, D_2O): $\delta = 181.2$ (Lac-C1), 167.8 (Bz-CO), 136.0, 134.9, 134.0, 129.7, 129.1, 128.8, 128.8, 128.6 (8C, Ar-C), 98.7, 98.7 (2C, Gal-C1, Fuc-C1), 83.6 (MeCy-C2), 82.6 (Lac-C2), 80.5 (MeCy-C2), 80.9 (Gal-C3), 79.7 (MeCy-C1), 74.2 (Gal-C5), 72.1 (2C, Gal-C2, Fuc-C3), 69.3 (Fuc-C4), 68.2 (Fuc-C2), 66.5 (Fuc-C5), 66.4 (Gal-C4), 61.7 (Gal-C6), 39.2 (Lac-C3), 38.5 (MeCy-C3), 32.9 (MeCy-C4), 30.8 (MeCy-C6), 22.3 (MeCy-C5), 20.0 (MeCy-Me), 18.0 (MeCy-C6), 15.0 (Fuc-C6); HR-MS: m/z : Calcd for $\text{C}_{36}\text{H}_{47}\text{Na}_2\text{O}_{14}[\text{M}+\text{Na}]^+$: 749.2761, found: 749.2760.

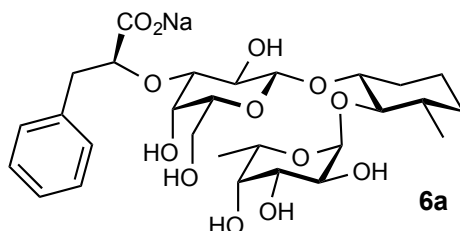
(1R,2R,3S)-2-[(α -L-Fucopyranosyl)oxy]-3-methyl-cyclohex-1-yl 2-O-benzoyl-3-O-[sodium (*S*)-1-carboxy-2-(3,4-difluorophenyl)-ethyl]- β -D-galactopyranoside (5c**)**



A suspension of **17c** (48 mg, 0.040 mmol) and $\text{Pd}(\text{OH})_2/\text{C}$ (50 mg, 10% Pd) in MeOH (5 mL) was hydrogenated (2 bar H_2) at rt. After 1 h TLC (silica: petroleum ether/EtOAc, 3:2; C-18: $\text{H}_2\text{O}/\text{MeOH}$, 1:1) indicated completion of the reaction. The reaction mixture was filtered through celite and concentrated *in vacuo*. The crude product was purified by reversed phase column chromatography (C-18, $\text{H}_2\text{O}/\text{MeOH}$, gradient 0-100%) followed by a Dowex 50 (Na^+ form) ion exchange column, a Sephadex G15 column, and lyophilization from water to yield the product **5c** as colorless foam (18.8 mg, 0.025 mmol, 63%). $R_f =$ (C-18, $\text{H}_2\text{O}/\text{MeOH}$, 1:1) 0.45; $[\alpha]_{\text{D}}^{22} -80.7$ (c 0.24, MeOH); ^1H NMR (500.1 MHz, CDCl_3): $\delta = 7.86$ (d, $J = 7.7$ Hz, 2H, Ar-H), 7.72 (t, $J = 7.2$ Hz, 1H, Ar-H), 7.54 (t, $J = 7.7$ Hz, 2H, Ar-H), 6.95 (m, 1H, Ar-H), 6.77 (m, 1H, Ar-H), 6.58 (m, 1H, Ar-H), 5.16 (t, $J = 8.8$ Hz, 1H, Gal-H2), 5.03 (br s, 1H, Fuc-H1), 4.89 (dd, $J = 6.6, 13.0$ Hz, 1H, Fuc-H5), 4.70 (d, $J = 8.1$ Hz, 1H, Gal-H1), 4.16 (d, $J = 10.8$ Hz, 1H, Lac-H2), 4.08 (br s, 1H, Gal-H4), 3.90 - 3.65 (m, 7H, Gal-H6, Gal-H5, Fuc-H2, Fuc-H3, Fuc-H4, Gal-H3), 3.56 (t, $J = 11.6$ Hz, 1H, MeCy-H1), 3.05 (t, $J = 9.3$ Hz, 1H, MeCy-H2), 2.97 (d, $J = 14.1$ Hz, 1H, Lac-H3), 2.72 (m, 1H, Lac-H3b), 1.90 (m, 1H, MeCy-H6), 1.57 - 1.34 (m, 3H, MeCy-H3, MeCy-H4, MeCy-

H5), 1.30 (d, $J = 6.2$ Hz, 3H, Fuc-H6), 1.14 (m, 1H, MeCy-H5'), 1.01 (d, $J = 6.2$ Hz, 3H, MeCy-Me), 1.96 - 0.70 (m, 2H, MeCy-H4', MeCy-H6'); ^{13}C NMR (125.8 MHz, D_2O): $\delta = 180.7$ (Lac-C1), 167.4 (Bz-CO), 150.2 - 146.4 (m, 2C, ArF-C), 135.1, 134.0, 129.5, 128.6, 128.5, 124.8 (8C, Ar-C), 117.1 (d, $J = 17.2$ Hz, Ar-C), 116.6 (d, $J = 17.3$ Hz, Ar-C), 98.8, 98.7 (2C, Gal-C1, Fuc-C1), 83.6 (MeCy-C2), 82.2 (Lac-C2), 80.9 (MeCy-C2), 80.9 (Gal-C3), 79.8 (MeCy-C1), 74.3 (Gal-C5), 72.1, 72.0 (2C, Gal-C2, Fuc-C3), 69.3 (Fuc-C4), 68.2 (Fuc-C2), 66.6, 66.5 (Fuc-C5, Gal-C4), 61.7 (Gal-C6), 38.7 (Lac-C3), 38.5 (MeCy-C3), 32.9 (MeCy-C4), 30.7 (MeCy-C6), 22.3 (MeCy-C5), 18.0 (MeCy-Me), 15.4 (Fuc-C6); HR-MS: m/z : Calcd for $\text{C}_{35}\text{H}_{43}\text{F}_2\text{NaO}_{14}[\text{M}+\text{Na}]^+$: 771.2416, found: 771.2415.

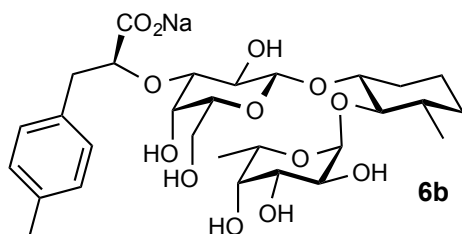
(1R,2R,3S)-2-[(α -L-Fucopyranosyl)oxy]-3-methyl-cyclohex-1-yl 3-O-[sodium (S)-1-carboxy-2-phenylethyl]- β -D-galactopyranoside (6a)



A suspension of **17a** (15 mg, 0.013 mmol) and $\text{Pd}(\text{OH})_2/\text{C}$ (50 mg, 10% Pd) in MeOH (5 mL) was hydrogenated (1 bar H_2) at rt. After 2 h TLC (silica: petroleum ether/EtOAc, 3:2; C-18: $\text{H}_2\text{O}/\text{MeOH}$, 1:1) indicated completion of the reaction. The reaction mixture was filtered through celite and concentrated *in vacuo*. The residue was dissolved in MeOH/ H_2O (4:1, 1 mL) and treated with LiOH (16 mg, 0.65 mmol) for 7 days. The mixture was neutralized with Dowex 50X8 (H^+ form), filtered through a Dowex 50 (Na^+ form) ion exchange column and concentrated. The residue was purified by reversed phase column chromatography (C-18, $\text{H}_2\text{O}/\text{MeOH}$, gradient 0-100%) followed by a Sephadex G15 column, and lyophilization from water to yield the product **6a** as white foam (1.5 mg, 0.0024 mmol, 18%). $R_f = (\text{C-18}, \text{H}_2\text{O}/\text{MeOH}, 1:1) 0.25$; ^1H NMR (500.1 MHz, CDCl_3): $\delta = 7.43 - 7.36$ (m, 4H, Ar-H), 7.35 - 7.30 (m, 1H, Ar-H), 5.10 (d, $J = 4.0$ Hz, 1H, Fuc-H1), 4.86 - 4.80 (m, 1H, Fuc-H5), 4.40 (d, $J = 8.0$ Hz, 1H, Gal-H1), 4.18 (dd, $J = 4.7, 8.7$ Hz, 1H, Lac-H2), 3.92 (d, $J = 2.8$ Hz, 1H, Gal-H4), 3.89 (dd, $J = 3.2, 10.6$ Hz, 1H, Fuc-H3), 3.87 - 3.78 (m, 2H, Fuc-H2, Fuc-H4), 3.76 - 3.65 (m, 3H, MeCy-H1, Gal-H6), 3.58 - 3.51 (m, 2H, Gal-H2, Gal-H5), 3.30 (dd, $J = 3.2, 9.5$ Hz, 1H, Gal-H3), 3.22 (t, $J = 9.6$ Hz, 1H, MeCy-H2), 3.15 (dd, $J = 4.6, 14.0$ Hz, 1H, Lac-H3), 2.98 (dd, $J = 8.8, 14.0$ Hz, 1H, Lac-H3'), 2.12 - 2.05 (m, 1H, MeCy-H6), 1.70 - 1.53 (m, 3H, MeCy-H3, MeCy-H5), 1.37 - 1.21 (m, 2H,

MeCy-H6', MeCy-H4), 1.18 (d, $J = 6.6$ Hz, 3H, Fuc-H6), 1.01 (d, $J = 6.3$ Hz, 3H, MeCy-Me), 1.14 – 1.04 (m, 1H, MeCy-H4'); ^{13}C NMR (125.8 MHz, D_2O): $\delta = 181.0$ (Lac-C1), 138.2, 129.4, 128.6, 126.7 (4C, Ar-C), 99.4 (Gal-C1), 98.9 (Fuc-C1), 84.1 (MeCy-C2), 82.7 (Gal-C3), 82.0 (Lac-C2), 78.6 (MeCy-C1), 74.2 (Gal-C2), 72.0 (Fuc-C2), 69.8 (Gal-C5), 69.3 (Fuc-C3), 68.2 (Fuc-C4), 66.5 (Fuc-C5), 66.2 (Gal-C4), 61.6 (Gal-C6), 39.3 (Lac-C3), 38.7 (MeCy-C3), 33.1 (MeCy-C4), 30.2 (MeCy-C6), 22.6 (MeCy-C5), 18.1 (MeCy-Me), 15.5 (Fuc-C6); HR-MS: m/z : Calcd for $\text{C}_{28}\text{H}_{42}\text{NaO}_{13}^+ [\text{M}+\text{H}]^+$: 609.2518, found: 609.2521.

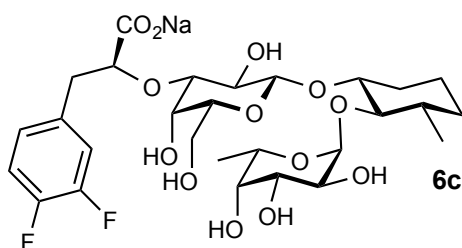
(1*R*,2*R*,3*S*)-2-[(α -L-Fucopyranosyl)oxy]-3-methyl-cyclohex-1-yl 3-*O*-[sodium (*S*)-1-carboxy-2-(4-methylphenyl)-ethyl]- β -D-galactopyranoside (6b**)**



A suspension of **17b** (40 mg, 0.035 mmol) and $\text{Pd}(\text{OH})_2/\text{C}$ (50 mg, 10% Pd) in MeOH (5 mL) was hydrogenated (2 bar H_2) at rt. After 2 h TLC (silica: petroleum ether/EtOAc, 3:2; C-18: $\text{H}_2\text{O}/\text{MeOH}$, 1:1) indicated completion of the reaction. The reaction mixture was filtered through celite and concentrated *in vacuo*. The residue was dissolved in MeOH/ H_2O (4:1, 1 mL) and treated with LiOH (41 mg, 1.73 mmol) for 32 h. The mixture was neutralized with Dowex 50X8 (H^+ form), filtered through a Dowex 50 (Na^+ form) ion exchange column and concentrated. The residue was purified by reversed phase column chromatography (C-18, $\text{H}_2\text{O}/\text{MeOH}$, gradient 0-100%) followed by a Dowex 50 (Na^+ form) ion exchange column, a Sephadex G15 column, and lyophilization from water to yield the product **6b** as colorless foam (8.9 mg, 0.014 mmol, 41%). $R_f =$ (C-18, $\text{H}_2\text{O}/\text{MeOH}$, 1:1) 0.31; $[\alpha]_{\text{D}}^{22} -76.6$ (c 0.26, MeOH); ^1H NMR (500.1 MHz, CDCl_3): $\delta = 7.28$ (d, $J = 8.0$ Hz, 2H, Ar-H), 7.23 (d, $J = 8.0$ Hz 1H, Ar-H), 5.10 (d, $J = 4.02$ Hz, 1H, Fuc-H1), 4.83 (m, 1H, Fuc-H5), 4.42 (d, $J = 8.0$ Hz, 1H, Gal-H1), 4.15 (dd, $J = 4.6, 8.9$ Hz, 1H, Lac-H2), 4.08 (br s, 1H, Gal-H4), 3.94 - 3.86 (m, 2H, Fuc-H3, Gal-H4), 3.83 - 3.78 (m, 2H, Fuc-H2, Fuc-H4), 3.75 - 3.63 (m, 3H, MeCy-H1, Gal-H6), 3.58 - 3.49 (m, 2H, Gal-H2, Gal-H5), 3.28 (dd, $J = 3.2, 9.5$ Hz, 1H, Gal-H3), 3.22 (t, $J = 9.6$ Hz, 1H, MeCy-H2), 3.12 (dd, $J = 4.5, 14.0$ Hz, 1H, Lac-H3a), 2.93 (dd, $J = 8.9, 14.0$ Hz, 1H, Lac-H3b), 2.09 (m, 1H, MeCy-H6), 1.74 - 1.52 (m, 3H, MeCy-H3, MeCy-H4, MeCy-H5), 1.31 - 1.20 (m, 2H, MeCy-H5'), 1.18 (d, $J = 6.6$ Hz, 3H, Fuc-H6), 1.01 (d, $J = 6.2$ Hz, 3H, MeCy-Me), 1.09 (m, 1H, MeCy-H4');

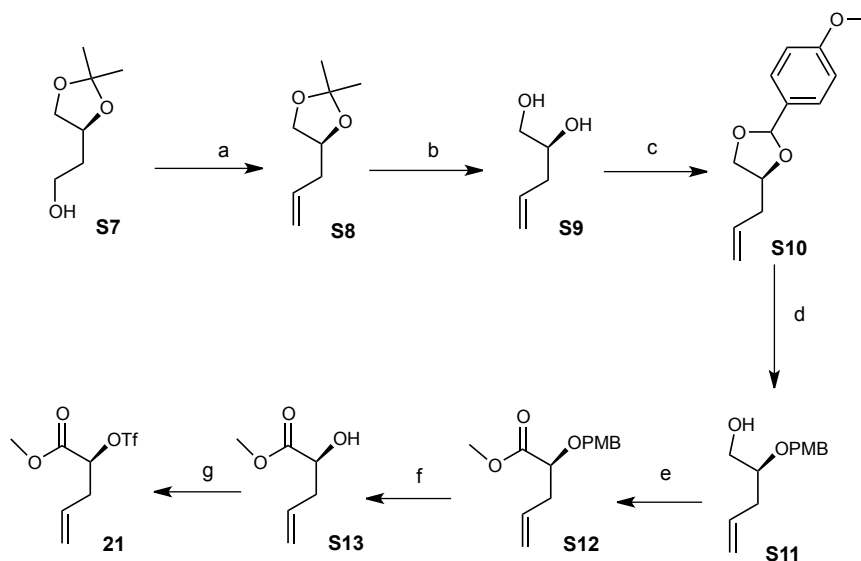
NMR (125.8 MHz, D₂O): δ = 180.6 (Lac-C1), 136.7, 135.1, 129.4, 129.1 (4C, Ar-C), 99.4, 98.9 (2C, Gal-C1, Fuc-C1), 84.1 (MeCy-C2), 82.7 (Gal-C3), 82.1 (Lac-C2), 78.6 (MeCy-C1), 74.2 (Gal-C2), 72.0 (Fuc-C2), 69.8 (Gal-C5), 69.3 (Fuc-C3), 68.3 (Fuc-C4), 66.5 (Fuc-C5), 66.2 (Gal-C4), 61.6 (Gal-C6), 38.9 (Lac-C3), 38.8 (MeCy-C3), 33.2 (MeCy-C4), 30.2 (MeCy-C6), 22.6 (MeCy-C5), 18.1 (MeCy-Me), 15.5 (Fuc-C6); HR-MS: m/z : Calcd for C₃₅H₄₃F₂Na₂O₁₄[M+Na]⁺: 645.2499, found: 645.2496.

(1*R*,2*R*,3*S*)-2-[(α -L-Fucopyranosyl)oxy]-3-methyl-cyclohex-1-yl 3-*O*-[sodium (*S*)-1-carboxy-2-(3,4-difluorophenyl)-ethyl]- β -D-galactopyranoside (6c**)**



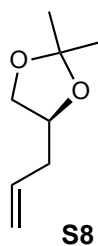
A suspension of **17c** (12 mg, 0.010 mmol) and Pd(OH)₂/C (50 mg, 10% Pd) in MeOH (5 mL) was hydrogenated (2 bar H₂) at rt. After 3 h TLC (silica: petroleum ether/EtOAc, 3:2; C-18: H₂O/MeOH, 1:1) indicated completion of the reaction. The reaction mixture was filtered through celite and concentrated *in vacuo*. The residue was dissolved in MeOH/H₂O (4:1, 1 mL) and treated with LiOH (12 mg, 0.51 mmol) for 7 days. The mixture was neutralized with Dowex 50X8 (H⁺ form), filtered through a Dowex 50 (Na⁺ form) ion exchange column and concentrated. The residue was purified by reversed phase column chromatography (C-18, H₂O/MeOH, gradient 0-100%) followed by a Sephadex G15 column, and lyophilization from water to yield the product **6c** as white foam (3.5 mg, 0.0054 mmol, 59%). R_f = (C-18, H₂O/MeOH, 1:1) 0.25; ¹H NMR (500.1 MHz, CDCl₃): δ = 7.33 – 7.25 (m, 1H, Ar-H), 7.24 – 7.15 (m, 1H, Ar-H), 7.15 – 7.03 (m, 1H, Ar-H), 5.11 (d, *J* = 4.0 Hz, 1H, Fuc-H1), 4.90 – 4.80 (m, 1H, Fuc-H5), 4.43 (d, *J* = 8.0 Hz, 1H, Gal-H1), 4.15 (dd, *J* = 4.8, 8.2 Hz, 1H, Lac-H2), 3.93 (d, *J* = 2.9 Hz, 1H, Gal-H4), 3.90 (dd, *J* = 3.2, 10.6 Hz, 1H, Fuc-H3), 3.87 – 3.78 (m, 2H, Fuc-H2, Fuc-H4), 3.76 – 3.66 (m, 3H, MeCy-H1, Gal-H6), 3.59 – 3.53 (m, 2H, Gal-H2, Gal-H5), 3.33 (dd, *J* = 3.1, 9.5 Hz, 1H, Gal-H3), 3.22 (t, *J* = 9.6 Hz, 1H, MeCy-H2), 3.09 (dd, *J* = 4.6, 14.1 Hz, 1H, Lac-H3), 2.97 (dd, *J* = 8.2, 14.1 Hz, 1H, Lac-H3'), 2.15 – 2.08 (m, 1H, MeCy-H6), 1.72 – 1.54 (m, 3H, MeCy-H3, MeCy-H5), 1.33 – 1.22 (m, 2H, MeCy-H6', MeCy-H4), 1.19 (d, *J* = 6.6 Hz, 3H, Fuc-H6), 1.10 (d, *J* = 6.3 Hz, 3H, MeCy-Me), 1.14 – 1.04 (m, 1H, MeCy-H4'); ¹³C NMR (125.8 MHz, D₂O): δ = 180.2 (Lac-C1), 135.1 (m, Ar-C), 125.7 (m, Ar-C), 118.3 (d, *J* = 17.0 Hz, Ar-C), 116.8 (d, *J* = 16.7

Hz, Ar-C), 99.7 (Gal-C1), 98.9 (Fuc-C1), 84.2 (MeCy-C2), 82.4 (Gal-C3), 81.6 (Lac-C2), 78.7 (MeCy-C1), 74.1 (Gal-C2), 72.0 (Fuc-C2), 69.8 (Gal-C5), 69.3 (Fuc-C3), 68.2 (Fuc-C4), 66.5 (Fuc-C5), 66.2 (Gal-C4), 61.7 (Gal-C6), 38.7 (Lac-C3), 38.4 (MeCy-C3), 33.2 (MeCy-C4), 30.3 (MeCy-C6), 22.6 (MeCy-C5), 18.1 (MeCy-Me), 15.5 (Fuc-C6); HR-MS: m/z : Calcd for $C_{28}H_{39}F_2Na_2O_{13}^+ [M+H]^+$: 667.2149, found: 667.2150.



Scheme S3: Synthesis of triflate **21**: a) (i) PCC, DCM; (ii) PPh_3BrMe , LiHMDS, THF (33% over two steps); b) $TsOH \cdot H_2O$, MeOH (97%); c.) CSA, para-anisaldehyde dimethyl acetal (55%); d) DIBAL-H, toluene; e) (i) DMSO, $(COCl)_2$, Et_3N , DCM; (ii) $tBuOH$, NaH_2PO_4 , 2-methyl-2-butene, $NaClO_2$; (iii) $TMSCHN_2$, MeOH/toluene (81% over 3 steps); f) DDQ, DCM, H_2O (87%); g) Tf_2O , DTBMP, DCM (87%).

4-Allyl-2,2-dimethyl-1,3-dioxolane (S8)

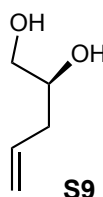


(4*S*)-(+)-4-(2-Hydroxyethyl)-2,2-dimethyl-1,3-dioxolane **S7** (1.00 g, 6.87 mmol) was dissolved in DCM (60 mL) and pyridiniumchlorochromate (7.40 g, 34.4 mmol) was added at 0°C. The reaction mixture was poured into Et_2O (100 mL) and the resulting mixture was filtered through a pad of celite. The solvent was removed *in vacuo* and the crude aldehyde used without further purification in the next step.

Methyltriphenylphosphonium bromide (3.70 g, 10.3 mmol) was suspended in THF (30 mL) and LiHMDS (1.0 M solution in THF, 8.93 ml, 8.93 mmol) was added dropwise at -78°C . The mixture was stirred for 30 min at this temperature and additional 45 min at 0°C . The crude aldehyde of the previous step was added and the mixture was stirred at room temperature for 16 h. Water (30 mL) was added and the solution was washed with DCM (3 x 200 mL). The combined organic layers were dried (Na_2SO_4), filtrated and the solvent was removed *in vacuo*. The crude product was purified by flash chromatography (petroleum ether/EtOAc 20:1) to yield **S8** as a colorless syrup (321 mg, 2.26 mmol, 33%).

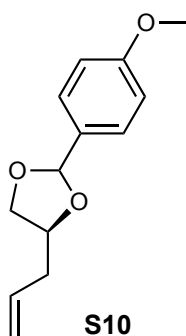
R_f (petroleum ether/EtOAc 10:1) 0.67; $[\alpha]_D^{22}$ 1.3 (c 4.90, CHCl_3); ^1H NMR (500.1 MHz, CDCl_3): δ = 5.82 - 5.73 (m, 1H, R-H4), 5.13 - 5.05 (m, 2H, R-H5), 4.17 - 4.11 (m, 1H, R-H2), 4.01 (dd, J = 6.0, 8.0, 1H, R-H1), 3.56 (dd, J = 7.2, 7.9, 1H, R-H1'), 2.42 - 2.36 (m, 1H, R-H3), 2.30 - 2.23 (m, 1H, R-H3'), 1.40 (s, 1H, Me), 1.34 (s, 1H, Me). ^{13}C NMR (125.8 MHz, CDCl_3): δ = 133.8 (R-C4), 117.81 (R-C5), 109.2 ($\text{C}(\text{CH}_3)_2$), 75.4 (R-C2), 69.1 (R-C1), 38.2 (R-C3), 27.1 (Me), 25.8 (Me).

(S)-Pent-4-ene-1,2-diol (**S9**)



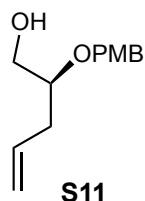
Acetal **S8** (673 mg, 4.73 mmol) was dissolved in MeOH (25 mL) and *p*-toluenesulfonic acid monohydrate (450 mg, 2.37 mmol) was added. The solution was stirred for 2.5 h at rt. The solvent was removed *in vacuo* and the crude product was purified by flash chromatography (DCM/MeOH 10:1) to yield **S9** as colorless oil (465 mg, 4.55 mmol, 96%).

R_f (petroleum ether/EtOAc 5:3) 0.14; $[\alpha]_D^{22}$ 11.4 (c 2.30, CHCl_3); ^1H NMR (500.1 MHz, CDCl_3): δ = 5.82 (ddt, J = 17.2, 10.2, 7.2 Hz, 1H, R-H5), 5.19 - 5.10 (m, 2H, R-H4), 3.81 - 3.73 (m, 1H, R-H2), 3.67 (dd, J = 11.2, 3.1 Hz, 1H, R-H1'), 3.48 (dd, J = 11.2, 7.2 Hz, 1H, R-H1), 2.30 - 2.18 (m, 2H, R-H3); ^{13}C NMR (125.8 MHz, CDCl_3): δ = 134.2 (R-C5), 118.4 (R-C4), 71.3 (R-C2), 66.4 (R-C1), 38.0 (R-C3).

(4S)-4-Allyl-2-(4-methoxyphenyl)-1,3-dioxolane (S10)

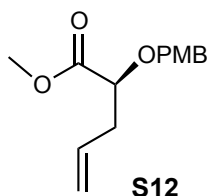
Diol **S9** (465 mg, 4.55 mmol) was dissolved in DCM (25 mL) and camphorsulfonic acid and *p*-methoxybenzaldehyde dimethyl acetal (1.66 g, 9.11 mmol) were added. The solution was stirred at 50°C for 15 h. The reaction was quenched with Et₃N and the solvent was removed in vacuo. The residue was purified by flash chromatography (petroleum ether/EtOAc 19:1) to give **S10** as mixture of diastereomers (1.0 g, 4.55 mmol, quant.) which was used without further purification in the next step.

R_f (petroleum ether/EtOAc 5:1) 0.75.

(S)-2-((4-Methoxybenzyl)oxy)pent-4-en-1-ol (S11)

Acetal **S10** (1.00 g, 4.55 mmol) was dissolved in 50 ml toluene and DIBAL-H (1M in toluene, 9.1 mL) was added at 0°C. The reaction mixture was allowed to reach room temperature over 4 h and was filtered through a short pad of silica. The filtrate was evaporated and the crude product was purified by flash chromatography (petroleum ether/EtOAc 10:1 to 4:1) to yield **S11** as a colorless liquid (0.72 g, 3.23 mmol, 71%).

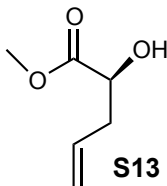
R_f (petroleum ether/EtOAc 5:1) 0.12; [α]_D²² 24.1 (*c* 1.40, CHCl₃); ¹H NMR (500.1 MHz, CDCl₃): δ = 7.30 - 7.26 (m, 2H, Ar), 6.90 - 6.88 (m, 2H, Ar), 5.85 - 5.77 (m, 1H, R-H4), 5.15 - 5.05 (m, 2H, R-H5), 4.60 (d, *J* = 11.7 Hz, 2H, Ph-CH₂), 4.47 (d, *J* = 11.2 Hz, 1H, Ph-CH₂), 3.80 (s, 3H, Me), 3.68 - 3.65 (m, 1H, R-H1), 3.58 - 3.51 (m, 2H, R-H1', R-H2), 2.42 - 2.37 (m, 1H, R-H3), 2.34 - 2.25 (m, 1H, R-H3'); ¹³C NMR (125.8 MHz, CDCl₃): δ = 134.3 (R-C4), 128.9 (Ar), 117.7 (R-C5), 114.1 (Ar), 78.9 (R-C2), 71.4 (Ph-CH₂), 65.2 (R-C1), 55.5 (Me), 35.5 (R-C3).

(S)-1-Methoxy-1-oxopent-4-en-2-yl 4-methoxybenzoate (S12)

Oxalylchloride (556 μ L, 6.56 mmol) was dissolved in DCM (25 mL) and dimethylsulfoxid (622 μ L, 8.75 mmol) was added at -78°C and the mixture was stirred for 15 min at this temperature. A solution of alcohol **S11** (487 mg, 2.19 mmol) DCM (5 mL) was added dropwise. The suspension was stirred for 35 min at -78°C , Et_3N (1.83 mL, 13.1 mmol) was added and the resulting solution was stirred for 1 h at rt. The reaction was quenched with water (80 mL) and the mixture was extracted with DCM (3 x 50 mL). The combined organic layers were dried, evaporated and used without further purification for the next step. The crude aldehyde was dissolved in *t*BuOH (17 mL) and NaH_2PO_4 (420 mg, 3.5 mmol, in 2.8 mL H_2O), 2-methylbut-2-ene (40 mL, 2M in THF) and NaClO_2 (633 mg, 7.00 mmol) were added successively. The resulting solution was stirred for 2.5 h and the volatiles evaporated under reduced pressure. The residue was solved in DCM (60 mL) and the solution was dried (Na_2SO_4), concentrated under reduced pressure and used without further purification in the next step.

