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

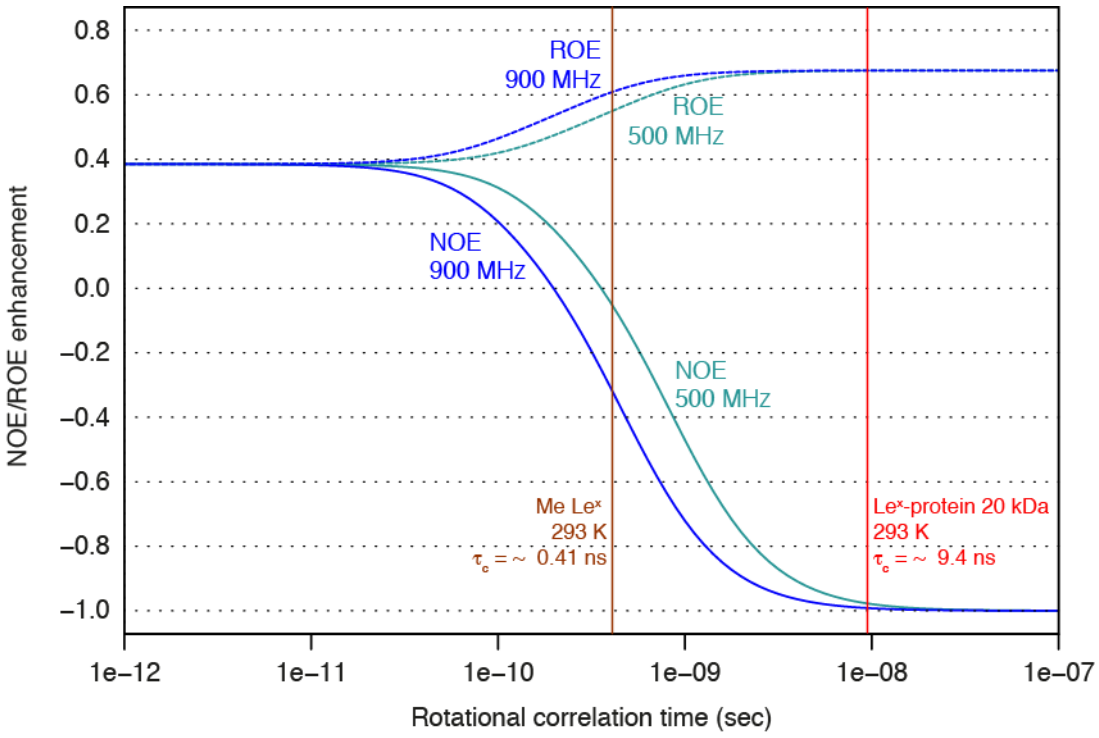


Figure S1. Maximal NOE and ROE enhancements calculated for a transient NOE experiment at two different field strengths in D_2O.[S1]

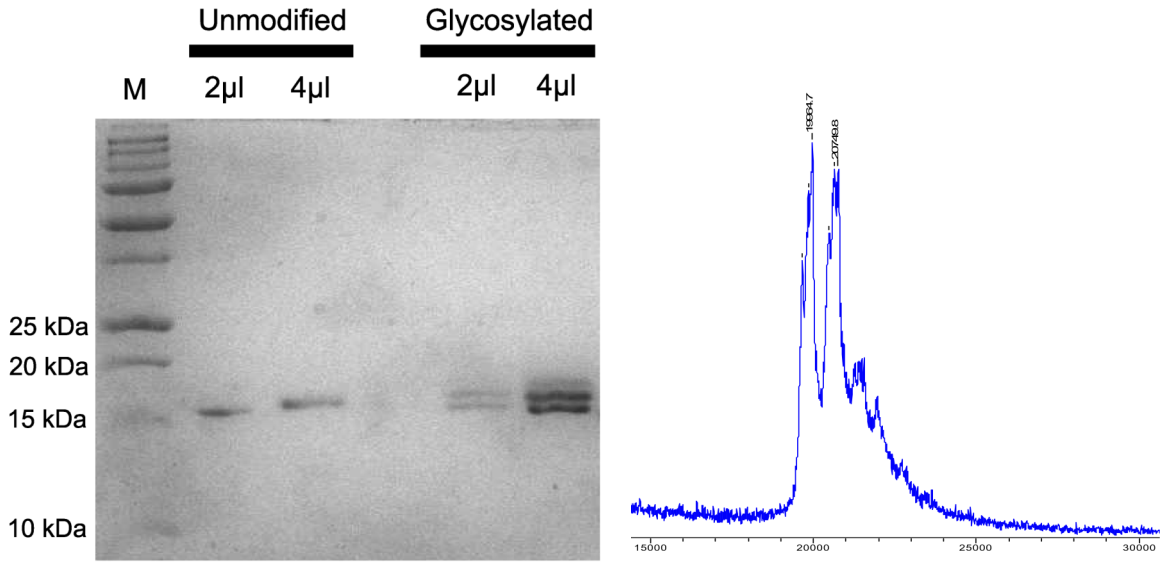


Figure S2. a.) SDS-PAGE analysis of unmodified and glycosylated uniformly $^{13}C,^{15}N$-labeled S78C-FimH. The glycosylated samples contained also unmodified protein, however this does not effect the measurements as the protein signals are suppressed in the NMR experiments. b.) MALDI-TOF MS analysis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

