J|A|C|S
ARTICLES
Published on Web 02/20/2007

A Simple Model System for the Study of Carbohydrate−Aromatic Interactions

Giancarlo Terraneo,[†] Donatella Potenza,[†] Angeles Canales,[‡] Jesus Jiménez-Barbero,*[,‡] Kim K. Baldridge,*[,§] and Anna Bernardi*[,†]

Contribution from the Universita' di Milano, Dipartimento di Chimica Organica e Industriale e Centro di Eccellenza CISI, via Venezian 21, 20133 Milano, Italy, Centro de Investigaciones Biológicas, Consejo Superior de Investigaciones Científicas, Ramiro de Maeztu 9, 28040 Madrid, Spain, and Institute of Organic Chemistry, University of Zürich, Winterthurerstrasse 190, CH-8057 Zürich, Switzerland

Received September 14, 2006; E-mail: anna.bernardi@unimi.it; jjbarbero@cib.csic.es; kimb@oci.unizh.ch

Abstract: A molecular scaffold was identified which enables the establishment of intramolecular interactions between a monosaccharide and a nearby phenyl ring. A group of molecules containing four different monosaccharides (glucose, galactose, *N*-acetyl-glucosamine, and *N*-acetyl-galactosamine) was synthesized and used to investigate the extent and nature of this carbohydrate−arene interaction, as well as the effect on the overall 3D structure of the molecules involved. The sugar−aromatic distance was evaluated by rigorous NMR studies supported by molecular modeling and found to be constant throughout the series, independent of the nature of the sugar and of the conformational behavior of the fragment connecting the two elements. Ab initio calculations at the B3LYP/DZV(2d,p) level of theory enable the analysis of the electronic nature of the interaction. The study shows that, given the opportunity, persistent intramolecular aromatic−sugar interactions can be established and can significantly influence overall molecular shape and energetics. These results have important implications in the design of structural mimics of oligosaccharides.

Introduction.

Sugar−protein interactions play an important role in a wide range of biological processes, from regulatory processes such as fertilization to pathologies such as tumor spread. These interactions can be responsible for mediating diverse cellular activities, such as cell recognition, growth, and apoptosis.[1,2] A large variety of proteins, with very different functions and topologies, are involved in carbohydrate recognition, including enzymes, periplasmic receptors, antibodies, and lectins.[3] Because of the amphiphilic character of oligosaccharides, a variety of forces mediate the recognition process, with many types of interactions revealed in the available X-ray structures of sugar−protein complexes. Polar groups of the sugars are bound by H-bond donors and acceptors in the protein backbone and polar side chains. Apolar regions (formed primarily by the pyranose CH groups) are complemented by nonpolar surfaces of the proteins and, in particular, tend to pack against aromatic residues in the receptor side chains.[4] Convincing arguments have been proposed that suggest that desolvation of such apolar patches is the driving force for intermolecular interactions.[5] NMR studies have shown the presence of interactions between hydrogens belonging to the apolar faces of sugar rings and aromatic residues of protein side chains[6] and have linked the strength of the interaction to the size and the electron richness of the arene.[7] In addition, studies concerning the role of *N*-linked oligosaccharides during protein folding have revealed[8] that the incidence of aromatic amino acids in proximity to glycans during the folding process is higher than the standard levels occurring on the surface or within the protein core. Together, these data provide significant implications for molecular recognition of carbohydrates in water solution, and indeed, artificial carbohydrate receptors that exploit sugar/aromatic interactions have been described.[9] Carbohydrate−aromatic complexes have also been investigated using computational methods, suggesting the

[†] Universita' di Milano.
[‡] Consejo Superior de Investigaciones Científicas.
[§] University of Zürich.

(1) Lis, H.; Sharon, N. *Chem. Rev.* **1998**, *98*, 637−674.
(2) Simanek, E. E.; McGarvey, G. J.; Jablonowski, J. A.; Wong, C. H. *Chem. Rev.* **1998**, *98*, 833−862.
(3) (a) Wies, W. L.; Drickamer, K. *Ann. Rev.* **1996**, *65*, 441−473. (b) Vyas, N. K. *Curr. Opin. Struct. Biol.* **1991**, *1*, 732−740.
(4) (a) Quiocho, F. A.; *Biochem. Soc. Trans.* **1993**, 21, 442−448. (b) Elgavish, S.; Shaanan, B. *J. Mol. Biol.* **1998**, *277*, 917−932.
(5) Lemieux, R. U. *Acc. Chem. Res.* **1996**, *29*, 373−380.
(6) (a) Jiménez-Barbero, J.; Peters, T. *NMR Spectroscopy of Glycoconjugates*; Wiley-VCH: Weinheim, Germany, 2003. (b) Wormald, M. R.; Petrescu, A. J.; Pao, Y. L.; Glithero, A.; Elliott, T.; Dwek, R. A. *Chem. Rev.* **2002**, *102*, 371−386. (c) Jiménez-Barbero, J.; Asensio, J. L.; Cañada, F. J.; Poveda, A. *Curr. Opin. Struct. Biol.* **1999**, *9*, 549−555.
(7) Chavez, M. I.; Andreu, C.; Vidal, P.; Aboitiz, N.; Freire, F.; Groves, P.; Asensio, J. L.; Asensio, G.; Muraki, M.; Cañada, F. J.; Jiménez-Barbero, J. *Chem.—Eur. J.* **2005**, *11*, 7060−7074.
(8) Petrescu, A. J.; Milac, A. L.; Petrescu, S. M.; Dwek, R. A.; Wormald, M. R. *Glycobiology* **2004**, *14*, 103−114.
(9) (a) Davis, A. P.; Wareham, R. S. *Angew. Chem., Int. Ed.* **1999**, *38*, 2978−2996 and references therein. (b) Mazik, M.; Cavga, H.; Jones, P. G. *J. Am. Chem. Soc.* **2005**, *127*, 9045−9052. (c) Klein, E.; Crump, M. P.; Davis, A. P. *Angew. Chem., Int. Ed.* **2005**, *44*, 298−302. (d) Vacca, A.; Nativi, C.; Cacciarini, M.; Pergoli, R.; Roelens, S. *J. Am. Chem. Soc.* **2004**, *126*, 16456−16465.

 10.1021/ja066633g CCC: $37.00

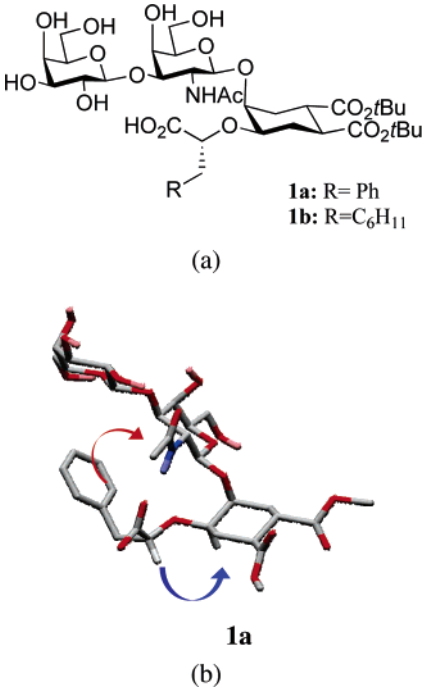


Figure 1. Aromatic−sugar interactions stabilize the bioactive conformation of glycomimetic **1a** (see ref 14). (a) Structure of **1**. (b) Preferred conformation of **1a** in water. The arrows show schematic representations of the closed contact between aromatic and sugar residues as inferred from NOE measurements, which were observed in the NMR spectrum of **1a** in D_2O.

interaction to be primarily intermolecular van der Waals contacts and/or CH−π interaction.[10,11,12]

Carbohydrate−arene interactions have also been shown to direct conformational equilibrium of glycophanes[13] and oligosaccharide mimics[14] in water solution. We have recently suggested that aromatic/sugar interactions may find application in the design of glycomimetics, since *intramolecular* sugar/aromatic interactions can help to stabilize bioactive conformations of small-molecule oligosaccharide mimics. In fact in the course of our studies directed toward the synthesis of glycomimetic ligands of the cholera toxin (CT),[14] we found clear NMR evidence of a close spatial proximity between the phenyl ring and the *N*-acetylgalactosamine residue in compound **1a** (Figure 1), which binds to CT with micromolar affinity. Comparison with a series of analogues bearing different alkyl groups on the ether side chain suggested that such an interaction biases the conformational behavior of **1a** by restricting the conformational freedom of the chain. As a result, the side chain of **1a** is "preorganized" in a suitable spatial orientation that enables optimal interaction of the binding determinants to the binding region of CT (Figure 1).[14]

Sugar−aromatic interactions may have a broader application as a key element in conformational-based design of sugar mimics and therefore warrant further investigation. To this end,

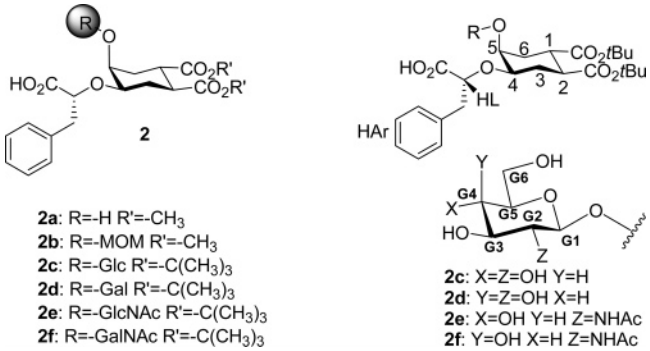


Figure 2. Compounds used in this study and relative numbering convention.

compound **2** (Figure 2) was chosen as a suitable model to investigate intramolecular interactions between various monosaccharide fragments (**2c**−**f**, R = monosaccharide) and the phenyl ring appended to the ether side chain.