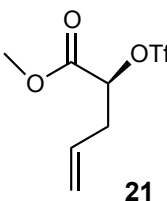
The crude acid was dissolved MeOH/toluene (1:2, 30 mL) and TMSCHN_2 (1.42 mL, 2 M in hexane) was added dropwise. The persistence of the yellow color indicated the end of the reaction. The solution was stirred for 1 h at rt, quenched with a few drops of Ac_2O and evaporated in vacuo. The crude ester was purified by flash chromatography (petroleum ether/EtOAc 6:1) to yield **S12** (444 mg, 1.77 mmol, 81 %) as a colorless liquid.

R_f (petroleum ether/EtOAc 5:1) 0.60; $[\alpha]_D^{22}$ -78.8 (c 0.90, CHCl_3); ^1H NMR (500.1 MHz, CDCl_3): δ = 7.27 (d, J = 8.7 Hz, 2H, Ar), 6.87 (d, J = 8.6 Hz, 2H, Ar), 5.80 (ddt, J = 17.2, 10.2, 7.0 Hz, 1H, R-H4), 5.15 – 5.04 (m, 2H, R-H5), 4.63 (d, J = 11.4 Hz, 1H, PhCH_2), 4.38 (d, J = 11.4 Hz, 1H, PhCH_2), 4.02 – 3.96 (m, 1H, R-H2), 3.80 (s, 3H, Me), 3.74 (s, 3H, COOMe), 2.53 – 2.48 (m, 2H, H-3); ^{13}C NMR (125.8 MHz, CDCl_3): δ = 133.2 (R-C4), 129.9 (Ar), 129.8 (Ar), 129.6 (Ar), 118.1 (R-C5), 113.9 (Ar), 77.6 (R-C2), 72.12 (PhCH_2), 55.4 (Me), 52.0 (COOMe), 37.5 (R-C3).

(S)-Methyl 2-hydroxypent-4-enoate (S13)

Ester **S12** (100 mg, 400 μmol) was dissolved in DCM (9 mL) and H_2O (0.4 mL) and DDQ (95.3 mg, 420 μM) were added. The reaction mixture was stirred for 16 h at rt. The volatiles were removed under reduced pressure and the crude product was purified by flash chromatography (petroleum ether/EtOAc 5:1) to yield **S13** (45.5 mg, 350 μmol , 87%) as a colorless liquid.

R_f (petroleum ether/EtOAc 5:1) 0.27; $[\alpha]_D^{22}$ 10.5 (c 4.40, CHCl_3); ^1H NMR (500.1 MHz, CDCl_3): δ = 5.80 (ddt, J = 17.2, 10.2, 7.1 Hz, 1H, R-H4), 5.20 – 5.11 (m, 2H, R-H5), 4.28 (dd, J = 6.3, 4.6 Hz, 1H, R-H2), 3.79 (s, 3H, COOMe), 2.62 – 2.54 (m, 1H, R-H3), 2.48 – 2.41 (m, 1H, R-H3'); ^{13}C NMR (125.8 MHz, CDCl_3): δ = 154.4 (COOMe), 132.7 (R-C4), 119.1 (R-C5), 70.2 (R-C2), 52.8 (Me), 38.9 (R-C3).

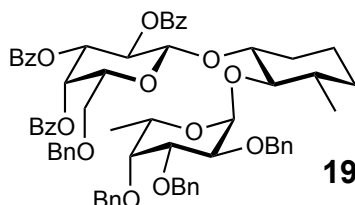
(S)-Methyl 2-(((trifluoromethyl)sulfonyl)oxy)pent-4-enoate (21)

Alcohol **S13** (35.2 mg, 270 μM) was dissolved in DCM (1.5 mL) and DTBMP (222 mg, 108 mM) and triflic anhydride (123 μL , 730 μmol) were added dropwise at -20°C - -30°C . The resulting solution was stirred for 45 min at this temperature and further 45 min at 0°C . The reaction was diluted with DCM (50 mL) and washed with ice cold KH_2PO_4 (40 mL, 1M in H_2O). The aqueous layer was washed with DCM (2 x 50 mL). The organic phase was dried (Na_2SO_4) and the solvent removed *in vacuo*. The residue was purified by flash chromatography (petroleum ether/EtOAc 6:1) to yield triflate **21** (61.6 mg, 235 μM) as a colorless liquid.

R_f (petroleum ether/EtOAc 5:1) 0.61; $[\alpha]_D^{22}$ -14.1 (c 2.00, CHCl_3); ^1H NMR (500.1 MHz, CDCl_3): δ = 5.80 – 5.69 (m, 1H, R-H4), 5.28 - 5.22 (m, 2H, R-H5), 5.16 (dd, J = 7.5, 4.5 Hz, 1H, R-H2), 3.85 (s, 3H, Me), 2.83 – 2.68 (m, 2H, R-H3); ^{13}C NMR (125.8 MHz,

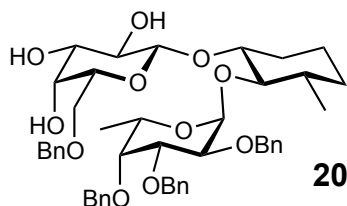
CDCl₃): δ = 167.1 (COO), 129.5 (R-C4), 121.3 (R-C5), 123.85 – 112.31 (q, J = 319.9 Hz, CF₃), 82.6 (R-C2), 53.4 (Me), 36.4 (R-C3).

(1*R*,2*R*,3*S*)-2-[(2,3,4-Tri-*O*-benzyl-6-deoxy- α -L-galactopyranosyl)oxy]-3-methyl-cyclohex-1-yl 2,3,4-tri-*O*-benzoyl-6-*O*-benzyl- β -D-galactopyranoside (19**)**

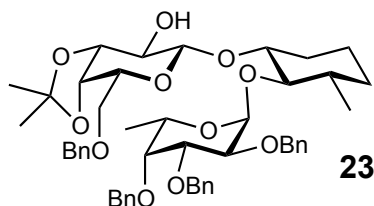


Compound **15**² (871 mg, 1.59 mmol) and thioglycoside **18** (1.10 g, 1.75 mmol) were dissolved in dry DCM (30 mL) and stirred together with powdered 4 Å activated molecular sieves (3 g) for 4 h at rt. DMTST (1.23 g, 4.77 mmol) was dissolved in DCM (10 mL) and stirred together with powdered 4 Å activated molecular sieves (1 g) for 3.5 h at rt as well. Both suspensions were combined and stirred for 3 days at rt. The mixture was filtered over a short pad of celite, washed with a sat. solution of NaHCO₃ (40 mL) and water (40 mL). The combined aqueous phases were extracted with DCM (3 x 30 mL). The combined organic layers were dried (Na₂SO₄) and the solvent was removed *in vacuo*. The crude product was purified by flash chromatography (petroleum ether/EtOAc 4:1) to afford **19** (1.40 g; 1.26 mmol; 79 %) as a white solid. Analytical data were in accordance with literature.²

(1*R*,2*R*,3*S*)-2-[(2,3,4-Tri-*O*-benzyl-6-deoxy- α -L-galactopyranosyl)oxy]-3-methyl-cyclohex-1-yl 6-*O*-benzyl- β -D-galactopyranoside (20**)**



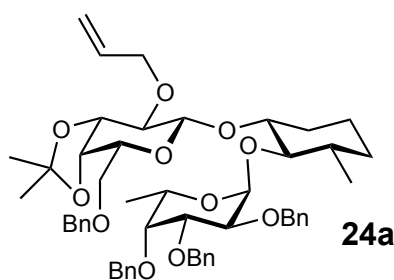
Benzoate **19** was suspended in a freshly prepared solution of NaOMe (12.6 mL, 0.02 M). The solution formed after a few minutes was stirred for 16 h at rt. The reaction mixture was neutralized with Dowex ion exchange resin (50x8), filtered and concentrated *in vacuo*. The crude product was purified by flash chromatography (CH₂Cl₂/MeOH 35/1) to yield **20** as a white solid (815 mg, 1.02 mmol, 81 %). Analytical data were in accordance with literature.²

(1R,2R,3S)-2-[(2,3,4-Tri-O-benzyl-6-deoxy- α -L-galactopyranosyl)oxy]-3-methyl-cyclohex-1-yl 6-O-benzyl-3,4-O-isopropylidene- β -D-galactopyranoside (23)

Alcohol **20** (810 mg, 1.01 mmol) was dissolved in acetone (50 mL) and 2,2-dimethoxypropane (497 μ L, 4.06 mmol), CuSO₄ (2.43 g, 15.2 mmol) and PPTS (25.5 mg, 0.10 mmol) were added. The mixture was stirred for 16 h and additional 2,2-dimethoxypropane (248 μ L, 505 μ mol) and PPTS (50.9 mg, 0.20 mmol) were added. The suspension was stirred for 24 h and filtered over a short pad of Al₂O₃. The solvent was removed *in vacuo* and the crude product was purified by flash chromatography (petroleum ether/EtOAc 3:2) to yield **23** as a white solid (708 mg, 844 μ mol, 83 %).

R_f (petroleum ether/EtOAc 1:1) 0.50; [α]_D²² -47.4 (*c* 0.70, CHCl₃); ¹H NMR (500.1 MHz, CDCl₃): δ = 7.35 – 7.21 (m, 20H, Ar-H), 5.08 (d, *J* = 3.6 Hz, 1H, Fuc-H1), 4.95 (d, *J* = 11.6 Hz, 1H, Ph-CH₂), 4.81 (d, *J* = 11.7 Hz, 1H, Ph-CH₂), 4.74 (d, *J* = 11.7 Hz, 2H, 2x Ph-CH₂), 4.67 (d, *J* = 11.5 Hz, 1H, Ph-CH₂), 4.63 – 4.55 (m, 3H, 2x Ph-CH₂, Fuc-H5), 4.51 (d, *J* = 11.9 Hz, 1H, Ph-CH₂), 4.24 (d, *J* = 8.3 Hz, 1H, Gal-H1), 4.18 (dd, *J* = 5.5, 2.1 Hz, 1H, Gal-H4), 4.09 – 4.00 (m, 3H, Fuc-H2, Gal-H3, Fuc-H3), 3.89 (td, *J* = 6.2, 2.0 Hz, 1H, Gal-H5), 3.78 (dd, *J* = 9.7, 6.2 Hz, 1H, Gal-H6), 3.76 – 3.68 (m, 2H, Gal-H6', MeCy-H1), 3.65 – 3.63 (m, 1H, Fuc-H4), 3.46 (t, *J* = 7.8 Hz, 1H, Gal-H2), 3.24 (t, *J* = 9.4 Hz, 1H, MeCy-H2), 2.14 – 2.07 (m, 1H, MeCy-H6), 1.67 – 1.58 (m, 3H, MeCy-H3, MeCy-H5, MeCy-H4), 1.44 (s, 3H, Me), 1.36 (s, 3H, Me), 1.34 – 1.14 (m, 2H, MeCy-H6', Cy-H5'), 1.11 – 1.07 (m, 6H, Fuc-H6, MeCy-Me), 1.06 – 0.95 (m, 1H, MeCy-H4'); ¹³C NMR (125.8 MHz, CDCl₃) δ 139.2, 139.0, 138.6, 138.3, 128.4 - 127.5 (24C, Ar-C), 110.1 (C(CH₃)₂), 99.5 (Gal-C1), 98.2 (Fuc-C1), 83.6 (MeCy-C2), 80.1 (Fuc-C3), 79.0 (Gal-C3), 78.8 (MeCy-C1), 78.2 (Fuc-C4), 76.5 (Fuc-C2), 74.9 (Ph-CH₂), 74.3 (Ph-CH₂), 73.7 (Ph-CH₂), 73.7 (Ph-CH₂), 73.5 (Gal-4), 72.9 (Gal-C2), 72.4 (Gal-C5), 69.5 (Gal-C6), 66.5 (Fuc-C5), 39.0 (MeCy-C3), 33.7 (MeCy-C4), 31.1 (MeCy-C6), 28.4 (Me), 26.4 (Me), 23.2 (MeCy-C5), 19.1 (MeCy-Me), 17.1 (Fuc-C6); ESI-MS: *m/z*: Calcd for C₅₀H₆₂NaO₁₁ [M+Na]⁺: 861.42, found: 861.35.

(1R,2R,3S)-2-[(2,3,4-Tri-O-benzyl-6-deoxy- α -L-galactopyranosyl)oxy]-3-methyl-cyclohex-1-yl 2-O-allyl-6-O-benzyl-3,4-O-isopropylidene- β -D-galactopyranoside (24a)

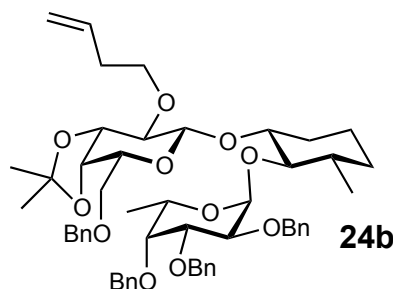


Alcohol **23** (204 mg, 243 μmol) was solved in 8 mL DMF and sodium hydride (60 % oil dispersion, 19.5 mg, 486 μmol) was added at 0°C. The mixture was stirred at this temperature for 1 h. The reaction was allowed to reach rt and allylbromide (63.0 μL , 723 μmol) was added. The suspension was stirred for 16 h and diluted with MeOH (30 mL) and Et₂O (30 mL). The solution was extracted with H₂O (40 mL) and the aqueous phase was washed with Et₂O (2 x 40 mL). The combined organic layers were dried (Na₂SO₄), concentrated and the crude product was purified by flash chromatography (petroleum ether/EtOAc 3:1) to give **24a** as a white foam (182 mg, 207 μmol , 85 %).

R_f (petroleum ether/EtOAc 2:1) 0.67; $[\alpha]_{\text{D}}^{22}$ -30.2 (*c* 1.20, CHCl₃); ¹H NMR (500.1 MHz, CDCl₃): δ = 7.34 - 7.19 (m, 20H, Ar-H), 5.94 - 5.84 (m, 1H, CH₂=CH-CH₂), 5.32 - 5.25 (m, 1H, CH₂=CH), 5.15 - 5.10 (m, 1H, CH₂=CH), 5.10 - 5.07 (m, 1H, Fuc-H1), 4.95 (d, *J* = 11.6 Hz, 1H, Ph-CH₂), 4.81 (d, *J* = 11.7 Hz, 1H, Ph-CH₂), 4.79 - 4.75 (m, 1H, Ph-CH₂), 4.75 - 4.71 (m, 2H, 2x Ph-CH₂), 4.66 (d, *J* = 11.4 Hz, 1H, Ph-CH₂), 4.62 (d, *J* = 11.6 Hz, 1H, Ph-CH₂), 4.58 (d, *J* = 12.0 Hz, 1H, Ph-CH₂), 4.48 (d, *J* = 12.0 Hz, 1H, Ph-CH₂), 4.33 - 4.27 (m, 1H, CH₂=CH-CH₂), 4.26 (d, *J* = 8.3 Hz, 1H, Gal-H1), 4.22 - 4.15 (m, 2H, CH₂=CH-CH₂, Gal-H4), 4.07 - 4.03 (m, 3H, Fuc-H2, Fuc-H3, Gal-H3), 3.85 - 3.79 (m, 1H, Gal-H5), 3.79 - 3.70 (m, 2H, Gal-H6, Gal-H6'), 3.68 - 3.65 (m, 1H, Fuc-H4), 3.64 - 3.58 (m, 1H, MeCy-H1), 3.27 (t, *J* = 9.1 Hz, 1H, MeCy-H2), 3.22 - 3.16 (m, 1H, Gal-H2), 2.08 - 2.01 (m, 1H, MeCy-H6), 1.66 - 1.58 (m, 3H, MeCy-H3, MeCy-H4, MeCy-H5), 1.41 (s, 3H, Me), 1.35 (s, 3H, Me), 1.33 - 1.14 (m, 2H, MeCy-H5', MeCy-H6'), 1.13 - 1.08 (m, 6H, MeCy-Me, Fuc-H6), 1.08 - 0.99 (m, 1H, MeCy-H4'); ¹³C NMR (125.8 MHz, CDCl₃): δ = ¹³C NMR (125.8 MHz, CDCl₃) δ 139.2, 139.13, 138.78, 138.30, 128.72 - 127.3 (24x Ar-C), 135.2 (CH₂=CH-CH₂), 116.7 (CH₂=CH-CH₂), 109.7 (C(CH₃)₂), 101.3 (Gal-C1), 98.2 (Fuc-C1), 82.5 (MeCy-C2), 80.3 (Fuc-C3), 80.2 (Gal-C2), 80.0 (MeCy-C1), 79.4 (Gal-C3), 78.2 (Fuc-C4), 76.6 (Fuc-C2), 74.8 (Ph-CH₂), 74.3 (Ph-CH₂), 73.7 (Ph-CH₂), 73.7 (Gal-C4), 73.0 (Ph-CH₂), 72.8 (Ph-CH₂), 71.8 (Gal-H5), 69.3 (Gal-C6), 66.1 (Fuc-C5), 39.1 (MeCy-C3),

33.6 (MeCy-C4), 23.0 (MeCy-C5), 19.1 (Fuc-C6), 17.3 (MeCy-Me); ESI-MS: m/z : Calcd for $C_{53}H_{66}NaO_{11}$ $[M+Na]^+$: 901.45, found: 901.47.

(1*R*,2*R*,3*S*)-2-[(2,3,4-Tri-*O*-benzyl-6-deoxy- α -L-galactopyranosyl)oxy]-3-methyl-cyclohex-1-yl 6-*O*-benzyl-2-*O*-(but-3-en-1-yl)-3,4-*O*-isopropylidene- β -D-galactopyranoside (24b)

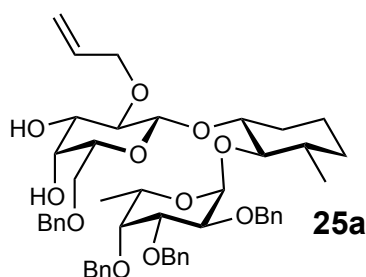


Alcohol **23** (716 mg, 854 μ mol) was dissolved in DMF (8 mL) and sodium hydride (60 % oil dispersion, 68.3 mg, 1.71 mmol) was added at 0°C. The mixture was stirred at this temperature for 1 h. The reaction was allowed to reach rt and triflate **21** (375 mg, 1.84 mmol) was added. The suspension was stirred for 16 h and diluted with MeOH (100 mL) and Et₂O (100 mL). The solution was extracted with H₂O (80 mL) and the aqueous phase was washed with Et₂O (2x 80 mL). The combined organic layers were dried (Na₂SO₄), concentrated and the crude product was purified by flash chromatography (petroleum ether/EtOAc 3:1) to give **24b** as a white foam (258 mg, 289 μ mol, 34 % (58%)). Reactant **23** was recovered in 42% yield (301 mg, 359 μ mol).

R_f (petroleum ether/EtOAc 1:1) 0.36; $[\alpha]_D^{22}$ -44.4 (c 0.50, CHCl₃); ¹H NMR (500.1 MHz, CDCl₃): δ = 7.34 – 7.17 (m, 20H, Ar-H), 5.86 – 5.74 (m, 1H, CH₂=CH-CH₂), 5.10 – 5.03 (m, 2H, Fuc-H1, CH₂=CH-CH₂), 5.02 – 4.97 (m, 1H, CH₂=CH-CH₂), 4.94 (d, J = 11.7 Hz, 1H, Ph-CH₂), 4.82 – 4.75 (m, 2H, Ph-CH₂, Fuc-H5), 4.72 (d, J = 10.0 Hz, 2H, 2x Ph-CH₂), 4.66 (d, J = 11.4 Hz, 1H, Ph-CH₂), 4.61 (d, J = 11.7 Hz, 1H, Ph-CH₂), 4.57 (d, J = 12.0 Hz, 1H, Ph-CH₂), 4.47 (d, J = 12.0 Hz, 1H, Ph-CH₂), 4.22 (d, J = 8.2 Hz, 1H, Gal-H1), 4.17 – 4.13 (m, 1H, Gal-H4), 4.07 – 4.04 (m, 2H, Fuc-H2, Fuc-H3), 4.03 – 3.99 (m, 1H, Gal-H3), 3.83 – 3.63 (m, 6H, 2x CH₂-CH₂-O, Gal-H5, Gal-H6, Gal-H6', Fuc-H4), 3.63 – 3.56 (m, 1H, MeCy-H1), 3.25 (t, J = 9.1 Hz, 1H, MeCy-H2), 3.13 – 3.08 (m, 1H, Gal-H2), 2.33 – 2.26 (m, 2H, CH₂=CH-CH₂), 2.05 – 1.99 (m, 1H, MeCy-H6), 1.65 – 1.54 (m, 3H, MeCy-H3, MeCy-H5, MeCy-H4), 1.41 (s, 3H, Me), 1.34 (s, 3H, Me), 1.31 – 1.12 (m, 2H, MeCy-H6', MeCy-H5'), 1.12 – 0.99 (m, 7H, Fuc-H6, MeCy-Me, MeCy-H4); ¹³C NMR (125.8 MHz, CDCl₃): δ = 139.2, 139.1, 138.8, 138.3, 128.6 – 127.3 (24C, Ar-C), 135.4 (CH₂=CH-CH₂),

116.4 (CH₂=CH-CH₂), 109.7 (Gal-C1), 101.1 (Fuc-C1), 98.1 (C(CH₃)₂), 82.3 (MeCy-C2), 81.1 (Gal-C2), 80.3 (Fuc-C3), 79.8 (MeCy-C1), 79.4 (Gal-C3), 78.2 (Fuc-C4), 76.6 (Fuc-C2), 74.8 (Ph-CH₂), 74.3 (Ph-CH₂), 73.7 (Ph-CH₂), 73.7 (Gal-C4), 72.8 (Ph-CH₂), 71.9 (CH₂-CH₂-O), 71.8 (Gal-C5), 69.4 (Gal-C6), 66.1 (Fuc-C5), 39.1 (MeCy-C3), 34.8 (CH₂=CH-CH₂), 33.6 (MeCy-C4), 31.0 (MeCy-C6), 28.3 (Me), 26.4 (Me), 21.0 (MeCy-C5), 19.1 (Fuc-C6), 17.3 (MeCy-Me); ESI-MS: *m/z*: Calcd for C₅₄H₆₈NaO₁₁ [M+Na]⁺: 915.47, found: 915.62.

(1*R*,2*R*,3*S*)-2-[(2,3,4-Tri-*O*-benzyl-6-deoxy- α -L-galactopyranosyl)oxy]-3-methyl-cyclohex-1-yl 2-*O*-allyl-6-*O*-benzyl- β -D-galactopyranoside (25a**)**



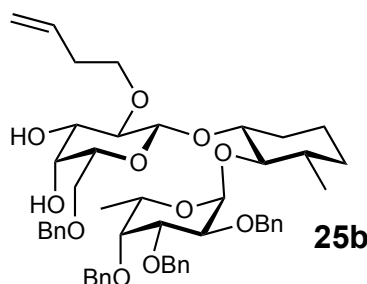
Acetic acid (1 mL, 80 %) was added to acetone **24a** (55.0 mg, 62.6 μmol) and the reaction mixture was stirred for 3 d at rt. The acetic acid was removed *in vacuo* and coevaporated with toluene. The crude product was purified by flash chromatography (petroleum ether/EtOAc 0 to 2:3) to yield **25a** in (41.0 mg, 48.9 μmol, 78%) as a white solide.

R_f (petroleum ether/EtOAc 2:1) 0.13; [α]_D²² -45.6 (*c* 3.64, CHCl₃); ¹H NMR (500.1 MHz, CDCl₃): δ = 7.36 – 7.21 (m, 20H, Ar-H), 5.96 – 5.86 (m, 1H, CH₂=CH), 5.30 – 5.24 (m, 1H, CH₂=CH), 5.19 – 5.14 (m, 1H, CH₂=CH), 5.08 (d, *J* = 3.1 Hz, 1H, Fuc-H1), 4.93 (d, *J* = 11.5 Hz, 1H, PhCH₂), 4.87 – 4.79 (m, 2H, Fuc-H5, Ph-CH₂), 4.76 – 4.71 (m, 2H, 2x Ph-CH₂), 4.66 (d, *J* = 11.4 Hz, 1H, Ph-CH₂), 4.58 (d, *J* = 11.5 Hz, 1H, Ph-CH₂), 4.50 (s, 2H, 2x Ph-CH₂), 4.49 – 4.44 (m, 1H, CH₂=CH-CH₂), 4.35 (d, *J* = 7.7 Hz, 1H, Gal-H1), 4.15 – 4.09 (m, 2H, Ph-CH₂, CH₂=CH-CH₂), 4.08 – 4.02 (m, 3H, Fuc-H3, Gal-H4, Fuc-H2), 3.77 (dd, *J* = 9.4, 6.9 Hz, 1H, Gal-H6), 3.73 – 3.71 (m, 1H, Fuc-H4), 3.70 – 3.59 (m, 2H, Gal-H6', MeCy-C1), 3.57 – 3.51 (m, 2H, Gal-H3, Gal-H5), 3.34 (dd, *J* = 9.2, 7.9 Hz, 1H, Gal-H2), 3.25 (t, *J* = 9.2 Hz, 1H, MeCy-H2), 2.11 – 2.05 (m, 1H, MeCy-H6), 1.70 – 1.59 (m, 3H, MeCy-H3, MeCy-H4, MeCy-H5), 1.37 – 1.17 (m, 2H, MeCy-H6', MeCy-H5'), 1.14 (d, *J* = 6.5 Hz, 3H, Fuc-H6), 1.10 (d, *J* = 6.5 Hz, 3H, MeCy-Me), 1.09 - 1.00 (m, 1H, MeCy-H4');

¹³C NMR (125.8 MHz, CDCl₃): δ = 139.3, 139.2, 138.7, 137.9, 128.8 – 127.3 (Ar-C), 135.0 (CH₂=CH), 117.3 (CH₂=CH), 101.4 (Gal-C1), 98.4 (Fuc-C1), 82.7 (MeCy-C2), 80.3 (Fuc-

C3), 79.7 (MeCy-C1), 79.1 (Gal-C1), 78.5 (Fuc-C4), 76.6 (Fuc-C2), 75.1 (Ph-CH₂), 74.4 (Ph-CH₂), 73.8 (Ph-CH₂), 73.7 (Gal-C3), 73.5 (Gal-C5), 72.6 (Ph-CH₂), 69.1 (Gal-C6), 68.5 (Gal-C4), 66.3 (Fuc-C5), 39.2 (MeCy-C3), 33.6 (MeCy-C4), 31.1 (MeCy-C6), 23.1 (MeCy-C5), 19.0 (MeCy-Me), 17.2 (Fuc-C6); ESI-MS: *m/z*: Calcd for C₅₀H₆₂NaO₁₁ [M+Na]⁺: 861.42, found: 861.35.

(1*R*,2*R*,3*S*)-2-[(2,3,4-Tri-*O*-benzyl-6-deoxy- α -L-galactopyranosyl)oxy]-3-methyl-cyclohex-1-yl 6-*O*-benzyl-2-*O*-(but-3-en-1-yl)- β -D-galactopyranoside (25b**)**

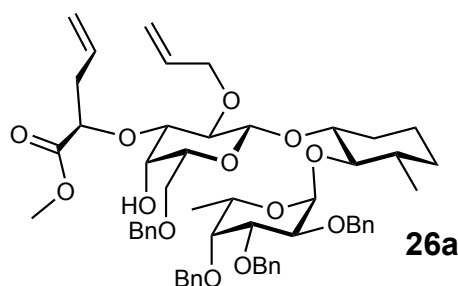


Acetic acid (2 ml, 80 %) was added to acetone **24b** (23.5 mg, 26.3 μ mol) and the reaction mixture was stirred for 3 d at rt. The acetic acid was removed *in vacuo* and coevaporated with toluene. The crude product was purified by flash chromatography (petroleum ether/EtOAc 0 to 60%) to yield **25b** in 91 % (20.5 mg, 24.0 μ mol).