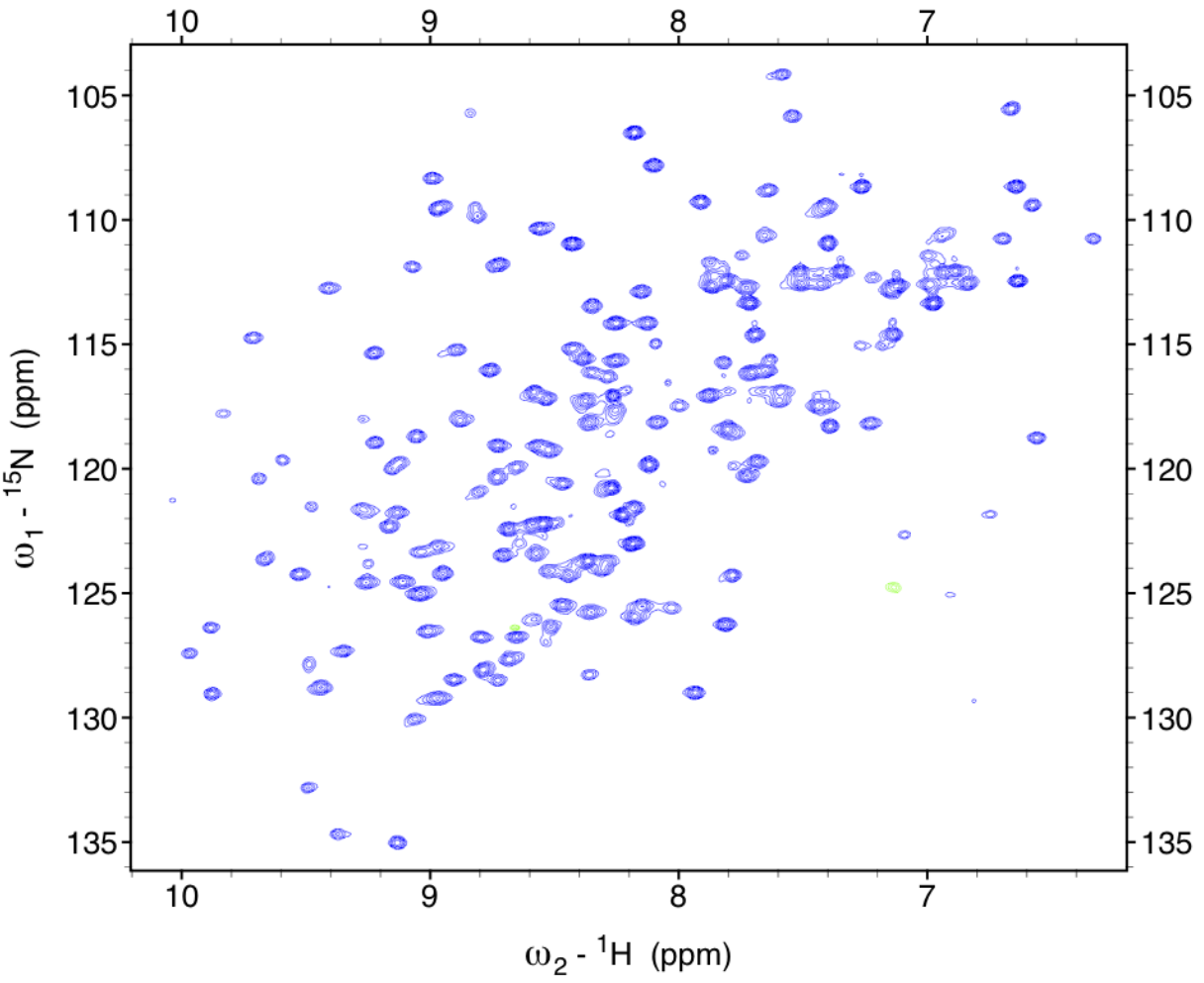


Figure S3. $^{1}H^{15}N$-HSQC spectrum of the glycosylated $^{13}C,^{15}N$-labeled S78C-FimH (0.7 mM in 93% H_2O/7% D_2O) recorded at 900 MHz and 293K with 2 scans and 256 increments.

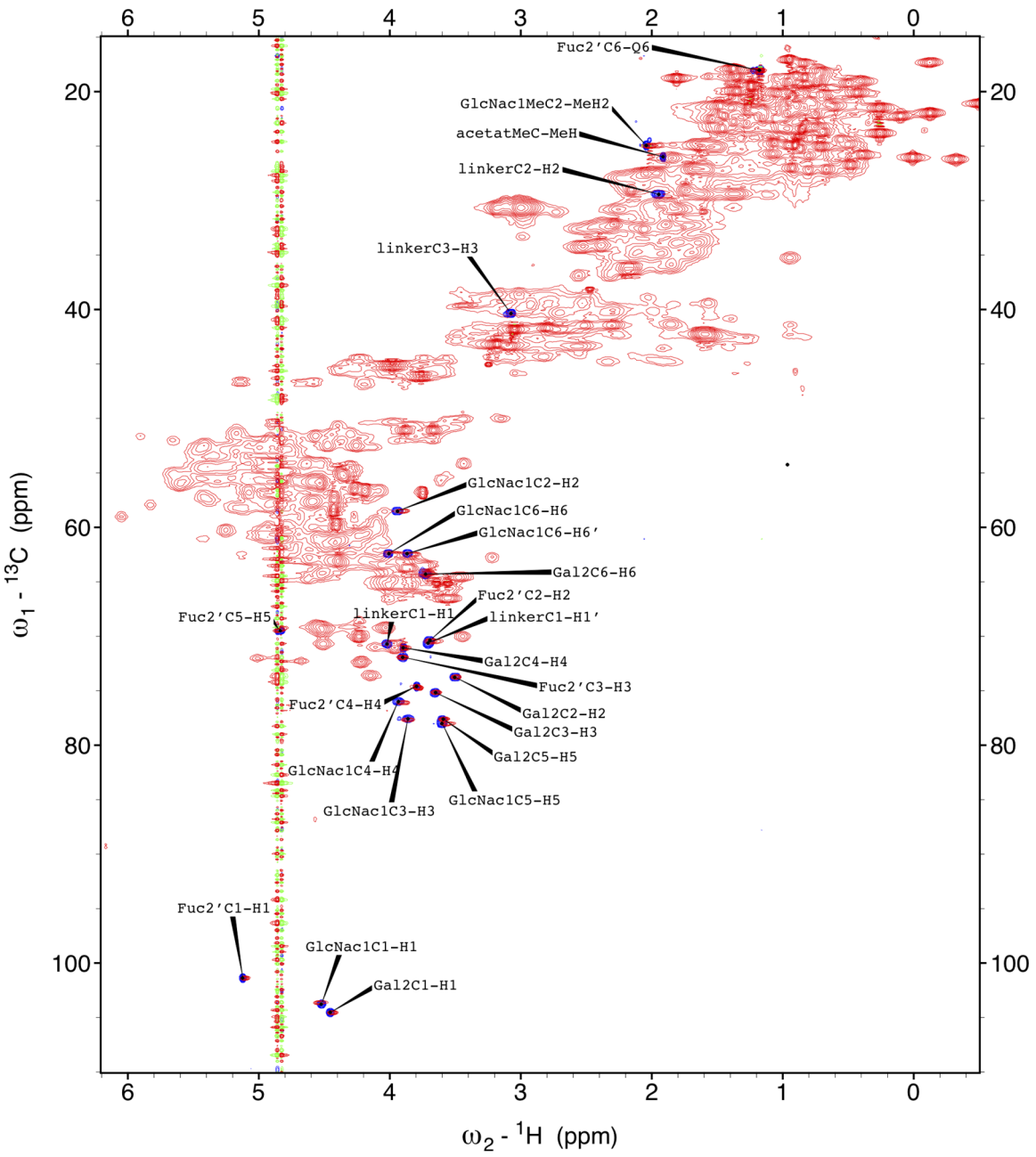


Figure S4. $^{1}H^{13}C$-HSQC spectra of the glycosylated $^{13}C,^{15}N$-labeled S78C-FimH (red) and the free Le^{x} trisaccharide (blue). Spectra were recorded at 500 MHz and 293 K with 8 scans and 512 increments or 4 scans and 400 increments, respectively. The concentrations of the glycosylated protein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

