JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY
Article
pubs.acs.org/JACS

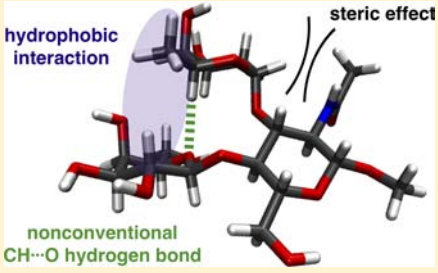
Stabilization of Branched Oligosaccharides: Lewis^x Benefits from a Nonconventional C−H···O Hydrogen Bond

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 Le^x (Gal[Fucα(1−3)]β(1−4)Glc*N*Ac). Mainly because of the unfavorable tumbling regime, the elucidation of the solution conformation of Le^x by NMR has only been partially successful so far. Le^x was therefore attached to a ^{13}C,^{15}N-labeled protein. ^{13}C,^{15}N-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 Le^x structure (exoanomeric effect, steric compression induced by the *N*HAc 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 Le^x core represents a novel element of the glycocode. 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 (Neu5Acα(2−3)Galβ(1−4)[Fucα(1−3)]Glc*N*Ac, 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

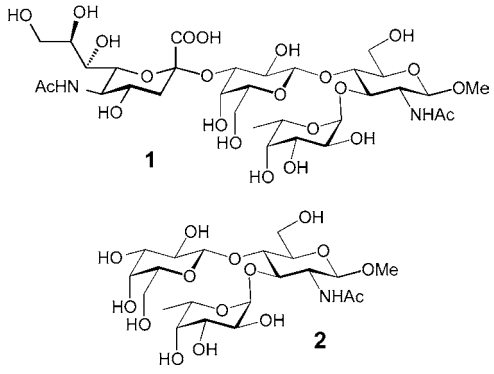


Figure 1. Methyl sialyl Lewis^x (sLe^x, **1**) and methyl Lewis^x (Le^x, **2**).

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 NMR[7] and later confirmed by X-ray crystallography.[8] The analysis of the solution conformation of

Received: May 31, 2013
Published: August 12, 2013

   dx.doi.org/10.1021/ja4054702 | J. Am. Chem. Soc. 2013, 135, 13464−13472

sLe^x (**1**) can be divided into two parts: (i) the conformation of the Le^x core Galβ(1−4)[Fucβ(1−3)]Glc*N*Ac (**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 Glc*N*Ac moiety or equatorial alkyl groups in the 2-position of carbocyclic Glc*N*Ac 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 crystallography[12] (Figure 2).

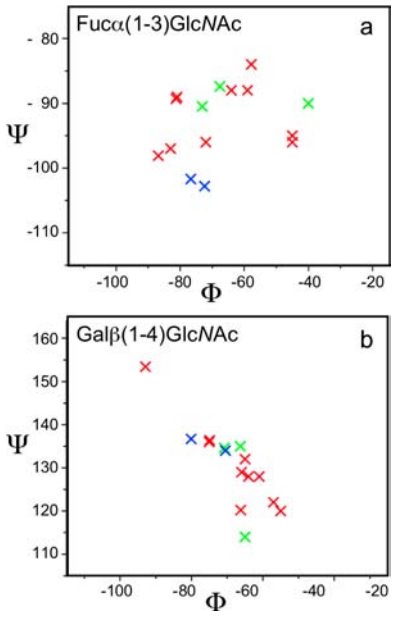


Figure 2. Previously reported Le^x structures/substructures: ϕ/ψ angles of the Fucα(1,3)Glc*N*Ac linkage (a) and the Galβ(1,4)Glc*N*Ac linkage (b). Torsion angles based on NMR data and MD simulations are shown in red,[11a,e,g−j] those of structures based on residual dipolar coupling (RDC) data in green,[11i] and those from the crystal structure of Le^x in blue.[12] The torsion angles are defined as follows: ϕ, O5−C1−O1−C′$_x$, ψ, C1−O1−C′$_x$−C′$_{x-1}$. 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°[11h] to −93/153°[11g] for the glycosidic bond Galβ(1−4)Glc*N*Ac. 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 Galβ(1−4)Glc*N*Ac. Differences in the crystal packing may explain these deviations.