R_f (petroleum ether/EtOAc 2:1) 0.15; $[\alpha]_D^{22}$ -77.9 (*c* 0.40, CHCl₃); ¹H NMR (500.1 MHz, CDCl₃): δ = 7.37 – 7.19 (m, 20H, Ar-H), 5.87 – 5.77 (m, 1H, CH₂=CH), 5.16 – 5.10 (m, 1H, CH₂=CH), 5.10 – 5.06 (m, 2H, CH₂=CH, Fuc-H1), 4.94 (d, *J* = 11.5 Hz, 1H, Ph-CH₂), 4.88 – 4.79 (m, 2H, Fuc-H5, Ph-CH₂), 4.77 – 4.71 (m, 2H, 2x Ph-CH₂), 4.67 (d, *J* = 11.4 Hz, 1H, Ph-CH₂), 4.59 (d, *J* = 11.5 Hz, 1H, Ph-CH₂), 4.50 (s, 2H, 2x Ph-CH₂), 4.32 (d, *J* = 7.7 Hz, 1H, Gal-H1), 4.09 – 3.98 (m, 4H, Fuc-H2, Fuc-H4, Gal-H4, CH₂=CH-CH₂), 3.77 (dd, *J* = 9.4, 7.0 Hz, 1H, Gal-H6), 3.75 – 3.72 (m, 1H, Fuc-H4), 3.71 – 3.59 (m, 3H, Gal-H6', MeCy-H2, CH₂=CH-CH₂), 3.56 – 3.49 (m, 2H, Gal-H5, Gal-H3), 3.31 – 3.22 (m, 2H, Gal-H2, MeCy-H2), 2.37 – 2.26 (m, 2H, O-CH₂-CH₂), 2.12 – 2.04 (m, 1H, MeCy-H6), 1.69 – 1.55 (m, 3H, MeCy-H5, MeCy-H3, MeCy-H4), 1.35 – 1.18 (m, 2H, MeCy-H6', MeCy-H5'), 1.15 (d, *J* = 6.5 Hz, 3H, Fuc-H6), 1.11 (d, *J* = 6.5 Hz, 3H, MeCy-Me), 1.09 – 1.00 (m, 1H, MeCy-H4'); ¹³C NMR (125.8 MHz, CDCl₃): δ = 139.3, 139.2, 138.7, 138.0, 128.8 – 127.3 (Ar-C), 136.0 (CH₂=CH), 117.1 (CH₂=CH), 101.3 (Gal-C1), 98.3 (Fuc-C1), 82.6 (MeCy-C2), 80.2 (Fuc-C3), 79.5 (MeCy-C1), 79.5 (Gal-C2), 78.5 (Fuc-C4), 76.5 (Fuc-C2), 75.1 (Ph-CH₂), 74.4 (Ph-CH₂), 73.7 (Ph-CH₂), 73.6 (Gal-C3), 72.7 (Gal-C5), 72.6 (Ph-CH₂), 71.8 (CH₂=CH-CH₂), 69.0 (Gal-C6), 68.2 (Gal-C4), 66.23 (Fuc-C5), 39.2 (Cy-C3), 34.8 (O-CH₂-

CH₂), 33.6 (MeCy-C4), 31.1 (MeCy-C6), 23.1 (MeCy-C5), 19.0 (MeCy-Me), 17.2 (Fuc-C6); ESI-MS: *m/z*: Calcd for C₅₁H₆₄NaO₁₁ [M+Na]⁺: 875.43, found: 875.47.

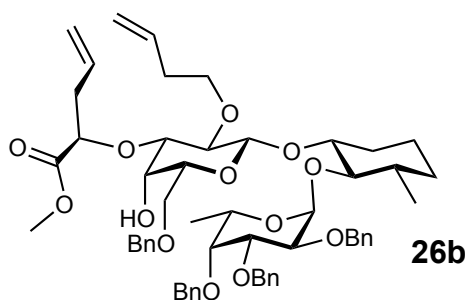
(1*R*,2*R*,3*S*)-2-[(2,3,4-Tri-*O*-benzyl-6-deoxy- α -L-galactopyranosyl)oxy]-3-methyl-cyclohex-1-yl 2-*O*-allyl-6-*O*-benzyl-3-*O*-((*R*)-1-methoxy-1-oxo-pent-4-en-2-yl)- β -D-galactopyranoside (26a**)**



Alcohol **25a** (38.5 mg, 45.9 μ mol) and Bu₂SnO (34.3 mg, 138 μ mol) were dried for 16 h at rt, suspended in MeOH (3 mL) and refluxed for 2 h. The resulting solution was concentrated and coevaporated with toluene and the tin acetal was dried for 16 h under reduced pressure. The acetal was dissolved in freshly dried (Al₂O₃ column) DME (3 mL) and added to CsF (dried for 2 h at high vacuo at 100°C, 12.5 mg, 82.6 μ mol). Triflate **21** (18.0 mg, 68.8 μ mol) was added and the suspension was stirred for 16 h at rt. A 10 % solution of KF (in aq. 1M KH₂PO₄, 3 mL) was added. After stirring for 1 h at rt DCM (4 mL) was added and the aq. phase was extracted with DCM (2 x 10 mL). The combined organic layers were dried (Na₂SO₄) and concentrated under reduced pressure. Column chromatography on silica (petroleum ether/EtOAc 2:1) afforded **26a** (25.0 mg, 26.3 μ mol, 57%) as a white solid. *R*_f (petroleum ether/EtOAc 2:1) 0.49; [α]_D²² -39.2 (*c* 2.90, CHCl₃); ¹H NMR (500.1 MHz, CDCl₃): δ = 7.39 – 7.15 (m, 20H, Ar-H), 5.93 – 5.75 (m, 2H, R-H4, O-CH₂CH=CH₂), 5.25 – 5.13 (m, 3H, O-CH₂-CH=CH₂, 2x R-H5), 5.12 – 5.04 (m, 2H, O-CH₂-CH=CH₂), 4.94 (d, *J* = 11.5 Hz, 1H, Ph-CH₂), 4.88 – 4.80 (m, 2H, Fuc-H5, Ph-CH₂), 4.76 – 4.66 (m, 3H, 3x Ph-CH₂), 4.60 (d, *J* = 11.5 Hz, 1H, Ph-CH₂), 4.49 (s, 2H, 2x Ph-CH₂), 4.40 – 4.34 (m, 1H, O-CH₂-CH=CH₂), 4.29 (d, *J* = 6.9 Hz, 1H, Gal-H1), 4.21 (dd, *J* = 8.0, 4.7 Hz, 1H, R-H2), 4.10 – 4.03 (m, 3H, O-CH₂CH=CH₂, Fuc-H2, Fuc-H3), 4.00 – 3.96 (m, 1H, Gal-H4), 3.79 – 3.75 (m, 2H, Fuc-H4, Gal-H6), 3.71 (s, 3H, Me), 3.66 – 3.57 (m, 2H, Gal-H6', MeCy-H1), 3.47 – 3.43 (m, 1H, Gal-H5), 3.43 – 3.35 (m, 2H, Gal-H2, Gal-H3), 3.24 (t, *J* = 9.2 Hz, 1H, MeCy-H2), 2.56 – 2.43 (m, 2H, R-H3), 2.08 – 1.98 (m, 1H, MeCy-H6), 1.64 – 1.53 (m, 3H, MeCy-H3, MeCy-H4, MeCy-H5), 1.37 – 1.23 (m, 2H, MeCy-H6', MeCy-H5'), 1.15 (d, *J* = 6.5 Hz, 3H, Fuc-H6), 1.12 – 0.98 (m, 4H, MeCy-Me, Cy-H4'); ¹³C NMR (125.8 MHz, CDCl₃): δ =

172.0 (COO), 139.4, 139.3, 138.8, 138.1, 128.8 – 127.2 (24C, Ar-C), 135.5 (CH₂-CH₂-CH=CH₂), 133.3 (R-C4), 119.1 (R-C5), 116.2 (CH₂-CH₂-CH=CH₂), 101.3 (Gal-C1), 98.3 (Fuc-C1), 82.3 (Cy-C2), 81.5 (Gal-C3), 80.3 (Fuc-C3), 79.4 (Cy-C1), 78.6 (Fuc-C4), 77.9 (Gal-C2), 76.5 (Fuc-C2), 75.1 (Ph-CH₂), 74.4 (Ph-CH₂), 73.8 (Ph-CH₂), 72.6 (Ph-CH₂), 72.2 (CH₂-CH₂-CH=CH₂), 72.1 (Gal-C5), 68.7 (Gal-C6), 66.4 (Gal-C4), 66.2 (Fuc-C5), 52.1 (Me), 39.2 (MeCy-C3), 37.9 (R-C3), 34.8 (CH₂-CH₂-CH=CH₂), 33.7 (MeCy-C4), 31.1 (MeCy-C6), 23.1 (MeCy-C5), 19.0 (MeCy-Me), 17.2 (Fuc-C6); ESI-MS: *m/z*: Calcd for C₅₆H₇₀NaO₁₃ [M+Na]⁺: 973.47, found: 973.48.

(1*R*,2*R*,3*S*)-2-[(2,3,4-Tri-*O*-benzyl-6-deoxy- α -L-galactopyranosyl)oxy]-3-methyl-cyclohex-1-yl 2-*O*-(but-3-en-1-yl)-6-*O*-benzyl-3-*O*-((*R*)-1-methoxy-1-oxo-pent-4-en-2-yl)- β -D-galactopyranoside (26b**)**

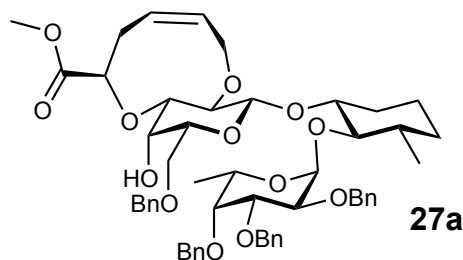


Alcohol **25b** (24.5 mg, 28.7 μ mol) and Bu₂SnO (21.4 mg, 86.2 μ mol) were dried for 16 h at rt, suspended in MeOH (1.5 mL) and refluxed for 2 h. The resulting solution was concentrated and coevaporated with toluene and the tin acetal was dried for 16 h under reduced pressure. The acetal was dissolved in freshly dried (Al₂O₃ column) DME (1.5 mL) and added to CsF (dried for 2 h at high vacuo at 100°C, 7.85 mg, 51.7 μ mol). Triflate **21** (11.3 mg, 43.1 μ mol) was added and the suspension was stirred for 16 h at rt. A 20% solution of KF (in aq. 1M KH₂PO₄, 2 mL) was added. After stirring for 1 h at rt DCM (4mL) was added and the aqueous phase was extracted with DCM (2 x 10 mL). The combined organic layers were dried (Na₂SO₄) and concentrated under reduced pressure. Column chromatography on silica (petroleum ether/EtOAc 2:1) afforded **26b** (25.9 mg, 26.9 μ mol, 93%) as a white solid.

*R*_f (petroleum ether/EtOAc 2:1) 0.65; [α]_D²² -51.9 (*c* 0.60, CHCl₃); ¹H NMR (500.1 MHz, CDCl₃): δ = 7.37 – 7.20 (m, 20H, Ar-H), 5.86 – 5.75 (m, 2H, R-H4, CH₂-CH₂-CH=CH₂), 5.21 – 5.14 (m, 2H, R-H5, R-H5'), 5.08 – 4.98 (m, 3H, 2x CH₂-CH₂-CH=CH₂, Fuc-H1), 4.94 (d, *J* = 11.5 Hz, 1H, Ph-CH₂), 4.88 – 4.80 (m, 2H, Fuc-H5, Ph-CH₂), 4.75 – 4.67 (m,

3H, 3x Ph-CH₂), 4.60 (d, $J = 11.5$ Hz, 1H, Ph-CH₂), 4.49 (s, 2H, 2x Ph-CH₂), 4.28 (d, $J = 7.2$ Hz, 1H, Gal-H1), 4.24 (dd, $J = 7.9, 4.8$ Hz, 1H, R-H2), 4.08 – 4.02 (m, 2H, Fuc-H2, Fuc-H3), 4.00 – 3.96 (m, 1H, Gal-H4), 3.90 – 3.84 (m, 1H, CH₂-CH₂-CH=CH₂), 3.80 – 3.74 (m, 2H, Gal-H6, Fuc-H4), 3.73 (s, 3H, Me), 3.67 – 3.55 (m, 3H, Gal-H6', MeCy-H1, CH₂-CH₂-CH=CH₂), 3.46 – 3.41 (m, 1H, Gal-H5), 3.38 – 3.31 (m, 2H, Gal-H2, Gal-H3), 3.24 (t, $J = 9.2$ Hz, 1H, MeCy-H2), 2.58 – 2.43 (m, 2H, R-H3), 2.30 – 2.23 (m, 2H, 2x CH₂-CH₂-CH=CH₂), 2.07 – 2.01 (d, 1H, MeCy-H6), 1.70 – 1.53 (m, 3H, MeCy-H3, MeCy-H5, MeCy-H4), 1.32 – 1.18 (m, 2H, MeCy-H6', MeCy-H5'), 1.15 (d, $J = 6.5$ Hz, 3H, Fuc-H6), 1.12 – 0.97 (m, 4H, MeCy-Me, MeCy-H4'); ¹³C NMR (125.8 MHz, CDCl₃): $\delta = 172.0$ (COO), 139.4, 139.3, 138.8, 138.1, 128.8 – 127.2 (24C, Ar-C), 135.5 (CH₂-CH₂-CH=CH₂), 133.3 (R-C4), 119.1 (R-C5), 116.2 (CH₂-CH₂-CH=CH₂), 101.3 (Gal-C1), 98.3 (Fuc-C1), 82.3 (Cy-C2), 81.5 (Gal-C3), 80.3 (Fuc-C3), 79.4 (Cy-C1), 78.6 (Fuc-C4), 77.9 (Gal-C2), 76.5 (Fuc-C2), 75.1 (Ph-CH₂), 74.4 (Ph-CH₂), 73.8 (Ph-CH₂), 72.6 (Ph-CH₂), 72.2 (CH₂-CH₂-CH=CH₂), 72.1 (Gal-C5), 68.7 (Gal-C6), 66.4 (Gal-C4), 66.2 (Fuc-C5), 52.1 (Me), 39.2 (MeCy-C3), 37.9 (R-C3), 34.8 (CH₂-CH₂-CH=CH₂), 33.7 (MeCy-C4), 31.1 (MeCy-C6), 23.1 (MeCy-C5), 19.0 (MeCy-Me), 17.2 (Fuc-C6); ESI-MS: m/z : Calcd for C₅₇H₇₂NaO₁₃ [M+Na]⁺: 987.49, found 987.65.

Compound 27a

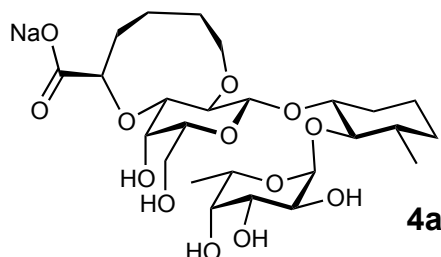


Compound **26a** (10.0 mg, 10.5 μ mol) was dissolved in DCM (2 mL) and Grubbs 2nd generation catalyst (0.89 mg, 1.05 μ mol) was added and the solution stirred for 2 h at rt. The solvent was removed *in vacuo* and the crude product purified by flash chromatography (petroleum ether/EtOAc 3:1) to yield **27a** (8.30 mg, 8.99 μ mol, 86%) as a white solid.

R_f (petroleum ether/EtOAc 3:1) 0.27; $[\alpha]_D^{22} -25.1$ (c 0.57, CHCl₃); ¹H NMR (500.1 MHz, CDCl₃): $\delta = 7.36 - 7.17$ (m, 20H, Ar-H), 5.60 – 5.52 (m, 2H, R-H4, R-H5), 5.07 (d, $J = 2.6$ Hz, 1H, Fuc-H1), 4.93 – 4.86 (m, 2H, Ph-CH₂, Fuc-H5), 4.82 (d, $J = 11.7$ Hz, 1H, Ph-CH₂), 4.75 – 4.68 (m, 2H, 2x Ph-CH₂), 4.65 (d, $J = 11.4$ Hz, 1H, Ph-CH₂), 4.59 – 4.46 (m, 4H, 3x Ph-CH₂, R-H5), 4.41 (dd, $J = 12.6, 2.5$ Hz, 1H, R-H2), 4.26 (d, $J = 7.7$ Hz, 1H, Gal-H1),

4.20 – 4.11 (m, 2H, Gal-H4, R-H5'), 4.08 – 4.02 (m, 2H, Fuc-H2, Fuc-H3), 3.80 – 3.70 (m, 6H, Gal-H3, Me, Fuc-H4, Gal-H6), 3.68 – 3.60 (m, 2H, MeCy-H1, Gal-H6'), 3.55 – 3.46 (m, 2H, R-H3, Gal-H5), 3.24 (t, $J = 9.2$ Hz, 1H, MeCy-H2), 3.18 (t, $J = 8.2$ Hz, 1H, Gal-H2), 2.35 – 2.29 (m, 1H, R-H3'), 2.11 – 2.03 (m, 1H, MeCy-H6), 1.68 – 1.59 (m, 3H, MeCy-H3, MeCy-H4, MeCy-H5), 1.33 – 1.15 (m, 2H, MeCy-H5', MeCy-H6'), 1.13 (d, $J = 6.5$ Hz, 3H, Fuc-H6), 1.10 (d, $J = 6.5$ Hz, 3H, MeCy-Me), 1.08 – 0.97 (m, 1H, MeCy-H4'); ^{13}C NMR (125.8 MHz, CDCl_3): $\delta = 172.5$ (COO), 139.4, 139.3, 138.8, 138.2, 130.8, 128.8 – 127.3, 126.9 (26C, 24x Ar-C, CH=CH), 101.5 (Gal-C1), 98.3 (Fuc-C1), 82.5 (MeCy-C2), 80.5 (Gal-C2), 80.3 (Fuc-C3), 80.0 (MeCy-C1), 78.7 (Fuc-C4), 76.6 (R-C2), 76.5 (Fuc-C2), 76.2 (Gal-C3), 75.1 (Ph-CH₂), 74.4 (Ph-CH₂), 73.7 (Ph-CH₂), 72.6 (Gal-C5), 72.6 (Ph-CH₂), 71.9 (R-C5), 69.5 (Gal-C6), 69.2 (Gal-C4), 66.3 (Fuc-C5), 52.1 (Me), 39.1 (MeCy-C3), 33.7 (MeCy-C4), 31.6 (R-C3), 31.0 (MeCy-C6), 23.0 (MeCy-C5), 19.0 (MeCy-Me), 17.1 (Fuc-C6); ESI-MS: m/z : Calcd for $\text{C}_{54}\text{H}_{66}\text{NaO}_{13}$ $[\text{M}+\text{Na}]^+$: 945.44, found: 945.47.

Compound 4a



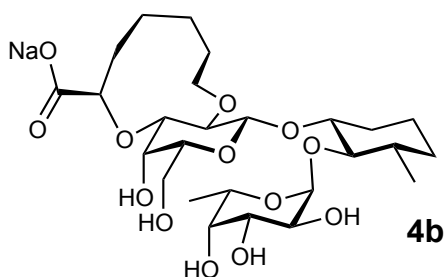
Compound **27a** (12.2 mg, 13.2 μmol) was dissolved in dioxane/water (4:1, 1 mL) and $\text{Pd}(\text{OH})_2/\text{C}$ (1.5 mg, 10% $\text{Pd}(\text{OH})_2$) was added. The suspension was stirred for 16 h under an atmosphere of hydrogen. Solvent was removed *in vacuo*.

The resulting ester was dissolved in H_2O (1 mL) and LiOH (0.63 mg, 26.4 μmol) was added. The turbid reaction mixture was stirred for 24 h at rt and the solvent was removed under reduced pressure. The residue was purified *via* RP chromatography ($\text{MeOH}/\text{H}_2\text{O}$), eluted through a sodium exchange column (Dowex 50/8 sodium form) and finally purified *via* size exclusion chromatography. Lyophilization from water/dioxane gave **4a** (4.30 mg, 7.51 μmol , 57 %) as a white fluffy foam.

R_f ($\text{DCM}/\text{MeOH}/\text{H}_2\text{O}$ 10:5:0.4) 0.30; $[\alpha]_D^{22}$ -68.1 (c 1.10, H_2O); ^1H NMR (500.1 MHz, D_2O) $\delta = 5.13$ (d, $J = 4.0$ Hz, 1H, Fuc-H1), 4.94 - 4.89 (m, 1H, Fuc-H5), 4.49 (d, $J = 7.8$ Hz, 1H, Gal-H1), 4.32 - 4.26 (m, 1H, R-H2), 4.18 - 4.15 (m, 1H, Gal-H5), 3.97 - 3.90 (m, 2H, Fuc-H3, R-H5), 3.84 - 3.78 (m, 3H, Fuc-H4, Fuc-H2, R-H5'), 3.77 - 3.65 (m, 4H, Gal-H6',

Gal-H6, Gal-H3, Cy-H1), 3.53 (t, $J = 5.9$ Hz, 1H, Gal-H4), 3.35 (dd, $J = 9.2, 8.2$ Hz, 1H, Gal-H2), 3.24 (t, $J = 9.4$ Hz, 1H, Cy-H2), 2.20 - 2.09 (m, 3H, R-H4, R-H2, MeCy-H6), 1.86 - 1.73 (m, 3H, R-H2', R-H3, R-H3'), 1.73 - 1.58 (m, 3H, MeCy-H5, MeCy-H4, MeCy-H3), 1.46 - 1.39 (m, 1H, R-H4), 1.37 - 1.24 (m, 2H, MeCy-H6', MeCy-H5'), 1.22 (d, $J = 6.6$ Hz, 3H, Fuc-H6), 1.11 (d, $J = 6.4$ Hz, 4H, MeCy-Me, MeCy-H4'); ^{13}C NMR (125.8 MHz, D_2O): $\delta = 180.3$ (COO), 100.1 (Gal-C1), 98.7 (Fuc-C1), 83.8 (MeCy-C2), 81.0 (R-C1), 79.0 (MeCy-C1), 77.6 (Gal-C2), 75.3 (Gal-C3), 74.6 (Gal-C4), 74.0 (R-C5), 72.0 (Fuc-C4), 69.3 (Fuc-C3), 69.0 (Gal-C5), 68.2 (Fuc-C2), 66.5 (Fuc-C5), 61.7 (Gal-C6), 38.5 (MeCy-C3), 33.1 (MeCy-C4), 30.4 (MeCy-C6), 29.0 (R-C2), 26.1 (R-C3), 25.3 (R-C4), 22.4 (MeCy-C5), 18.2 (MeCy-Me), 15.7 (Fuc-C6); ESI-MS: m/z : Calcd for $\text{C}_{25}\text{H}_{42}\text{NaO}_{13}$ $[\text{M}+\text{H}]^+$: 573.2523, found: 573.2525.

Compound 4b



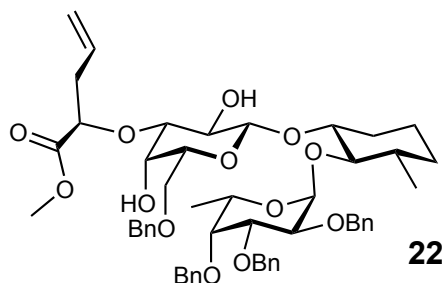
Compound **26a** (25.5 mg, 26.4 μmol) was dissolved in DCM (mL) and Grubbs 2nd generation catalyst (4.49 mg, 5.28 μmol) was added. The solution was stirred for 2 h at rt. The solvent was removed *in vacuo* and the crude product purified by short column of silica (petroleum ether/EtOAc 3:1) to give **27a** (20.4 mg) as a white solid.

3.7 mg of the olefin **27a** (3.95 μmol) was dissolved in dioxane/water (4:1) and $\text{Pd}(\text{OH})_2/\text{C}$ (1 mg, 10% $\text{Pd}(\text{OH})_2$) was added. The suspension was stirred for 16 h under an atmosphere of hydrogen. The solvent was removed under reduced pressure. The residue was dissolved H_2O (1 mL) and LiOH (0.19 mg, 7.90 μmol) was added. The resulting turbid reaction mixture was stirred for 24 h at rt. The solvent was removed under reduced pressure and the crude product was purified by RP chromatography (MeOH/ H_2O), eluted through a sodium exchange column (Dowex 50/8 sodium form) and finally purified by size exclusion chromatography. Lyophilization from water/dioxane afforded **4b** (800 μg , 1.42 μmol , 25 %) as a white fluffy foam.

R_f (DCM/MeOH/ H_2O 10:5:0.4) 0.28; ^1H NMR (500.1 MHz, CDCl_3): $\delta = 5.14$ (d, $J = 4.0$ Hz, 1H, Fuc-H1), 4.99 – 4.93 (m, 1H, Fuc-H5), 4.52 (d, $J = 7.3$ Hz, 1H, Gal-H1), 4.40 –

4.35 (m, 1H, R-H2), 4.24 – 4.19 (m, 1H, R-H7), 4.11 – 4.08 (m, 1H, Gal-H5), 3.92 (dd, $J = 10.5, 3.3$ Hz, 1H, Fuc-H3), 3.82 (d, $J = 22.7$ Hz, 4H, Fuc-H2, Gal-H6, R-H7', Fuc-H4), 3.70 (m, 4H, Gal-H2, Gal-H6', MeCy-H1, Gal-H3), 3.54 (t, $J = 6.0$ Hz, 1H, Gal-H4), 3.25 (t, $J = 9.5$ Hz, 1H, MeCy-H2), 2.20 - 2.13 (m, 1H, MeCy-H6), 1.87 – 1.49 (m, 10H, MeCy-H3, MeCy-H5, MeCy-H4, R-H3, MeCy-H3', MeCy-H4, MeCy-H4', MeCy-H5, MeCy-H5', MeCy-H6), 1.44 - 1.28 (m, 3H, MeCy-H6', R-H6', MeCy-H5'), 1.27 (d, $J = 6.6$ Hz, 3H, Fuc-H6), 1.11 (d, $J = 6.4$ Hz, 3H, MeCy-Me); ^{13}C NMR (125.8 MHz, D_2O): $\delta = 180.3$ (COO), 100.1 (Gal-C1), 98.7 (Fuc-C1), 83.8 (MeCy-C2), 81.0 (R-C1), 79.0 (MeCy-C1), 77.6 (Gal-C2), 75.3 (Gal-C3), 74.6 (Gal-C4), 74.0 (R-C5), 72.0 (Fuc-C4), 69.3 (Fuc-C3), 69.0 (Gal-C5), 68.2 (Fuc-C2), 66.5 (Fuc-C5), 61.7 (Gal-C6), 38.5 (MeCy-C3), 33.1 (MeCy-C4), 30.4 (MeCy-C6), 29.0 (R-C2), 26.1 (R-C3), 25.3 (R-C4), 22.4 (MeCy-C5), 18.2 (MeCy-Me), 15.7 (Fuc-C6); ESI-MS: m/z : Calcd for $\text{C}_{26}\text{H}_{44}\text{NaO}_{13}$ $[\text{M}+\text{H}]^+$: 587.2680, found: 587.2681.

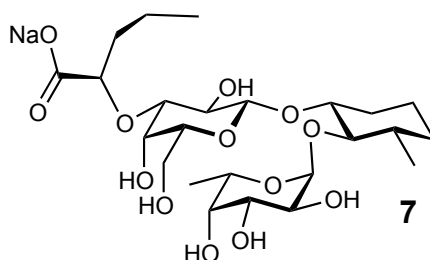
(1*R*,2*R*,3*S*)-2-[(2,3,4-Tri-*O*-benzyl-6-deoxy- α -*L*-galactopyranosyl)oxy]-3-methyl-cyclohex-1-yl 6-*O*-benzyl-3-*O*-((*R*)-1-methoxy-1-oxo-pent-4-en-2-yl)- β -*D*-galactopyranoside (22)



Alcohol **20** (30.8 mg, 38.6 μmol) and Bu_2SnO (28.8 mg, 116 μmol) were dried for 8 h at rt, suspended in MeOH (2 mL) and refluxed for 2 h. The resulting solution was concentrated and coevaporated with toluene. The resulting tin acetal was dried for 16 h under reduced pressure. The residue was dissolved in DME (1.5 mL) and added to CsF (dried for 2 h at high vacuo at 100°C , 10.5 mg, 69.4 μmol). Triflate **21** (15.2 mg, 57.8 μmol) was added and the suspension was stirred for 16 h at rt. A 20% solution of KF (in 1M KH_2PO_4 solution, 3 mL) was added. After stirring for 1 h at rt DCM (10 mL) was added. The aqueous phase was extracted with DCM (2 x 10 mL). The combined organic layers were dried (Na_2SO_4), concentrated under reduced pressure and purified by column chromatography on silica (petroleum ether/EtOAc 2:1) to yield **22** (20.1 mg, 22.1 μmol , 57%) as a white solid.