A series of molecules **2a**−**f** with different substituents at the axial hydroxy group of the dicarboxy cyclohexanediol[15] (R = H, MOM, Glcβ1-, Galβ1-, GlcNAcβ1-, GalNAcβ1-), were synthesized. Compounds **2a**−**f** can be exploited as models for deducing chemical features and structural requirements of the C−H vector responsible for optimal interaction with the aromatic rings. An NMR-based approach was used to study these features. The chemical nature of the compounds enabled the sugar/aromatic interaction to be studied in D_2O by analyzing the nuclear Overhauser effect (NOE) observed between diagnostic protons belonging to the two fragments. Moreover, the mobility of the ether chain could be studied within the series by NOE-based methods. Spectral variations taking place in the different glycan fragments were analyzed.

Results and Discussion.

Synthesis. The synthesis of compounds **2** followed procedures that have been previously described for the synthesis of **1a**.[14] Starting from diol **3**,[15] the monoether **5** was obtained by regioselective alkylation with the triflate **4**[14] (45% yield; Scheme 1). The alcohol **2a** was obtained from **5a** by deprotection of the benzyl ester (H_2/Pd). The MOM derivative **2b** was obtained by protection of **5a** (MOMCl/TBAI/DIPEA in CH_2Cl_2), followed by removal of the benzyl ester (Scheme 1). Compounds **2c**−**f** were synthesized following a common glycosylation procedure using the sugar trichloroacetamidate as the donor and the monoether **5b** as the acceptor. The glycosylation reaction was carried out at −30 °C and catalyzed by trimethylsilyl triflate for glucose and galactose (**2c** and **2d**). For the two 2-acetamido sugars (GlcNAc and GalNAc in **2e** and **2f**, respectively), the glycosylation reaction was catalyzed by triflic acid and stirred at room temperature for 3 h before refluxing overnight in CH_2Cl_2. For the entire series, the benzyl ester was removed using H_2/Pd and the acetyl groups were removed under standard Zemplen's condition (MeONa/MeOH). Experimental details are given as Supporting Information.

NMR Studies of Compounds 2a−f in D_2O. NMR spectra of compounds **2a**−**f** were recorded in D_2O at 400 MHz and 298−300 K. Chemical shifts and coupling constants are reported in the Supporting Information. NOESY experiments were carried out using a delay of 800 ms that was selected by applying

(10) Fernandez-Alonso, M.; Cañada, F. J.; Jiménez-Barbero, J.; Cuevas, G. *J. Am. Chem. Soc.* **2005**, *127*, 7379−7386.
(11) (a) Sujatha, M. S.; Sasidhar, Y. U.; Balaji, P. V. *Biochemistry* **2005**, *44*, 8554−8562. (b) Sujatha, M. S.; Sasidhar, Y. U.; Balaji, P. V. *Protein Sci.* **2004**, *13*, 2502−2514.
(12) Spiwok, V.; Lipovova, P.; Skalova, T.; Vondrackova, E.; Dohnalek, J.; Hasek, J.; Kralova, B. *J. Comput.-Aided Mol. Des.* **2005**, *19*, 887−901.
(13) Morales, J. C.; Penadés, S. *Angew. Chem., Int. Ed.* **1998**, *37*, 654−657.
(14) Bernardi, A.; Arosio, A.; Potenza, D.; Sanchez-Medina, I.; Mari, S.; Cañada, F. J.; Jiménez-Barbero, J. *Chem.−Eur. J.* **2004**, *10*, 4395−4406.
(15) Bernardi, A.; Arosio, D.; Manzoni, L.; Micheli, F.; Pasquarello, S.; Seneci, P. *J. Org. Chem.* **2001**, *66*, 6209−6216.

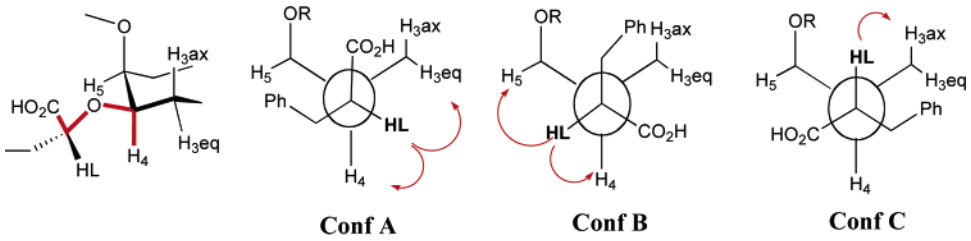
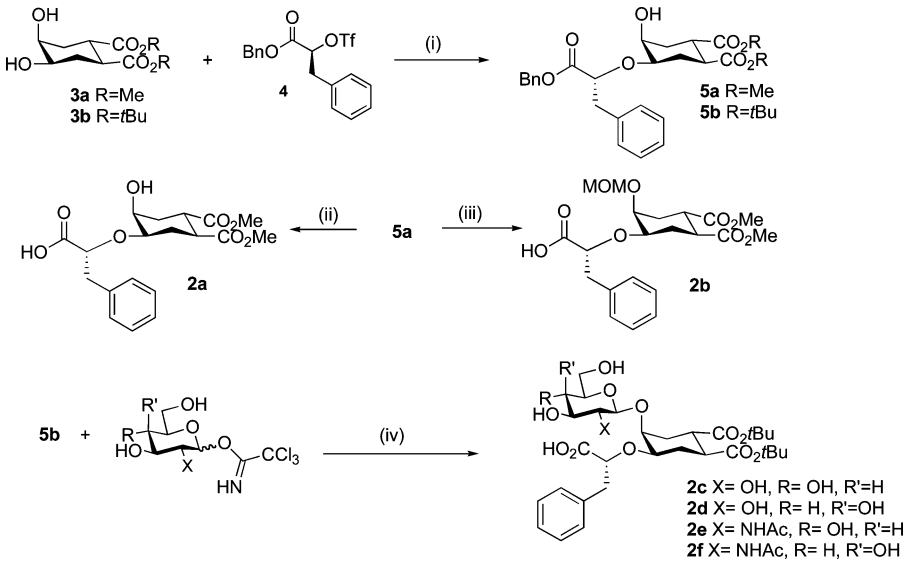


Figure 3. Idealized staggered conformations describing the orientation of HL relative to the diol moiety (CHD). The improper dihedral angle χ (C(O)−Cα−C4CHD−H4CHD) is used to define the conformations.

Scheme 1. Synthesis of **2a**−**f**[a]



[a] (i) Bu_2SnO reflux, then **4** at rt, **5a**: 33%, **5b:** 45%; (ii) H_2/Pd, 1 atm, rt, **2a**: 95%; (iii) 1. MOMCl/TBAI/DIPEA in CH_2Cl_2 at rt, 2. (ii) **2b**: 60% over two steps; (iv) 1. DCM/cat; 2. (ii); 3. MeONa in MeOH at rt, **2**c: 61%, **2d**: 69%, **2e**: 64%, **2f**: 70% over three steps.

the inversion recovery experiments (t1ir, Bruker library) from 50 ms to 3 s. The 800 ms delay was found to be the most appropriate mixing time to evaluate aromatic/sugar interaction in molecules **2c**−**f**. NOE contact tables are collected in the Supporting Information (Tables SI-1 to SI-6). All compounds gave satisfactory NOESY spectra.

The relation between NOE signals and proton−proton distances is well established[16] and can be worked out in a semiquantitative manner using an interproton relaxation matrix. The NOE intensities reflect the conformer population, and therefore information concerning the population distribution in solution can be obtained by focusing on the diagnostic, mutually exclusive, NOE interactions that characterize the different possible conformations.[17] A comparative analysis of compounds **2c**−**f** reveals a series of common features throughout the series. For all compounds, analysis of vicinal proton−proton coupling constants for the sugar fragment indicates that all the pyranoses are in the usual $^4C_1$ conformation. The cyclohexanediol moiety (CHD) also adopts a chair conformation, with the ester groups in equatorial position as shown by the exclusive NOE contact between the H2 and H4 protons on the CHD ring (Figure 2). The orientation around the sugar−diol linkage corresponds to the *syn*:*syn* global minimum conformation for all the monosaccharides examined, as defined by the G1/H5CHD NOE contact observed between the anomeric protons of the sugars and the proton in position 5 of the diol (Figure 2).[18] Presence of the alternative *anti* conformation would give rise to exclusive G1/H6axCHD NOE contact, which is never observed in any of the molecules **2c**−**f**.