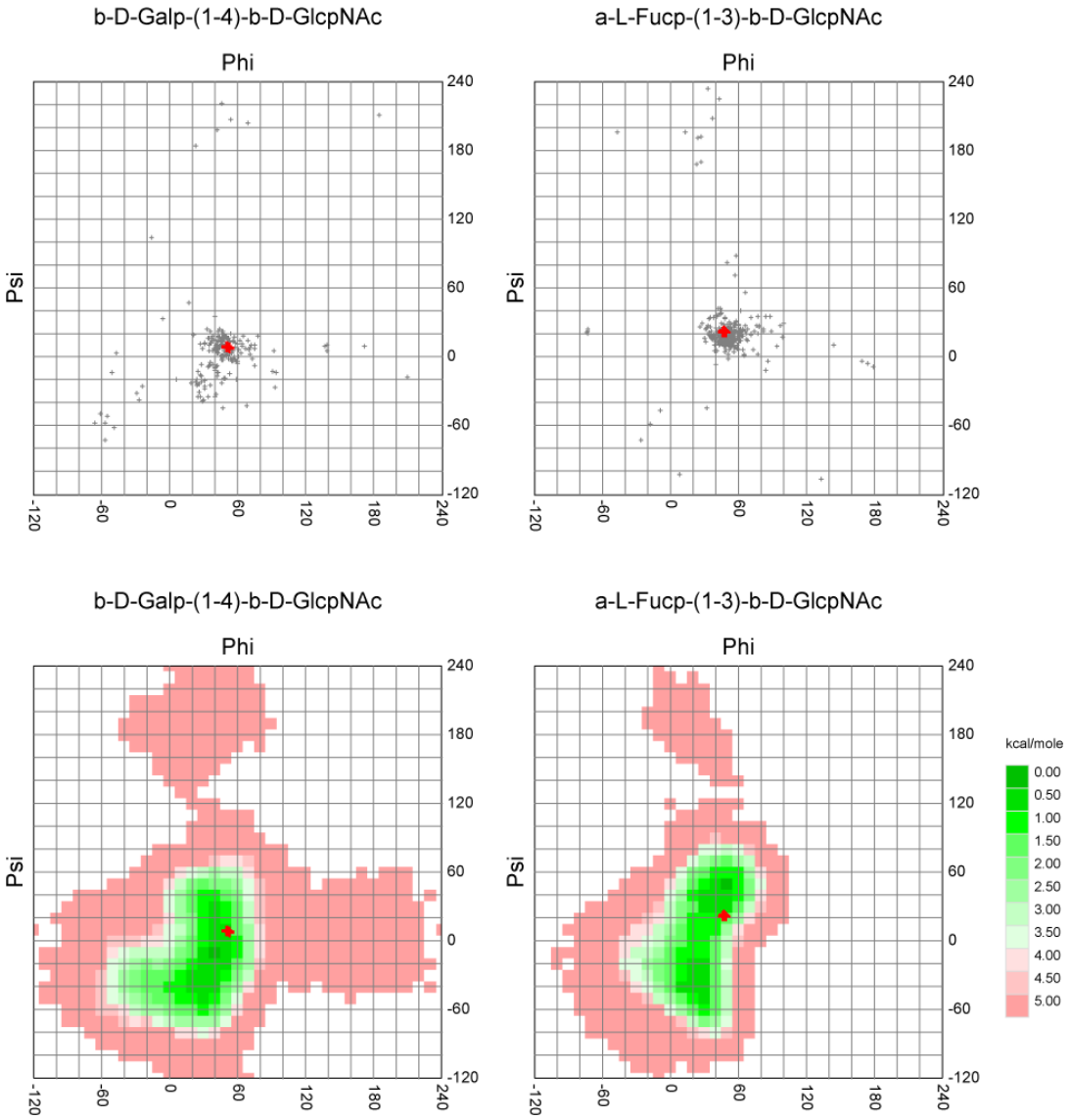


Figure S5. Phi-psi plots of the solution structure of Le^x (attached to FimH) superimposed with all entries in the PDB database containing the corresponding saccharide linkages are displayed in the upper panel whereas the bottom panel shows superimpositions on energy landscapes. Plots were generated 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.

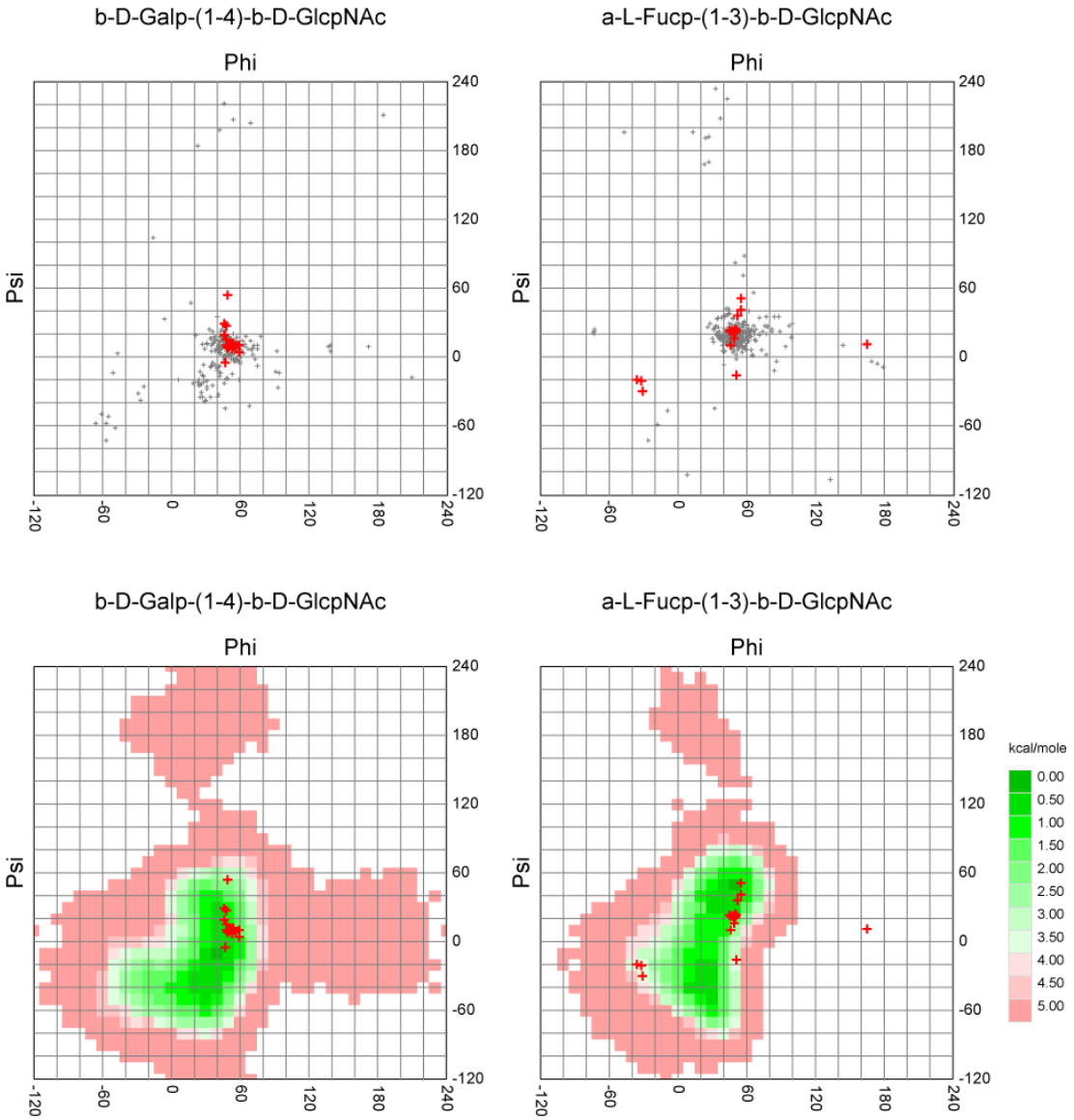


Figure S6. phi-psi plots of the solution structure of methyl Le^x (**2**) based on a 2D NOESY experiment. Note that these data were insufficient to derive a well-defined structure due to the unfavorable tumbling time. Superimpositions of phi-psi angles of all entries in the PDB database for a given linkage 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.