R_f (petroleum ether/EtOAc 1:1) 0.3; $[\alpha]_D^{22}$ -21.5 (*c* 0.64, CHCl₃); ¹H NMR (500.1 MHz, CDCl₃): δ = 7.36 – 7.21 (m, 20H, Ar-H), 5.91 – 5.80 (m, 1H, R-H4), 5.26 – 5.17 (m, 2H, R-H5, R-H5'), 5.07 (d, *J* = 3.5 Hz, 1H, Fuc-H1), 4.94 (d, *J* = 11.5 Hz, 1H, Ph-CH₂), 4.81 (d, *J* = 11.7 Hz, 1H, Ph-CH₂), 4.77 – 4.70 (m, Ph-CH₂, 3H), 4.68 (d, *J* = 11.4 Hz, 1H, Ph-CH₂), 4.60 (d, *J* = 11.5 Hz, 1H, Ph-CH₂), 4.52 (s, 2H, 2x Ph-CH₂), 4.30 (d, *J* = 7.7 Hz, 1H, Gal-H1), 4.13 (dd, *J* = 9.1, 3.7 Hz, 1H, R-H2), 4.06 (dd, *J* = 10.3, 3.5 Hz, 1H, Fuc-H2), 4.02 (dd, *J* = 10.3, 2.5 Hz, 1H, Fuc-H3), 3.97 (d, *J* = 2.8 Hz, 1H, Gal-H4), 3.77 (dd, *J* = 9.3, 7.4 Hz, 1H, Gal-H6), 3.75 (s, 3H, Me), 3.73 – 3.63 (m, 4H, MeCy-H1, Fuc-H4, Gal-H2, Gal-H6'), 3.52 – 3.47 (m, 1H, Gal-H5), 3.30 – 3.24 (m, 2H, Gal-H3, MeCy-H2), 2.63 – 2.56 (m, 1H, R-H3), 2.52 – 2.44 (m, 1H, R-H3'), 2.11 – 2.05 (m, 1H, MeCy-H6), 1.68 – 1.55 (m, 3H, MeCy-H5, MeCy-H3, MeCy-H4), 1.38 – 1.24 (m, 1H, MeCy-H6'), 1.24 – 1.16 (m, 1H, MeCy-H5'), 1.14 (d, *J* = 6.5 Hz, 3H, Fuc-H6), 1.09 (d, *J* = 6.4 Hz, 3H, MeCy-Me), 1.07 – 0.97 (m, 1H, MeCy-H4'); ¹³C NMR (125.8 MHz, CDCl₃): δ = 173.2 (COO), 139.3, 139.2, 138.7, 138.1, 128.7 – 127.3 (24x Ar-C), 133.5 (R-C4), 119.3 (R-C5), 100.5 (Gal-C1), 98.3 (Fuc-C1), 83.0 (MeCy-C2), 82.9 (Gal-C3), 80.2 (Fuc-C3), 79.0 (MeCy-C1), 78.6 (Fuc-C4), 76.5 (Fuc-C2), 75.0 (Ph-CH₂), 74.6 (Ph-CH₂), 73.8 (Ph-CH₂), 72.7 (Gal-C5), 72.6 (Ph-CH₂), 69.9 (Gal-C2), 68.7 (Gal-C6), 66.4 (Fuc-C5), 65.7 (Gal-C4), 52.6 (Me), 39.1 (MeCy-C3), 37.6 (R-C3), 33.7 (MeCy-C4), 30.9 (MeCy-C6), 23.1 (MeCy-C5), 19.1 (MeCy-Me), 17.0 (Fuc-C6); ESI-MS: *m/z*: Calcd for C₅₃H₆₆NaO₁₃ [M+Na]⁺: 933.44, found: 933.53.

(1*R*,2*R*,3*S*)-2-[(α -L-Fucopyranosyl)oxy]-3-methyl-cyclohex-1-yl 3-*O*-[sodium (*R*)-1-carboxybutane]- β -D-galactopyranoside (7)

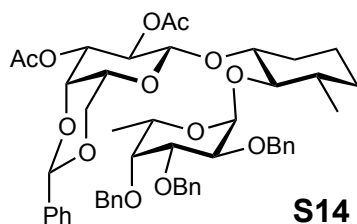


Compound **22** (14.0 mg, 15.4 μ mol) was dissolved in dioxane/water (4:1) and Pd(OH)₂/C (2.0 mg, 10% Pd(OH)₂) was added. The suspension was stirred for 16 h under an atmosphere of hydrogen. The solvent was removed *in vacuo*. The residue was dissolved in H₂O (1 ml) and LiOH (0.74 mg, 30.8 μ mol) was added. The turbid mixture was stirred for 16 h at ambient temperature. The solvent was removed under reduced pressure and the

residue was purified *via* RP chromatography (MeOH/H₂O), eluated through a sodium exchange column (Dowex 50/8 sodium form) and finally purified by size exclusion chromatography. Lyophilization from water/dioxane gave **7** (4.30 mg, 7.98 μmol, 52 %) as a white fluffy foam.

R_f (DCM/MeOH/H₂O 10:5:0.4) 0.22; [α]_D²² -7.6 (*c* 0.48, H₂O); ¹H NMR (500.1 MHz, D₂O): δ = 5.13 (d, *J* = 4.1 Hz, 1H, Fuc-H1), 4.90 – 4.85 (m, 1H, Fuc-H5), 4.49 (d, *J* = 7.8 Hz, 1H, Gal-H1), 4.16 (dd, *J* = 7.6, 5.0 Hz, 1H, R-H2), 4.08 (d, *J* = 3.1 Hz, 1H, Gal-H4), 3.92 (dd, *J* = 10.5, 3.3 Hz, 1H, Fuc-H3), 3.84 – 3.79 (m, 2H, Fuc-H2, Fuc-H4), 3.79 – 3.70 (m, 3H, MeCy-H1, Gal-H6, Gal-H6'), 3.65 (dd, *J* = 9.7, 7.9 Hz, 1H, Gal-H2), 3.57 (dd, *J* = 7.8, 4.1 Hz, 1H, Gal-H5), 3.48 (dd, *J* = 9.8, 3.2 Hz, 1H, Gal-H3), 3.25 (t, *J* = 9.6 Hz, 1H, MeCy-H2), 2.19 – 2.12 (m, 1H, MeCy-H6), 1.78 – 1.70 (m, 2H, R-H2, R-H2', MeCy-H5), 1.70 – 1.57 (m, 3H, MeCy-H4, R-H3', MeCy-H3), 1.51 – 1.41 (m, 2H, R-H4, R-H4'), 1.35 – 1.24 (m, 2H, MeCy-H6', MeCy-H5'), 1.21 (d, *J* = 6.6 Hz, 3H, Fuc-H6), 1.11 (d, *J* = 6.4 Hz, 4H, MeCy-Me), 0.94 (t, *J* = 7.4 Hz, 3H, R-H5); ¹³C NMR (125.8 MHz, D₂O): δ = 179.1 (COO), 99.9 (Gal-C1), 98.8 (Fuc-C1), 84.0 (MeCy-C2), 81.2 (Gal-C3), 78.9 (R-C2), 78.7 (Cy-C1), 74.8 (Gal-C5), 72.1 (Fuc-C4), 69.6 (Gal-C2), 69.3 (Fuc-C3), 68.2 (Fuc-C2), 66.5 (Fuc-C5), 66.0 (Gal-C4), 61.6 (Gal-C6), 38.7 (MeCy-C3), 34.7 (R-C3), 33.2 (MeCy-C4), 30.3 (MeCy-C6), 22.6 (MeCy-C5), 18.2 (MeCy-Me), 18.2 (R-C4), 15.5 (Fuc-C6), 12.9 (R-C5); ESI-MS: *m/z*: Calcd for C₂₄H₄₂NaO₁₃ [M+H]⁺: 561.2523, found: 561.2523.

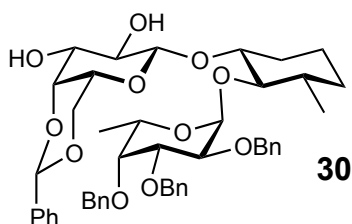
(1*R*,2*R*,3*S*)-2-[(2,3,4-Tri-*O*-benzyl-6-deoxy-α-*L*-galactopyranosyl)oxy]-3-methyl-cyclohex-1-yl 2,3-di-*O*-acetyl-4,6-*O*-benzylidene-β-*D*-galactopyranoside (S14)



Compound **15** (300 mg, 549 μmol) and thioglycoside **29** (283 mg, 713 μmol) were dissolved in dry DCM (10 mL) and stirred together with 1 g powdered 4 Å activated molecular sieves for 4 h at rt. DMTST (425 mg, 1.65 mmol) was dissolved in DCM (5 mL) and stirred together with 0.5 g powdered 4 Å activated molecular sieves for 3.5 h at rt as well. Both suspensions were combined and stirred for 3 days at ambient temperature. The mixture was filtered through a short pad of celite, washed with an aq. sat. solution of

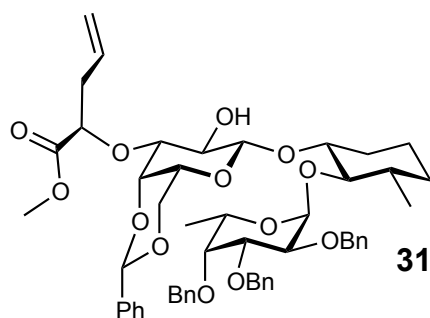
NaHCO₃ (40 mL) and water (40 mL). The combined organic phases were extracted with DCM (3 x 30 mL). The combined organic layers were dried (Na₂SO₄) and the solvent was removed *in vacuo*. The crude product was purified by flash chromatography (petroleum ether/EtOAc 3:2) to yield **S14** (266 mg, 302 μmol, 55%) as a white solid. Analytical data were in accordance with literature.²

(1R,2R,3S)-2-[(2,3,4-Tri-*O*-benzyl-6-deoxy-α-L-galactopyranosyl)oxy]-3-methyl-cyclohex-1-yl 4,6-*O*-benzylidene-β-D-galactopyranoside (30**)**



Acetate **S14** (156 mg, 178 μmol) was suspended in MeOH (5 mL) and sodium was added. The solution, formed after a few minutes was stirred for 16 h rt. The reaction mixture was neutralized by adding a few drops of glacial acetic acid and concentrated under reduced pressure. The crude product was purified by flash chromatography to yield **30** as a white solid (132 mg, 166 μmol, 94 %). Analytical data were in accordance with literature.²

(1R,2R,3S)-2-[(2,3,4-Tri-*O*-benzyl-6-deoxy-α-L-galactopyranosyl)oxy]-3-methyl-cyclohex-1-yl 4,6-*O*-Benzylidene-3-*O*-((*R*)-1-methoxy-1-oxo-pent-4-en-2-yl)-β-D-galactopyranoside (31**)**

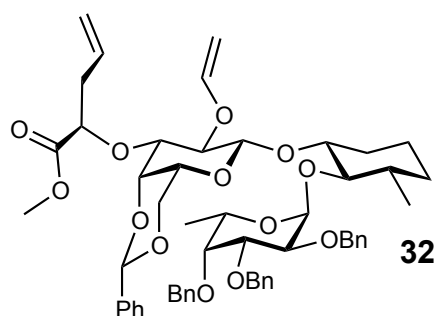


Alcohol **30** (60.0 mg, 75.3 μmol) and Bu₂SnO (56.2 mg, 226 μmol) were dried for 2.5 h at high vacuo, suspended in MeOH (4 mL) and refluxed for 2 h. The resulting solution was concentrated and coevaporated with toluene and the stannyl acetal was dried for 16 h at high vacuo. The residue was solved in DME (2 mL) and added together with triflate **21** (29.6 mg, 113 μmol) to CsF (dried for 2 h at high vacuo at 100°C, 20.6 mg, 136 μmol). The suspension was stirred for 16 h at rt. A 10% solution of KF (in 1M KH₂PO₄ solution, 3 mL)

was added and the suspension was stirred for additional 1 h at rt. DCM (10 mL) was added and the aqueous phase was separated. The aqueous phase was extracted with DCM (2 x 12 mL). The combined organic layers were dried (Na_2SO_4), concentrated under reduced pressure and purified by flash chromatography (petroleum ether/EtOAc 2:1) to afford **31** (58.4 mg, 64.2 μmol , 85 %) as a white solid.

R_f (petroleum ether/EtOAc 1:1) 0.35; $[\alpha]_D^{22}$ -18.5 (c 0.10, CHCl_3); ^1H NMR (500.1 MHz, CDCl_3): δ = 7.61 – 7.14 (m, 20H, Ar-H), 6.00 – 5.85 (m, 1H, R-H4), 5.58 (s, 1H, Ph-CH), 5.17 – 5.11 (m, 1H, R-H5), 5.04 – 4.99 (m, 1H, R-H5'), 4.97 (d, J = 3.4 Hz, 1H, Fuc-H1), 4.94 – 4.88 (m, 1H, Fuc-H5), 4.81 (d, J = 11.7 Hz, 1H, Ph- CH_2), 4.70 (d, J = 11.7 Hz, 1H, Ph- CH_2), 4.60 (s, 2H, 2x Ph- CH_2), 4.37 – 4.31 (m, 2H, Gal-H6, Gal-H1), 4.25 (d, J = 11.3 Hz, 1H, Ph- CH_2), 4.20 – 4.17 (m, 1H, Gal-H4), 4.14 (t, 1H, R-H2), 4.10 – 4.04 (m, 1H, Gal-H6'), 3.98 – 3.88 (m, 3H, Fuc-H3, Fuc-H2, Gal-H2), 3.76 (s, 3H, Me), 3.65 – 3.58 (m, 2H, MeCy-H1, Ph- CH_2), 3.41 (dd, J = 9.6, 3.3 Hz, 1H, Gal-H3), 3.34 – 3.30 (m, 1H, Gal-H5), 3.28 – 3.20 (m, 2H, MeCy-H2, Fuc-H4), 2.59 – 2.53 (m, 2H, R-H3, R-H3'), 2.12 – 2.04 (m, 1H, MeCy-H6), 1.70 – 1.48 (m, 3H, MeCy-H5, MeCy-H3, MeCy-H4), 1.45 – 1.34 (m, 1H, MeCy-H6'), 1.29 – 1.15 (m, 1H, MeCy-H5'), 1.12 – 0.99 (m, 7H, Fuc-H6, MeCy-Me, MeCy-H4'); ^{13}C NMR (125.8 MHz, CDCl_3): δ = 173.2 (COO), 139.86, 139.68, 138.83, 138.25, 128.9 - 126.0 (24x Ar-C), 133.3 (R-C4), 118.3 (R-C5), 101.5 (Gal-C1), 99.8 (Ph-CH), 98.6 (Fuc-C1), 82.3 (MeCy-C2), 81.2 (Gal-C3), 80.3 (MeCy-C1), 79.9 (Fuc-C3), 78.9 (Fuc-C4), 77.2 (R-C2), 75.7 (Fuc-C2), 74.9 (Ph- CH_2), 74.6 (Ph- CH_2), 72.5 (Gal-C4), 71.4 (Ph- CH_2), 69.8 (Gal-C6), 68.9 (Gal-C2), 66.4 (Gal-C5), 66.2 (Fuc-C5), 52.5 (Me), 39.7 (MeCy-C3), 37.6 (R-C3), 33.9 (MeCy-C4), 31.5 (MeCy-C6), 23.5 (MeCy-C5), 19.0 (Fuc-H6), 16.7 (MeCy-Me); ESI-MS: m/z : Calcd for $\text{C}_{53}\text{H}_{64}\text{NaO}_{13}$ $[\text{M}+\text{Na}]^+$: 931.42, found: 931.50.

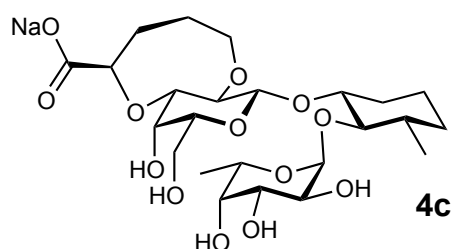
(1R,2R,3S)-2-[(2,3,4-Tri-*O*-benzyl-6-deoxy- α -L-galactopyranosyl)oxy]-3-methyl- cyclohex-1-yl 4,6-*O*-Benzylidene-3-*O*-((*R*)-1-methoxy-1-oxo-pent-4-en-2-yl)-2-*O*-vinyl- β -D-galactopyranoside (32**)**



Alcohol **31** (14.1 mg, 15.5 μmol) and sodium carbonate (822 mg, 7.76 μmol) were dissolved in toluene (1 mL) and vinyl acetate (1 mL) and chloro(1,5-cyclooctadiene)iridium (I) dimer was added. The mixture was refluxed at 80 °C for 48 h. The reaction mixture was diluted with DCM (20 mL) and washed with sat. aq. sodium bicarbonate solution (30 mL). The aqueous phase was washed with DCM (2 x 10 mL). The combined organic layers were dried (Na_2SO_4) and concentrated *in vacuo*. The crude product was purified by flash chromatography (0 to 50% petroleum ether/EtOAc + 0.5% TEA) to yield **32** in 70 % (10.1 mg, 10.8 μmol).

R_f (petroleum ether/EtOAc 2:1) 0.38; $[\alpha]_D^{22}$ -4.4 (c 0.195, CHCl_3); ^1H NMR (500.1 MHz, CDCl_3): δ = 7.61 – 7.56 (m, 2H, Ar-H), 7.37 – 7.13 (m, 18H, Ar-H), 6.38 (dd, J = 13.7, 6.3 Hz, 1H, $\text{CH}_2=\text{CHO}$), 5.93 – 5.82 (m, 1H, R-H4), 5.59 (s, 1H, Ph-CH), 5.17 – 5.09 (m, 1H, R-H5), 5.06 – 5.01 (m, 1H, R-H5'), 4.99 – 4.91 (m, 2H, Fuc-H1, Fuc-H5), 4.81 (d, J = 11.7 Hz, 1H, Ph- CH_2), 4.70 (d, J = 11.7 Hz, 1H, Ph- CH_2), 4.64 – 4.57 (m, 2H, 2x Ph- CH_2), 4.37 (d, J = 7.7 Hz, 1H, Gal-H1), 4.36 – 4.29 (m, 2H, $\text{CH}_2=\text{CHO}$, Gal-H6), 4.26 – 4.21 (m, 2H, Gal-H4, Ph- CH_2), 4.17 (dd, J = 7.3, 5.6 Hz, 1H, R-H2), 4.11 – 4.04 (m, 1H, Gal-H6'), 3.99 – 3.86 (m, 4H, $\text{CH}_2=\text{CHO}$, Fuc-H3, Fuc-H2, Gal-H2), 3.70 (s, 3H, Me), 3.66 (dd, J = 9.6, 3.5 Hz, 1H, Gal-H3), 3.61 – 3.54 (m, 2H, Ph- CH_2 , MeCy-H1), 3.34 – 3.31 (m, 1H, Gal-H5), 3.26 – 3.20 (m, 2H, Fuc-H4, MeCy-H2), 2.58 – 2.49 (m, 2H, R-H3, R-H3'), 2.04 – 1.96 (m, 1H, MeCy-H6), 1.68 – 1.57 (m, 3H, MeCy-H3, MeCy-H5, MeCy-H4), 1.40 – 1.11 (m, 2H, MeCy-H6', MeCy-H5'), 1.12 – 0.95 (m, 7H, MeCy-H4', Fuc-H6, MeCy-Me); ^{13}C NMR (125.8 MHz, CDCl_3): δ = 153.8 (COO), 133.6 (R-C4), 129.0 - 125.9 (24x Ar-C), 118.1 (R-C5), 100.9 (Gal-C1), 99.9 (Ph-CH), 98.5 (Fuc-C1), 88.7 ($\text{CH}_2=\text{CHO}$), 82.0 (MeCy-C2), 81.5 (MeCy-C1), 79.8 (Fuc-C3), 78.9 (Gal-C3), 78.9 (Fuc-C4), 78.8 (Gal-C2), 77.9 (R-C2), 77.4 (Fuc-C2), 75.8 (Ph- CH_2), 75.0 (Ph- CH_2), 74.6 (Gal-C4), 73.7 (Ph- CH_2), 71.5 (Ph- CH_2), 69.6 (Gal-C6), 66.2 (Gal-C5), 66.1 (Fuc-C5), 52.1 (Me), 39.7 (MeCy-C3), 37.8 (R-C3), 37.8 (MeCy-C4), 31.2 (MeCy-C6), 23.5 (MeCy-C5), 19.9 (Fuc-C6), 16.8 (MeCy-Me); ESI-MS: m/z : Calcd for $\text{C}_{55}\text{H}_{66}\text{NaO}_{13}$ $[\text{M}+\text{Na}]^+$: 957.44, found: 957.41.

Compound 4c

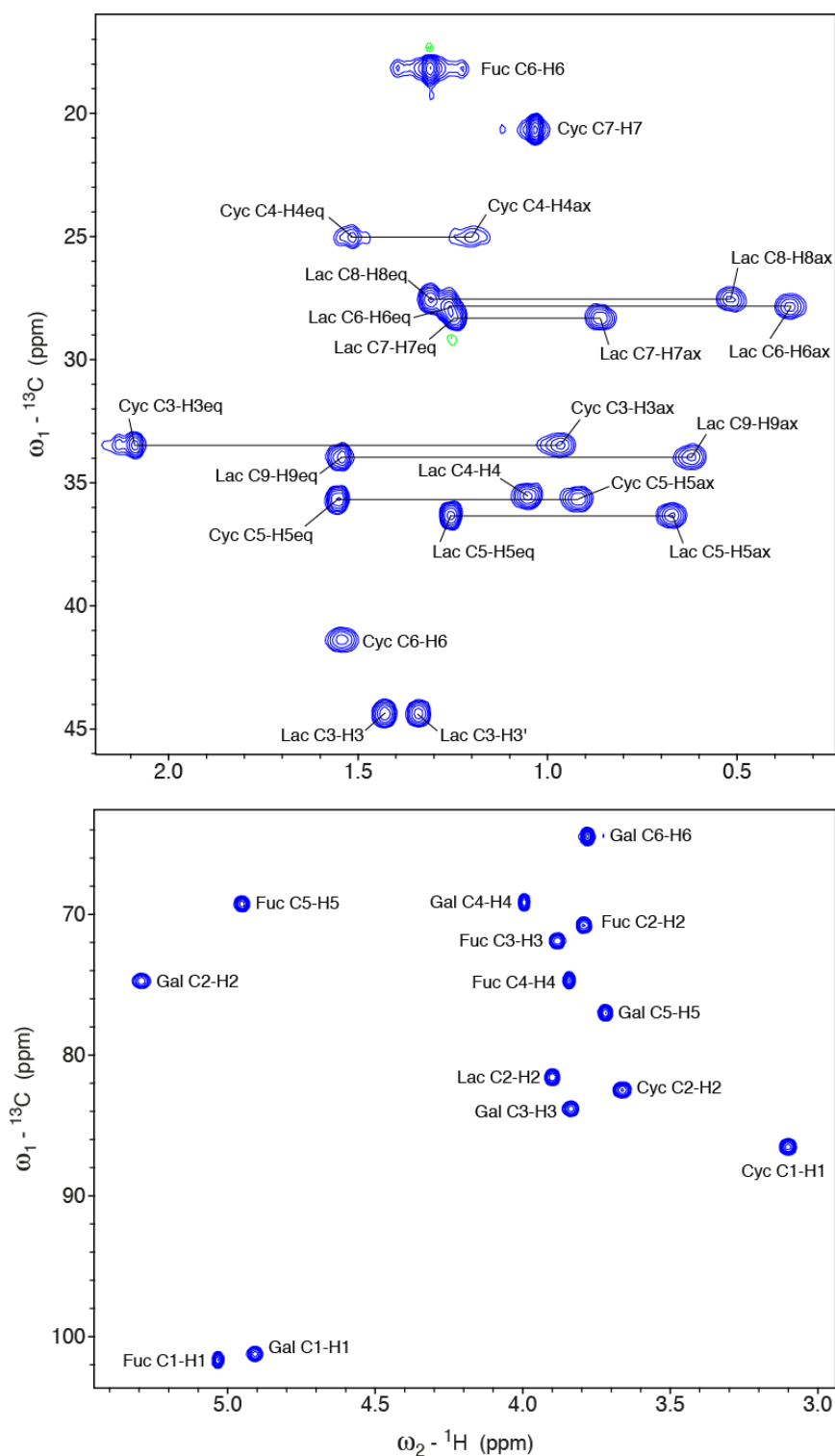


Compound **32** (4.2 mg, 4.49 μmol) was dissolved in DCM (1 mL) and Grubbs 2nd generation catalyst (0.76 mg, 0.90 μmol) was added. The solution was stirred for 6 h at rt. Additional Grubbs 2nd generation catalyst (0.76 mg, 0.90 μmol) was added and the solution stirred for another 10 h at ambient temperature. The solvent was removed under reduced pressure and the crude product purified by a short column of silica (petroleum ether/EtOAc 3:1). The olefin was dissolved in dioxane/water (4:1; 1 mL) and Pd(OH)₂/C (0.5 mg, 10% Pd(OH)₂) was added. The suspension was stirred for 16 h under an atmosphere of hydrogen, filtered and concentrated.

The residue was solved in an aqueous LiOH solution (2 mL, 2.4 mM, 4.85 μmol) was added and the turbid mixture was stirred for 24 h at rt. The residue was purified *via* RP chromatography (MeOH/H₂O), eluted through a sodium exchange column (Dowex 50/8 sodium form) and finally purified *via* size exclusion chromatography. Lyophilization from water/dioxane gave **4c** (1.46 mg, 2.62 μM , 58 %) as a white fluffy foam.

R_f (DCM/MeOH/H₂O 10:5:0.4) 0.30; $[\alpha]_{\text{D}}^{22}$ -74.7 (*c* 0.28, H₂O); ¹H NMR (500.1 MHz, D₂O): δ = 5.13 (d, *J* = 4.0 Hz, 1H, Fuc-H1), 4.89 - 4.75 (m, Fuc-H5), 4.62 (d, *J* = 7.9 Hz, 1H, Gal-H1), 4.27 (dd, *J* = 12.4, 2.3 Hz, 1H, R-H1), 4.23 - 4.19 (m, *J* = 3.1 Hz, 1H, Gal-H4), 4.19 - 4.13 (m, 1H, R-H4), 4.02 (dd, *J* = 9.7, 3.2 Hz, 1H, Gal-H3), 3.93 (dd, *J* = 10.6, 3.3 Hz, 1H, Fuc-H3), 3.90 - 3.71 (m, 6H, R-H4', Fuc-H4, Fuc-H2, MeCy-H1, Gal-H6, Gal-H6'), 3.62 (t, *J* = 5.9 Hz, 1H, Gal-H5), 3.40 (dd, *J* = 9.5, 8.0 Hz, 1H, Gal-H2), 3.27 (t, *J* = 9.6 Hz, 1H, MeCy-H2), 2.20 - 2.06 (m, 2H, MeCy-H6, R-H2), 2.03 - 1.87 (m, 2H, R-H3, R-H2'), 1.73 - 1.55 (m, 4H, R-H3, MeCy-H4, MeCy-H5, MeCy-H3), 1.34 - 1.24 (m, 2H, MeCy-H5', MeCy-H6'), 1.21 (d, *J* = 6.6 Hz, 3H, Fuc-H6), 1.11 (d, *J* = 6.3 Hz, 3H), 1.16 - 1.06 (m, MeCy-H4'); ¹³C NMR (125.8 MHz, D₂O): δ = 99.6 (Fuc-C1), 98.5 (Gal-C1), 84.6 (MeCy-C2), 80.8 (R-C1), 79.9 (Gal-C2), 79.2 (Gal-C3), 79.2 (Gal-C4), 76.4 (Gal-C3), 75.9 (Gal-C5), 72.9 (Fuc-C4), 70.6 (R-C4), 70.2 (Fuc-C3), 69.1 (Fuc-C2), 67.3 (Fuc-C5), 62.5 (Gal-C6), 39.8 (MeCy-C3), 34.0 (Cy-C4), 30.9 (MeCy-C6), 28.0 (R-C3), 25.2 (R-C2), 23.5 (MeCy-C5), 19.2 (MeCy-Me), 16.5 (Fuc-C6); ESI-MS: *m/z*: Calcd for C₂₄H₄₀NaO₁₃ [M+H]⁺: 559.2367, found: 559.2365.

Supplementary Figures



SFigure 1. ^{13}C - ^1H HSQC of **3** recorded at 750 MHz and 275 K with assignments. The aliphatic region is shown on top, the carbohydrate region at the bottom.

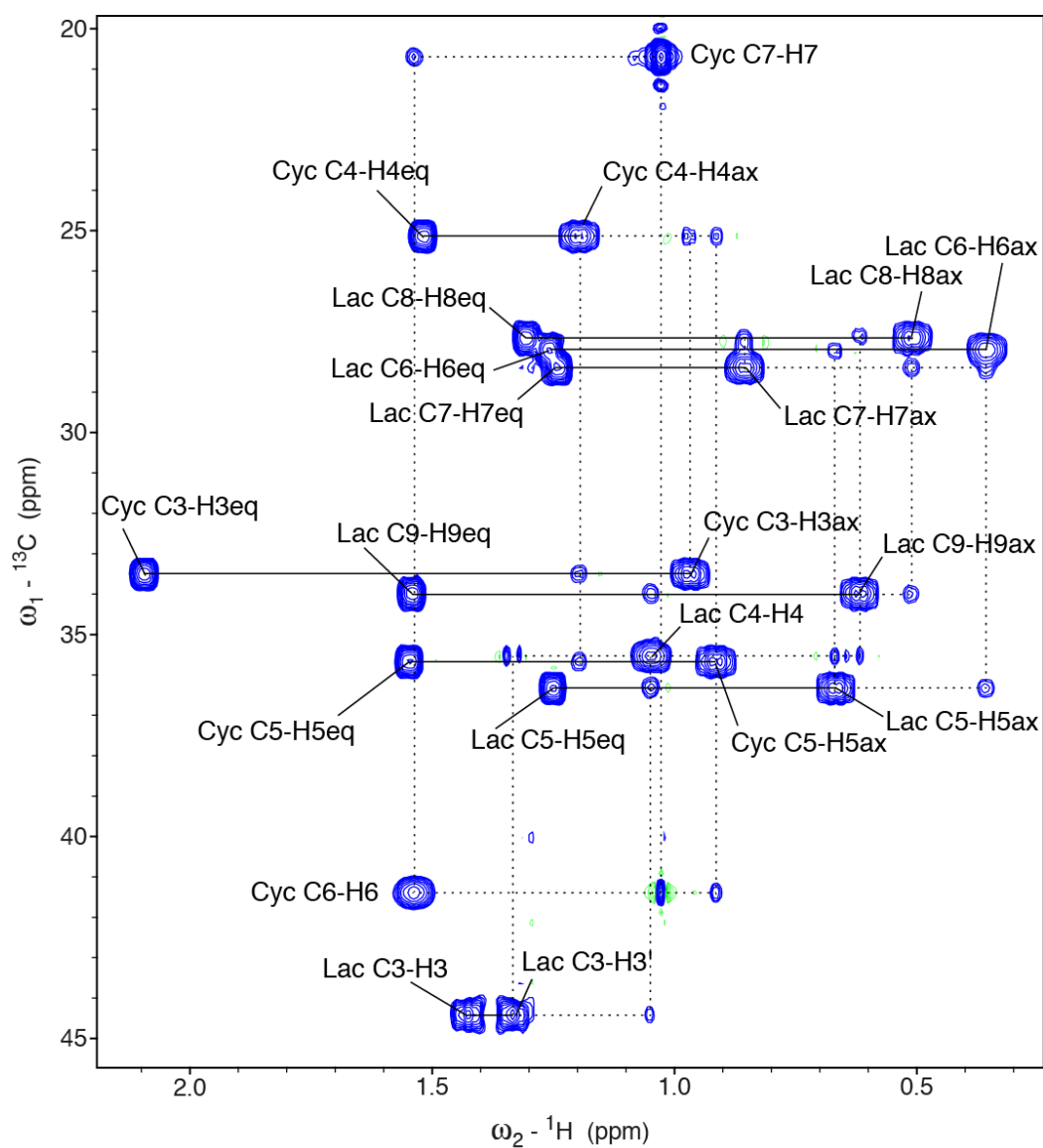


Figure 2. Chemical shift assignment of cyclohexane resonances using a ${}^{13}\text{C}$ - ${}^1\text{H}$ HMQC-COSY. The small unassigned signals originating from two-bond scalar couplings ${}^2J_{\text{CH}}$ are ideally suited to assign the cyclohexane resonances by following the dotted lines. Resonances from axial protons were distinguished from equatorial ones by their strength of ${}^3J_{\text{HH}}$ scalar coupling constants and NOE signal intensities (data not shown). As is typical the axial resonances are more upfield than the equatorial ones.

