Ab Initio Study of Lowest-Energy Conformers of Lewis X (Le^x) Trisaccharide

Gábor I. Csonka*

Department of Inorganic Chemistry, Budapest University of Technology, H-1521 Budapest, Hungary

Carlos P. Sosa

Silicon Graphics Inc., 655 E. Lone Oak Drive Eagan, Minnesota 55123

Imre G. Csizmadia

Department of Chemistry, University of Toronto, Toronto, Ontario, Canada M5S 3H6

Received: September 30, 1999; In Final Form: January 4, 2000

This paper presents the first ab initio conformational study of the Gal-β-1,4-[Fuc-α-1,3]-GlcNAc-β-OCH3 and Gal-β-1,4-[Fuc-α-1,3]-GlcNAc-β-OH trisaccharides (Lewis x) in the gas phase. Their lowest-energy conformers were selected first by the MM2*-SUMM conformational search technique. MM2* relative energies do not follow the same order for the two similar compounds. The molecular geometries and energies of the lowest-energy rotamers (7 of the acetal and 11 of the hemiacetal) were further analyzed at the HF/6-31G(d) level of theory. The ab initio method yields the same energetic order for the rotamers of the two molecules with considerably larger energetic differences for the first 7 rotamers: the MM2* method provides 0.3−0.5 kcal/mol, whereas the HF/6-31G(d) method provides 4.5 kcal/mol. In the most stable MM2* structures the hydrogen-bonded chains of galactose (in counterclockwise direction) and fucose (in clockwise direction) are not connected. The Gal(O6H) is a hydrogen bond donor (in clockwise direction) to the O3 glycosidic oxygen of GlcNAc. The Fuc(O2H)→(O═C)GlcNAc interaction connects the fucose and GlcNAc. In contrast, the most stable HF/6-31G(d) structure has a long chain of seven ordered hydrogen bonds including a Gal(O6H)→(O3)Fuc interaction (with clockwise hydrogen-bonded chain in galactose and fucose). The torsion angles for Fuc-α-1,3-GlcNAc and Gal-β-1,4-GlcNAc glycosidic bonds agree well in the solid, liquid, and gas phases. For example there is a rather good overlap between the GlcNAc moiety of one of the X-ray structures and the most similar HF/6-31G(d) structure. The stacking of the fucose and galactose moieties is similar. The orientations of the hydroxyl groups are usually different, as they are influenced by intramolecular hydrogen bonding in the gas-phase Hartree−Fock structure versus intermolecular hydrogen bonding in the solid-phase X-ray structure.

1. Introduction

Carbohydrates play a vital role in the process of cell recognition. An example of cell recognition may be seen in the case of sialyl Lewis^x (sLe^x, NeuAc-α-2,3-Gal-β-1,4-[Fuc-α-1,3]-GlcNAc) with its antigen properties. This tetrasaccharide is found on the terminus of glycolipids that are present on the surface of human white blood cells. A Ca^{2+}-dependent sLe^x−E−selectin recognition occurs after tissue injury and leads to acute and chronic inflammations.[1] Oligosaccharides such as sLe^x are recognized in the cell by their structural properties. Several studies have been reported on the conformation[2,3] as well as structural features of sLe^x.[4,5] The 2-, 3-, and 4-hydroxyl groups of α-L-fucose, the 4-, and 6-hydroxyl groups of D-galactose, and the carboxylate group of sialic acid play an essential role in the E-, P-, and L-selectin recognition of sLe^x.[6] Molecular mechanics (MM), molecular dynamics, and NMR nuclear Overhauser effects (NOE) studies of the free tetrasaccharide agree that the rigid Le^x (Gal-β-1,4-[Fuc-α-1,3]-GlcNAc) part of the molecule is connected to sialic acid moiety through a flexible linkage.[6] NMR data alone are frequently insufficient for the determination of the conformation of an oligosaccharide. Also, NMR NOE experimental values represent an average of the conformational space of the more or less flexible oligosaccharides, and thus the modeling work is essential to discover the elements of the conformational space. Together with suitable modeling studies, NMR NOE experiments are powerful tools in the determination of the bound bioactive conformation of oligosaccharides.[7,8]