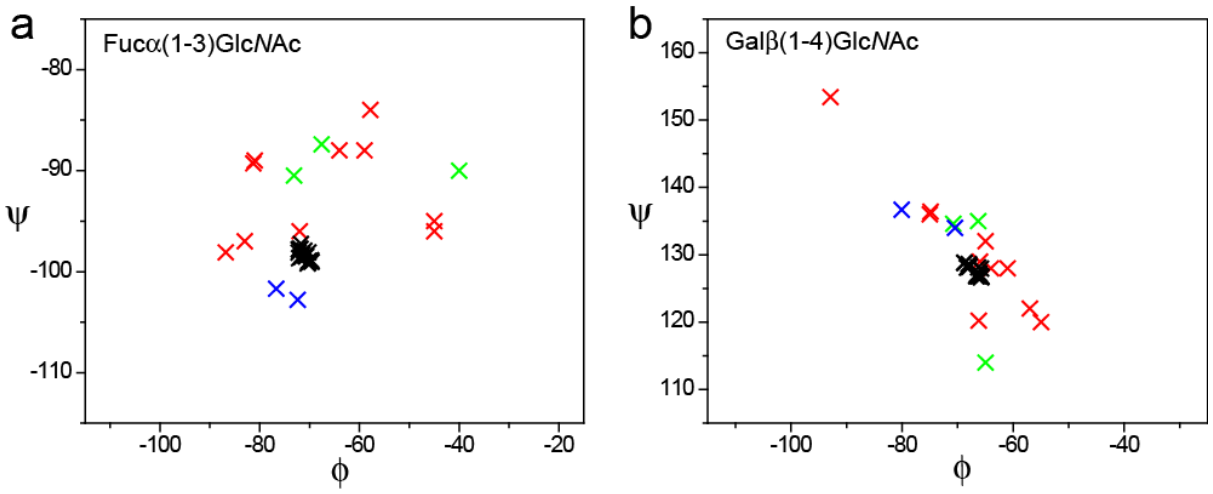
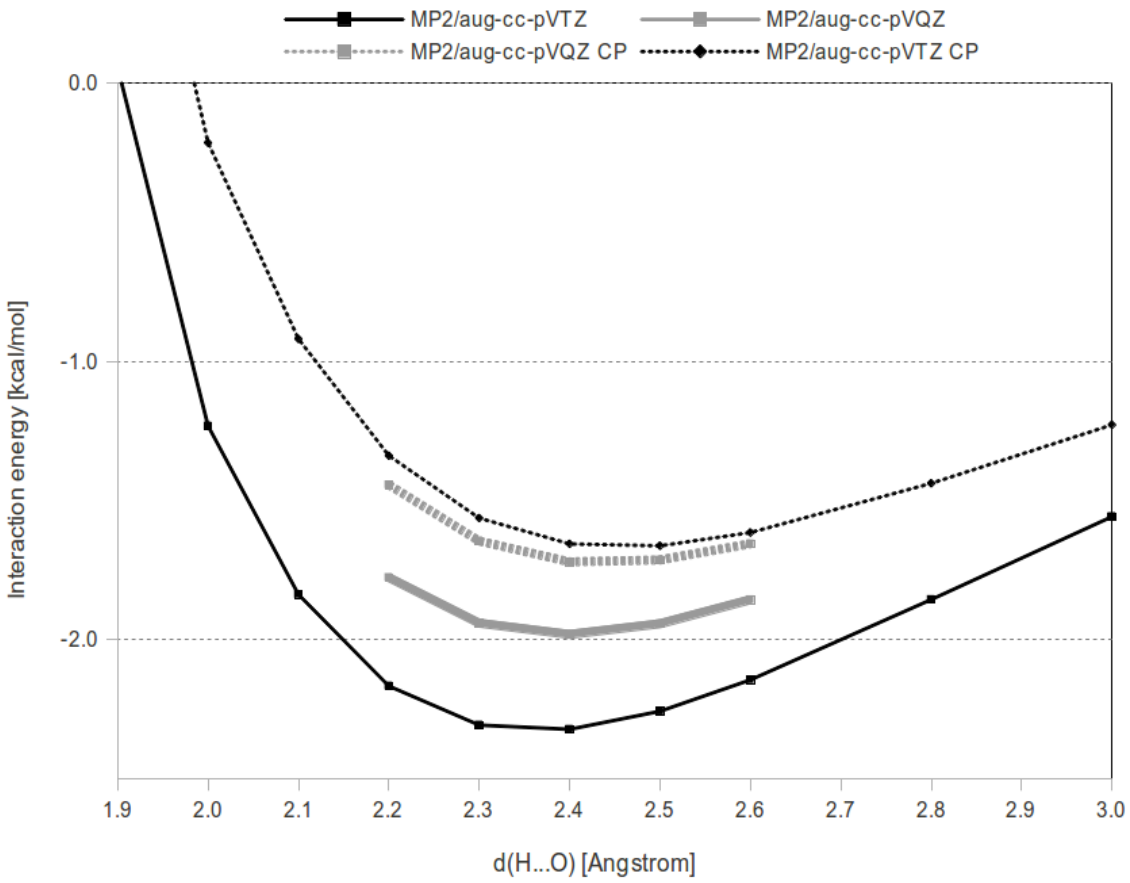


Figure S7. Comparison of glycosidic angles of the obtained NMR ensemble with previously reported Le^x structures/substructures. ϕ/ψ-angles of the Fucα(1,3)Glc*N*Ac linkage (a) and Galβ(1,4)Glc*N*Ac linkage (b). Torsion angles of the here presented NMR ensemble are shown in black, angles of structures 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}.



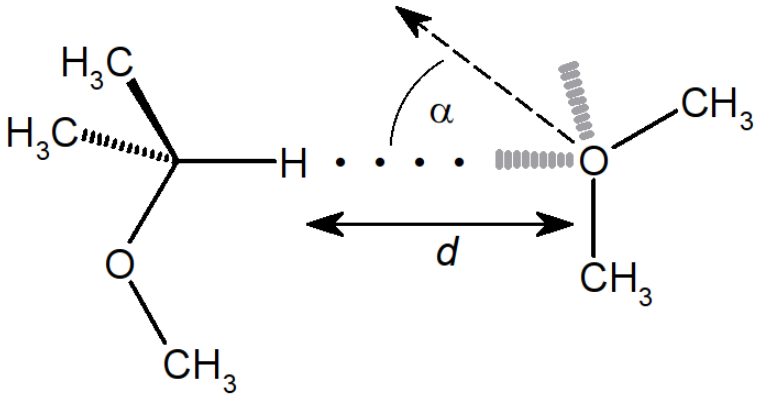


Figure S8. Schematic representation of the model system for quantification of the non-conventional C-H···O hydrogen bond between H5 of L-fucose (resembling an isopropyl methyl ether) and O5 of D-galactose (resembling a dimethyl ether) of Le^x and corresponding energy potential curves as a function of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

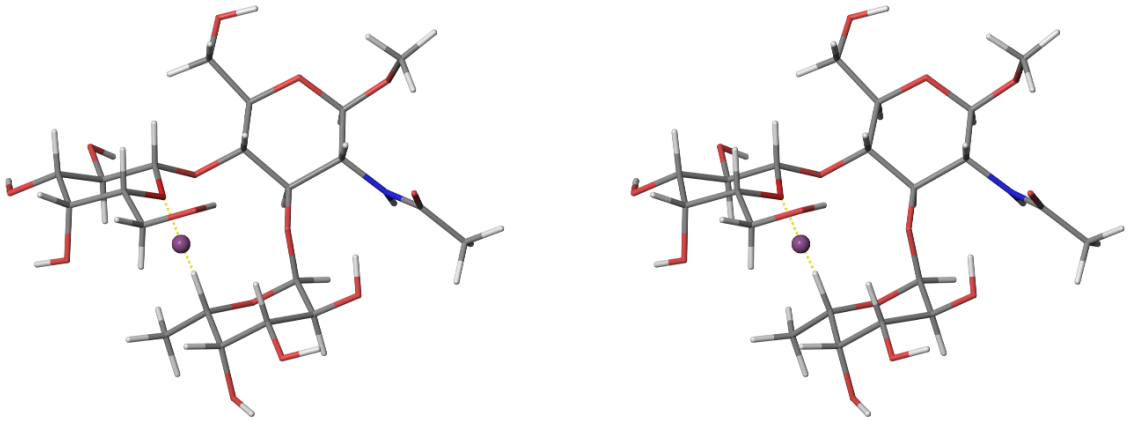


Figure S9. Stereo view of the bond critical point (signature: 3,-1; density ρ_b = 0.01327 au; displayed as a blue sphere) identified for the C-H···O hydrogen bond between C5-H5 of L-Fuc and O5 of D-Gal of Me Le^x (**2**) based on the B3LYP/6-31G(d,p) wavefunction.