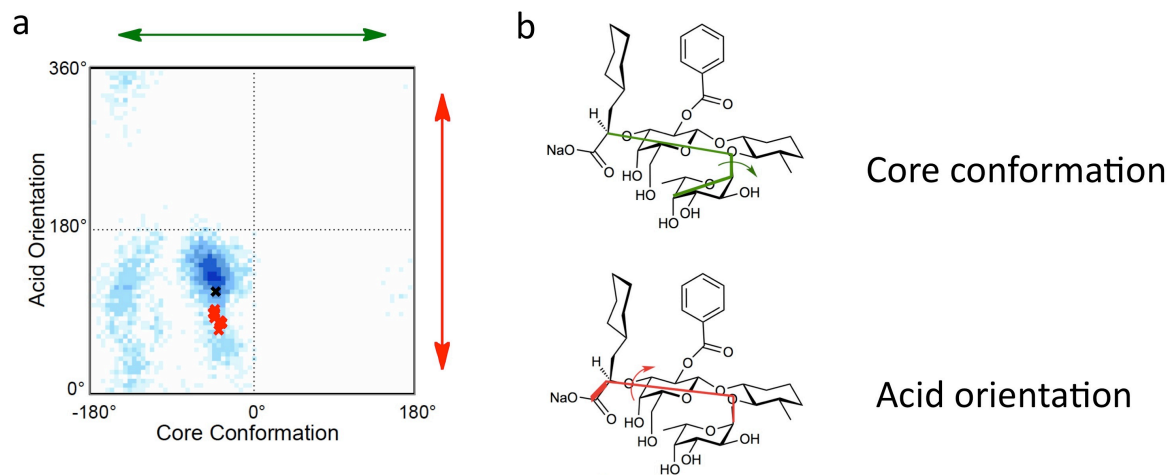


Figure 3. a) Acid orientation / core conformation blot of **3**: crystal structure of sLe^x (black cross),³ molecular dynamics simulations (blue) and the ensemble of 20 solution conformations determined by NMR at 277 K at 900 MHz high field (red crosses). b) Schematic overview of core conformation (= relative orientation of D-galactose and L-fucose) and acid orientation (= tilting angle of the carboxylic acid relative to the core).

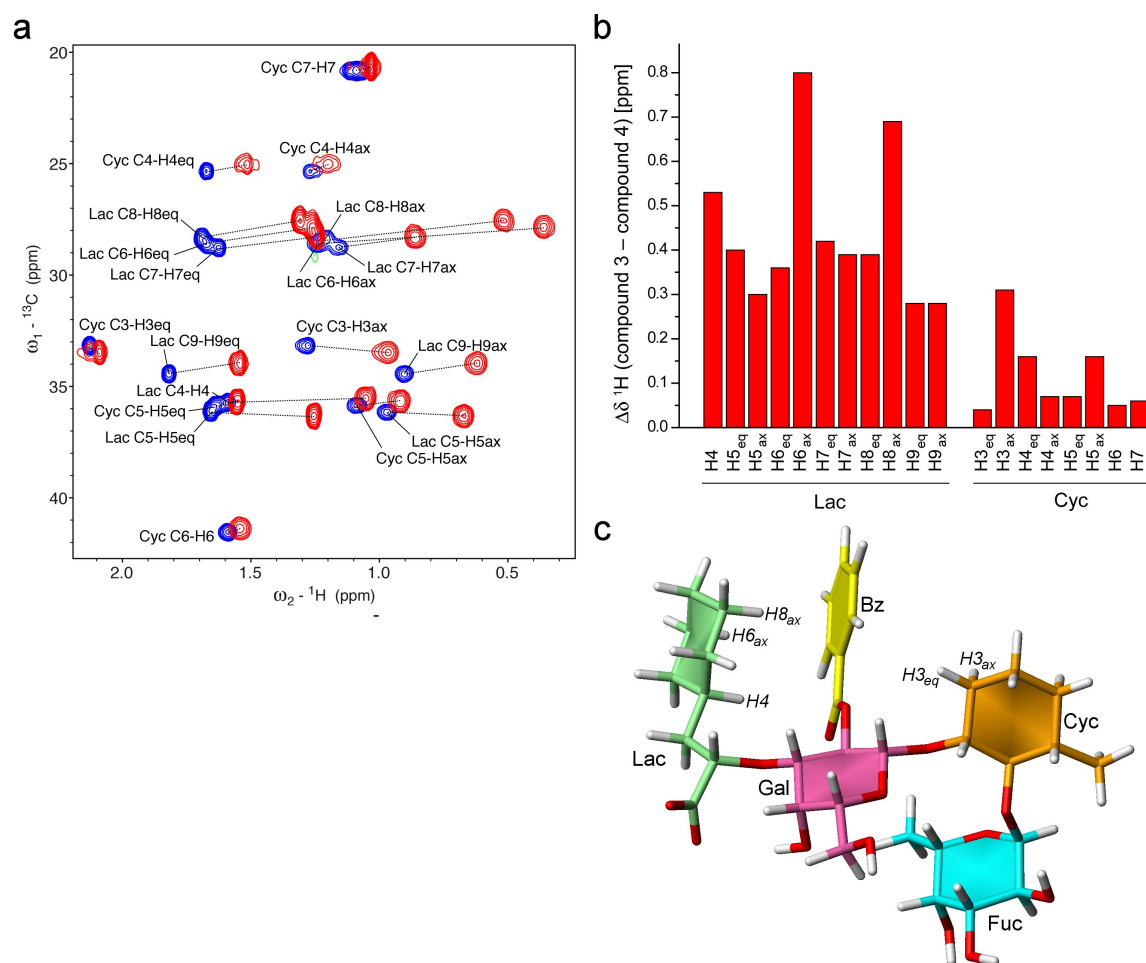


Figure 4. Confirmation of the three-dimensional solution structure by chemical shift deviations caused by the benzoyl group. a) Overlay of HSQC spectra of compound **3** (red) and compound **2** (blue). The region with signals of the methyl-cyclohexane diol mimic and the (*S*)-cyclohexan lactic acid are shown. Corresponding signals are connected by dotted lines. The benzoyl group causes a large variety of upfield chemical shift changes, which are typically caused by ring current effects on protons perpendicular to the aromatic ring plane. b) Histogramm of the chemical shift deviations of cyclohexane protons caused by the benzoyl group. c) Representative of the obtained structural ensemble by solution NMR spectroscopy. Relevant protons in the proximity of the aromatic ring are labeled. The closest protons perpendicular to the aromatic plane are expected to experience the largest ring current effect and indeed they show the largest chemical shift perturbation.

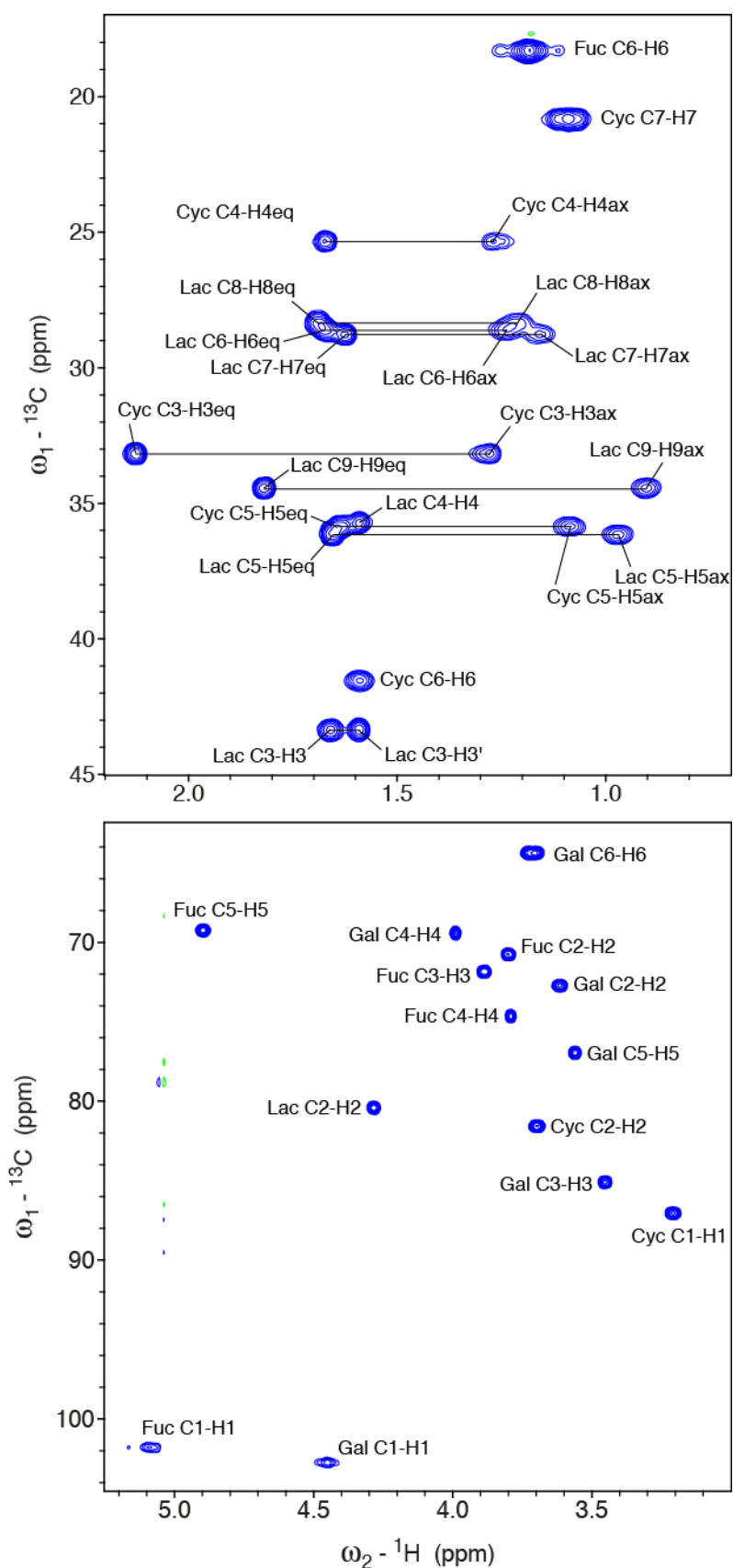
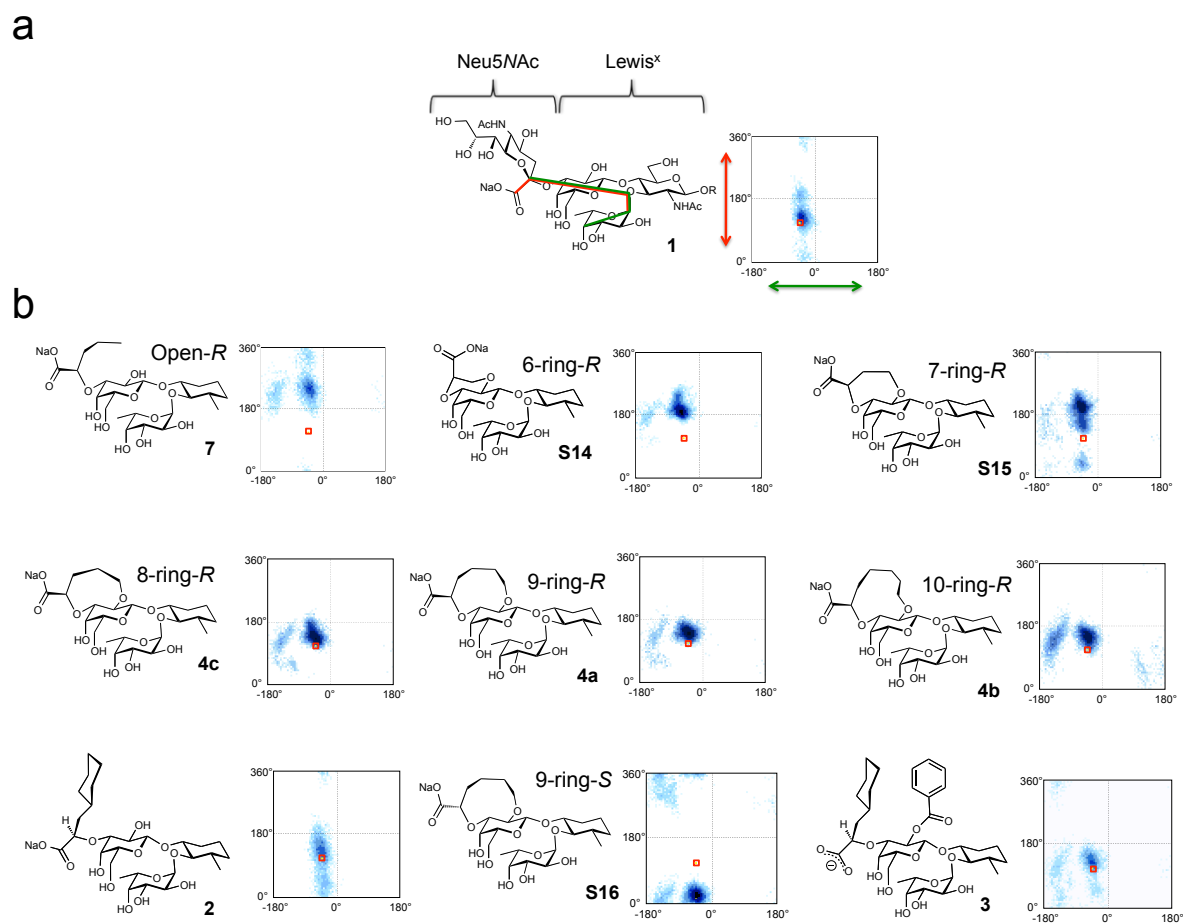


Figure 5. ^{13}C - ^1H HSQC of **2** recorded at 750 MHz and 275 K with assignments. The aliphatic region is shown on top, the carbohydrate region at the bottom.



SFigure 6. The predicted pre-organization of pharmacophores of **1**, **2**, **3**, **4a-c**, **7** and **S14-S16** is determined in acid orientation/core conformation plots (from light blue (not likely) to dark blue (most likely)). The acid orientation of sLe^x (**1**) in the crystal structure is highlighted by a red cross.³

Supplementary Tables

Table 1. Intra and interresidual NOEs of **3** at 275K with the corresponding calculated ^1H - ^1H distances.

proton pair	S/N of NOEs cross peaks	Corresponding ^1H - ^1H distance ^b	proton pair	S/N of NOEs cross peaks	Corresponding ^1H - ^1H distance ^b
Intra			Intra		
Lac H2-H3ax	476	2.6	Cyc H3eq-H5ax	167	3.1
Lac H2-H3eq	943	2.3	Cyc H3ax-H4eq	848	2.4
Lac H2-H4eq	187	3.0	Cyc H5ax-H7	303 ^f	2.8
Lac H3ax-H4eq	441	2.6	Fuc H1-H2	798	2.4
Lac H4-H6ax	268 ^a	2.9	Fuc H1-H3	113 ^a	3.3
Lac H4-H8ax	231 ^a	2.9	Fuc H3-H5	962	2.3
Lac H3ax-H5ax	312	2.8	Fuc H4-H5	1494	2.1
Lac H3eq-H5ax	407	2.7	Inter		
Lac H5ax-H6ax	341	2.7	Lac H9eq-Gal H3	223	2.9
Lac H7ax-H8eq	413 ^a	2.7	Lac H4-Bz H2	88 ^{a,e}	3.4
Lac H2-H9eq	1012	2.3	Lac H6ax-Bz H2	56 ^{a,e}	3.7
Lac H2-H9ax	261	2.9	Lac H8eq-Bz H2	84 ^{a,e}	3.5
Lac H3ax-H9ax	102	3.4	Lac H8ax-Bz H2	114 ^{a,e}	3.3
Lac H7ax-H9ax	335	2.7	Lac H9eq-Bz H2	80 ^{a,e}	3.5
Gal H1-H5	1271	2.2	Gal H1-Cyc H2	1221	2.2
Gal H1-H2	277	2.8	Gal H1-Cyc H3a	969	2.3
Gal H1-H3	541	2.5	Gal H1-Cyc H3b	167	3.1
Gal H2-H3	198 ^a	3.0	Gal H2-Fuc H5	387	2.7
Gal H4-H5	1568 ^a	2.1	Gal H2-Fuc H6	241 ^f	2.9
Cyc H1-H2	290	2.8	Gal H6-Fuc H6	62 ^{e,f}	3.6
Cyc H1-H3eq	146	3.2	Gal H1-Bz H2	57 ^{a,e}	3.7
Cyc H1-H3ax	712	2.4	Bz H2-Cyc H3eq	157 ^e	3.1
Cyc H1-H5ax	612	2.5	Bz H2-Cyc H3ax	123 ^{a,e}	3.2
Cyc H1-H7	424 ^f	2.6	Fuc H1-Cyc H1	1576	2.1
Cyc H2-H3	929	2.3	Fuc H1-Cyc H6	173	3.1
Cyc H2-H4ax	625 ^a	2.5	Fuc H1-Cyc H7	777 ^f	2.4
Cyc H3eq-H3ax	4648 ^d	1.77	Fuc H5-Cyc H1	173 ^a	3.1
Cyc H3eq-H4eq	894	2.3	Fuc H5-Cyc H2	140	3.2
Cyc H3eq-H4ax	742	2.4			

^a Only one cross-peak was used because of artifacts or overlap.

^b The ^1H - ^1H distances were calculated from experimentally obtained NOE intensities using the H3a-H3b cross-peak of Cyc as a reference with a distance of 1.77 Å assuming a r^{-6} dependence of the NOE intensities.

^d Reference restraints for the ^{13}C -filtered-filtered NOESY.

^{e,f} Signal to noise ratios from cross-peaks involving methyl- or methylene protons with coinciding f frequencies were divided by 3 or 2, respectively

STable 2. NMR structure determination statistics of **3** in solution.

Structure of 3	
NMR distance and dihedral restraints	
Total NOE restraints	56
Intra-residue	36
Inter-residue	20
Sequential ($ i - j = 1$)	10
Nonsequential ($ i - j > 1$)	10
Structure statistics *	
Violations (mean and s.d.)	
Number of distance constraint violations > 0.1 Å	0±0
Max. distance constraint violation [Å]	0.10±0.01
Deviations from idealized geometry	
Bond lengths [Å]	0.0125±0.0002
Bond angles [°]	1.54±0.03
Heavy atom RMSD to mean [Å]	0.79±0.30
Glycosidic linkage phi / psi angles **	
Fuca(1,1)CycMe	-68.9±0.8/-100.5±0.4 (crystal: -78.0/-101.4)
Galb(1,2)CycMe	-75.1±4.8/151.3±2.3 (crystal: -99.8/142.4)
Acid Orientation	82.0±6.4 (crystal: 120.5)
Core	-37.8±3.6 (crystal: -41.6)

* for an ensemble of 20 refined structures

** phi is defined as $O_5-C_1-O_x-C'_x$ and psi as $C_1-O_x-C'_x-C'_{x-1}$

† extracted by XtalView

References

- (1) Gottlieb, H. E.; Kotlyar, V.; Nudelman, A. *J. Org. Chem.* **1997**, *62*, 7512.
- (2) Binder, F. P. C., PhD Thesis, Basel, **2011**.
- (3) Somers, W. S.; Tang, J.; Shaw, G. D.; Camphausen, R. T. *Cell* **2000**, *103*, 467.

3.2.1.2 Additional experiments: Ring closing metathesis trials for amide and ester series

Contributions

- Manuscript preparation
- Compound synthesis

Abstract

Analogue to the ether macrocycles (chapter 3.2.1), we planned to synthesize and evaluate a series of lactons and lactams. The rigidity of those compounds should be even higher than that of the corresponding ether series. Furthermore, molecular dynamic (MD) calculations predicted the orientation of the acid pharmacophore in the bioactive conformation. To take advantage of our experience obtained from the cyclic ether derivatives, we decided to follow a ring closing metathesis synthesis route. But in contrast to the ether series the ring closing metathesis (RCM) strategy was not successful and led to oligomerization. MD simulations led to the conclusion, that an unfavorable orientation of the intramolecular olefins in the reactants was responsible for failure of the RCM, and instead intermolecular cross metathesis (CM) and acyclic diene metathesis polymerization reactions (ADMET) were preferred.

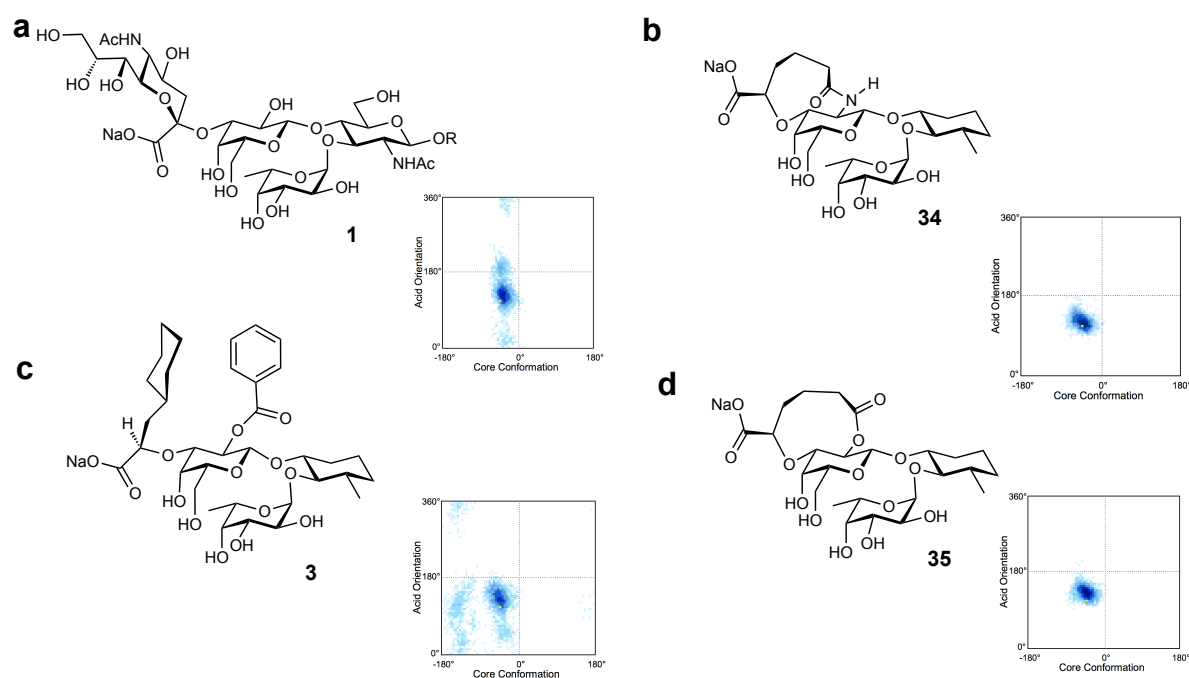


Figure 1. Acid orientation (y-axis) and core conformation (x-axis) plots for sLe^x (**1**) (a), compound **3** (c) and lactam **34** (b) and lactone **35** (d).

Introduction

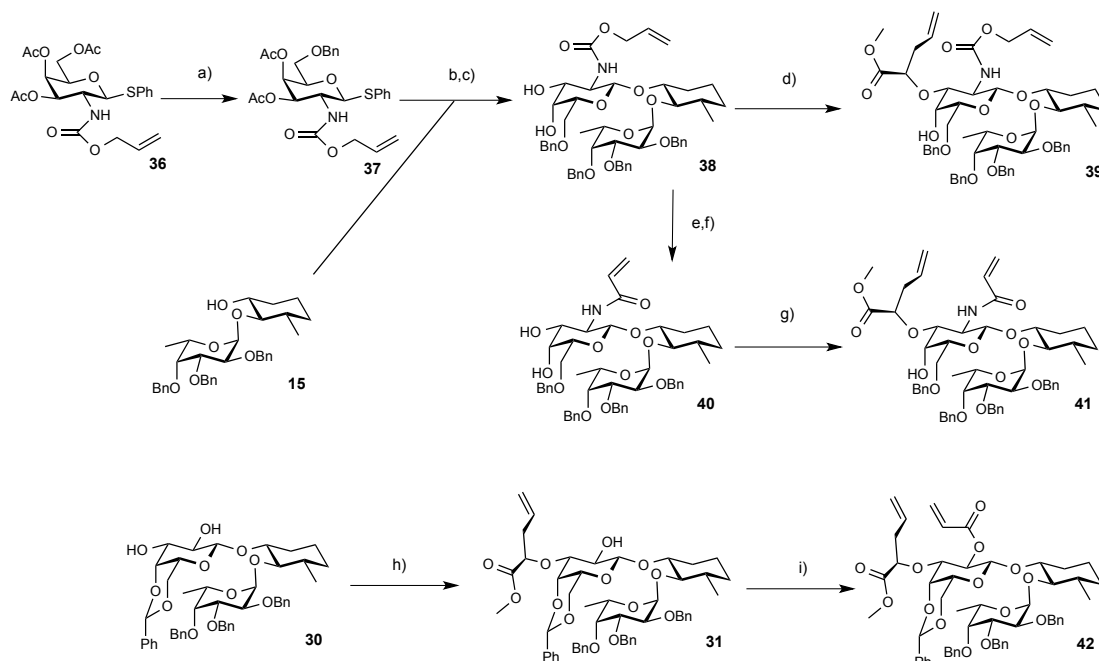
In chapter 3.2.1 we synthesized a series of ether macrocycles, where the acid pharmacophore is incorporated in a ring system and thus, locked in the bioactive conformation. Due to the formation of the 8- to 10-membered rings to the D-galactose moiety with its defined chair

conformation, the conformational flexibility is restricted. A further stabilizing effect originated from an amide bond or ketone functionality in the ring system could further improve pre-organization and thus, guarantee a maximal entropic gain upon binding. But this almost total loss of flexibility could also turn out as disadvantage, when the acid cannot adjust properly to the selectin upon binding. The additional hydrogen bond acceptors and donors contributed by the amide (or ketone) functionality could also turn out to be a disadvantage regarding the pharmacokinetic properties, e.g. oral availability. However, MD analysis clearly indicated that acid orientation and also core conformation were excellent displayed (Figure 1).

Results and Discussion

Thioglycoside **36**¹ was deacetylated using Zemplén conditions. For benzylation in the 6-position, a benzylidene acetal was introduced, followed by selective ring opening with MeNH₂•BH₃ and subsequent acetylation. This gave glycosyldonor **37** in 60 % over 4 steps. Glycosylation of pseudodisaccharide **15**² by DMTST mediated activation of glycosyldonor **37**, followed by deacetylation gave diol **38**, which serves as precursor for RCM substrate **39**. Bu₃SnH mediated alloc deprotection of **38** and subsequent reaction of the amino group with acryloyl chloride gave acrylamide **40**. A selective tin-mediated alkylation of the hydroxy group in the 3-position of the D-galactose moiety with triflate **21** (chapter 3.2.1) afforded olefin **41**, which was used as substrate for RCM trials. As control the alloc protected diol **38** was alkylated with triflate **21**. In the resulting compound **39** the terminal olefin of the alloc protection group served as substrate for the RCM. The prolonged size, accompanied with increased flexibility of the olefin moiety and the absence of an electron withdrawing group in immediate proximity to the alkene, should allow a successful RCM.

The substrate for the lactone series was synthesized from starting material **30** (see chapter 3.2.1). In analogy to the amide series, the 3'-position of the D-galactose residue of **30** was alkylated with triflate **21**. Alcohol **31** was converted to acrylate **42** by esterification with acryloyl chloride.