Additional degrees of freedom available to compounds **2a**−**f** correspond to the hydroxy acid side-chain bonds and to the relative orientation between the monosaccharides and the aromatic ring in compounds **2c**−**f**. Due to the absence of anomeric effects, the ether linkage that connects the CHD ring to the hydroxy acid moiety is more flexible than the glycosidic linkage. The NOE contacts observed in this region can arise from multiple φ,ψ combinations that provide similar through-space proton−proton interactions for the reporter HL (Figure 2). Thus, the experimental interproton distances obtained by NMR spectroscopy can be analyzed more conveniently using the improper dihedral angle descriptor, χ(C(O)−Cα−CHDC4−CHDH4), defined using the Newman-type projections shown in Figure 3. This descriptor defines the orientation of HL (and of the hydroxy acid carboxy group) relative to the CHD ring.

In principle, although rotations around the O−C bond may also produce quasi-eclipsed orientations, three idealized staggered orientations are possible involving the ether connector, identified in Figure 3 as Conf. A (χ = 180°), Conf. B (χ = +60°), and Conf. C (χ = −60°), each associated with specific

(16) Hwang, T. L.; Shaka, A. J. *J. Am. Chem. Soc.* **1992**, *114*, 3157−3158.
(17) Dabrowski, J.; Kozar, T.; Grosskurth, H.; Nifant'ev, N. E. *J. Am. Chem. Soc.* **1995**, *117*, 5534−5539.
(18) (a) Bernardi, A.; Potenza, D.; Capelli, A. M.; Garcìa-Herrero, A.; Cañada, F. J.; Jiménez-Barbero, J. *Chem.—Eur. J.* **2002**, *8*, 4597−4612. (b) Bernardi, A.; Arosio, D.; Manzoni, L.; Monti, D.; Posteri, H.; Potenza, D.; Mari, S.; Jiménez-Barbero, J. *Org. Biomol. Chem.* **2003**, *1*, 785−792.

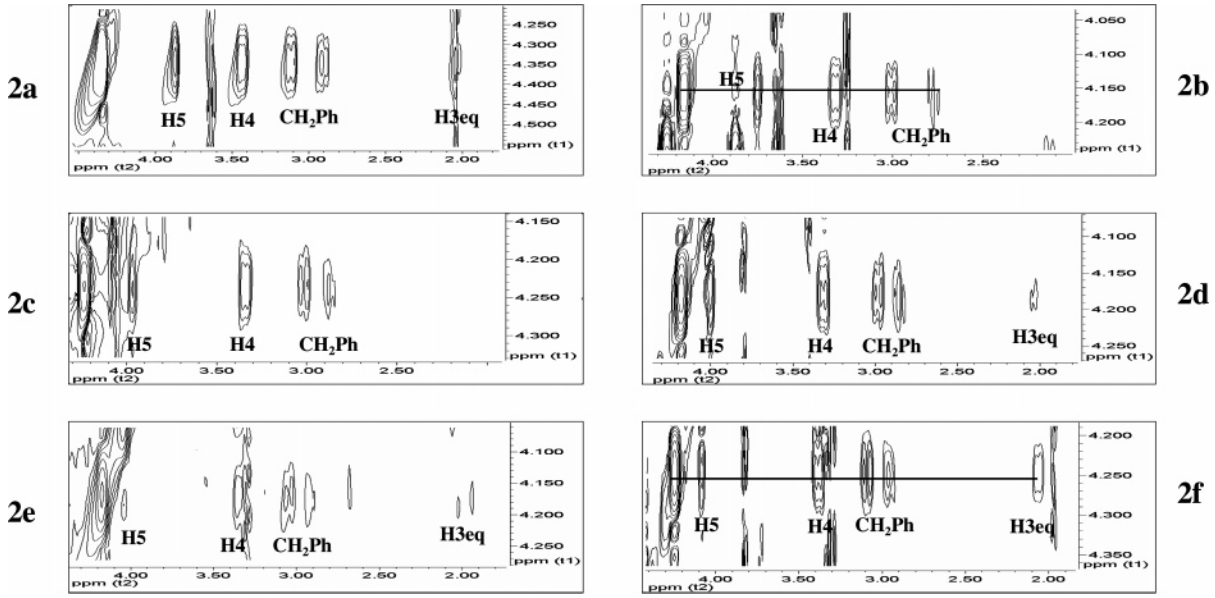


Figure 4. NOESY spectra of **2a**−**f** in D_2O: NOE contacts of HL.

Table 1. Diagnostic NOESY Cross-peaks for the HL Proton and Conformational Distribution of the Side Chain[a]

	calculated NOE intensity[b] (%) (experimental NOE intensity)[c]					
proton pairs	**2a**	**2b**	**2c**	**2d**	**2e**	**2f**
HL−H4	6.4 (6.6)	6.5 (6.9)	8.2 (7.3)	7.7 (6.7)	7.7 (2.2)	6.5 (6.4)
HL−H5	7.9 (5.6)	6.5 (6.7)	6.1 (6.2)	5.2 (4.8)	3.8 (0.4)	2.3 (2.0)
HL−H3eq	1.8 (1.6)	0.2 (0.0)	0.1 (0.0)	0.6 (0.6)	1.6 (0.2)	1.6 (1.8)
A/B/C ratio[d] (set)[e]	36:64:0 (fitting)	6:94:0 (MM)	0:100:0 (fitting)	12:88:0 (fitting)	27:73:0 (MM)	28:72:0 (fitting)
av error %[f]	2.8	0.9	1.7	1.4	10.0	0.5

[a] Spectra recorded in D_2O at 400 MHz and 300 K. [b] Calculated using Noeprom. [c] The NOE % is expressed as the ratio between the cross-peak volume and the volume of the corresponding diagonal peak. [d] Relative population of the limit conformations of Figure 3 which allowed the best reproduction of the NOE cross-peak intensity. [e] Set of conformations that produced the lowest error. MD, molecular dynamics; MM, multiminimization; fitting, populations from NOE data fitting with MM geometries. [f] Total error obtained for the three diagnostic contacts of Table 1.

interactions between HL and the protons on the C4, C3, and C5 carbon atoms of the CHD ring. The primary orientations of the hydroxy acid relative to the ring can thus be defined by focusing on the NOE cross-peaks of the HL proton to H4, H5, H3eq, and H3ax of the cyclohexanediol: conformer B is associated with the exclusive HL/H5 NOE contact; conformer A, with the HL/H3eq contact; and conformer C, with the HL/H3ax cross-peak (Figure 3). Finally, the orientation of the aromatic ring relative to the pyranose moiety in **2c**−**f** can be investigated by examining the NOE contacts between the hydrogens on the α face of the sugar and the aromatic protons.

A first examination of the NOESY spectra shows, for most compounds, the simultaneous presence of NOE contacts between the HL/H5 and HL/H3eq proton pairs (Figure 4 and Table 1). This fact suggests a fair amount of conformational freedom within the side chain, with a general prevalence of conformer B. These observations contrast with the spectrum of **1a**, which shows no NOE contacts between HL and H5 and a strong NOE cross-peak between HL and H3eq.[14]

A quantitative evaluation of the NMR results was attempted with the assistance of molecular modeling techniques. Both molecular mechanics (MC/EM[19]) and dynamics (MC/SD[20]) calculations were used. The determination of the conformational distribution involving the ether chain of each compound, however, involves small differences in the conformational behavior of a flexible fragment, which is quite difficult to describe well by computational methods. Our previous experience with similar compounds[18a] indicated that molecular mechanics and dynamics calculations can provide a qualitative reproduction of the side-chain behavior but cannot quantitatively reproduce the experimental data. In an effort to obtain the most accurate description of the conformational distribution of this elusive fragment, conformer populations were generated using three different methods and subsequently evaluated by comparing predicted NOE intensities with experimentally obtained values.[21] Three sets of populations were obtained, as described in detail in the Experimental Section, from molecular dynamics ensemble averaging (MD set), Boltzmann distributions obtained from molecular mechanics conformational searches (MM set), and NOE-fitting procedures (Fitting set). The experimental NOE intensity was measured as the ratio between the volume of the cross-peak and the volume of the corresponding diagonal peak. The expected NOE intensity was calculated for each set of conformations with the Noeprom[22] software, which uses the full interproton relaxation matrix approach.[23] The accuracy of the calculations was evaluated by measuring the percentage of NOE intensity |exptl − theor| errors. The set of conformers that produced the most accurate Noeprom results was then taken as the best description of the conformational behavior of the side chain. The results obtained for all compounds across all methods are reported in the Supporting Information. The optimal description obtained for each molecule is reported in Table 1.