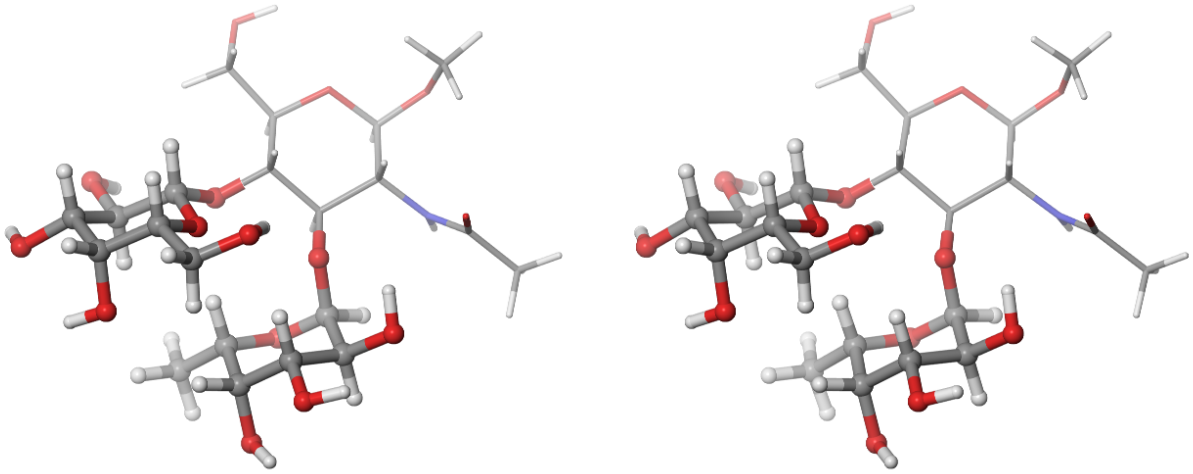


Figure S10. Stereo view of methyl Le^x (**2**) optimized using the ONIOM method.[S3] Atoms displayed as balls and sticks were assigned to the high layer and atoms displayed as thin sticks were assigned in the low layer.

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	Fucα1,3Glc*N*Ac torsion angles [a,b]				Galβ1,4Glc*N*Ac 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^x)	-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 ^1H-^1H distance[b] [Å]	average S/N of NOEs cross peaks	Corresponding ^1H-^1H 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		
Glc*N*Ac H1-H3	207	3.0[f]		
Glc*N*Ac H1-H5	474	2.6	290[a]	2.4
Glc*N*Ac H2-Q8	75[a]	4.2[e]		
Glc*N*Ac H5-H62	435	2.6	192[a]	2.6
Glc*N*Ac H61-H62[c]	1004[a]	1.77	3209[a]	1.77
Glc*N*Ac 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]		
Glc*N*Ac H1-HN2	60[a]	3.6		
Glc*N*Ac H3-HN2	72[a]	3.5		
Glc*N*Ac Q8-HN2	115[a]	3.9[e]		
inter				
Gal H1 - Glc*N*Ac H4	910	2.3	206	2.6
Gal H1 - Glc*N*Ac H62	438	2.6	206[a]	2.6
Gal H1 - Glc*N*Ac 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]		
Glc*N*Ac H3 - Fuc H1	286[a]	2.8		
Glc*N*Ac Q8 - Fuc H1	142	3.8[e]		
Glc*N*Ac 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 Glc*N*Ac 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 ^{15}N-filtered-filtered NOESY.
d) Reference restraints for the ^{13}C-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)Glc*N*Ac	–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)Glc*N*Ac	–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 O_5-C_1-O_x-C'_x and psi as C_1-O_x-C'_x-C'_{x-1}

† extracted by XtalView [S11]

¶ NMR dihedral angles extracted automatically by CARP [S2], phi is defined as H_1-C_1-O_1-C'_X and psi as C1-O1-C'_X-H'_X

Table S4. ^{1}H-^{1}H 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 ^{1}H-^{1}H distances[a] [Å]	Applied upper distances restraint[b] [Å]	^{1}H-^{1}H 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
Glc*N*Ac H1-H3	3.0	4.0[e]	2.60
Glc*N*Ac H1-H5	2.6	3.1	2.52
Glc*N*Ac H2-Q8	4.2	4.7	3.80[f]
Glc*N*Ac H5-H62	2.6	3.1	3.06
Glc*N*Ac H61-H62[c]	1.77	–	1.77
Glc*N*Ac 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]
Glc*N*Ac H1-HN2	3.6	4.1	2.64
Glc*N*Ac H3-HN2	3.5	4.0	2.59
Glc*N*Ac Q8-HN2	3.9	4.4	2.12[f]
inter			
Gal H1 - Glc*N*Ac H4	2.3	2.8	2.45
Gal H1 - Glc*N*Ac H62	2.6	3.1	2.74
Gal H1 - Glc*N*Ac 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]
Glc*N*Ac H3 - Fuc H1	2.8	3.3	2.52
Glc*N*Ac Q8 - Fuc H1	3.8	4.3	3.01[f]
Glc*N*Ac 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 ^{1}H-^{1}H 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 Fucα(1-3)Glc*N*Acβ methyl glycoside and Galβ(1-4)Glc*N*Acβ 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]	Fucα(1-3) Glc*N*Ac (exp)[d]	Galβ(1-4) Glc*N*Ac (exp)[d]	Fucα(1-3) Glc*N*Ac (theo)	Galβ(1-4) Glc*N*Ac (theo)
Glc*N*Ac H1	4.52	4.18	4.31	4.31	4.47	4.51	4.24	4.24
Glc*N*Ac H2	3.94	3.92	3.98	3.99	3.82	3.76	3.94	3.83
Glc*N*Ac H3	3.86	3.67	3.55	3.55	3.66	3.72	3.82	3.58
Glc*N*Ac H4	3.94	3.36	3.42	3.43	3.51	3.72	3.33	3.10
Glc*N*Ac H5	3.60	3.53	3.39	3.38	3.50	3.60	3.59	3.54
Glc*N*Ac H61	3.87	3.80	3.69	3.69	3.77	3.84	3.80	3.80
Glc*N*Ac H62	4.01	4.29	4.10	4.09	3.95	4.00	4.28	4.10
Glc*N*Ac 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 D_2O.

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 Fucα1,3Glc*N*Acβ methyl glycoside and Galβ(1-4)Glc*N*Acβ propanolamine aglycone in D_2O at 293 K.

Table S6. Structural parameters of the C-H···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···O) [Å]	2.50 ± 0.01	2.464	2.333	2.220	2.290, 2.311
r(C···O) [Å]	3.56 ± 0.01	3.483	3.424	3.277	3.269, 3.304
a(C-H···O) [°]	165.0 ± 0.9	156.0	175.2	163.7	163.1, 176.4

Table S7. Stacking and C-H···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** ··· 1-deoxy-Gal	-4.83	-2.62	-7.81 (-3.49)	-6.58 (-4.52)	n/a
Me-O-**iPro** ··· Me-O-Me	-1.23	-1.34	-2.86 (-1.23)	-2.52 (-1.76)	-2.14 (-1.84)
1-deoxy-**Ara** ··· 1-deoxy-Gal	-4.48	-2.75	-6.76 (-3.16)	-5.83 (-4.03)	n/a
Me-O-**Et** ··· 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
E.coli		
BL21(DE3)	F^- *ompT hsdS*$_B$(r_B^- m_B^-) *gal dcm* (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
Oligonucleotide Primers (5'→3')[a]		
FimH-Extern Forward	CC **TCT AGA** ATG ATT GTA ATG AAA CGA GTT ATT ACC CTG	This study
FimH-Extern Revers	CC **AAG CTT** TCG GGC TTT GTT AGC AGC CGG ATC TCA GTG	This study
FimH-S78C Forward	ACC GTA AAA ***TAT*** TGT GGC AGT AGC TAT	This study
FimH-S78C Revers	ATA GCT ACT GCC ***ACA*** 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.