Slynko 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 ^{13}C,^{15}N-labeled recombinantly expressed protein, ^{13}C,^{15}N-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]

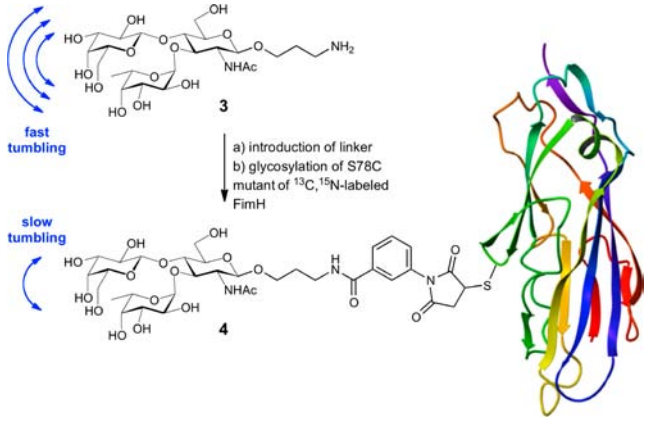
In this paper, we applied this approach to the structural analysis of Le^x chemically linked to a ^{13}C,^{15}N-labeled bacterial lectin (MW of ∼20 kD). The resulting increase of the correlation time τ_c together with the high resolution obtained at 900 MHz enabled the observation of numerous inter-residual 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 ^{13}C,^{15}N-labeled protein (Scheme 1).

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

Scheme 1. [a]

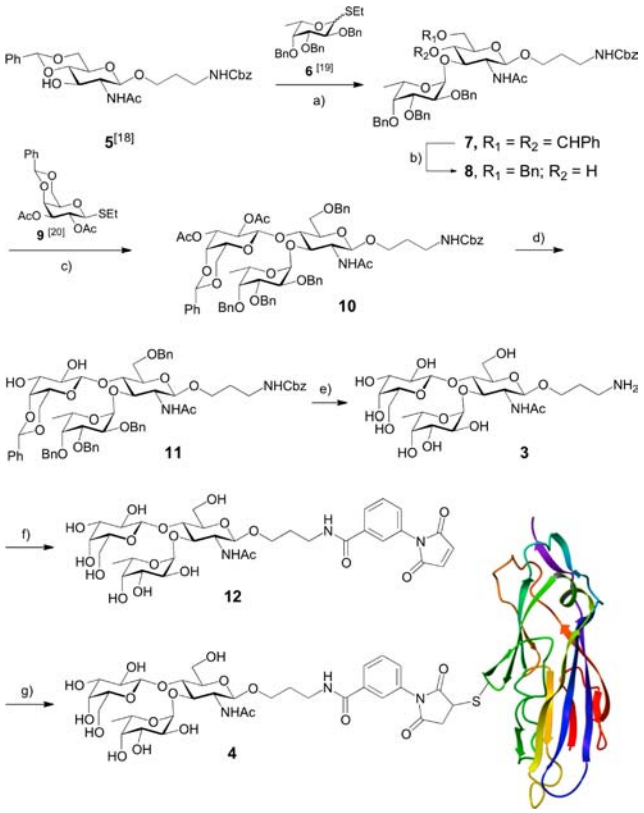


[a]For the improvement of the tumbling properties and consequently the extractable NMR spectroscopic information, low-molecular-weight Le^x (**3**) was linked to the bacterial protein FimH (→**4**).

dx.doi.org/10.1021/ja4054702 | J. Am. Chem. Soc. 2013, 135, 13464−13472

Ligand Synthesis and Chemical Glycosylation of Carrier Protein. To link oligosaccharides to carriers, 3-propanolamine is typically used.[17] 3-Aminopropanyl Le^x (**3**) was obtained by glycosylating the Glc*N*Ac derivative **5**[18] first with the L-fucose building block **6**[19] and then, after deprotection of the 4-position, with the thiogalactoside **9**[20] (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]