(19) Chang, G.; Guida, W. C.; Still, W. C. *J. Am. Chem. Soc.* **1989**, *111*, 4379−4386.
(20) Guarnieri, F.; Still, W. C. *J. Comput. Chem.* **1994**, *15*, 1302−1310.
(21) Mari, S.; Cañada, J. F.; Jiménez-Barbero, J., Bernardi, A.; Marcou, G.; Motto, I.; Velter, I., Nicotra, F.; La Ferla, B. *Eur. J. Org. Chem.* **2006**, 2925−2933.
(22) Martin-Pastor, M. http://desoft03.usc.es/mmartin/software.html
(23) Cumming, A. C.; Carver, J. P. *Biochemistry* **1987**, *26*, 6664−6676.

Table 2. Correlation Time (τ_c) and Interproton Distances (r) Obtained by 1D NOE and 1D ROE Experiments for the Pairs HL/H5, HL/H4, and HL/H3eq

	2c		2d		2e		2f	
proton pair	τ_c[a] (ps)	r[b] (Å)	τ_c (ps)	r (Å)	τ_c (ps)	r (Å)	τ_c (ps)	r (Å)
HL/H4	210	2.63	164	2.61	231	2.71	168	2.64
HL/H5	213	2.87	221	2.74	233	3.18	168	3.08
HL/H3eq	187	3.53	162	3.53	296	3.07	207	2.98
average τ_c	203		182		253		181	

[a] Correlation time in ps. [b] Distance in Å.

Table 3. Conformational Distribution of the Side Chain Conformers Estimated from 1D NOE and 1D ROE Experiments

compounds	ratio Conf. A/Conf. B
2c	18:82
2d	22:78
2e	36:64
2f	32:68

The errors obtained, defined as the difference between the experimental and calculated NOE cross-peak intensities, are consistently lower than 3% for compounds **2b**, **2c**, **2d**, and **2f** and 10% for compound **2e**, for which the 2D-NOESY spectrum is noisier than the others.

In order to further address the description of the ether chain behavior in the sugar series (**2c**−**2f**), we also performed 1D NOE and 1D ROE experiments in D_2O. The buildup curves for the NOE and ROE effect were measured at different mixing times, and effective correlation times (τ_c) were estimated for each of the proton couples involved (see Experimental Section). It is known[24] that the ratio between the longitudinal and transversal relaxation rates for a given proton pair and the related NOE gives a mathematical function that is independent of the distance and only depends on the effective correlation time (τ_c) for that particular proton pair. Using this function, it is possible to estimate the spectral density $J(n\omega_0)$ from which the interproton distances can be rigorously derived (see Methods section). This approach allows one to account for the existence of distinct motions in different parts of the molecule and to estimate the correlated distances for any given proton pair without using an internal reference.

According to this analysis, all average effective correlation times for the vectors that interconnect proton pairs of the ether side chain with the CHD moiety are in the region 217 ± 32 ps, which indicates that the analysis shown in Table 1 can be safely used to address the conformational distribution. No major differences exist among the four studied derivatives. Using the average interproton distances for the MM clusters A and B and the experimental interproton distances of Table 2, the conformational distributions shown in Table 3 could be estimated.

In summary, for all compounds and experimental approaches used, the α-hydroxyacid chain is found to be flexible, with conformer C ($\chi = -60°$) represented very little, if at all. The majority of the conformations belong to the B cluster ($\chi = +60°$), and the A/B ratio varies between 20:80 (**2c** and **2d**) and 40:60 (**2e** and **2f**) depending on the sugar (Figure 5). Stabilization of the A cluster ($\chi = 180°$) in the GlcNHAc- and GalNHAc-substituted compounds **2e** and **2f** is likely related to the establishment of an electrostatic interaction between the hexosamine NH group and the carboxy group on the ether side chain, which in the A conformation is directed toward the sugar ring.

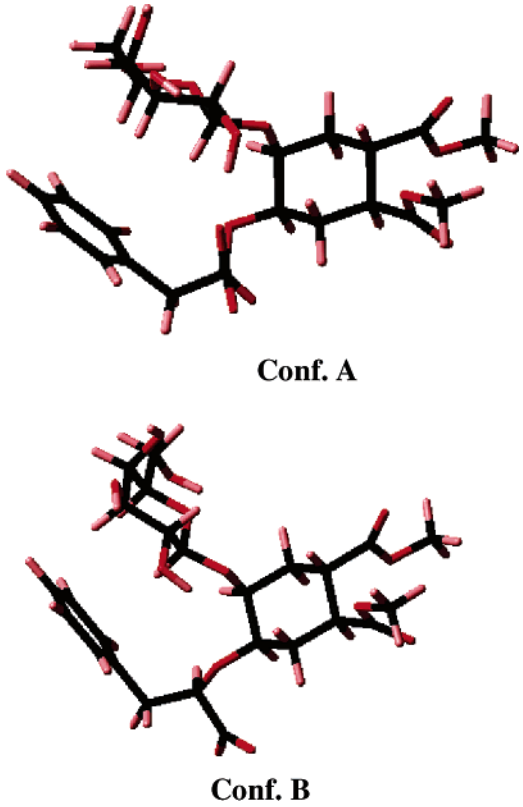


Figure 5. Low energy conformations of **2d** showing the two main side-chain conformations A ($\chi = 180°$) and B ($\chi = +60°$). The interaction between the phenyl ring and the sugar is similar in both conformers.

The distance and orientation of the aromatic ring relative to the monosaccharides can also be directly determined by examining the NOE contacts between the two fragments. The cross-peaks observed in D_2O are shown in Figure 6, and their relative intensities are reported in Table 4. The reported NOE intensities are measured relative to the intensity of the cross-peak between the geminal protons of the benzyl group, which is assigned a 100% value. The protons of the phenyl ring appear as a broad multiplet, and in the aromatic line for all compounds, many NOE contacts are detected. Cross-peaks between the aromatic ring, the side chain protons (HL), and the cyclohexane ring protons (H5) are clearly observed for all compounds examined. It is worth noting that the HAr/H5 cross-peak is consistently more intense in the spectra of **2e** and **2f**, containing GlcNAc and GalNAc, respectively. Various contacts are also detected with the protons belonging to the α face of the sugars. For all compounds **2c**−**f**, HAr/G1, HAr/G3, and HAr/G5 cross-peaks are present, all with the same intensity (ca. 10%) and independent of the nature of the sugar. In addition, the NOESY spectra of **2d** and **2f**, which are substituted with galactose and galactosamine moieties, respectively, also display an NOE contact between the aromatic protons and the α-proton on the sugar 4 position (HAr/G4).

(24) (a) Poveda, A.; Santamaria, M.; Bernabe, M.; Rivera, A.; Corzo, J.; Jiménez-Barbero, J. *Carbohydr. Res.* **1997**, *304*, 219−228. (b) Poveda, A.; Asensio, J. L.; Martin-Pastor, M.; Jiménez-Barbero, J. *J. Biomol. NMR* **1997**, *10*, 29−43. (c) Poveda, A.; Asensio, J. L.; Martin-Pastor, M.; Jiménez-Barbero, J. *Carbohydr. Res.* **1997**, *300*, 3−10.

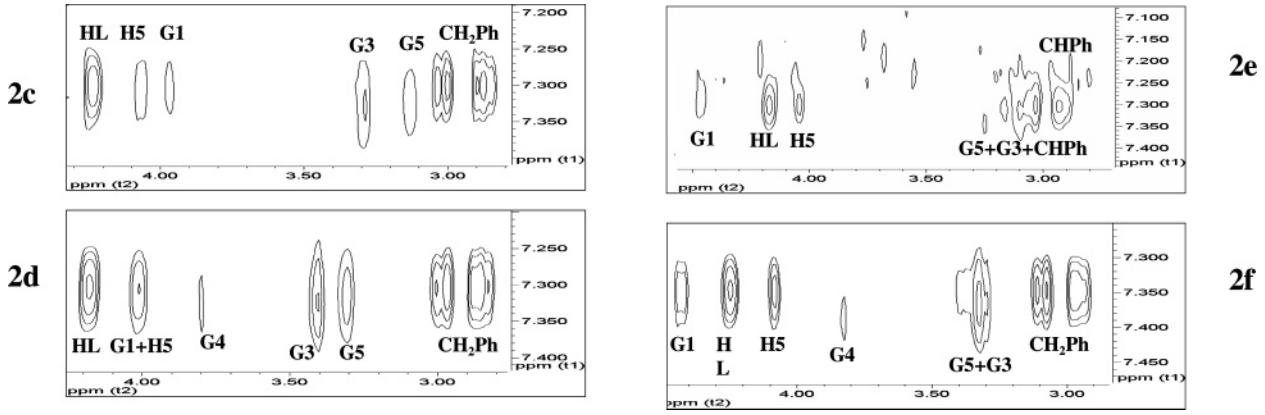


Figure 6. NOESY spectra of **2c**−**f** (400 MHz): the line of the aromatic protons (HAr) in D_2O.