Scheme 1. a.) (i) NaOMe, MeOH; (ii) PhCH(OMe)₂, MeCN; (iii) MeNH₂·BH₃, AlCl₃, H₂O, THF; (iv) Ac₂O, pyridine, 47 % over 4 steps; b.) DMTST, DCM, 4Å mol.sieves, ; c.) NaOMe, MeOH, 25% over 2 steps; d.) (i) Bu₂SnO, MeOH; (ii) CsF, **21**, DCM 88%; e.) (i) Bu₄SnH, DCM, ; (i) Acroyloylchloride, 69%; f.) (i) Bu₂SnO, MeOH; (ii) CsF, triflate **21**, DCM, 65%; h.) (i) Bu₂SnO, MeOH; (ii) CsF, triflate **21**, DCM, 85%; i.) Acryloyl chloride, Et₃N, 44%.

For RCM reaction with amide **41** different catalysts, reaction times and temperatures were explored (Table 1). The olefin metathesis with Grubbs 1st generation catalyst **43**³ was not successful in DCM, although 0.5 equivalents of catalyst at 40°C were used. The Grubbs' 2nd generation catalyst **44**⁴, with higher affinity did also not lead to product formation with different amounts of catalyst at 22°C. However, at 40°C, substrate **41** reacted to a series of side products, which were verified as oligomers of **41** by mass spectrometry. A similar result was observed with the Grubbs-Hoveyda catalyst **45**⁵. Oligorimization was also observed with acrylate **42** and 2nd generation catalyst **43** at 40°C.

Surprisingly, the alloc protected trisaccharide **39**, which was assumed to be optimally suited for RCM, decomposed. According to MS analysis, no product formation occurred.

a

Reactant	Catalyst / ratio	Time [h]	Temperature [°C]	Product
41	43 / 0.5 eq	24	40	no reaction
41	44 / 0.1 eq	24	22	no reaction
41	44 / 0.25 eq	24	22	no reaction
41	44 / 0.1 eq	3	40	polymerisation
41	45 / 0.1 eq	3	40	polymerisation
42	44 / 0.1 eq	3	40	polymerisation
39	44 / 0.1 eq	3	40	decomposition

b

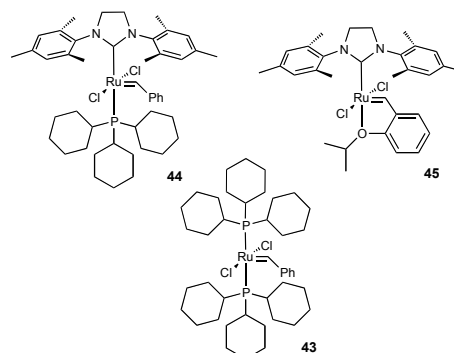
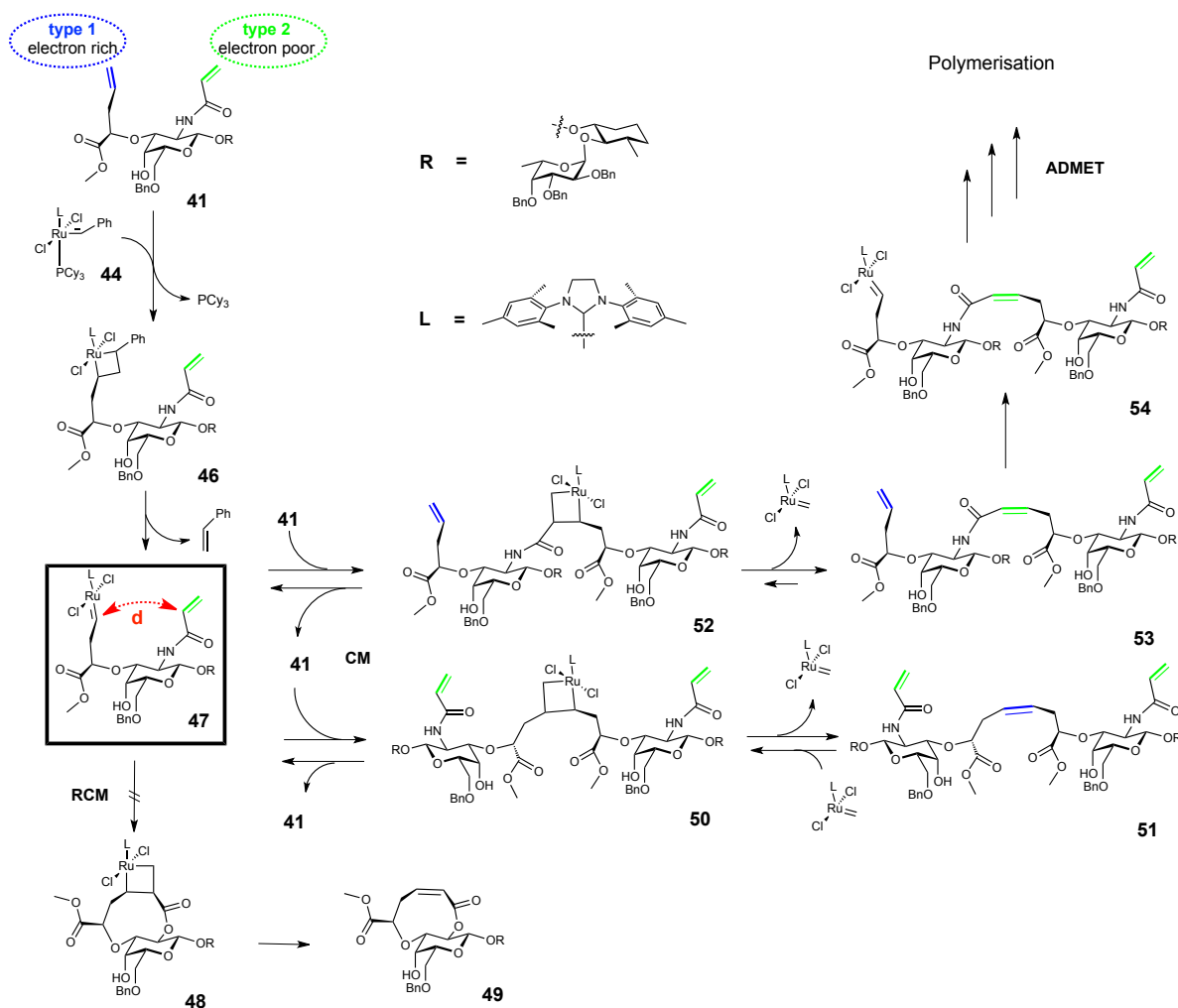


Table 1. (a) Reaction conditions for RCM trials in DCM; (b) Ruthenium(II) based Grubbs' 1st generation catalyst 43³, Grubbs' 2nd generation catalyst 44⁴ and 2nd generation Grubbs-Hoveyda catalyst 44⁵.

Discussion

It is known that acyclic diene metathesis polymerization (ADMET) is a competitive reaction to the RCM.⁶ To explain the failure of RCM in our ester and amide trials we



propose a mechanism for the olefin metathesis for amide **41** with Grubbs 2nd generation catalyst **4** (Scheme 2). The substrate **41** is characterized by an electron rich type 1 olefin and a electron poor type 2 olefin, classified by the empirical model for olefin reactivity in cross metathesis published by Chatterjee *et al.*⁷ Type 1 olefins are categorized as those able to undergo a rapid homodimerization and which homodimers can participate in cross metathesis (CM) as well as their terminal olefin counterpart, whereas type 2 olefins homodimerize slowly, and unlike Type I olefins, their homodimers can only be sparingly consumed in subsequent metathesis reactions.⁷ The initial step is characterized by trialkyl phosphine dissociation of catalysator **44**, followed by coordination of the electron rich type 1 olefin of the substrate and formation of a metallacyclobutane intermediate **46**. Cycloreversion yields the key intermediate, the alkylidene **47**. Compound **47** can either act as substrate for RCM or an ADMET reaction. We propose that the reason for failure of RCM is the reduced flexibility of the amide bond. In contrast, the flexible ether group readily underwent RCM (chapter 3.2.1). Thus, we propose that the probability for the two double bonds to be in RCM-distance is low, leading predominantly to the intermolecular CM in this compound class. The dimerisation of two electron rich type 1 alkene (**47**→**51**) is more likely but reversible.

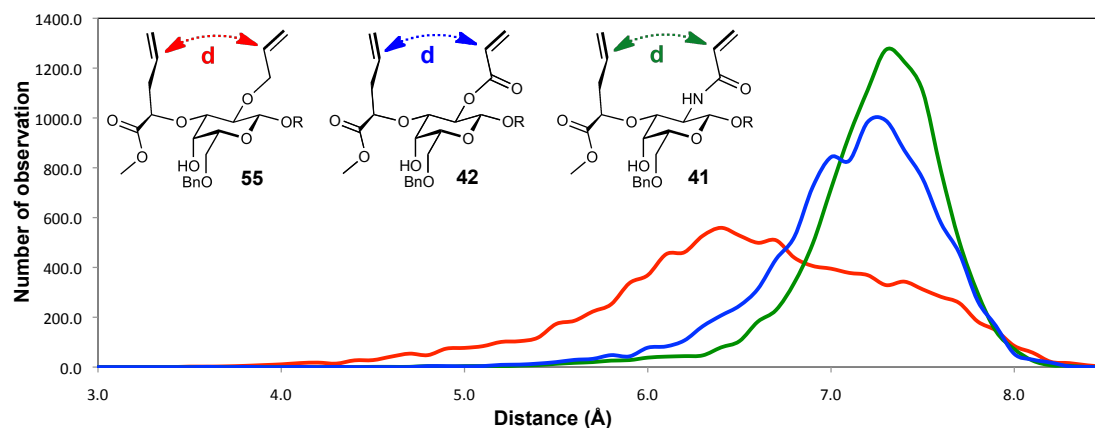


Figure 2. MD simulation of distance distribution between highlighted sp²-hybridized atoms involved in olefin metathesis reaction in allylether **55** (red), acrylate **42** (blue) and acrylamide **41** (green) in octane at 303 K.

Dimer **53** is rarely formed by CM between a type 1 and type 2 olefin, and would rather act as substrate for continuous CM than do the reverse reaction to **47**, as the terminal type 1 olefin is more reactive than the electron poor type 2 olefin undergoing dimerization (Scheme 2), the first step in the acyclic diene metathesis polymerization. We proposed a similar

mechanism for the ester **42**, caused by rigification of the metathesis substrate due to the carbonyl group.

To support our hypothesis, that the probability of vicinal proximity between the intramolecular olefins is unfavorable for RCM, we performed a series of MD studies. We analyzed the distance of the carbon atoms involved in the metathesis reaction at 303 K in octanol (the dielectricity constant for octanol and DCM is similar) for ether **55**, ester **42** and amide **41** (Figure 4). As resumed, the distance distribution of the ether **55** is significantly shifted to shorter distances. On the other hand, distances closer than 5.5 Å are hardly observed for compound **42** and **41**, probably due to the rigidifying carbonyl and amide bond. Therefore, the energy barrier to arrange the two olefins in an optimal distance for RCM, would be higher for **41** and **42** compared to more flexible substrates like **55**, and thus benefits CM and ADMET.

Methods

Molecular modeling. The 3D structure of the carbohydrate mimic molecules was built in the Maestro⁸ molecular modeling environment and a conformational analysis was performed using Macromodel⁹ by performing 5000 steps of the mixed torsional/low-mode sampling in combination with the OPLS-2005 force-field in implicit solvent conditions (water).

Global minimum identified in the conformational search was then placed in a periodic boundary system (cube, side length 40.0 Å) filled with explicit water molecules (TIP3P parametrization). The system was minimized and a simulated annealing simulation (gradual heating of the system to the temperature of 400 K in 330 ps followed by slow gradual cooling to 300 K in 500 ps) was performed in Desmond^{10,11} in order to bring the system into equilibrium. The production phase molecular dynamics simulation of the total duration of 48 ns was performed. Two critical structural descriptors – the core conformation and the acid orientation – were monitored every 4.8 ps (10 000 data points were saved).

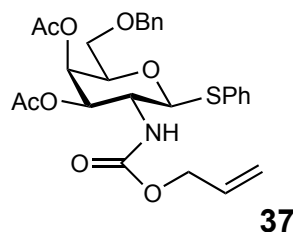
The pre-organization of the mimic molecule was evaluated by plotting the frequency of the two critical parameters in a two-dimensional color-coded plot using an in-house computer program.

General methods. Commercial materials (Sigma-Aldrich) were used without further purification, solvents were reagent grade (Acros). CH₂Cl₂ and MeOH were dried by passing

through an Al₂O₃ (Fluka, type 5016 A basic) column. DMF extra dry (Acros) was used as is. All reactions were performed in oven dried glassware under an atmosphere of argon.

¹H and ¹³C NMR spectra were recorded on a Bruker Avance DMX-500 at room temperature. Chemical shifts are reported in ppm and referenced to TMS using residual solvent peaks.¹² For complex molecules the following prefixes were used: Fuc (fucose), Gal (galactose) and GlcNAc (*N*-acetylglucosamine). The coupling constants (*J*) are reported in Hertz (Hz). Analytical TLC was performed on Merck silica gel 60 F₂₅₄ glass plates and visualized by UV light and charring with a molybdate solution (a 0.02 M solution of ammonium cerium sulfate dihydrate and ammonium molybdate tetrahydrate in aq. 10% H₂SO₄) by heating for 5 min at 140°C. Column chromatography was performed on a CombiFlash Companion (Teledyne-ISCO, Inc.) using RediSep® normal phase disposable flash columns (silica gel). Reversed phase chromatography was carried out with LiChroprep®RP-18 (Merck, 40-63 μm). Optical rotations were determined on a Perkin-Elmer Polarimeter 341. Low resolution mass spectra were measured on a Waters micromass ZQ. High resolution mass spectra (HRMS) were obtained on a micrOTOF spectrometer (Bruker Daltonics, Germany) equipped with a TOF hexapole detector. MALDI-TOF and ESI-MS glycoprotein analyses were recorded by the Functional Genomic Center Zurich (FGCZ). VIVASPIN® 500 ultrafiltration tubes with 10000 MWCO, 10 kDa molecular weight cutoff PES membrane, and ZelluTrans/Roth dialysis membranes MWCO 8000-10000 were used for glycoprotein concentration and dialysis.

Phenyl 3,4-*O*-acetyl-6-*O*-benzyl-2-deoxy-2-[(2-propen-1-yl-oxy)carbonyl]amino]-1-thio-β-D-galactopyranoside (37)

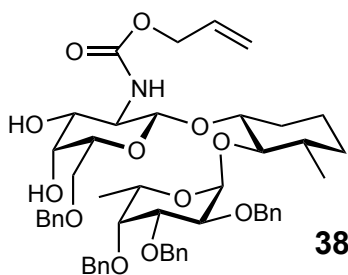


Compound **36** (7.00 g, 14.5 μmol) was suspended in MeOH (120 mL). Sodium (50.0 mg, 2.17 μmol) was added and the resulting solution was stirred for 18 h at rt. AcOH was added (250 μL) and the solvent was removed in vacuo. The residue was suspended MeCN (450 mL) and camphorsulfonic acid (1.25 g, 5.38 mmol) and C₅H₆CH(OMe)₂ (6.54 ml, 43.6 μmol) were added. The resulting solution was stirred for 24 h at rt and quenched with 1.3 ml

Et₃N. Volatiles were evaporated and the crude product was purified by flash chromatograph (petroleum ether/EE 2:1→1:3). 2.25 g (5.07 μmol) of the resulting benzylidene acetal were dissolved in THF (100 mL) and Me₃N•BH₃ (1.48 g, 20.3 mmol), AlCl₃ (4.09 g, 30.4 mmol) and H₂O (183 μl, 10.2 mmol) were added. The turbid reaction mixture was stirred for 16 h at ambient. An aq. solution of HCl (200 mL, 0.5 M) was added and extracted with EtOAc (2 x 200 mL). The combined organic layers were washed with brine (150 mL), dried and evaporated *in vacuo*. The crude product was purified by flash chromatography (petroleum ether/EtOAc 1:2). The resulting alcohol was dissolved in pyridine (125 mL) and acetic anhydride (13.8 mmol, 1.31 ml) was added. The solution was stirred for 48 h at rt. The reaction mixture was quenched with MeOH and the solvents removed under reduced pressure. Column chromatography (petroleum ether/EtOAc 3:2) gave **37** (1.39 g, 3.1 mmol, 47 % over 4 steps) as a white solid.

R_f (petroleum ether/EtOAc 3:2) 0.35; [α]_D²² -24.2 (*c* 2.3, CHCl₃); ¹H NMR (500.1 MHz, CDCl₃): δ = 7.55 – 7.50 (m, 2H, Ar-H), 7.37 – 7.26 (m, 8H, Ar-H), 5.95 – 5.86 (m, 1H, CH=CH₂), 5.48 – 5.43 (m, 1H, Gal-H4), 5.33 – 5.27 (m, 1H, CH=CH₂), 5.24 – 5.19 (m, 1H, CH=CH₂), 5.15 – 5.08 (m, 1H, Gal-H3), 4.90 – 4.82 (m, 1H, Gal-H1), 4.53 (d, *J* = 11.8 Hz, 1H, Ph-CH₂), 4.42 (d, *J* = 11.8 Hz, 1H, Ph-CH₂), 3.94 (m, 1H, Gal-H2), 3.88 (t, *J* = 6.3 Hz, 1H, Gal-H5), 3.60 (dd, *J* = 9.7, 6.3 Hz, 1H, Gal-H6), 3.50 (dd, *J* = 9.7, 6.3 Hz, 1H, Gal-H6'), 2.05 (s, 3H, Me), 1.99 (s, 3H, Me); ¹³C NMR (126 MHz, CDCl₃): δ = 170.5, 170.3, 155.7, 137.8, 132.7, 132.2, 132.2, 129.1 – 128.0 (8H), 117.8, 87.9, 76.2, 73.7, 71.6, 68.1, 67.6, 66.0, 51.4, 20.8, 20.8; ESI-MS: *m/z*: Calcd for C₅₃H₆₂NaO₁₃ [M+Na]⁺: 929.41, found: 929.45.

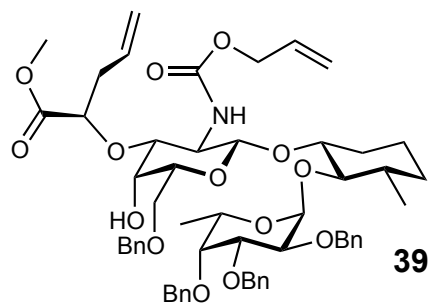
(1*R*,2*R*,3*S*)-2-[(2,3,4-Tri-*O*-benzyl-6-deoxy-α-*L*-galactopyranosyl)oxy]-3-methyl-cyclohex-1-yl 6-*O*-benzyl-2-deoxy-2-[(2-propen-1-yl-oxy)carbonyl]amino]-β-*D*-galactopyranoside (38**)**



Disaccharide mimic **15** (979 mg, 1.79 mmol), galactosyldonor **37** (1.13 g, 2.15 mmol) and 4 Å molecular sieves (8 g) were suspended in DCM (50 mL) and stirred for 4 h at rt. DMTST (1.39 g, 5.38 mmol) and 4 Å molecular sieves (4 g) was suspended in DCM (20 mL) and stirred for 4 h in a second reaction vessel for 4 h at rt and subsequently added via syringe to the reaction mixture. The suspension was stirred for 20 h at rt, filtered and extracted with a sat. aq. NaHCO₃ solution (50 mL) and water (50 mL). The combined aq. phases were extracted with DCM (3 x 30 mL). The combined organic layers were dried, evaporated and purified by flash chromatography (petroleum ether/EtOAc 3:2). The resulting trisaccharide mimic was suspended in MeOH (20 mL) and sodium (1.2 mg, 53.1 μmol) was added. The resulting solution was stirred for 16 h, quenched with a few drops of AcOH and the solvent was removed under reduced pressure. The crude product was purified by flash chromatography to give **38** (505 mg, 572 μmol, 32 % over 2 steps) as a white solide.

R_f (petroleum ether/EtOAc 1:4) 0.62; [α]_D²² -34.4 (c 0.90, CHCl₃); ¹H NMR (500.1 MHz, CDCl₃): δ = 7.36 – 7.19 (m, 20H, Ar-H), 5.94 – 5.84 (m, 1H, CH₂=CH-CH₂-O), 5.33 – 5.27 (m, 1H, CH₂=CH-CH₂-O), 5.24 – 5.19 (m, 1H, CH₂=CH-CH₂-O), 5.06 (d, *J* = 3.3 Hz, 1H, Fuc-H1), 4.92 (d, *J* = 11.5 Hz, 1H, Ph-CH₂), 4.82 (d, *J* = 11.6 Hz, 1H, Ph-CH₂), 4.76 – 4.64 (m, 4H, Fuc-H5, 3x Ph-CH₂), 4.60 – 4.49 (m, 5H, 3x Ph-CH₂, 2x CH₂=CH-CH₂-O), 4.45 (d, *J* = 6.8 Hz, 1H, Gal-H1), 4.08 – 3.98 (m, 2H, Fuc-H2, Fuc-H3), 3.98 – 3.94 (m, 1H, Gal-H4), 3.87 – 3.74 (m, 2H, Gal-H3, Gal-H6), 3.74 – 3.66 (m, 3H, Gal-H6', Fuc-H4, MeCy-H1), 3.62 – 3.56 (m, 1H, Gal-H5), 3.45 – 3.38 (m, 1H, Gal-H2), 3.22 (t, *J* = 9.1 Hz, 1H, MeCy-H1), 2.02 – 1.94 (m, 1H, MeCy-H6), 1.65 – 1.53 (m, 3H, MeCy-H3, MeCy-H4, MeCy-H5), 1.23 – 1.14 (m, 2H, MeCy-H6', MeCy-H5'), 1.12 (d, *J* = 6.5 Hz, 3H, Fuc-H6), 1.09 (d, *J* = 6.5 Hz, 3H, MeCy-Me), 1.06 – 0.95 (m, 1H, MeCy-H4'); ¹³C NMR (125.8 MHz, CDCl₃): δ = 171.3, 139.2, 139.1, 138.7, 137.9, 132.5, 128.6 - 127.4 (20C), 118.4, 98.4, 82.7, 80.2, 78.5, 76.5, 75.1, 74.4, 73.7, 73.2, 72.6, 69.3, 68.1, 66.4, 66.4, 60.5, 55.8, 39.0, 33.4, 30.7, 22.9, 21.1, 18.9, 16.9, 14.3; ESI-MS: *m/z*: Calcd for C₅₁H₆₃NNaO₁₂ [M+Na]⁺: 904.42, found: 904.28.

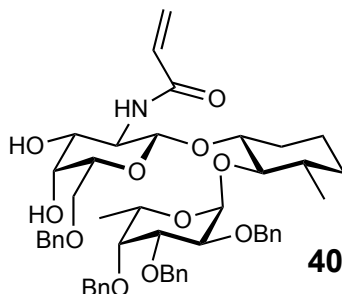
(1*R*,2*R*,3*S*)-2-[(2,3,4-Tri-*O*-benzyl-6-deoxy-α-*L*-galactopyranosyl)oxy]-3-methyl-cyclohex-1-yl 6-*O*-benzyl-2-deoxy-2-[(2-propen-1-yl-oxy)carbonyl]amino]-3-*O*-((*R*)-1-methoxy-1-oxo-pent-4-en-2-yl)-β-*D*-galactopyranoside (39**)**



Alcohol **38** (21.0 mg, 23.8 μmol) and Bu_2SnO (17.8 mg, 71.4 μmol) were dried for 16 h at rt, suspended in MeOH (2 mL) and refluxed for 2 h. The resulting solution was concentrated and coevaporated with toluene. The tin acetal was dried for 16 h at high vacuo, solved in freshly dried (Al_2O_3 column) DME (1.5 mL) and added to CsF (dried for 2 h at high vacuo at 100°C , 6.5 mg, 42.9 μmol). Triflate **21** (18.7 mg, 71.4 μmol) was added and the suspension was stirred for 16 h at rt. A 20% solution of KF (in 1M KH_2PO_4 solution, 2 mL) was added. After stirring for 1 h at rt DCM (10 mL) was added and the aq. phase was extracted with DCM (2 x 10 ml). The combined organic layers were dried (Na_2SO_4) and evaporated under reduced pressure. Column chromatography on silica (petroleum ether/EtOAc 2:1) afforded **39** (20.1 mg, 20.8 μmol , 88 %) as a white solid.

R_f (petroleum ether/EtOAc 3:2) 0.63; ^1H NMR (500.1 MHz, CDCl_3): δ = 7.39 – 7.20 (m, 20H, Ar-H), 5.97 – 5.86 (m, 1H, R-H4), 5.86 – 5.78 (m, 1H, $\text{CH}_2=\text{CH}-\text{CH}_2-\text{O}$), 5.33 – 5.27 (m, 1H, $\text{CH}_2=\text{CH}-\text{CH}_2-\text{O}$), 5.23 – 5.12 (m, 4H, $\text{CH}_2=\text{CH}-\text{CH}_2-\text{O}$, R-H5, R-H5', Gal-H1), 5.05 (d, J = 3.5 Hz, 1H, Fuc-H1), 4.94 (d, J = 11.5 Hz, 1H, Ph- CH_2), 4.85 – 4.76 (m, 2H, Fuc-H5, Ph- CH_2), 4.76 – 4.56 (m, 6H, 4x Ph- CH_2 , Gal-H2, $\text{CH}_2=\text{CH}-\text{CH}_2-\text{O}$), 4.52 – 4.47 (m, 3H, 2x Ph- CH_2 , $\text{CH}_2=\text{CH}-\text{CH}_2-\text{O}$), 4.09 – 3.96 (m, 4H, R-H2, Gal-H3, Fuc-H2, Fuc-H3), 3.79 – 3.73 (m, 1H, Gal-H6), 3.73 – 3.69 (m, 4H, Fuc-H4, Me), 3.65 – 3.60 (m, 1H, Gal-H6'), 3.60 – 3.54 (m, 1H, MeCy-H1), 3.54 – 3.48 (m, 1H, Gal-H5), 3.22 (t, J = 9.1 Hz, 1H, MeCy-H2), 2.59 – 2.51 (m, 1H, R-H3), 2.50 – 2.42 (m, 1H, R-H3'), 1.94 – 1.87 (m, 1H, MeCy-H6), 1.68 – 1.60 (m, 3H, MeCy-H3, MeCy-H4, MeCy-H5), 1.33 – 1.16 (m, 2H, MeCy-H5', MeCy-H6'), 1.13 (d, J = 6.5 Hz, 3H, Fuc-H6), 1.08 (d, J = 6.5 Hz, 3H, MeCy-Me), 1.06 – 0.95 (m, 1H, MeCy-H4'); ^{13}C NMR (125.8 MHz, CDCl_3): δ = 172.2, 160.3, 139.4, 139.2, 138.8, 138.1, 133.6, 133.1, 128.7 - 127.4 (20C), 119.5, 117.6, 98.3, 82.5, 80.4, 79.6, 78.6, 76.5, 75.1, 74.4, 73.7, 72.6, 72.2, 68.6, 66.3, 65.5, 64.7, 52.3, 39.1, 37.5, 33.5, 33.4, 32.1, 31.1, 31.1, 30.8, 24.9, 22.8, 19.0, 17.0; ESI-MS: m/z : Calcd for $\text{C}_{57}\text{H}_{71}\text{NNaO}_{14}$ $[\text{M}+\text{Na}]^+$: 1016.48, found: 1016.62.

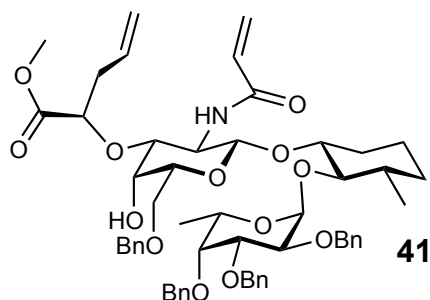
(1*R*,2*R*,3*S*)-2-[(2,3,4-Tri-*O*-benzyl-6-deoxy- α -L-galactopyranosyl)oxy]-3-methyl-cyclohex-1-yl 6-*O*-benzyl-2-deoxy-2-[(1-oxo-prop-2-en-1-yl)amino]- β -D-galactopyranoside (40)



Compound **38** (63.8 mg, 72.3 μ mol) was dissolved in DCM (10 mL) and PPh_4 (1.7 mg, 1.44 μ mol) and Bu_3SnH (23.2 mg, 79.6 μ mol) in DCM (1 mL) was added. The solution was stirred for 30 min and the product purified by a short column of silica (DCM/MeOH 10:3). The resulting amine was dissolved in DCM (10 mL) and Et_3N (37 μ l, 265 μ mol) and acryloylchloride (8.6 μ l, 106 μ mol) were added at 0°C. The reaction mixture was stirred for 2.5 h at this temperature. The solvent was removed under reduced pressure and purified by flash chromatography (DCM/MeOH 10:1) to yield **40** (58.2 mg, 68.3 μ mol, 57 %) as a white solid.