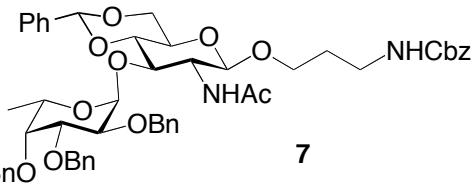
Supplementary Methods

General methods. Commercial materials (Sigma-Aldrich) were used without further purification, solvents were reagent grade (Acros). 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 Glc*N*Ac (*N*-acetyl glucosamine). The coupling constants (*J*) are reported in Hertz (Hz). Analytical TLC was performed on Merck silica gel 60 F_{254} 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% H_2SO_4) 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)

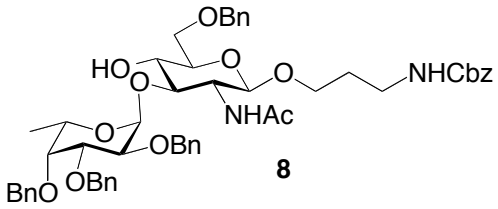


7

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, 3× 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}H NMR (500.1 MHz, $CDCl_3$) δ 7.50 – 7.23 (m, 25H, Ar-H), 5.96 (d, *J* = 6.8 Hz, 1H, N*H*Ac), 5.49 (s, 1H, benzylidene C*H*), 5.15 – 5.03 (m, 3H, Fuc-H1, 2xCH_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, Glc*N*Ac-H1, 3× CH_2-Ph), 4.58 (d, *J* = 11.4 Hz, 1H, CH_2-Ph), 4.31 (dd, *J* = 10.4, 4.9 Hz, 1H, Glc*N*Ac-H6), 4.17 – 4.03 (m, 3H, Glc*N*Ac-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, Glc*N*Ac-H6'), 3.64 – 3.54 (m, 3H, Fuc-H4, Glc*N*Ac-H2, -H4), 3.52 – 3.39 (m, 2H, O-CH_2, Glc*N*Ac-H5), 3.39 – 3.25 (m, 1H, CH_2-NH), 3.23 – 3.12 (m, 1H, CH_2NH), 1.81 – 1.64 (m, 5H, 2xCH$_2$-CH_2-CH$_2$, CO-CH_3), 0.86 (d, *J* = 6.4 Hz, 3H, Fuc-H6) ppm. ^{13}C NMR (125.8 MHz, $CDCl_3$) δ 171.2 (*C*O-CH$_3$), 156.7 (O*C*O-NH), 138.8 – 126.3 (25C, Ar-C), 101.6 (benzylidene *C*H), 101.58 (Glc*N*Ac-C1), 98.51 (Fuc-C1), 80.9 (Fuc-C4), 79.9 (Fuc-C3), 77.7 (Glc*N*Ac-C4), 77.1 (Fuc-C2), 75.7 (Glc*N*Ac-C3), 75.1 (*C*H$_2$-Ph), 74.2 (*C*H$_2$-Ph), 72.8 (*C*H$_2$-Ph), 68.9 (Glc*N*Ac-C6), 67.5 (Fuc-C5), 67.1 (O-*C*H$_2$), 66.7 (Glc*N*Ac-C5), 66.5 (*C*H$_2$-Ph), 57.6 (Glc*N*Ac-C2), 38.0 (*C*H$_2$-NH), 29.6 (CH$_2$-*C*H$_2$-CH$_2$), 23.2 (CO-*C*H$_3$), 16.5 (Fuc-C6) ppm. ESI-MS Calcd for $C_{53}H_{60}N_2O_{12}$ [M+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)

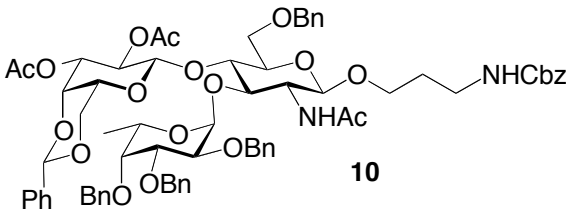


8

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, $CHCl_3$); ^{1}H 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, Glc*N*Ac-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, Glc*N*Ac-H6), 3.70 – 3.53 (m, 4H, Fuc-H4, Glc*N*Ac-H6', Glc*N*Ac-H2, O-CH_2), 3.53 – 3.43 (m, 2H, Glc*N*Ac-H3, Glc*N*Ac-H5), 3.43 – 3.34 (m, 2H, Glc*N*Ac-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 (O*C*O-NH_2), 138.7 – 127.7 (25C, Ar-C), 100.8 (Glc*N*Ac-C1), 99.7 (Fuc-C1), 85.1 (Glc*N*Ac-C3), 79.2 (Fuc-C3), 77.4 (Fuc-C4), 76.3 (Fuc-C2), 75.2 (*C*H_2-Ph), 75.1 (Glc*N*Ac-C5), 74.3 (*C*H_2-Ph), 73.5 (*C*H_2-Ph), 73.2 (*C*H_2-Ph), 70.71 (Glc*N*Ac-C4), 69.7 (Glc*N*Ac-C6), 68.3 (Fuc-C5), 67.0 (O-*C*H_2), 66.7 (*C*H_2-Ph), 55.7 (Glc*N*Ac-C2), 37.8 (*C*H_2-NH), 29.6 (CH_2-*C*H_2-CH_2), 23.0 (CO-*C*H_3), 16.8 (Fuc-C6) ppm. ESI-MS Calcd for $C_{53}H_{62}NaN_2O_{12}$ [M+Na]+: 941.42; Found: 941.47.

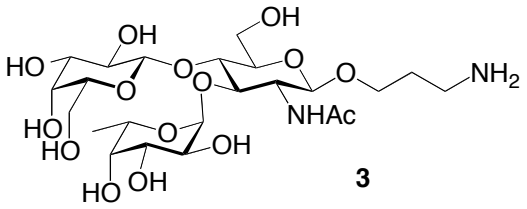
(3-*N*-Benzyloxycarbonylamino)propyl (2,3-di-*O*-acetyl-4,6-*O*-benzylidene-β-D-galactopyranosyl)-(1-4)-[(2,3,4-tri-*O*-benzyl-6-deoxy-α-L-galactopyranosyl)-(1-3)]-2-acetamido-6-*O*-benzyl-2-deoxy-β-D-glucopyranoside (10)