Table 4. Intensity of the NOESY Cross-peaks between the Aromatic Protons and Those on the Monosaccharide Ring in Compounds **2c**−**f** (D_2O, 400 MHz)[a]

crosspeak	2c	2d	2e	2f
HL-HAr	23	23	25	20
G1-HAr	8		10	9
H5-HAr	5	12[b]	10	10
G4-HAr	n.d.[c]	3	n.d.[c]	2
G3-HAr	13	14	47[d]	18[e]
G5-HAr	10	10		
CHPh-HAr	18	19		19
CHPh-HAr	24	24	30	24

[a] NOE % relative to the intensity of the cross-peak between the geminal protons of the benzyl group, which is assigned 100% intensity. [b] G1 and H5 overlap. [c] Not detected. [d] G3, G5, and one of the benzyl proton overlaps. [e] G3 and G5 overlap.

Table 5. Correlation Time (τ_c) and Distances (r) between Aromatic Protons (HAr) and Sugar Protons Obtained by 1D NOE and 1D ROE Experiments

	2c		2d		2e		2f	
cross-peak	τ_c[a] (ps)	r[b] (Å)	τ_c[a] (ps)	r[b] (Å)	τ_c[a] (ps)	r[b] (Å)	τ_c[a] (ps)	r[b] (Å)
G1−HAr	124	3.19	ov[c]		158	3.13	122	3.05
G3−HAr	129	3.62	229	3.59	n.e.[d]		ov[c]	
G5−HAr	123	3.26	ov[c]		n.e.[d]		ov[c]	

[a] Correlation time (τ_c) in ps. [b] Distance in Å. [c] Not detected because signal overlaps. [d] Not estimated.

The nature and the intensity of the contacts suggest that the conformational freedom of the ether chain does not affect the position of the aromatic ring to a measurable degree. This in turn suggests that either the differences are too small to be identified or the phenyl−sugar interaction is insensitive to the conformational flexibility of the side chain, since the aromatic ring can be accommodated below the sugar in both the A and B conformations assumed by the α-hydroxyacid chain. This latter hypothesis is supported by molecular mechanics calculations, which reveal a low-energy minimum featuring a stacked conformation of the two rings in both the A and B clusters. Structures found for **2d** are shown in Figure 5 for illustration. In both conformation A and B, the aromatic ring is at an average distance of 4.2−4.4 Å from the sugar (distance between the rings' centroids) consistently in all compounds. The interproton distances have be evaluated quantitatively using the 1D experiments described above (see Methods) and are reported in Table 5.

The data in Table 5 clearly show that the G1−HAr distance is indeed similar in all compounds of the series. The G3−HAr and G5−HAr distances are also similar across all cases where estimation could be made. The corresponding effective correlation times are 170 ± 50 ps for all cases in which they were estimated. These correlation times are somewhat shorter than those measured between the side chain and the CHD moiety, although not to a significant degree. These values provide an indication that the different parts of these molecules are moving in solution at different frequencies. Thus, all the structural and dynamical data seem to point to the aromatic/sugar interaction as a predominant feature in these model compounds in water solution, occurring independent of the conformational motions of the appended ether chain.

Computational Studies of Compound 2c. To further analyze the nature of the carbohydrate−aromatic interaction, ab initio calculations were performed. The optimized structure of the glucose-substituted compound **2c**, chosen as the model compound, was determined at the B3LYP/DZV(2d,p) level of theory, starting from the lowest energy B conformation located by molecular mechanics and validated by NMR for **2c**. The hydrogen bond network resulting from the hydroxyl groups in position G2, G3, and G4 was adopted in a counterclockwise arrangement in order to preserve the "cascade" donor−acceptor between the groups. Because the β-glucose structure adopts an all-equatorial hydroxyl position, the hydroxyl in G4 acts as a donor for the acceptor G3, which in turns acts as a donor to the hydroxyl in position G2, which then enables further hydrogen bonding with the ether oxygen. In the glucose moiety the C5−C6 bond, characterized by the torsion angle ω (O6−C6−C5−O5), has three possible staggered conformations: *gauche-trans* (*gt*), *trans-gauche* (*tg*), and *gauche-gauche* (*gg*) (Figure 7). In the *gt* and *gg* conformations, the G6 hydroxyl acts as a donor to the intraring oxygen; in the *tg* conformation, the G6 hydroxyl acts as a donor to the G4 hydroxyl group. To better understand the implications of this conformational freedom on the sugar−aromatic interaction, we investigated the *gt*, *tg*, and *gg* conformers of **2c**. Their energy differences are shown in Table 6, together with the distance between the aromatic ring and the

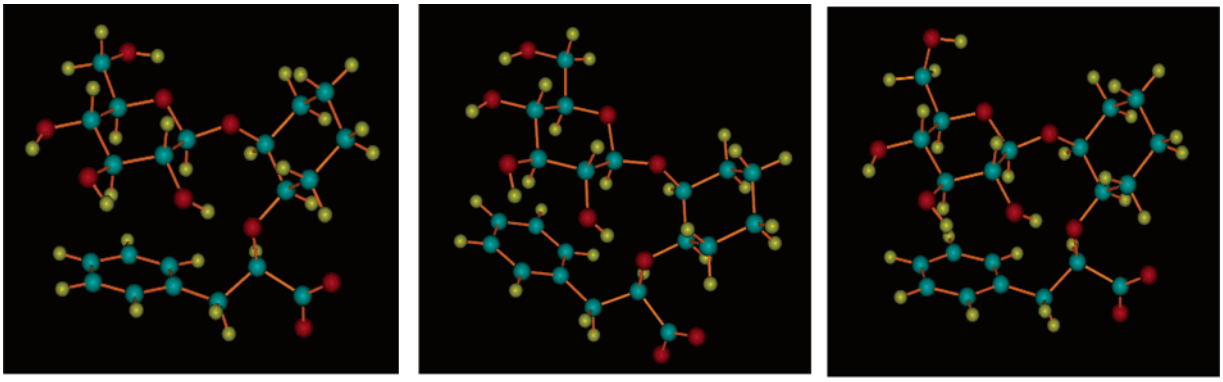


Figure 7. Three staggered (ω-angle) calculated conformations of **2c**.

Table 6. B3LYP/DZV(2d,p) and MP2/DZV(2d,p)//B3LYP/DZV(2d,p) Calculated Relative Energies for the Three Staggered Conformations (ω-Angle) of **2c**

ω-angle	ΔE^a B3YLP/DZV(2d,p)	distance $C_{Glc}-C_{ar}$[b] (Å)	ΔE^c MP2/DZV(2d,p)// B3LYP/DZV(2d,p)
gt	0.0 (0.0)	4.43	0.12 (0.03)
tg	2.60 (0.62)	4.42	3.24 (0.77)
gg	1.13 (0.21)	4.42	0.0 (0.0)

[a] kJ/mol; values in parentheses are kcal/mol at B3LYP/DZV(2d,p). [b] Distance between the glucose centroid and centroid of aromatic ring at optimized geometry at B3YLP/DZV(2d,p). [c] kJ/mol; values in parentheses are kcal/mol at MP2/DZV(2d,p)//B3LYP/DZV(2d,p) and correspond to a 46:13:48 population distribution at 300 K.

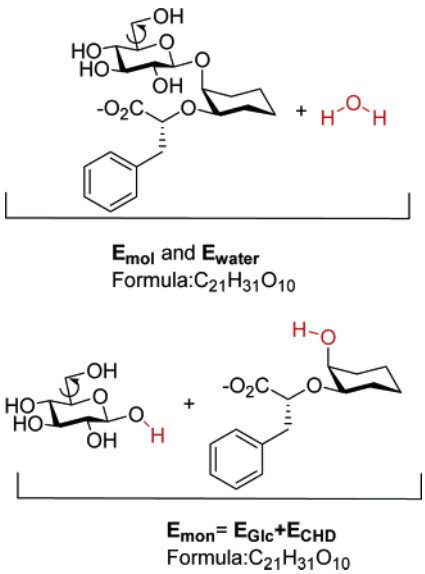


Figure 8. Molecular components exemplified as used for calculation of interaction energies, E_{int}. Elements in red indicate placement of bonds that are made and broken.

glucose moiety. The calculated energy differences for the three conformers of **2c** (*gt*, *tg*, *gg*, see Table 6) are in qualitative agreement with the experimental rotamer distribution in the solid state,[25] in solution,[26] and with ab initio calculations performed on β-methyl glucoside.[27] The distance between the aromatic ring and the sugar ring does not appear to be affected by the C5−C6 rotamer.

Single-point MP2/DZV(2d,p)//B3LYP/DZV(2d,p) calculations provide more accurate estimates of the relative energies for such interactions and are reported in Table 6. The rotamer distribution obtained at this level of theory (*gt*/*tg*/*gg*, 46:13:48 at 300 K) is close to the experimental data[25,26] and consistent with similar population of the *gg* and *gt* rotamers and low contribution of the *tg* conformer.