[a]:Legend: (a) TBAB, $CuBr_2$, 4 Å molecular sieves, DCM/DMF, 77%; (b) $NaBH_3(CN)$, HCl, THF, 85%; (c) DMTST, 4 Å molecular sieves, DCM, 60%; (d) NaOMe, MeOH; (e) $Pd(OH)_2$, H_2, DCM/MeOH/H_2O/AcOH, 77% over two steps; (f) 3-maleimidobenzoic acid *N*-hydroxy-succinimide ester (MBS), DMSO, H_2O, 62%; (g) ^{13}C,^{15}N-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 (→**8**), the 4-hydroxy group of the Glc*N*Ac 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 ^{13}C,^{15}N-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 ^{13}C,^{15}N-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,^{15}N-HSQC spectrum of the glycoconjugate see Figures S2 and S3 in the Supporting Information). Although the Michael addition proceeded only to approximately 50% conversion, 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 ^{13}C,^{15}N-labeled glycoprotein (see above) linked to unlabeled carbohydrate to detect distance-related NOE cross-peaks within the carbohydrate by recording 2D ^{13}C F1-filtered F2-filtered NOESY[22] (in D_2O) and 2D ^{15}N F1-filtered F2-filtered NOESY[15] spectra (in H_2O) (Figure 3a,b). In these types of NOESY experiments, the

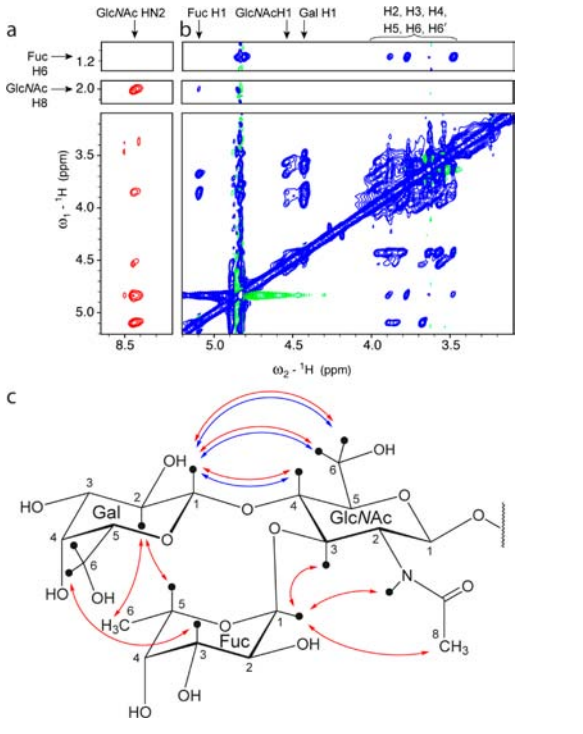


Figure 3. NOE cross-peaks between protons of the Le^x trisaccharide: (a) ^{15}N filtered-filtered 2D NOESY recorded with 96 scans in 16 h; (b) ^{13}C filtered-filtered 2D NOESY recorded with 96 scans in 16 h for Le^x-FimH (**4**); (c) schematic overview of interresidual NOEs of Le^x-FimH (**4**) (red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

signals of the ^{13}C,^{15}N-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 ^{13}C F1-filtered TOCSY and ^1H,^{13}C-HSQC spectra (Figure S4, Supporting Information). Since no carbohydrate−protein NOE cross-peaks were present

dx.doi.org/10.1021/ja4054702 | J. Am. Chem. Soc. 2013, 135, 13464−13472

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 (Le^x-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 Le^x-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 Le^x-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 Cyana[23] 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 Le^x-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 chemical shifts of Fucα(1−3)GlcNAcβ-OMe and Galβ(1−4)GlcNAcβ-O-$(CH_2)_3NH_2$[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 Fucα(1−3)GlcNAcβ-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 Le^x-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

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

	Le^x-FimH (**4**)		methyl Le^x (**2**)	
proton pair	*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]

[a]The $^1H-^1H$ 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. [b]Signal to noise ratios (*S/N*) from cross-peaks involving methyl or methylene protons were divided by 3 or 2, respectively. [c]Only one cross-peak was used because of artifacts or overlap. [d]Reference restraint for the ^{15}N-filtered-filtered 2D NOESY. [e]Reference restraint for the ^{13}C-filtered-filtered 2D NOESY.