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 $NaHCO_3$ (50 ml) and water (50 ml). The aqueous phases were extracted with DCM (3× 30 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 (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; $[\alpha]_D^{22}$ -24.6 (*c* 0.82, $CHCl_3$); $[\alpha]_D^{22}$ -24.6 (*c* 0.82, $CHCl_3$); ^{1}H NMR (500.1 MHz, $CDCl_3$) δ 7.58 – 7.13 (m, 30H, Ar-H), 5.74 (d, *J* = 6.4 Hz, 1H, N*H*Ac), 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-C*H*$_2$, Glc*N*Ac-H1), 4.94 (d, *J* = 2.7 Hz, 1H, Fuc-H1), 4.84 (d, *J* = 11.8 Hz, 1H, Ph-C*H*$_2$), 4.76 – 4.67 (m, 3H, Ph-C*H*$_2$, Gal-H3, Fuc-H5), 4.67 – 4.59 (m, 3H, Ph-C*H*$_2$, Gal-H6), 4.56 (d, *J* = 8.2 Hz, 1H, Gal-H1), 4.38 (d, *J* = 12.1 Hz, 1H, Ph-C*H*$_2$), 4.33 – 4.17 (m, 3H, 2x Ph-C*H*$_2$, Gal-H4), 4.12 (t, *J* = 9.4 Hz, 1H, Glc*N*Ac-H3), 4.03 – 3.93 (m, 3H, Ph-C*H*$_2$, Glc*N*Ac-H4, Fuc-H2), 3.92 – 3.80 (m, 3H, Fuc-H3, O-C*H*$_2$, Glc*N*Ac-H6), 3.79 – 3.71 (m, 1H, Glc*N*Ac-H6'), 3.56 (d, *J* = 11.2 Hz, 1H, Ph-C*H*$_2$), 3.53 – 3.47 (m, 1H, O-C*H*$_2$), 3.43 – 3.36 (m, 1H, Glc*N*Ac-H5), 3.29 – 3.17 (m, 4H, 2x CH_2-NH_2, Glc*N*Ac-H_2, Fuc-H4), 3.04 – 2.99 (m, 1H, Gal-H5), 2.12 (s, 3H, CO-C*H*$_3$), 2.02 (s, 3H, CO-C*H*$_3$), 1.76 – 1.67 (m, 2H, CH_2-C*H*$_2$-CH_2), 1.59 (s, 3H, CO-C*H*$_3$), 1.12 (d, *J* = 6.5 Hz, 3H, Fuc-H6). ^{13}C NMR (125.8 MHz, $CDCl_3$) δ 170.8 (*C*O-CH_3), 169.9 (*C*O-CH_3), 168.9 (*C*O-CH_3), 156.7 (O*C*O-NH), 139.6-125.9 (36C, Ar-C), 99.9 (benzylidene-CH), 99.8 (Gal-C1), 99.6 (Glc*N*Ac-C1), 98.2 (Fuc-C1), 79.6 (Fuc-C3), 78.9 (Fuc-C4), 76.0 (Fuc-C2), 75.1 (Ph-*C*H_2), 75.0 (Glc*N*Ac-C5), 74.4 (Glc*N*Ac-C4), 74.3 (Ph-*C*H_2), 73.8 (Glc*N*Ac-C3), 73.5 (Ph-*C*H_2), 73.4 (Gal-C4), 72.2 (Gal-C3), 71.5 (Gal-C6), 69.2 (Ph-*C*H_2), 68.8 (Gal-C2), 68.0 (Glc*N*Ac-C6), 67.0 (O-*C*H_2), 66.6 (Ph-CH_2), 66.5 (Fuc-C5), 66.4 (Gal-C5), 59.4 (Glc*N*Ac-C2), 37.9 (*C*H_2-NH_2), 29.4 (CH_2-*C*H_2-CH_2), 23.3 (NH-CO-*C*H_3), 21.0 (CO-*C*H_3), 20.9 (CO-*C*H_3), 16.2 (Fuc-C6) ppm. ESI-MS Calcd for $C_{70}H_{80}N_2NaO_{19}$ [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)

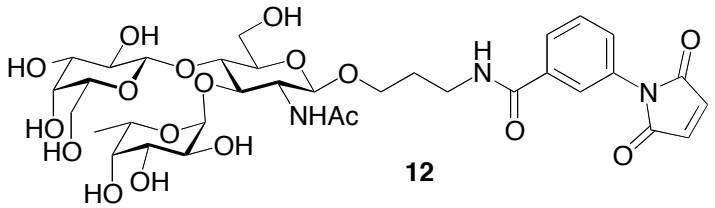


3

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/H_2O (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).

$[\alpha]_D^{22}$ -68.6 (*c* 0.96, D_2O); ^{1}H NMR (500.1 MHz, D_2O) 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, Glc*N*Ac-H1), 4.44 (d, *J* = 7.8 Hz, 1H, Gal-H1), 4.05 – 3.97 (m, 2H, O-CH_2, Glc*N*Ac-H6), 3.96 – 3.81 (m, 7H, Glc*N*Ac-H2, -H4, Fuc-H3, Gal-H4, Glc*N*Ac-H3, -H6'), 3.80 – 3.77 (m, 1H, Fuc-H4), 3.76 – 3.66 (m, 4H, Gal-H6, -H6', O-CH_2, Fuc-H2), 3.64 (dd, *J* = 9.8, 3.2 Hz, 1H, Gal-H3), 3.62 – 3.56 (m, 2H, Glc*N*Ac-H5, Gal-H5), 3.51 – 3.46 (m, 1H, Gal-H2), 3.10 – 3.04 (m, 2H, CH_2-NH_2), 2.03 (s, 3H, CO-CH_3), 1.98 – 1.90 (m, 2H, CH_2-CH_2-CH_2), 1.17 (d, *J* = 6.6 Hz, 3H, Fuc-H6) ppm; ^{13}C NMR (125.8 MHz, $CDCl_3$) δ 177.1 (*C*O-Ac), 104.5 (Gal-C1), 103.7 (Glc*N*Ac-C1), 101.3 (Fuc-C1), 78.0 (Glc*N*Ac-C5), 77.6 (Glc*N*Ac-C3), 77.5 (Gal-C5), 76.0 (Glc*N*Ac-C4), 75.2 (Gal-C3), 74.6 (Fuc-C4), 73.7 (Gal-C2), 71.9 (Fuc-C3), 71.0 (Gal-C4), 70.7 (O-*C*H_2), 70.4 (Fuc-C2), 69.4 (Fuc-C5), 64.2 (Gal-C6), 62.4 (Glc*N*Ac-C6), 58.5 (Glc*N*Ac-C2), 40.3 (*C*H_2-NH_2), 29.4 (CH_2-*C*H_2-CH_2), 24.9 (CO-*C*H_3/Acetate), 18.0 (CO-*C*H_3), 18.0 (Fuc-C6) ppm; HR-MS (ESI) Calcd for $C_{23}H_{42}N_2NaO_{15}$ [M+Na]+: 609.2483; Found: 609.2484.