R_f (EtOAc) 0.32; $[\alpha]_D^{22}$ -30.9 (c 0.33, CHCl_3); ^1H NMR (500.1 MHz, CDCl_3): δ = 7.33-7.22 (m, 20H, Ar-H), 6.28 (d, 1H, J = 16.9, 0.8 Hz, 1H, $\text{CH}_2=\text{CHCO}$), 6.09 (dd, J = 10.3, 16.9 Hz, 1H, $\text{CH}_2=\text{CH-CO}$), 5.70 (dd, J = 10.4, 0.8 Hz, 1H, $\text{CH}_2=\text{CH-CO}$), 5.03 (d, J = 3.3, 1H, Fuc-H1), 4.91 (d, J = 11.4 Hz, 1H, PhCH_2), 4.81 (d, J = 11.6 Hz, 1H, Ph-CH_2), 4.73-4.64 (m, 4H, 3x Ph-CH_2 , Fuc-H5), 4.59-4.48 (m, 3H, 3x Ph-CH_2), 4.47 (d, J = 8.3 Hz, 1H Gal-H1), 4.05 (dd, J = 10.3 Hz, 3.4 Hz, 1H, Fuc-H2), 4.01 (dd, J = 10.3 Hz, 2.4 Hz, 1H, Fuc-H3), 3.96 (d, J = 2.4 Hz, 1H, Gal-H4), 3.85-3.78 (m, 2H, Gal-H3, Gal-H6), 3.76-3.68 (m, 3H, Fuc-H4, MeCy-H1, Gal-H6'), 3.63-3.53 (m, 2H, Gal-H2, Gal-H5), 3.2 (t, 1H, J = 9.3 Hz, MeCy-H1), 2.05-1.98 (m, 1H, MeCy-H6), 1.67-1.55 (m, 3H, MeCy-H3, MeCy-H4, MeCy-H5), 1.30-1.16 (m, 2, MeCy-H6', MeCy-H5'), 1.10-1.06 (d, J = 6.4 Hz, 6H, Fuc-H6, MeCy-Me), 1.06 - 0.86 (m, 1H, MeCy-H4') ppm. ^{13}C NMR (125.8 MHz, CDCl_3): δ = 168.1, 139.2, 139.2, 138.7, 138.2, 130.3, 128.8 – 127.4 (21C), 98.6, 97.7, 82.9, 80.3, 78.6, 78.4, 76.4, 75.2, 74.6, 73.9, 73.7, 72.7, 69.4, 67.9, 66.6, 56.6, 39.1, 33.5, 31.0, 29.9, 23.1, 19.0, 17.0, 14.4; ESI-MS: m/z : Calcd for $\text{C}_{50}\text{H}_{61}\text{NNaO}_{11}$ $[\text{M}+\text{Na}]^+$: 874.41, found: 874.46.

(1*R*,2*R*,3*S*)-2-[(2,3,4-Tri-*O*-benzyl-6-deoxy- α -L-galactopyranosyl)oxy]-3-methyl-cyclohex-1-yl 6-*O*-benzyl-2-deoxy-2-[(1-oxo-prop-2-en-1-yl)amino]-3-*O*-((*R*)-1-methoxy-1-oxo-pent-4-en-2-yl)- β -D-galactopyranoside (41**)**

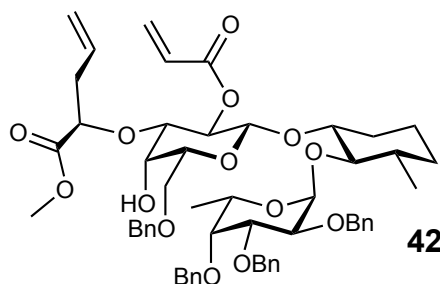


Compound **40** (82.4 mg, 96.7 μ mol) and Bu_2SnO (72.2 mg, 290 μ mol) were dried for 4 h at rt and suspended in MeOH (6 mL). The reaction mixture was refluxed for 2 h, concentrated under reduced pressure and coevaporated with toluene. The stannyl acetal was dried for 16 h under high vacuum, dissolved in freshly dried DME (1.5 mL) and added to CsF (dried for 2 h at high vacuo at 100°C, 26.4 mg, 174 μ mol). Triflate **21** (76.0 mg, 290 μ mol) was added and the suspension was stirred for 18 h at rt. A 20% solution of KF (in 1M KH_2PO_4 solution, 2 mL) was added and the resulting mixture stirring for 1 h at rt. DCM (10 mL) was added and the aqueous phase was extracted with DCM (2 x 15 mL). The combined organic layers were dried (Na_2SO_4) and evaporated under reduced pressure. The crude product was purified by flash chromatography (petroleum ether/EtOAc 2:1) and afforded **41** (49.0 mg, 50.8 μ mol (15.7 mg of reactant were recovered), 65 %) as a white solid.

R_f (petroleum ether/EtOAc 1:1) 0.61; $[\alpha]_D^{22}$ -23.1 (c 0.45, CHCl_3); ^1H NMR (500.1 MHz, CDCl_3): δ = 7.38 – 7.20 (m, 20H), 6.25 – 6.20 (m, 1H, $\text{CH}_2=\text{CHCO}$), 6.14 – 6.07 (m, 1H, $\text{CH}_2=\text{CH-CO}$), 6.00 (d, J = 6.7 Hz, 1H, NH), 5.86 – 5.76 (m, 1H, R-H4), 5.64 – 5.60 (m, 1H, $\text{CH}_2=\text{CH-CO}$), 5.23 – 5.13 (m, 2H, R-H5, R-H5'), 5.04 (d, J = 3.4 Hz, 1H, Fuc-H1), 4.98 – 4.92 (m, 2H, Gal-H1, Ph- CH_2), 4.86 – 4.77 (m, 3H, Fuc-H5, Ph- CH_2), 4.75 – 4.62 (m, 5H, 3x Ph- CH_2), 4.60 (d, J = 11.5 Hz, 1H, Ph- CH_2), 4.51 (s, 2H, 2x Ph- CH_2), 4.24 (dd, J = 10.5, 2.8 Hz, 1H, Gal-H3), 4.11 – 3.96 (m, 3H, Fuc-H2, Fuc-H3, Gal-H4), 3.87 – 3.70 (m, 3H, Gal-H6, Fuc-H4), 3.69 – 3.62 (m, 4H, Me, Gal-H6'), 3.62 – 3.51 (m, 2H, MeCy-H1, Gal-H5), 3.44 – 3.34 (m, 1H, Gal-H2), 3.19 (t, J = 9.1 Hz, 1H, MeCy-H2), 2.57 – 2.49 (m, 1H, R-H3), 2.48 – 2.39 (m, 1H, R-H3'), 1.87 (m, 1H, MeCy-H6), 1.65 – 1.49 (m, 3H, Cy-H3, MeCy-H4, MeCy-H5), 1.36 – 1.10 (m, 5H, MeCy-H5', MeCy-H6', Fuc-H6), 1.07 (d, J = 6.5 Hz, 3H, MeCy-Me), 1.03 – 0.92 (m, 1H, MeCy-H4'); ^{13}C NMR (125.8 MHz, CDCl_3): δ = 172.4, 166.4, 139.3, 139.2, 138.8, 138.1, 133.5, 131.5, 128.7 - 127.4 (21x C), 126.3, 119.5, 98.4, 97.7, 82.7, 80.3, 79.7, 78.6, 76.6, 76.5, 75.1, 74.4, 73.7, 72.6, 72.1, 68.6, 66.3, 64.8,

54.5, 52.4, 39.1, 37.6, 33.5, 30.8, 23.0, 19.0, 17.1; ESI-MS: m/z : Calcd for $C_{56}H_{69}NNaO_{11}$ $[M+Na]^+$: 986.47, found: 986.42.

(1*R*,2*R*,3*S*)-2-[(2,3,4-Tri-*O*-benzyl-6-deoxy- α -L-galactopyranosyl)oxy]-3-methyl-cyclohex-1-yl 6-*O*-benzyl-2-*O*-[(1-oxo-prop-2-en-1-yl)amino]-3-*O*-((*R*)-1-methoxy-1-oxo-pent-4-en-2-yl)- β -D-galactopyranoside (42**)**



Alcohol **31** (23.9 mg, 26.3 μ mol) was dissolved in DCM (2 mL) and Et_3N (9.1 μ L, 65.7 μ mol) and acryloylchloride (3.3 μ L, 39.4 μ L) were added at 0°C. The reaction mixture was allowed to reach rt over 16 h. The reaction mixture was quenched with a few drops of MeOH and the solvent was removed under reduced pressure. The crude product was purified by flash chromatography (petroleum ether/EtOAc 2:1) to give **42** (11.2 mg, 11.6 μ mol, 44%) as a white solide.

R_f (petroleum ether/EtOAc 1:1) 0.67; $[\alpha]_D^{22}$ -39.7 (c 0.27, $CHCl_3$); 1H NMR (500.1 MHz, $CDCl_3$): δ = 7.62 – 7.14 (m, 20H, Ar-H), 6.50 – 6.42 (m, 1H, $CH_2=CH-CO$), 6.22 – 6.12 (m, 1H, $CH_2=CH-CO$), 5.89 – 5.77 (m, 2H, $CH_2=CH-CO$), 5.58 (s, 1H, Ph-CH), 5.44 – 5.36 (m, 1H, Gal-H2), 5.11 - 5.02 (m, 1H, R-H5), 5.01 – 4.92 (m, 3H, R-H5', Fuc-H5, Fuc-H1), 4.80 (d, J = 11.7 Hz, 1H, Ph- CH_2), 4.69 (d, J = 11.7 Hz, 1H, Ph- CH_2), 4.58 – 4.50 (m, 3H, Gal-H6, Gal-H6', Gal-H1), 4.38 – 4.31 (m, 2H, Gal-H4, Ph- CH_2), 4.29 – 4.23 (m, 1H, R-H2), 4.20 (d, J = 11.3 Hz, 1H, Ph- CH_2), 4.09 (d, J = 12.0 Hz, 1H, Ph- CH_2), 3.99 – 3.88 (m, 2H, Fuc-H2, Fuc-H3), 3.87 – 3.81 (m, 1H, Gal-H3), 3.62 (d, J = 11.3 Hz, 1H, Ph- CH_2), 3.59 - 3.50 (m, 4H, MeCy-H1, Me), 3.39 – 3.35 (m, 1H, Gal-H5), 3.28 – 3.18 (m, 2H, Fuc-H4, MeCy-H2), 2.52 – 2.39 (m, 2H, R-H3), 1.95 – 1.88 (m, 1H, MeCy-H6), 1.65 – 1.50 (m, 3H, MeCy-H3, MeCy-H4, MeCy-H5), 1.35 – 1.12 (m, 5H, MeCy-H6', Fuc-H6, MeCy-H5'), 1.09 – 0.95 (m, 4H, MeCy-Me, Cy-H4'); ^{13}C NMR (125.8 MHz, $CDCl_3$): δ = 172.4, 164.8, 139.9, 139.7, 138.9, 137.9, 133.3, 131.0, 129.0 - 126.1 (21H), 118.0, 100.0, 99.9, 98.39, 81.50, 81.07, 79.9, 79.1, 77.3, 75.7, 75.7, 75.0, 74.6, 73.3, 71.3, 69.6, 69.3, 66.4, 51.9, 39.7, 37.9,

33.8, 31.4, 29.9, 23.5, 18.9, 16.5; ESI-MS: m/z : Calcd for $C_{56}H_{66}NaO_{14}$ $[M+Na]^+$: 985.44, found: 985.55.

References

- (1) Kindly provided by Chembiotek.
- (2) Binder, F. P. C., PhD Thesis, Basel, **2011**.
- (3) Schwab, P.; Grubbs, R. H.; Ziller, J. W. *J. Am. Chem. Soc.* **1996**, *118*, 100.
- (4) Scholl, M.; Ding, S.; Lee, C. W.; Grubbs, R. H. *Org. Lett.* **1999**, *1*, 953.
- (5) Garber, S. B.; Kingsbury, J. S.; Gray, B. L.; Hoveyda, A. H. *J. Am. Chem. Soc.* **2000**, *122*, 8168.
- (6) Vougioukalakis, G. C.; Grubbs, R. H. *Chem. Rev.* **2010**, *110*, 1746.
- (7) Chatterjee, A. K.; Choi, T.-L.; Sanders, D. P.; Grubbs, R. H. *J. Am. Chem. Soc.* **2003**, *125*, 11360.
- (8) Maestro, version 9.3, Schrödinger, LLC, New York, NY, **2012**.
- (9) MacroModel, version 9.9, Schrödinger, LLC, New York, NY, **2012**.
- (10) Desmond Molecular Dynamics System, version 3.1, D. E. Shaw Research, New York, NY, **2012**.
- (11) Bowers, K. J.; Chow, E.; Xu, H.; Dror, R. O.; Eastwood, M. P.; Gregersen, B. A.; Klepeis, J. L.; Kolossvary, I.; Moraes, M. A.; Sacerdoti, F. D.; Salmon, J. K.; Shan, Y.; Shaw, D. E. Proceedings of the 2006 ACM/IEEE Conference on Supercomputing (SC06), Tampa, FL, 11 to 17 November **2006** (ACM Press, New York, **2006**).
- (12) Gottlieb, H. E.; Kotlyar, V.; Nudelman, A. *J. Org. Chem.* **1997**, *62*, 7512.

3.2.2 Bioisosteric modification of the acid pharmacophore in cyclic selectin antagonists (Manuskript)

Contributions

- Manuscript preparation
- Compound synthesis

Abstract

In the mammalian immune system selectins are key players in inflammation. They initiate the first step in leukocyte extravasation from blood to infected and inflamed tissue. An excessive recruitment can be observed in several acute and chronic diseases. Blocking this selectin mediated rolling is a promising approach to treat diseases with inflammatory component.

The selectin antagonists investigated in chapter 3.2.1 showed an excellent pre-organization and a limited number of free rotatable bonds, which is an important parameter in the design of oral available drugs. But the high polar surface area and the low lipophilicity reduce the chance for oral availability. The replacement of polar pharmacophores by more lipophilic bioisosteres is a promising approach to improve the druglike properties. In this communication we analyzed the possibility of a replacement of the acid pharmacophore by bioisosteres. When the carboxylate was replaced by a methylamide a loss of affinity in an E-selectin binding assay was observed, which is probably due to repulsions between the hydrogen of the amide to the methylene groups of the macrocycle, or the *N*-methyl group to the amino acids of E-selectin. These spatial limitations in our cyclic antagonist will shift the focus to the hydroxy pharmacophores in further bioisosteric replacement studies.

Introduction

Selectins are mammalian lectins and involved in the initial step of the recruitment leukocytes to the inflamed tissue. Many inflammatory diseases like asthma,¹ psoriasis,² reperfusion injury³ and rheumatoid arthritis⁴ are associated with an excessive infiltration of leukocytes to the inflamed tissue. Furthermore, several metastatic cancers harness the selectin mediated extravasation mechanism from the blood vessels.^{5,6} Thus, selectins are an attractive therapeutic target.

To predict oral availability of test compounds, there are several physicochemical and structural rules and filters, e.g. the Lipinski rule of 5 (RO5)^{7,8} or the observation that compounds with less than 10 free rotatable bonds and a polar surface area (PSA) smaller than 140 Å² are beneficial for oral availability.⁹ The cyclic selectin antagonists **4a-c** described in chapter 3.2.1 are violating 2 of 4 rules of the RO5. The number of free rotatable bonds is smaller than 10, but the PSA exceed the limit of 140 Å². This is due to the high number of oxygen atoms in the mimics.

By replacing the hydrophilic groups by bioisosteres, lipophilicity can be increased and the PSA be reduced.¹⁰ The aim of this short communication is to evaluate bioisosteres of the carboxylate.

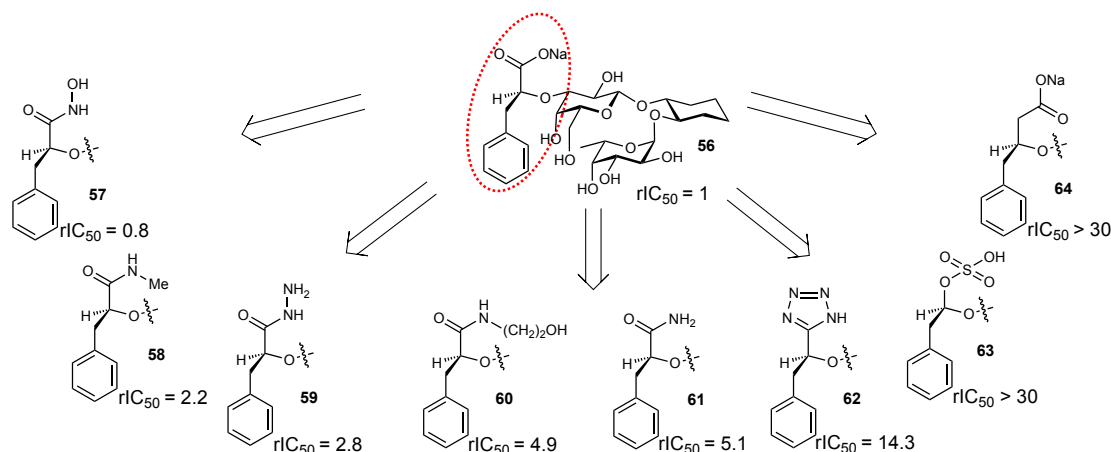


Figure 1. Bioisosteric replacements of the carboxylic acid of selectin antagonist **56**.¹¹

In previous work, the effects of bioisosteric replacements of the acid pharmacophore in sLe^x mimics were evaluated. Compound **56**¹² served as model system and the binding affinities of the acid bioisosteres **57** - **64** were compared in a competitive binding assay¹³ (Figure 1). An additional methylene group (\rightarrow **64**), the replacement of the acid functionality by sulfate (\rightarrow **63**), or terazole (\rightarrow **62**) moieties led to a loss in affinity exceeding a factor of 14. In contrast, a series of amides (\rightarrow **57-61**) showed rIC₅₀ values between 0.8 and 5.1.

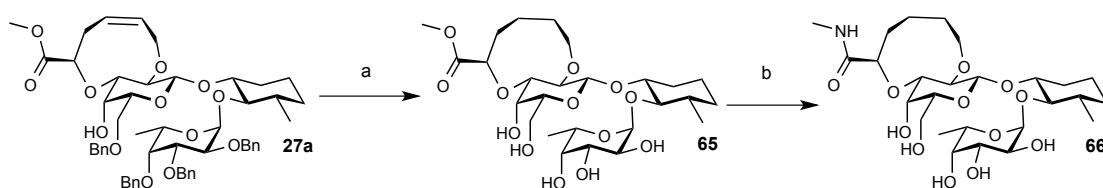
The methylamide fits best to our purpose to reduce the PSA on the one hand and to retain the binding affinity on the other hand. We therefore planed to adopt the methyl amide modification to our cyclic selectin antagonist.

Results and discussion

Synthesis of methyl amide **66** and biological evaluation.

To investigate the effects of the bioisostere methylamide modification on the affinity and the clogP, we converted the 9-member cyclic intermediate **27a** (chapter 3.2.1) to the methyl amide **66**.

The benzyl protecting groups and the double bond were hydrogenated with Pd(OH)₂/C in dioxane/water. Aminolysis of the resulting methylester **65** afforded methylamide **66**.

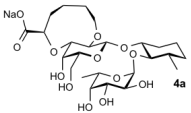
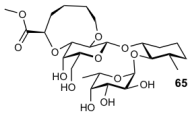
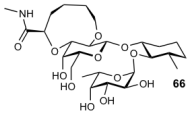


Scheme 1. a.) Pd(OH)₂/C, dioxane/H₂O (4:1), H₂; b.) MeNH₂ in THF and MeNH₂ in EtOH (70 % over 2 steps).

Compared to **4a** (Table 1; chapter 3.2.1), the experimental logD_{7.4} showed a minor gain of methylamide **66** and a more favorable value for methylester **65**. The calculated PSA values are reduced by approximately 15 Å² compared to acid **4a**.

But unfortunately amide **66** showed an unexpected drop in affinity by a factor of 10 compared to the acid **4a**. This is not in accordance with the only minor decrease in affinity observed for the selectin antagonists **56**→**58** (Figure 1).¹¹ One possible explanation is a different binding mode for amides **58** and **66**. While the cyclic antagonist **66** is rigid and the methylamide is locked in *trans* conformation, the amide functionality in the (*S*)-phenyl lactic acid antagonist **58** is rather flexible and can adjust upon binding. The affinity of the methylester **65** was even lower and could not be detected in our IC₅₀ assay.

Table 1. Calculated PSA,¹⁴ experimental logD_{7.4}, and IC₅₀ values of acid **4a**, methyl ester **65** and methyl amide **66**.

Compound	logD _{7.4} (mean ± SD)	PSA (Å ²)	IC ₅₀
	-1.3 ± 0.3	184	23 μM
	-0.2 ± 0.0	168	n.d.
	-0.8 ± 0.0	171	199 μM

Bulky acid isosteres are unfavorable in macrocycles.

To support this hypothesis, the low energy conformations of cyclic sLe^x mimic **66** was calculated and docked to E-selectin. The pharmacophores are oriented similar to the crystal structure of compound **3** (Figure 2; chapter 3.2.1).

In the preferred low-energy conformations (Figure 2c) the methyl group is oriented toward Tyr48 (red circle in Figure 2a) and causes steric repulsion. A conformer with favorably oriented amide group (Figure 2b) on the other hand is sterically unfavorable, because of repulsion to the methylene group of the macrocycle (3.6 kcal/mol above the global minimum). This would explain the significant drop in affinity from **4a** to **66**. The (*S*)-phenyl lactic acid analogue **58** on the other hand has more rotational freedom and could probably adjust the methylamide in the required orientation. This would explain the only minor drop in binding affinity from acid to amide in the (*S*)-phenyllactic acid series (**56**→**58**), in contrast to the major loss in the macrocyclic antagonists (**4a**→**66**).

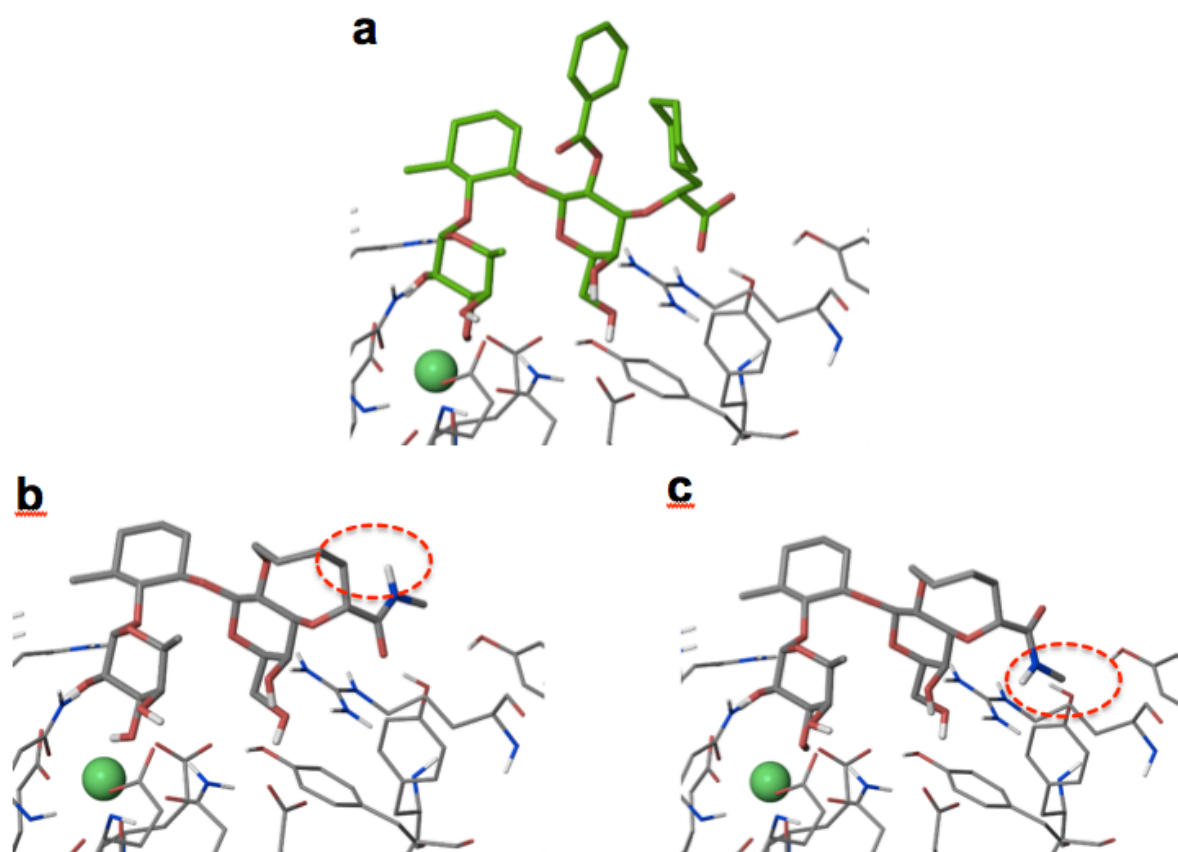


Figure 2. Crystal structure¹⁵ of **3** (chapter 3.2.1) bound to E-selectin (a) and high (b) and low (c) energy conformation of **66** docked to E-selectin.

Experimental part

Molecular modeling. The 3D structure of the reactant molecule (before ring closing) was built in the Maestro¹⁶ molecular modeling environment and a conformational analysis was performed using MacroModel¹⁷ by performing 5000 steps of the mixed torsional/low-mode sampling in combination with the OPLS-2005 force-field in implicit solvent conditions

(octanol). Unfortunately, in Macromodel there is no solvation model available for the dichloromethane, which was used in experimental setup for the ring-closing reactions. Therefore a solvent model for octanol ($\epsilon = 9.86$) with the dielectric constant similar to that of the dichloromethane ($\epsilon = 8.93$) was chosen from the list of available solvent models.

Global minimum identified in the conformational search was then placed in a periodic boundary system (cube, side length 40.0 Å) filled with explicit octanol molecules. The system was minimized and a simulated annealing simulation (gradual heating of the system to the temperature of 400 K in 330 ps followed by slow gradual cooling to 300 K in 500 ps) was performed in Desmond^{18,19} in order to bring the system into equilibrium. The production phase molecular dynamics simulation of the total duration of 48 ns was performed. The distance between the key atoms involved in the ring closing reaction was monitored every 4.8 ps (10 000 data points were saved).

The average distance of the key atoms was calculated and the frequency of conformations (with a short key atom distance) where ring closing reaction could eventually occur was evaluated.

logD_{7.4} determination.²⁰ The *in silico* prediction tool ALOGPS^{21,22} was used to estimate the logP values of the compounds. Depending on these values, the compounds were classified into three categories: hydrophilic compounds (logP below zero), moderately lipophilic compounds (logP between zero and one) and lipophilic compounds (logP above one). For each category, two different ratios (volume of 1-octanol to volume of buffer) were defined as experimental parameters (Table X):

Table 2.

Compound type	logP	Ratios (1-octanol: buffer)
hydrophilic	< 0	30:140, 40:130
moderately lipophilic	0 - 1	70:110, 110:70
lipophilic	> 1	3:180, 4:180

Equal amounts of phosphate buffer (0.1 M, pH 7.4) and 1-octanol were mixed and shaken vigorously for 5 min to saturate the phases. The mixture was left until separation of the two phases occurred, and the buffer was retrieved. Stock solutions of the test compounds were diluted with buffer to a concentration of 1 µM. For each compound, six determinations, i.e., 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 aluminium foil, shaken (1350 rpm, 25 $^{\circ}\text{C}$, 2 h) on a Heidolph Titramax 1000 plate-shaker (Heidolph Instruments GmbH & Co. KG, Schwabach, Germany) and centrifuged (2000 rpm, 25 $^{\circ}\text{C}$, 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).

$\log D_{7.4}$ was 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 equilibration:

$$\log D_{7.4} = \frac{1\mu\text{M} - c_B}{c_B} \times \frac{1}{o:b}$$

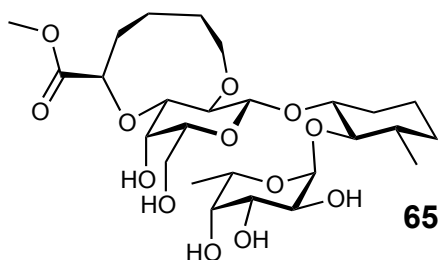
The average of the three $\log D_{7.4}$ values per 1-octanol:buffer ratio was calculated. If the two mean values obtained for a compound did not differ by more than 0.1 unit, the results were accepted.

General methods. Commercial materials (Sigma-Aldrich) were used without further purification, solvents were reagent grade (Acros). CH_2Cl_2 and MeOH were dried by passing through an Al_2O_3 (Fluka, type 5016 A basic) column. DMF extra dry (Acros) was used as is. All reactions were performed in oven dried glassware under an atmosphere of argon.

^1H and ^{13}C NMR spectra were recorded on a Bruker Avance DMX-500 at room temperature. Chemical shifts are reported in ppm and referenced to TMS using residual solvent peaks.^[S14] For complex molecules the following prefixes were used: Fuc (fucose), Gal (galactose) and GlcNAc (*N*-acetyl glucosamine). The coupling constants (*J*) are reported in Hertz (Hz). Analytical TLC was performed on Merck silica gel 60 F₂₅₄ glass plates and visualized by UV light and by charring with a molybdate solution (a 0.02 M solution of ammonium cerium sulfate dihydrate and ammonium molybdate tetrahydrate in aq. 10% H_2SO_4) by heating for 5 min at 140 $^{\circ}\text{C}$. Column chromatography was performed on a CombiFlash Companion (Teledyne-ISCO, Inc.) using RediSep® normal phase disposable flash columns (silica gel). Reversed phase chromatography was carried out with LiChroprep®RP-18 (Merck, 40-63 μm). Optical rotations were determined on a Perkin-Elmer Polarimeter 341.