An estimate of the interaction energy (E_{int}) between the Glc-fragment and the CHD-fragment in compound **2c** was approximated from the calculated values as

$$E_{int} = E_{mol} - E_{mon}$$

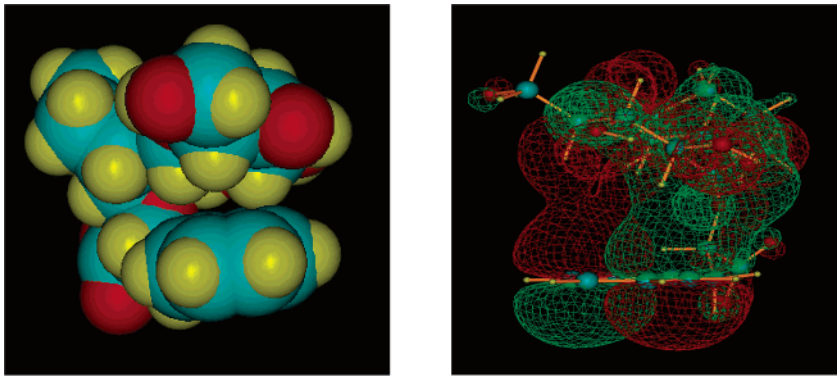
Here, E_{mol} is the calculated energy of compound **2c**, and E_{mon} is the calculated energy of the Glc-fragment (E_{Glc}) with the CHD-fragment (E_{CHD}). Full geometry optimizations were performed for all components, and the proper isodesmic relationship was established. To achieve this relationship, the two monomers were obtained by fragmenting at the glycosylic bond ($C1_{Glc}-O-C5_{CDH}$), followed by saturating the remaining structures with hydrogens, as appropriate (Figure 8). For glucose, we have kept the terminal bond as R−O*H*, and for the CHD-fragment, the fragmentation is fixed as CHD−*OH*. The isodesmic relatonship is balanced with one molecule of water (E_{water}). The resulting isodesmic relationship for **2c** together with the two monomers ($C_{21}H_{31}O_{10}$) provides the energy of interaction as

$$E_{int} = [(E_{mol} + E_{water}) - (E_{Glc} + E_{CHD})]$$

The components of E_{int} have been fully optimized at the B3LYP/DZV(2d,p) level of theory for each conformation of **2c**. All the energies and geometrical parameters are available in the Supporting Information.

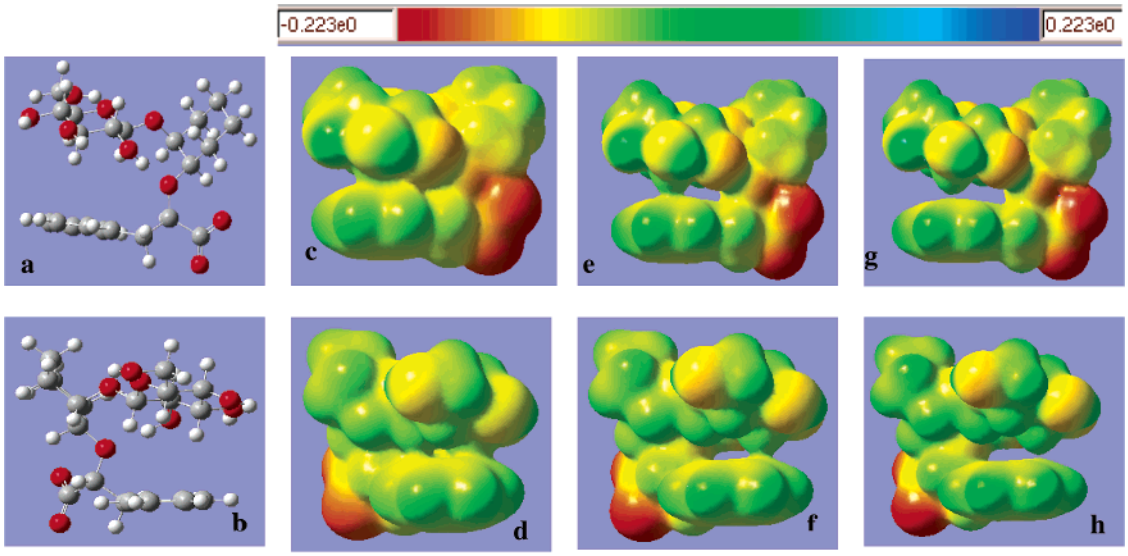
The calculated interaction energies (E_{int}) for each conformational variant of **2c** are shown in Table 7. The three interaction energies are very similar, all around 5 kcal/mol, favoring the complex E_{mon}. These results would suggest that loss of the glucose moiety from a position above the aromatic ring involves

(25) (a) Jeffrey, G. A.; McMullan, R. K.; Takagi, S. *Acta Crystallogr., Sect. B* **1977**, *33*, 728−737. (b) Marchessault, R. H.; Perez, S. *Biopolymers* **1979**, *18*, 2369−2374.
(26) Ohrui, H.; Nishida, Y.; Watanabe, M.; Hori, H.; Meguro, H. *Tetrahedron Lett.* **1985**, *26*, 3251−3254.
(27) Kirschner K. K.; Woods R. J. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 10521−10545.



9a 9b

Figure 9. (a) CPK surface. (b) Calculated molecular orbital, HOMO-6, displayed with isosurface ± 0.01 contours.



0.002 a.u 0.005 a.u 0.007 a.u

Figure 10. **2c**-*gt* conformation: views a and b. Electrostatic potential mapped on electron density isosurfaces of **2c** in the a and b view: c and d (isosurface 0.002 au), e and f (isosurface 0.005 au), and g and h (isosurface 0.007 au), respectively.

Table 7. Interaction Energy (E_{int}) for **2c** (*gt*, *tg*, and *gg*) Using Full Geometry Optimization at the B3LYP/DZV(2d,p) Level

2c ω-angle	E_{int} (kcal/mol)
gt	−5.42
tg	−4.72
gg	−5.18

Table 8. Relative Interaction Energy (E_{int}) for DFT and MP2 Methods

2c ω-angle	E_{int}[a] (kcal/mol)	ΔE_{int}[b] (kcal/mol)	ΔE_{int}[c] (kcal/mol)
gt	−5.42	0.0	0.0
tg	−4.72	0.70	0.74
gg	−5.18	0.23	0.21

[a] B3LYP/DZV(2d,p) interaction energy (E_{int}) at B3LYP/DZV(2d,p) optimized geometries. [b] B3LYP/DZV(2d,p) relative interaction energy (E_{int}) at B3LYP/DZV(2d,p) optimized geometries. [c] MP2/DZV(2d,p) relative interaction energy (E_{int}) at the B3LYP/DZV(2d,p) optimized geometries.

a decrease in stabilization for the entire system. This behavior reinforces our notion that sugar−aromatic interactions have an electrostatic character that provides energy stabilization on the order of roughly 5 kcal/mol, for these types of systems. Such interaction magnitudes constitute weak interactions of the wdV type and are consistent with previously reported values.[10−12] As well, one can deduce that 5 kcal/mol roughly accounts for the sum of the contributions between the hydrogens on the α-face of the sugar (G1, G3, and G5) and the π-system of the aromatic ring.

We have also evaluated this interaction energy using more conventional MP2 methods, as they have been shown to be a more accurate assessment of weak interactions in many examples. As such, we report MP2/DZV(2d,p)//B3LYP/DZV(2d,p) E_{int} as described above and reported in Table 8. In fact, for this particular analysis, the MP2 calculations are in agreement with the B3LYP values, with the same trend of relative energy interactions predicted.

The analysis of weak interactions is notoriously controversial.[28] In particular, verification of interaction and identification of the stabilizing atom−atom interactions between these two fragments is quite difficult; therefore we have considered several possibilities. For these investigations, we have concentrated on

(28) (a) Dunitz, J. D.; Gavezzotti, A. *Angew. Chem., Int. Ed.* **2005**, *44*, 1766−1787. (b) Gatti, C. *Z. Kristallogr.* **2005**, *220*, 399−457.

the structure with the lowest complexation energy, E_{int}, conformer *gt*. In Figure 9a, the CPK[29] surface of **2c**-*gt* is displayed, illustrating the G3 hydrogen to be within van der Waals contact of the aromatic ring and the G1 and G5 hydrogens to be extremely close to the van der Waals surface of the ring.

A further representation of the sugar/aromatic interaction for this model is provided by the analysis and visualization of the molecular orbital structure, displayed in Figure 9b. In fact, molecular orbital HOMO-6, where HOMO is the highest occupied molecular orbital, shows a bonding interaction between the sugar moiety and the aromatic ring, i.e., the G1, G3, and G5 hydrogens and the aromatic ring. The fact that this interaction is placed well below the HOMO ($E_{HOMO-6} = -4.580$ eV) suggests a strong contribution toward the stability of the system.

The intramolecular distances between the aromatic centroid (CAr) and the sugar protons for the compound **2c** are in the range of NOE contacts (CAr−G1 = 3.50 Å, CAr−G3 = 3.03 Å, and CAr−G5 = 3.87 Å).