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



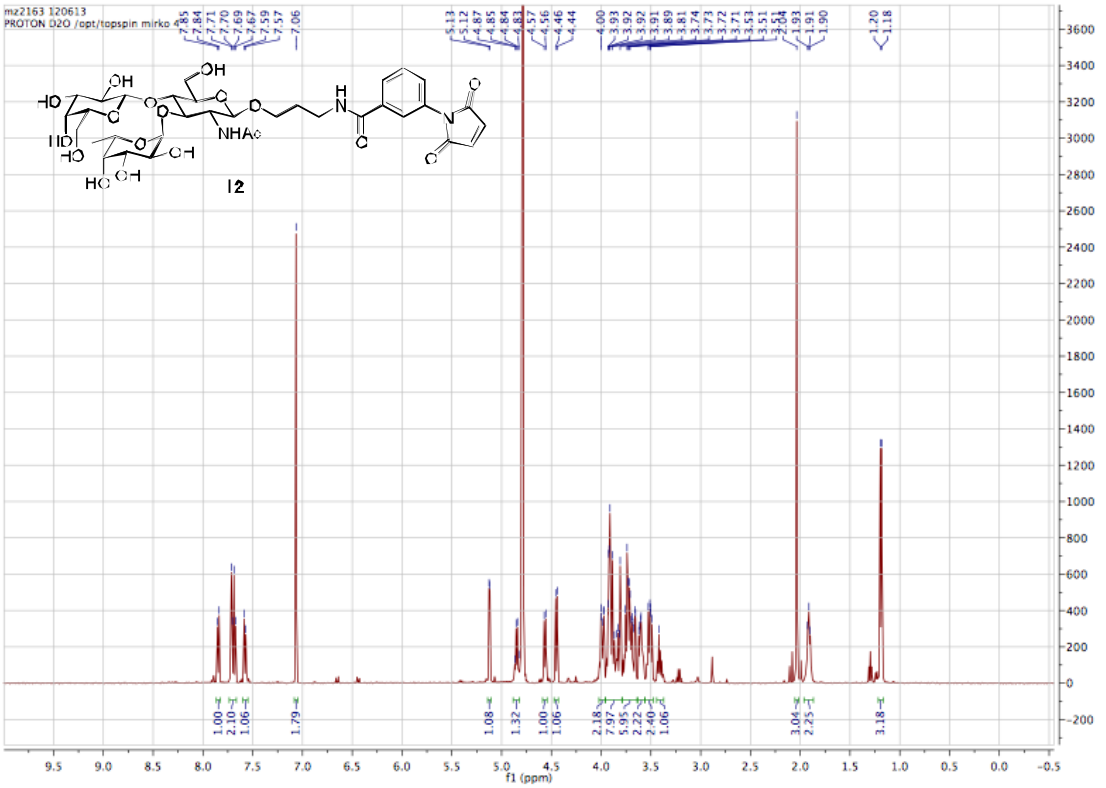
12

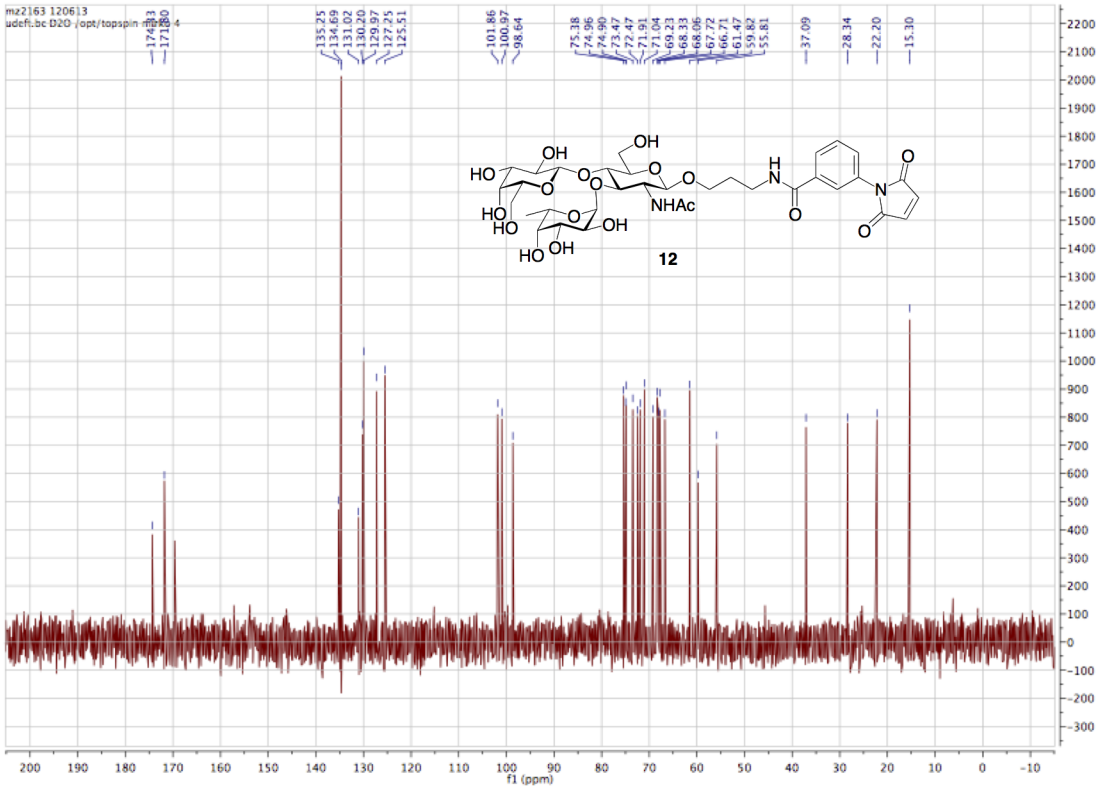
Compound **3** (11.0 mg, 18.8 µmol) was dissolved in H_2O (400 µl) and a solution of

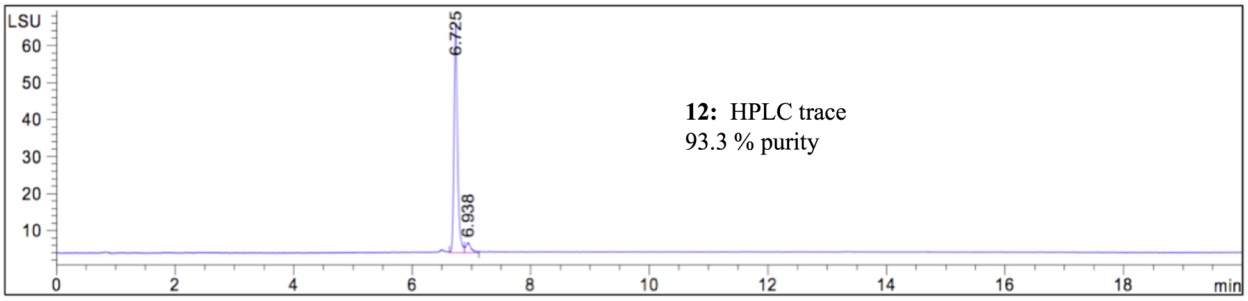
3-maleimidobenzoic acid *N*-hydroxysuccinimide ester (11.8 mg, 37.5 µmol) in DMSO (1 ml) was added. The solution was stirred for 1.5 h at r.t. The water was removed in vacuo and the crude product purified by flash chromatography (MeCN/MeOH + 0.01% TFA, 1 to 3/4) and lyophilized to give **12** as a white 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); ^{1}H 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, -C*H*=C*H*-), 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, Glc*N*Ac-H1), 4.45 (d, *J* = 7.8 Hz, 1H, Gal-H1), 4.03 – 3.96 (m, 2H, O-CH_2, Glc*N*Ac-H6), 3.96 – 3.64 (m, 14H, Glc*N*Ac-H2, Gal-H4, Fuc-H3, Glc*N*Ac-H4, -H3, -H6', Fuc-H4, Gal-H6, -H6', O-CH_2, Fuc-H2, Gal-H3), 3.64 – 3.55 (m, 2H, Glc*N*Ac-H5, Gal-H5), 3.55 – 3.47 (m, 2H, Gal-H_2, CH_2-NH), 3.45 – 3.37 (m, 1H, CH_2-NH), 2.04 (s, 3H, CO-CH_3), 1.95 – 1.88 (m, 2H, CH_2-CH_2-CH_2), 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, *C*O), 134.7 (2C, CO-*C*H), 135.3, 131.0, 130.2, 130.0, 127.3, 125.5 (6C, Ar-C), 101.9 (Gal-C1), 101.0 (Glc*N*Ac-C1), 98.6 (Fuc-C1), 75.4 (Glc*N*Ac-C5), 75.0 (Gal-C5), 74.9 (Glc*N*Ac-C3), 73.5 (Glc*N*Ac-C4), 72.5 (Fuc-C2), 71.9 (Fuc-C4), 71.0 (Gal-C2), 69.2 (Fuc-C3), 68.3 (Gal-C4), 68.1 (O-*C*H_2), 67.7 (Fuc-C2), 66.7 (Fuc-C5), 61.5 (Gal-H6), 59.8 (Glc*N*Ac-C6), 55.8 (Glc*N*Ac-C2), 37.1 (*C*H_2-NH), 28.3 (CH_2-*C*H_2-CH_2), 22.2 (CO-*C*H_3), 15.30(Fuc-C6) ppm; HR-MS (ESI) Calcd for $C_{34}H_{47}N_3NaO_{18}$ [M+Na]+: 808.2752; Found: 808.2752.

Experimental Data







12: HPLC trace
93.3 % purity

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.