dx.doi.org/10.1021/ja4054702 | J. Am. Chem. Soc. 2013, 135, 13464−13472

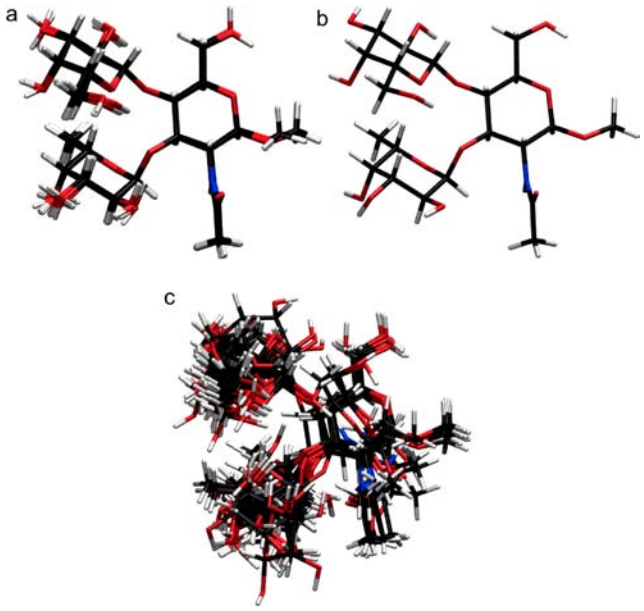


Figure 4. Calculated and refined structural ensembles of Le^x at 293 K using NOE restraints: (a) Le^x-FimH (**4**); (b) a representative structure thereof; (c) methyl Le^x (**2**).

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 Fucα(1-3)Glc*N*Ac and Galβ(1-4)Glc*N*Ac

	exptl			calcd		
proton	Le^x	$\Delta\delta$(Le^x-Fucα(1−3)Glc*N*Ac)	$\Delta\delta$(Le^x-Galβ(1−4)Glc*N*Ac)	Le^x	$\Delta\delta$(Le^x-Fucα(1−3)Glc*N*Ac)	$\Delta\delta$(Le^x-Galβ(1−4)Glc*N*Ac)
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_3 of L-Fuc	1.18	0.01		1.13	−0.05	
H2 of D-Gal	3.50		0.04	3.64		0.04

[a]Calculated using B3LYP/6-31G(d,p) in water with the polarizable conductor calculation model (CPCM).[29]

hydrogen bond.[33] The C−H···O hydrogen bond locks the Le^x conformation, resulting in a narrow cluster of ϕ/ψ torsion angles. Although C−H···O hydrogen bonds are only about half as strong as "classical" O−H···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···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···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	12
				3.304	
Me Le^x (**2**)	1UZ8[a]	1.8	Refmac 5.2	3.465	25a
				3.312	
Le^x-β(1−3)Gal	1SL5[a]	1.7	CNS 1.1	3.741	25b
Le^x-β(1−3)Galβ(1−4)Glc	3AP9[a]	1.33	Refmac 5.5	3.778	25d
	2OX9[a]	1.95	CNS 1.1	3.210	25c
				3.329	
				3.289	
				3.352	
Siaα(2−3)Le^x-OMe	1G1T[a]	1.5	CNS	3.434	30
Siaα(2−3)Le^x-OMe	2KMB[a]	2.0	X-PLOR 3.54	3.350	31
				3.256	
				3.343	
Siaα(2−3)Le^x-β(1−3Gal)-β(1−4)GalNAc-Thr	1G1S[a]	1.9	CNS	3.380	30
				3.374	