Low resolution mass spectra were measured on a Waters micromass ZQ. High resolution mass spectra (HRMS) were obtained on a micrOTOF spectrometer (Bruker Daltonics, Germany) equipped with a TOF hexapole detector. Purity of final compound was determined on an Agilent 1100 HPLC; detector: ELS, Waters 2420; column: Waters Atlantis dC18, 3 μm , 4.6 x 75 mm; eluents: A: water + 0.1% TFA; B: 90% acetonitrile + 10% water + 0.1% TFA; linear gradient: 0 - 1 min 5% B; 1 - 16 min 5 to 70% B; flow: 0.5 mL/min MALDI-TOF and ESI-MS glycoprotein analyses were recorded by the Functional Genomic Center Zurich (FGCZ). VIVASPIN® 500 ultrafiltration tubes with 10000 MWCO, 10 kDa molecular weight cutoff PES membrane, and ZelluTrans/Roth dialysis membranes MWCO 8000-10000 were used for glycoprotein concentration and dialysis.

Compound 65

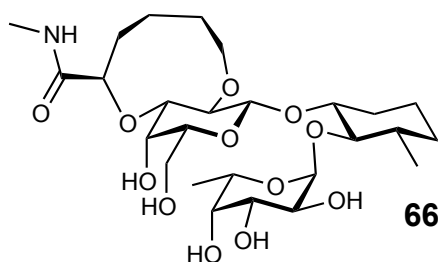


Compound **27a** (11.4 mg, 12.3 μmol) was dissolved in dioxane/water (4:1, 4 mL) and $\text{Pd}(\text{OH})_2/\text{C}$ (6 mg, 10% $\text{Pd}(\text{OH})_2$) was added. The suspension was stirred for 16 h under an atmosphere of hydrogen. Solvent was removed in vacuo. The residue was purified *via* RP chromatography (MeOH/ H_2O), eluted through a sodium exchange column (Dowex 50/8 sodium form) and finally purified *via* size exclusion chromatography. Lyophilization from water/dioxane gave **65** (4.9 mg, 868 μmol , 70%) as a white fluffy foam.

R_f (DCM/MeOH 10:1) 0.12; $[\alpha]_D^{22}$ -68.1 (c 1.10, H_2O); ^1H NMR (500.1 MHz, MeOD): δ = 4.95 (d, J = 4.0 Hz, 1H, Fuc-H1), 4.90 – 4.84 (m, 1H, Fuc-H5), 4.47 – 4.42 (m, 1H, R-H1), 4.25 (d, J = 7.6 Hz, 1H, Gal-H1), 4.08 (d, J = 3.1 Hz, 1H, Gal-H5), 3.81 – 3.72 (m, 2H, Fuc-H3, R-H5), 3.71 – 3.53 (m, 10H, Fuc-H2, Fuc-H4, R-H5', COOMe, Gal-H6, Gal-H6', Gal-H3, MeCy-H2), 3.34 – 3.20 (m, 2H, Gal-H4, Gal-H2), 3.13 (t, J = 9.2 Hz, 1H, MeCy-H1), 2.45 – 2.34 (m, 1H, R-H2), 2.25 – 2.13 (m, 1H, R-H4), 2.10 – 2.02 (m, 1H, MeCy-H6), 1.80 – 1.70 (m, 3H, R-H2, R-H3, R-H3'), 1.65 – 1.52 (m, 4H, MeCy-H5, MeCy-H5', MeCy-H4, MeCy-H3), 1.31 – 1.17 (m, 4H, R-H2, R-H3, R-H4', MeCy-H6'), 1.13 (d, J = 5.2 Hz, 3H, Fuc-H6), 1.07 (d, J = 6.4 Hz, 3H, MeCy-Me), 1.10 – 0.99 (m, 1H, MeCy-H4); ^{13}C NMR (125.8 MHz, MeOD) δ 174.8 (COO), 102.4 (Gal-C1), 100.2 (Fuc-C1), 84.1 (MeCy-C1),

80.2 (MeCy-C2), 79.9 (R-C1), 78.6 (Gal-C2), 77.7 (Gal-C3), 76.0 (Gal-C4), 74.8 (R-C5), 73.8 (Fuc-C4), 71.4 (Fuc-C3), 70.4 (Fuc-C2), 70.4 (Gal-C5), 67.4 (Fuc-C5), 63.0 (Gal-C6), 52.2 (COOMe), 40.2 (MeCy-C3), 34.8 (MeCy-C4), 32.1 (MeCy-C6), 30.2 (R-C2), 28.7 (R-C3), 26.8 (R-C4), 24.0 (MeCy-C5), 19.5 (MeCy-Me), 16.9 (Fuc-C6); HR-MS: m/z : Calcd for $C_{26}H_{44}NaO_{13}$ $[M+Na]^+$: 587.2680, found: 587.2684.

Compound 66



Compound **65** (3.90 mg, 6.9 μ mol) was dissolved in $MeNH_2$ in THF (2M, 1 ml) and $MeNH_2$ in EtOH (8M, 1ml) and stirred at rt for 24 h. The solvent was removed under reduced pressure and the crude product purified by RP flash chromatography and size exclusion chromatography to give **66** as a white solid (3.9 mg, 6.9 μ mol, quant.).

$[\alpha]_D^{22}$ -16.0 (c 0.28, H_2O); 1H NMR (500.1 MHz, D_2O): δ = 5.13 (d, J = 3.2 Hz, 1H, Fuc-H1), 4.98 – 4.90 (m, 1H, Fuc-H5), 4.55 – 4.48 (m, 2H, R-H2, Gal-H1), 3.99 – 3.95 (m, 1H, Gal-H4), 3.94 – 3.87 (m, 2H, R-H6, Fuc-H3), 3.87 – 3.75 (m, 4H, R-H6', Fuc-H2, Fuc-H4, Gal-H3), 3.74 – 3.62 (m, 3H, Cy-H1, Gal-H6, Gal-H6'), 3.59 – 3.48 (m, 2H, Gal-H2, Gal-H5), 3.24 (t, J = 9.4 Hz, 1H, MeCy-H2), 2.77 (s, 3H, NH- CH_3), 2.18 – 2.10 (m, 1H, MeCy-H6), 2.00 – 1.84 (m, 2H, R-H3, R-H3'), 1.83 – 1.54 (m, 8H, MeCy-H3, MeCy-H4, MeCy-H5, R-H4, R-H4', R-H5, R-H5'), 1.39 – 1.21 (m, 2H, MeCy-H6', MeCy-H5'), 1.24 (d, J = 6.4 Hz, 3H, Fuc-H6), 1.17 – 1.04 (m, 1H, Cy-H4'), 1.10 (d, J = 6.2 Hz, 3H, MeCy-Me); ^{13}C NMR (125.8 MHz, D_2O): δ = 176.3 (COO), 100.5 (Gal-C1), 98.6 (Fuc-C1), 83.8 (Cy-C2), 79.2 (Cy-C1), 76.4 (Gal-C3), 75.5 (R-C2), 74.8 (Gal-C5), 73.8 (Gal-C2), 72.0 (Fuc-C4), 71.9 (Fuc-C2), 69.3 (Fuc-C3), 69.0 (R-C6), 68.2 (Gal-C4), 66.5 (Fuc-C5), 61.4 (Gal-C6), 38.5 (Cy-C3), 33.1 (Cy-C4), 30.5 (Cy-C6), 29.9 (R-C2), 25.5 (Me), 23.8 (R-C3), 22.4 (Cy-C5), 21.1 (R-C4), 18.2 (Cy-Me), 15.7 (Fuc-C3); HR-MS: m/z : Calcd for $C_{26}H_{45}NNaO_{12}$ $[M+Na]^+$: 586.28, found: 586.31.

References

- (1) Romano, S. *Treatm. Respirat. Med.* **2005**, *4*, 85.
- (2) Bock, D.; Philipp, S.; Wolff, G. *Expert Opin. Invest. Drugs* **2006**, *15*, 963.
- (3) Lefer, D. J.; Flynn, D. M.; Phillips, M. L.; Ratcliffe, M.; Buda, A. J. *Circulation* **1994**, *90*, 2390.
- (4) Sfikakis, P. P.; Mavrikakis, M. *Clin. Rheum.* **1999**, *18*, 317.
- (5) Witz, I. In *The Link Between Inflammation and Cancer*; Dalglish, A., Haefner, B., Eds.; Springer US: **2006**; Vol. 130, p 125.
- (6) Witz, I. P. *Immun. Lett.* **2006**, *104*, 89.
- (7) Lipinski, C. A. *Drug Discov. Today: Technol.* **2004**, *1*, 337.
- (8) Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J. *Adv. Drug Deliv. Rev.* **2001**, *46*, 3.
- (9) Veber, D. F.; Johnson, S. R.; Cheng, H.-Y.; Smith, B. R.; Ward, K. W.; Kopple, K. *D. J. Med. Chem.* **2002**, *45*, 2615.
- (10) Meanwell, N. A. *J. Med. Chem.* **2011**, *54*, 2529.
- (11) Ernst, B. laboratory, unpublished results.
- (12) Kolb, H. C.; Ernst, B. *Chemi. – Eur. J.* **1997**, *3*, 1571.
- (13) Schwizer, D.; Patton, J. T.; Cutting, B.; Smiesko, M.; Wagner, B.; Kato, A.; Weckerle, C.; Binder, F. P. C.; Rabbani, S.; Schwardt, O.; Magnani, J. L.; Ernst, B. *Chem. – Eur. J.* **2012**, *18*, 1342.
- (14) QikProp, version 3.4, Schrödinger, LLC, New York, NY, **2011**.
- (15) Roland, P. G. c. Preston, R. PhD thesis, University of Basel, **2012**.
- (16) Maestro, version 9.3, Schrödinger, LLC, New York, NY, **2012**.
- (17) MacroModel, version 9.9, Schrödinger, LLC, New York, NY, **2012**.
- (18) Desmond Molecular Dynamics System, version 3.1, D. E. Shaw Research, New York, NY, **2012**.
- (19) Bowers, K. J.; Chow, E.; Xu, H.; Dror, R. O.; Eastwood, M. P.; Gregersen, B. A.; Klepeis, J. L.; Kolossvary, I.; Moraes, M. A.; Sacerdoti, F. D.; Salmon, J. K.; Shan, Y.; Shaw, D. E. Proceedings of the 2006 ACM/IEEE Conference on Supercomputing (SC06), Tampa, FL, 11 to 17 November **2006** (ACM Press, New York, **2006**).
- (20) Klein, T.; Abgottspon, D.; Wittwer, M.; Rabbani, S.; Herold, J.; Jiang, X.; Kleeb, S.; Lüthi, C.; Scharenberg, M.; Bezenç on, J.; Gubler, E.; Pang, L.; Smiesko, M.; Cutting, B.; Schwardt, O.; Ernst, B. *J. Med. Chem.* **2010**, *53*, 8627.

- (21) VCCLAB, *Virtual Computational Chemistry Laboratory*; <http://www.vcclab.org>, 2005.
- (22) Tetko, I.; Gasteiger, J.; Todeschini, R.; Mauri, A.; Livingstone, D.; Ertl, P.; Palyulin, V.; Radchenko, E.; Zefirov, N.; Makarenko, A.; Tanchuk, V.; Prokopenko, V. *J. Comput.-Aid. Mol. Design* **2005**, *19*, 453.

Chapter 4. – Summary and Outlook

Selectins are involved in several inflammatory diseases and cancer, and therefore, represent an interesting therapeutic target.¹ SLe^x is the natural ligand of selectins, however, due to its low binding affinity and the high polarity and complexity it is far from being druglike. In the past several sLe^x mimics were developed that showed improved binding affinities and more favorable physicochemical parameters.²

For complex ligands a certain degree of pre-organization in solution is crucial to reach adequate binding affinities to their target. Elucidating the structure of the Le^x trisaccharide core conformation in solution by NMR spectroscopy a nonconventional C–H···O hydrogen bond between H-C(5) of L-fucose and O(5) of D-galactose was identified as additional stabilizing element to the previous reported hydrophobic interaction between L-fucose and the b-face of D-galactose, the exo-anomeric effect and the steric repulsion between the L-fucose to the *N*-acetyl group of the adjacent GlcNAc (Figure 1).³ This nonconventional hydrogen bond in branched trisaccharides was not reported in literature yet, although it contributes with 40 % to the overall stabilization energy. It was shown that besides the Le^x trisaccharide the solution conformations of several other branched oligosaccharides are stabilized by these nonconventional C–H···O hydrogen bonds.

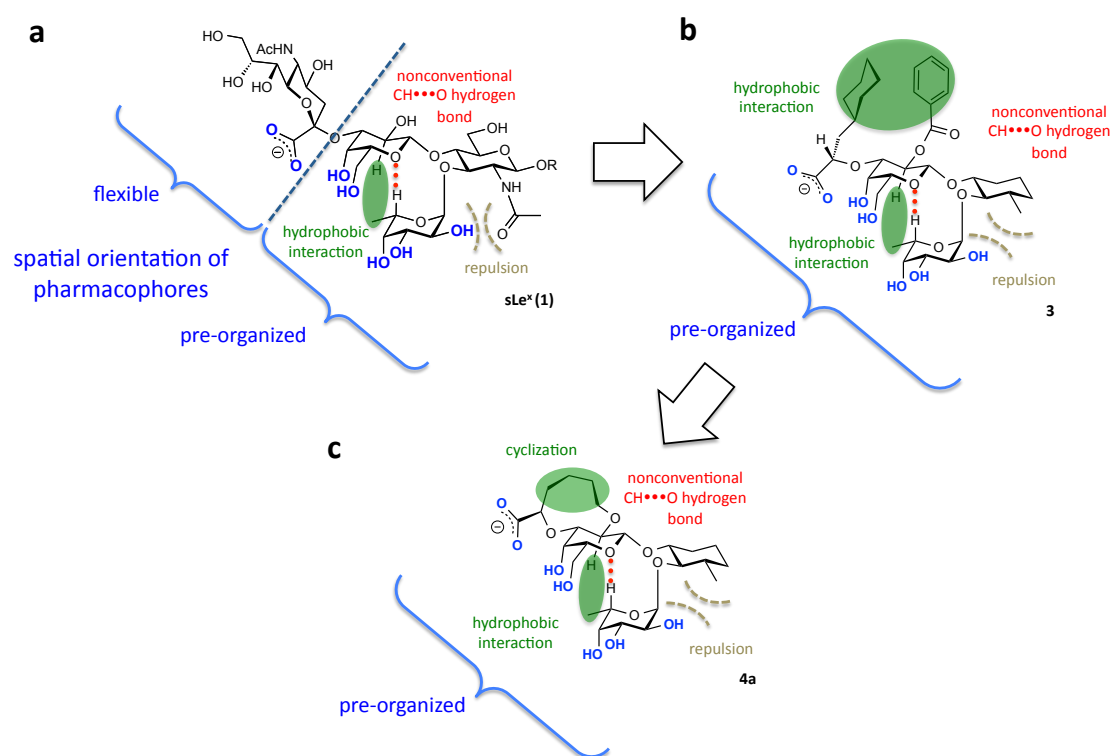


Figure 1. Stabilizing factors of the sLe^x tetrasaccharide **1** (a), mimic **3** (b) and the cyclic selectin antagonist **4a** (c) in solution. Pharmacophores are highlighted in blue (bold).

The recently published selectin antagonist **3**², mimics the conformation of the Le^x core and adopts all the stabilizing elements. Furthermore, the D-GlcNAc and D-Neu5NAc residues were replaced by hydrophobic moieties, which lead to favorable desolvation properties upon binding and an increase in affinity by a factor of almost 250 to the lead compound sLe^x.

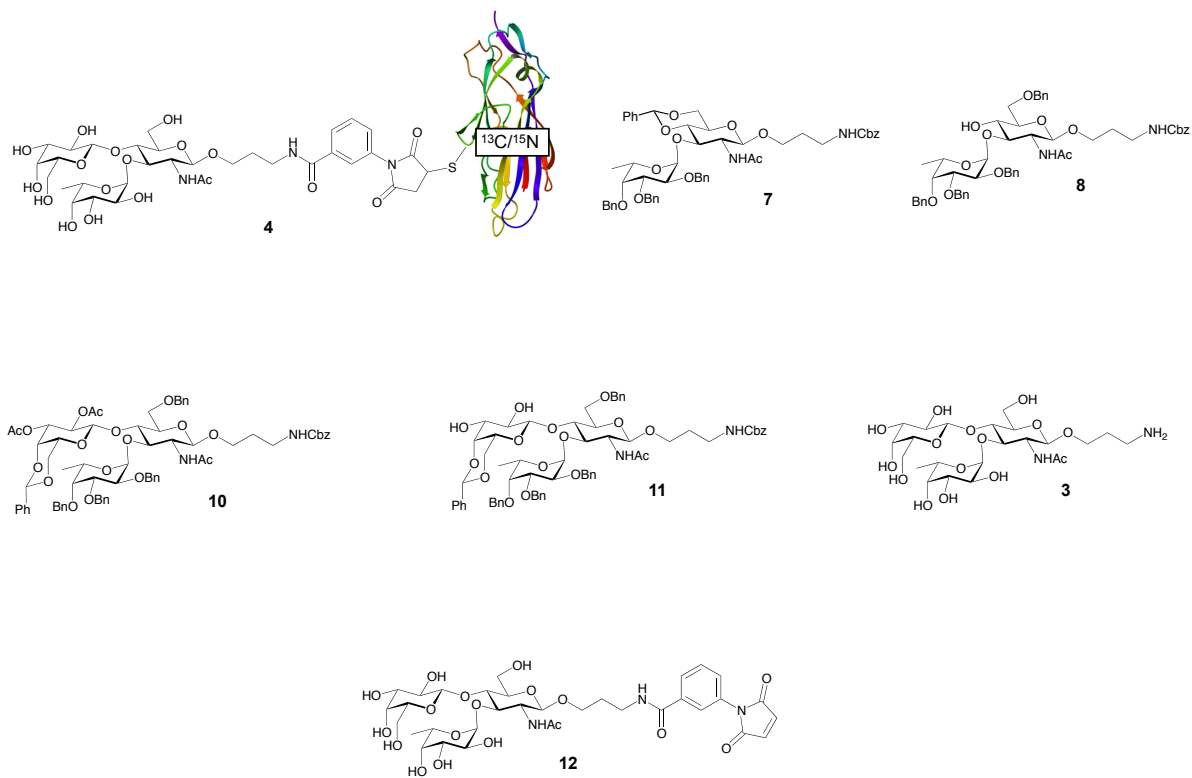
By solving the solution conformation of **3** an additional stabilizing factor was identified. An interaction between the benzoate and cyclohexyl lactic acid moiety contributes to a preferable orientation of the acid pharmacophore in the bioactive conformation (Figure 1b). Assuming, that the pre-organization of the acid functionality is a major factor for this strong binding affinity, we synthesized a series of cyclic compounds in which the acid pharmacophore is oriented in the bioactive conformation. Our most promising cyclic ligand **4a** has a similar affinity to compound **3**, while the molecular weight and number of free rotatable bonds is reduced, both important parameters for the design of oral available drugs. In the future the binding mode and the thermodynamic profile for E-selectin binding will be determined for the cyclic selectin ligands **4a-c**. To further improve the binding affinity hydrophobic substituents could be introduced in the ring structure. Furthermore, replacing hydroxy pharmacophores by bioisosteres could increase lipophilicity and reduce the polar surface area and thus, improve the druglike properties.

References

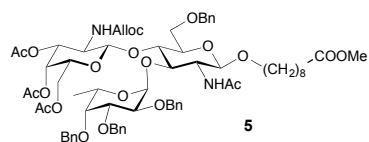
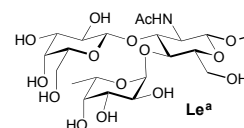
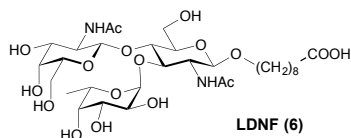
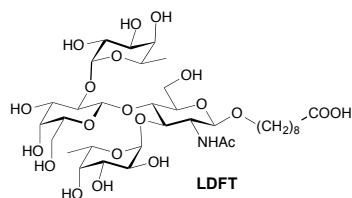
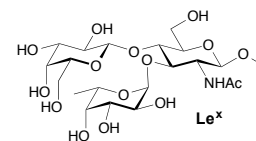
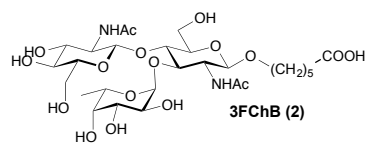
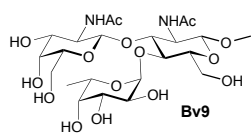
- (1) Ernst, B.; Magnani, J. L. *Nat. Rev. Drug. Discov.* **2009**, *8*, 661.
- (2) Schwizer, D.; Patton, J. T.; Cutting, B.; Smieško, M.; Wagner, B.; Kato, A.; Weckerle, C.; Binder, F. P. C.; Rabbani, S.; Schwardt, O.; Magnani, J. L.; Ernst, B. *Chem. – Eur. J.* **2012**, *18*, 1342; Kolb, H. C.; Ernst, B. *Chemi. – Eur. J.* **1997**, *3*, 1571; Thoma, G.; Magnani, J. L.; Patton, J. T.; Ernst, B.; Jahnke, W. *Angewandte Chemie International Edition* **2001**, *40*, 1941.
- (3) Imberty, A.; Mikros, E.; Koca, J.; Mollicone, R.; Oriol, R.; Pérez, S. *Glycoconj. J.* **1995**, *12*, 331.

Chapter 5. – Formula Index

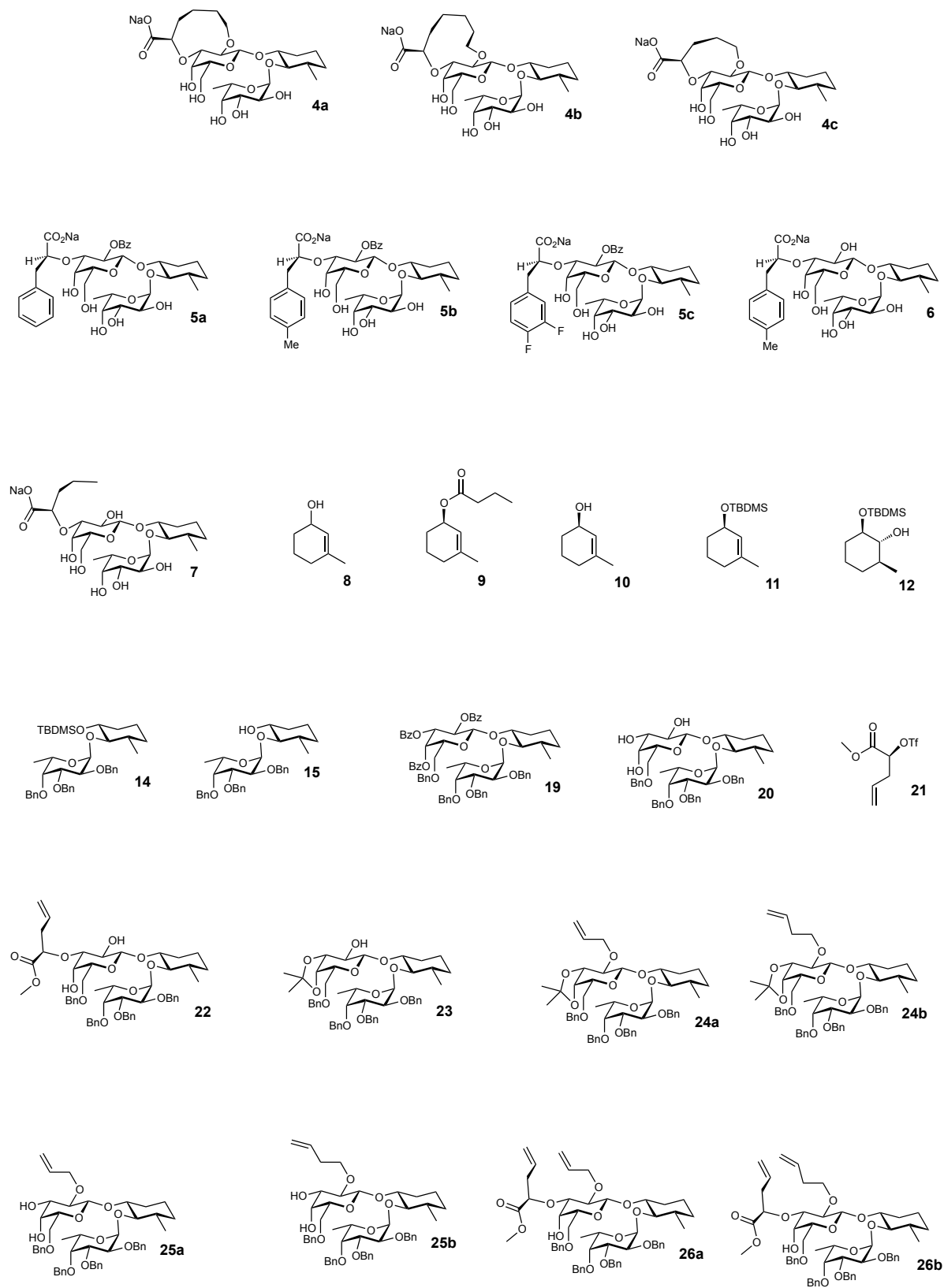
Chapter 3.1.1

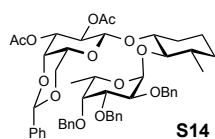
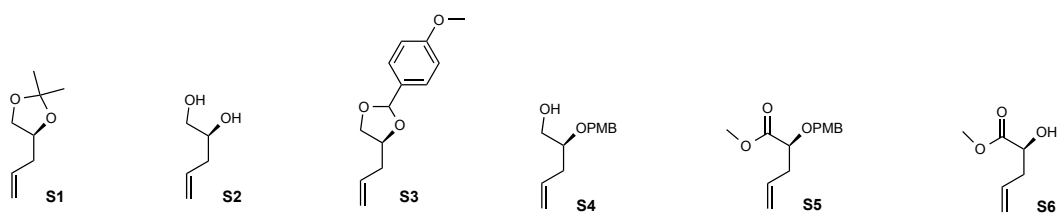
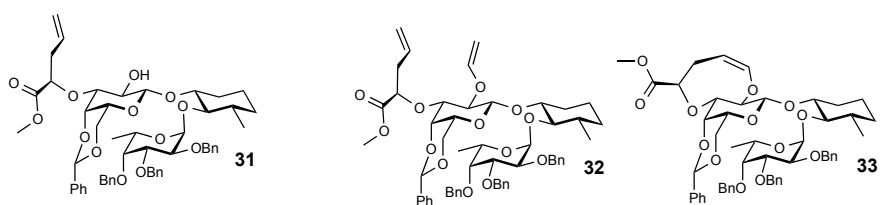
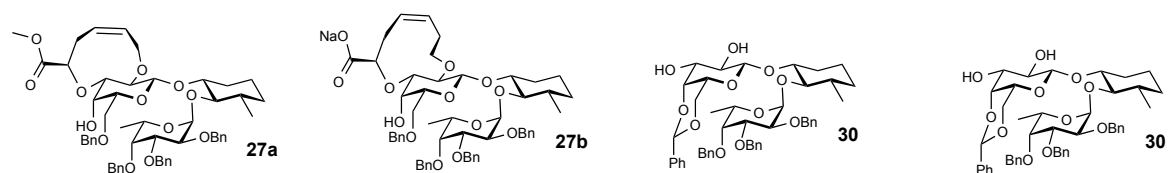
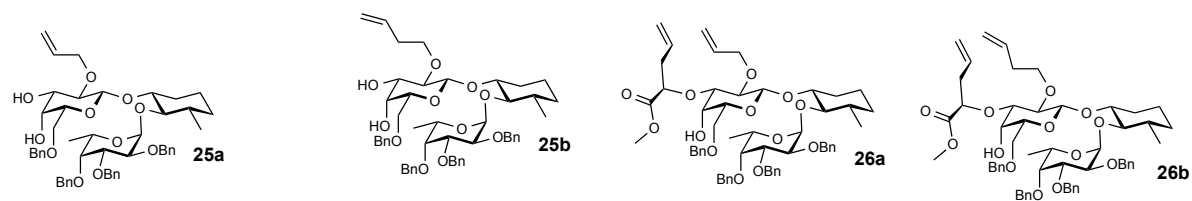


Chapter 3.1.2

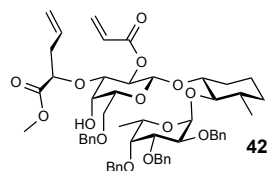
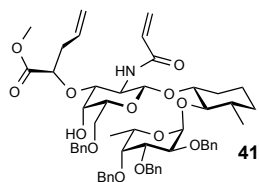
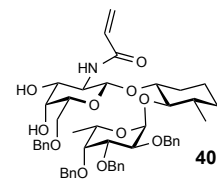
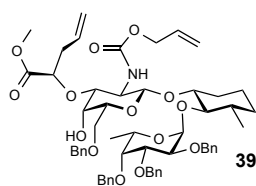
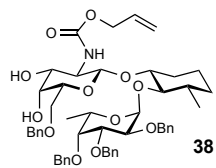
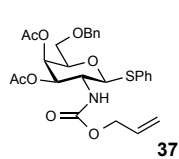


Chapter 3.2.1





Chapter 3.2.1.1



Chapter 3.2.2