Further assessment of the weak complex interaction can be evaluated using electrostatic potential contour maps. The electrostatic potential (ESP) of **2c** was mapped onto the total electron density as calculated from the ab initio methods and various contours evaluated in succession for illustrative purposes. The electron density isosurfaces evaluated begin from isosurface 0.002 au, representing >98% of the electronic charge, to isosurface 0.007 au, (i.e., closer to the nucleus), representing an increase of 4.39 kcal/mol for the electronic charge. All isosurfaces are mapped in the same range of ESP value (±0.223 au), and the contour colors represent the range including the most electron-rich areas (red) to the most electron-poor areas (blue). In this case, since there is a negative charge on the carboxyl group, the entire molecule shows negative character.

Figure 10a and b show two different views of the complex involving **2c**, with indicated levels of isosurface displayed for each in Figure 10c−h. Figure 10c and d show the isosurfaces 0.002 au and are analogous to the contour surface view offered by the CPK method. The relative ESP surface around the G1 and G3 hydrogens clearly indicates contact with the aromatic ring, suggesting and confirming the intermolecular bonding-type CH−π interaction, between the aromatic ring and the sugar moiety. Further illustration of this phenomena is observed by displaying the other surfaces, Figure 10e and f, at isosurface 0.005 au, encompassing a slightly smaller magnitude of electron density, and showing the G3/aromatic ring as still intact, again suggesting strong interaction between the two moieties at this hydrogen position. Moving yet closer to the nuclei, Figure 10g and 10h, at isocontour 0.007 au, now show no contact between any of the hydrogens on the α-face of the sugar with the aromatic ring. Together, this analysis enables an estimation of the extent of the weak interaction, placed between 0.002 au and 0.007 au or between 1.2 and 4.4 kcal/mol.

Conclusions

Compound **2** was chosen as a model for investigating the nature of intramolecular interactions between four different monosaccharides (glucose, galactose, *N*-acetyl-glucosamine, and *N*-acetyl-galactosamine in **2c**−**f**, respectively) and the phenyl ring appended to the ether side chain. NMR studies supported by molecular modeling and ab initio calculations led to several conclusions concerning the nature of the sugar/aromatic type interaction and the effect on the 3D structure of the molecules involved.

As revealed by NOE measurements, the phenyl ring and the monosaccharides appear to adopt a well-defined orientation, which is consistent across the series examined. NOESY spectra of **2c** did not depend on the solution pH in the range 3.6−7, and the intensity and position of the observed cross-peaks were identical for the corresponding methyl ester,[30] thus ruling out a major contribution of the carboxy group to the establishment of the molecular conformation. The sugar−aromatic distance was evaluated by rigorous NMR studies supported by molecular modeling and found to be constant throughout the series and independent of the nature of the sugar. Since all the aromatic proton signals strongly overlap in the spectra examined, the analysis of the NOE contacts may have missed some fine details of the possible differences in sugar−aromatic orientation. However, it is worth noting that the above conclusion is supported both by the molecular modeling results and by DFT calculations (B3LYP/DZV(2d,p)) not described here.[30] Thus it appears that, given the opportunity, persistent intramolecular aromatic sugar interactions are established in small molecules and can significantly influence the *overall* molecular shape. This is not to say that all the remaining internal degrees of freedom can be frozen or fully controlled by exploiting the formation of a sugar/aromatic intramolecular complex: indeed the work reported here shows that the mobility of the ether side chain in compounds **2c**−**f** does display some variation as a function of the carbohydrate involved. Furthermore, the dynamic behavior of the side chain in the unglycosylated compounds **2a**−**b** does not differ significantly from that observed in the glycosylated set (see Table 1), whereas our previous studies show that it is significantly influenced by the configuration of the included stereocenter.[18a] Rather it seems that the soft nature and low directionality of the sugar/aromatic interaction enables the structure to be resiliently conserved despite other molecular changes. This, in turn, allows the molecule to maintain its global shape. Such behavior can have important implications in the design of structural mimics of oligosaccharides, as we have found serendipitously in the course of previous studies.[14] This type of mechanism could also be at work in the formation of sugar/lectin complexes and fits well the current understanding of the sugar code as an exquisitely tunable one.[31]

Ab initio calculations suggest an energy in the range of 5 kcal/mol for the sugar/aromatic type interaction. This estimate is in agreement with previously reported values.[10−12] Aromatic−sugar contacts have been previously suggested as driven by CH/π,[10−12] van der Waals,[32] or hydrophobic[33] interactions. More generally, the nature and origin of weak interactions in molecules and crystals are the subject of intense and heated debate.[28] In our previous studies[14] we observed that **1b** (Figure 1), the cyclohexyl analogue of **1a**, had a dynamic behavior very different from that of **1a** and did not display any NOE contacts between the cyclohexyl ring and the carbohydrate portion. Visual inspection of the combined ab initio and density functional

(29) Gavezzotti, A. *J. Am. Chem. Soc.* **1983**, *105*, 5220−5225.
(30) Terraneo, G. Ph.D. thesis, Universita' di Milano, Milano, Italy, 2006.
(31) Gabius, H-J.; Siebert, H.-C.; André, S.; Jiménez-Barbero, J.; Rüdiger, H. *ChemBioChem* **2004**, *5*, 740−764.
(32) Toone, E. J. *Curr. Opin. Struct. Biol.* **1994**, *4*, 719−728 and references therein.
(33) Elgavish, S.; Shaanan, B. *Trends Biochem. Sci.* **1997**, *22*, 462−467.

model of **2c** in Figure 9 shows that the apolar hydrogen atoms on the α-face of the sugar are in proximity with the van der Waals surface of the aromatic residue, and the molecular orbital representation of the HOMO-6 orbital (Figure 9b) indicates a stabilizing bonding interaction between the sugar CH vectors and the aromatic ring. The electrostatic potential of the aromatic ring (Figure 10) appears slightly more positive on the face which is in contact with the sugar, suggesting that the sugar CHs are acting as weak acids on the π-electron cloud of the phenyl ring. Preliminary results obtained from NOESY spectra of **2c** in CD_3OD solution[30] appear to indicate a weakening of the NOE contacts and possibly suggest a role for hydrophobic packing in defining the molecule conformation. In light of the above observations, the interaction between the aromatic ring and the carbohydrate fragment in **2c**−**f** could be best described as a relatively extended contact between the two hydrophobic surfaces involving significant electron overlap.

Aromatic amino acid residues are often present in the binding sites of carbohydrate-binding proteins. The resulting carbohydrate−protein complexes are characterized by a placement of the sugar in a roughly parallel orientation relative to the plane of the aromatic ring, similar to the one observed in compounds **2c**−**f**. These molecules, therefore, can serve as simple and tunable models for study of the nature of carbohydrate−aromatic interactions and the implications in the molecular recognition of carbohydrates.

Experimental Section

Computational Methods. A. Molecular Mechanics Calculations and Generation of the Three Population Sets. All calculations were performed using the MacroModel/Batchmin 8.5[34] package (Maestro-version 6.0) and the AMBER* force field. Kolb's parameters were used for the hydroxyacid moiety.[35] Bulk water solvation was simulated by using MacroModel's generalized Born GB/SA continuum solvent model,[36] which treats the solvent as an analytical continuum starting near the van der Waals surface of the solute, and uses a dielectric constant (ϵ) of 78 for the bulk water and 1 for the molecule.

An initial conformational search was carried out using 10 000 steps of the usage-directed MC/EM procedure following previously established protocols.[37] Extended nonbonded cutoff distances (a van der Waals cutoff of 8.0 Å and an electrostatic cutoff of 20.0 Å) were used. The output structures were clustered based on the improper dihedral angle descriptor χ(C(O)−Cα−CHDC4−CHDH4), which describes the side-chain conformation. Three clusters were obtained, corresponding to Conf. A, Conf. B, and Conf. C of Figure 3. Conformations with values of the χ improper dihedral angle in the range 180° ± 60° were clustered in the A family, those with χ in the range 60° ± 60° were clustered in the B family, and those with χ in the range of −60° ± 60° were clustered in the C family. In this way all the space was investigated in all directions, without loosing information and no conformers were outside of the defined staggered conformations (A, B, or C). The above criteria were used also to define the A, B, and C populations of the following sets.

The cluster leaders were used as starting structures for the dynamics simulations.

The first conformer set (MD set) was obtained by running three MC/SD runs of 5 ns each, starting from three conformers corresponding to the three idealized geometries of Figure 3 (Conf. A, Conf. B, Conf. C, 15 ns total simulation) and saving 5000 structures for each run, using an electrostatic cutoff of 20 Å, a van der Waals cutoff of 8.0 Å, and a hydrogen bond cutoff of 4 Å. The simulation was performed at 300 K, with a dynamic time step of 1.5 fs; the Monte Carlo acceptance ratio was less 4%, and each accepted MC step was followed by an SD step. Structures were sampled every 1 ps and saved for later evaluation, monitoring both energetic and geometrical parameters. As expected, the MC/SD protocol allowed for fast interconversion of the starting conformers, so that the populations obtained from the individual runs appeared to be converged within 5% and the interproton distances differed by no more than 0.1 Å.