[a]PDB (Protein Data Bank). [b]CSD (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···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···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···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 DFT[36] optimized conformation served as starting point for a series of single point quantum mechanical calculations aimed at the quantification of the stacking interaction (Figure

dx.doi.org/10.1021/ja4054702 | J. Am. Chem. Soc. 2013, 135, 13464−13472

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

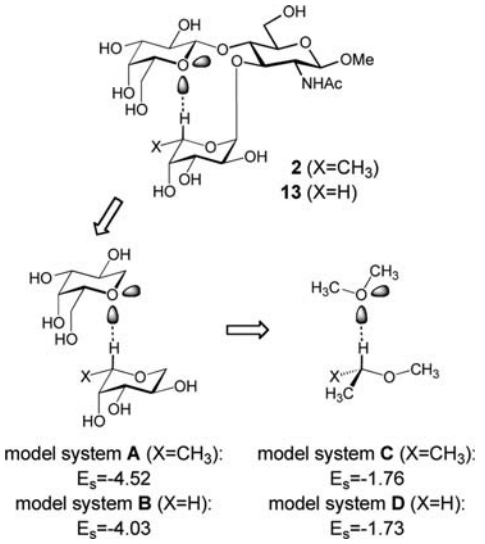


Figure 5. Model systems (**A**) 1-deoxy-galactose/1-deoxy-fucose, (**B**) 1-deoxy-galactose/1-deoxy-arabinose, (**C**) Me-OMe/Me-O-iPr, and (**D**) Me-OMe/Me-O-Et and the corresponding stabilization energies (E_s in kcal/mol).

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···O hydrogen bonding interaction in carbohydrates (Figure S8, Supporting Information) indicates an optimal H···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···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···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 effect[41] 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 *N*HAc 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···O hydrogen bond[33] between H−C(5) of L-fucose and O(5) of D-galactose was identified as an additional factor. This nonconventional C−H···O hydrogen bond contributes to a reduction of the conformational flexibility and exhibits a novel dimension of the glycocode.[43] We speculate that such interactions are widespread among glycoepitopes in mammalians. 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···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···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 (Galβ1,4-[Fucβ1,3]Glc*N*AcβOMe) and the disaccaride Fucα1,3Glc*N*AcβOMe 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 $^{13}C/^{15}N$- labeling *E. coli* BL21(DE3) cells harboring the FimH-CRD-6His (S78C) encoding plasmid were cultivated in M9 minimal medium[46] supplemented with 1% (v/v) of Basal Medium Eagle vitamin mix solution (Sigma, Buchs, Switzerland) and 100 μg/mL of ampicillin (Applichem, 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 $^{13}C_6$-glucose (2 g/L), ^{15}N-ammonium chloride (1 g/L) and 1% (v/v) Basal Medium Eagle (BME) vitamin mix solution. The cells were allowed to grow to an OD_{600} 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 HPLC[47] using BSA as standard.

FimH-Le^x Glycoprotein (4). A mixture of the $^{13}C,^{15}N$-labeled FimH mutant (5.7 mg, 0.29 nmol) and maleimide **12** (13.4 mg, 17.1 μmol) 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 ultrafiltation. 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 aquired on Bruker Avance or Avance III spectometers 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 filer unit (Vivaspin, Sartorius stedim, Goettingen, Germany, 10 kDa cutoff). Samples were prepared in either 93% H_2O/7% D_2O or 100% D_2O using lyophilization for the preparation of the latter. 2D ^{13}C F1-filtered F2-filtered NOESY[22] and 2D ^{15}N F1-filtered F2-filtered NOESY

dx.doi.org/10.1021/ja4054702 | J. Am. Chem. Soc. 2013, 135, 13464−13472

spectra[15] were recorded with a mixing time of 150 ms in D_2O and H_2O/D_2O, 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, $^1H-^{13}C$ HSQC, long-range $^1H-^{13}C$ HSQC, $^1H-^{13}C$ 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 Glc*N*Ac H61–H62 cross-peaks (1.77 Å) as reference. S/N ratios of signals involving CH_2 and CH_3 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 model[48] 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 DFT[36] 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**; $C_{21}H_{37}NO_{15}$) 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) CPCM[29] 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 (*i*Pro-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 ^{15}N-HSQC (Figure S3) of **4**, ^{13}C-HSQC overlay of glycosylated $^{13}C/^{15}N$-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) 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*, 1342−1351. (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.; Wyns, 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) Guntert, 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. *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.; Schlaeppi, 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) Møller, 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.