The second set (MM set) was obtained by Multiple Minimization of the MD set (i.e., of the total 15 000 snapshots saved during the molecular dynamics simulations) using the same force field and cutoff applied in the MD set. The third set of conformations (Fitting set) was obtained by fitting the NOE data, using a procedure that we have recently described.[21] Briefly, the MM set of conformations was clustered in the three clusters A, B, and C of Figure 3 using the χ(C(O)−Cα−CHDC4−CHDH4) descriptor as detailed above. Interproton distances and NOE intensities (%) were derived for each cluster using a full matrix relaxation calculation extended to all the members of each cluster. An initial value was assigned to the weight of each cluster, and the conformer distribution was estimated by fitting the Noeprom-based % overall NOE intensities to the experimental ones.

B. Ab initio Calculations. Full geometry optimizations, including structural, orbital, and various electrostatic property analysis, were carried out using the GAMESS software package.[38] Structural computations were performed using hybrid density functional theory (HDFT). The HDFT method employed Becke's three-parameter functional in combination with a nonlocal correlation provided by the Lee−Yang−Parr expression with both local and nonlocal terms, B3LYP.[39] The DZV(2d,p) basis set[40] was used. Full geometry optimizations were performed and uniquely characterized by calculating and diagonalizing the matrix of energy second derivatives (Hessian) to determine the number of imaginary frequencies (0 = minima; 1 = transition state). All energetics and Cartesian coordinates for the optimized geometries are available in the Supporting Information.

From the fully optimized structures, single-point energy calculations were performed using the MP2 dynamic correlation treatment for further analysis of energetics and properties. The optimal method was determined using several levels of theory to establish self-consistency in terms of basis sets as well as effects of dynamic correlation. Additionally, these methods have been previously shown to be reliable for the types of compounds considered here.[41] Molecular orbital contour plots and electrostatic potential plots, used as an aid in the discussion of the results, were generated using the program 3D-PLTORB[42] and GAMESS, respectively, and depicted using QMView[43] and MOLEKEL,[44] respectively.

NMR Methods. NMR spectra were recorded in a 5 mM D_2O solution using Bruker AVANCE 400, at 300 K. 1D, as well as 2D, COSY, HSQC, and NOESY experiments were recorded using the

(34) Mohamadi, F.; Richards, N. G. J.; Guida, W. C.; Liskamp, R.; Lipton, M.; Caufield, C.; Chang, G.; Hendrickson, T.; Still, W. C. *J. Comput. Chem.* **1990**, *11*, 440−467.
(35) Kolb, H. C.; Ernst, B. *Chem.−Eur. J.* **1997**, *3*, 1571−1578.
(36) Still, W. C.; Tempzyk, A.; Hawley, R.; Hendrickson, T. *J. Am. Chem. Soc.* **1990**, *112*, 6127−6129.
(37) (a) Brocca, P.; Bernardi, A.; Raimondi, L.; Sonnino, S. *Glycoconjugate J.* **2000**, *17*, 283−299. (b) Bernardi, A.; Raimondi, L.; Zuccotto, F. *J. Med. Chem.* **1997**, *40*, 1855−1862.
(38) Schmidt, M. W.; Baldridge, K. K.; Boatz, J. A.; Elbert, S. T.; Gordon, M. S.; Jensen, J. H.; Koseki, S.; Matsunaga, N.; Nguyen, K. A.; Su, S.; Windus, T. L.; Dupuis, M.; Montgomery, J. A. *J. Comput. Chem.* **1993**, *14*, 1347−1363 (www.msg.ameslab.gov/GAMESS/GAMESS.html).
(39) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648−5652.
(40) Dunning, T. H. *J. Chem. Phys.* **1970**, *53*, 2823−2833.
(41) (a) Milet, A.; Korona, T.; Moszynski, R.; Kochanski, E. *J. Chem. Phys.* **1999**, *111*, 7727−7735. (b) Perez-Jorda, J. M.; San-Fabian, E.; Perez-Jimenez, A. J. *J. Chem. Phys.* **1999**, *110*, 1916−192. (c) Guerra, C. F.; Bickelhaupt, F. M. *J. Chem. Phys.* **2003**, *119*, 4262−4273. (d) Zhang, Y.; Pan, W.; Yang, W. *J. Chem. Phys.* **1997**, *107*, 7921−7925.
(42) *3D-PLTORB*; San Diego, 3D version, 1997.
(43) Baldridge, K. K.; Greenberg, J. P. *J. Mol. Graphics* **1995**, *13*, 63−68.
(44) Flükiger, P.; Lüthi, H. P.; Portmann, S.; Weber, J. MOLEKEL 4.0; Swiss Center for Scientific Computing: Manno, Switzerland, 2000.

standard pulse sequences. NOESY experiments were carried out using a mixing time of 800 ms. Exclusive NOE contacts for both compounds were identified and integrated from the NOESY spectrum relative to the diagonal peak of the corresponding rows.

Cross-peaks from NOESY experiments were integrated using the MestReC software[45] obtaining the related NOE intensities (%) as the ratio between the diagonal peak and the cross-peak. Their intensities were compared with those estimated by Noeprom,[22] as described below. The theoretical NOE intensities were generated from the computational models using the Noeprom software[22] based on a full matrix relaxation approach[46] and using a rigid isotropic model.[47] The NOE volumes reported in Table SI-7 (Supporting Information) correspond to a 400 MHz spectrometer, using a 800 ms mixing time and a correlation time (τ_c) of 200 ps. The MD predictions were obtained as averages of all the conformations collected during the 15 ns dynamic runs The MM predictions were obtained as Boltzmann-weighted averages of all the conformations obtained by Multiple Minimization of the dynamics output, within 3 kcal/mol from the global minimum. The NOE Fitting prediction was obtained by using the Simplex algorithm.[48] An initial value was assigned to the weight of a given cluster (Conf. A, Conf. B, and Conf. C) and an estimation of the average % NOE intensities calculated; then, the deviation between the calculated and experimental values was iteratively minimized (up to 10 000 steps), and thus, the best-fitting population weights were obtained.

The Noeprom results were analyzed to estimate the % NOE intensity errors, which are collected in Table SI-7 of the Supporting Information, together with the associated population distributions and distances (Å).

1D Experiments. 1D NOE and 1D ROE NMR spectra were recorded in a 5 mM solution of D_2O using Bruker AVANCE 500, at 300 K using the standard selnogp and selrogp sequence from the Bruker library. The NOE buildup was measured at different mixing times, for 1D NOE 0.2, 0.4, 0.6, 0.8, 1 s and for 1D ROE 0.2, 0.3, 0.4, 0.5, 0.6, 0.75 s. The longitudinal relaxation rate (σ_{noe}) and the transversals relaxation rate (σ_{roe}) for each proton pair were calculated as the slope of the line for mixing time vs NOE intesity. The ratio between σ_{noe} and σ_{roe} gives a mathematical function (eq 1) that allows determining the τ_c for a proton pair without knowing the distance (ω_0 is the frequency of the magnet in rad/s).

$$\sigma_{noe}/\sigma_{roe} = (5 + \omega_0^2\tau_c^2 - 4\omega_0^4\tau_c^4)/(5 + 22\omega_0^2\tau_c^2 + 8\omega_0^4\tau_c^4) \quad (1)$$

After knowing the τ_c, σ_{noe}, and σ_{roe} for a proton pair, the interproton distance for this proton couple could be calculated. The density spectral function ($J(n\omega_0)$) correlates the molecular motion and the distance, and so, by using the eqs 2−4 the distance could be deduced.

$$J(n\omega_0) = \tau_c /1 + (n\omega_0\tau_c)^2 \quad (2)$$

$$\sigma_{noe} = C[6J(2\omega_0) - J(0)]$$

$$\sigma_{roe} = C[2J(0) + 3J(\omega_0)] \quad (3)$$

$$C = \left(\frac{\mu_0}{4\pi}\right)^2 \frac{\hbar^2\gamma_H^4}{10}\langle r^{-3}\rangle^2 \quad (4)$$

Acknowledgment. This project was supported in part by the European Union under Contract HPRN-CT-2002-00173 (Glycidic scaffolds Network). K.K.B. acknowledges the Swiss National Science Foundation (SNF-200021-107979) for support of this work. We thank Dr. Carlo Gatti for the many helpful comments and discussions.

Supporting Information Available: Experimental procedures, compound characterization, ^{1}H NMR and ^{13}C NMR of compounds **2a**−**f**, tables of NOE contacts and conformational analysis results for compounds **2a**−**f**, Cartesian coordinates and absolute energies from QM calculations of compound **2c** (*gt*, *tg*, and *gg*). This material is available free of charge via the Internet at http://pubs.acs.org.

JA066633G

(45) Cobas, J. C.; Sardina, F. J. *Concept. Magn. Reson*. **2003**, *19*, 80−96 (www.mestrec.com).
(46) Borgias, B. A.; James, T. L. *Methods Enzymol*. **1989**, *176*, 169−183.
(47) Lipari, G.; Szabo, A. *J. Am. Chem. Soc*. **1982**, *104*, 4546−4559.
(48) Dantzig, G. B. *Linear Programming and Extension*; Princeton University Press: 1998; Chapter 5: The Simplex Method.