The desialylated trisaccharide Le^x (Gal-β-1,4-[Fuc-α-1,3]-GlcNAc-β) is also of considerable biological interest. It is a stage-specific embryonic marker in cell−cell interactions and a signaling molecule for some host−pathogen recognition. Also, its relatively rigid structure has been the subject of many experimental studies, including NMR[3,6,9] and X-ray[10] diffraction. The X-ray results provided two different structures in the asymmetric unit. Both structures correspond approximately to the global lowest-energy conformation with slight, about 10°, variation in the torsion angles of the glycosidic linkages. This supports the limited flexibility predicted by MM studies. Three hydrogen bonds were observed between the two Le^x molecules in the asymmetric unit.[10] This phenomenon may provide some hints about the formation of glycosidic clusters at the cell's surface in carbohydrate-mediated cell−cell interactions,[11] and the different hydroxyl and hydroxymethylene orientations provide information about the possible hydrogen bonds.

10.1021/jp9935034 CCC: $19.00
Published on Web 03/22/2000

According to previous theoretical results there is a monotonic change in the relative energies of various conformers of monosaccharides (e.g., D-glucose) at Hartree−Fock (HF) level of theory as the basis set quality increases from 3 to 21G to cc-pVQZ.[12,13] The HF/3-21G results indicate that the $^1C_4$ chair hydroxyl and hydroxymethylene rotamers of β-D-glucose are about 7−8 kcal/mol more stable than the $^4C_1$ rotamers.[13] This is about a 14 kcal/mol error compared with the results of the most expensive MP2-based composed calculations.[12] The HF/6-31G(d) and cc-pVDZ results provide quite good relative energies that are close to the results of the most expensive MP2 calculations.[14] Further increase of the basis set with the HF level of theory up to cc-pVTZ or cc-pV^{T}QZ cause an opposite error. that is, destabilizing the $^1C_4$ conformers (up to 6−10 kcal/mol).[12] The origin of the above-mentioned errors is the following: The HF method supplemented with 3-21G basis set provides a rather ionic HO⇆H interaction and leads to an exaggerated stabilization of the hydrogen bonds. The HF method supplemented with good-quality basis sets tends to underestimate the HO→H interactions, because it overconcentrates the electron density around the atoms and in the normal covalent bonding regions and underconcentrates the electron density in the other regions of space. This latter error is inherent to the HF method, whereas the former error is a typical basis set error. The rela tive energies of $^1C_4$ and $^4C_1$ conformations are especially sensitive to the above-mentioned errors because of the differences in the nature of the HO⇆H interactions. In the $^4C_1$ conformation of D-glucose the 1−2 equatorial−equatorial interactions are dominant, and these interactions are relatively weak and constant (because of the fixed, large HO⇆H distances), quasi-independent of the method. In the $^1C_4$ conformation the 1−3 axial−axial interactions might be extra strong if a method erroneously provides rather short HO⇆H distances, and distorts the flexible ring torsion angle. The monotonic change in the relative stabilities between the two opposite errors as a function of basis set quality provides an opportunity to find a basis set for which the basis set truncation error compensates for the inherent deficiencies of the HF method. Using a double-ζ quality basis set [e.g., 6-31G(d) or cc-pVDZ] is close to the optimal choice for this type of energetic order at the HF level of theory.[12−17]

Barrows et al.[12] showed that including electron correlation at the MP2/6-31G(d), CCSD/6-31G(d), and MP2/cc-pVDZ levels of theory provides rather poor energetic order for the relative energies of $^1C_4$ and $^4C_1$ conformations of β-D-glucose. This is expected because the introduction of electron correlation increases the stability of the $^1C_4$ ring and the results become worse than the HF/6-31G(d) results by about 6−7 kcal/mol. Considerably larger basis sets (cc-pVTZ or larger) are required at the MP2 level of theory to approach the HF/6-31G(d) results or the theoretical limit for the relative energies.[12,14] Even this level of theory is not satisfactory and further basis set and correlation corrections are necessary.[12,14] This behavior limits the applicability of MP2 calculations for conformational studies of aldohexapyranoses. The density functional theory (DFT) methods (B3LYP or B3P) lead to results similar to those at the MP2 method when the 6-31G(d) basis set is used.[13] However, addition of the diffuse functions [e.g., 6-31+G(d) or aug-cc-pVDZ] improved the DFT relative energies considerably.[13] This is because the diffuse functions provide a space for the electrons far from the nuclei. Therefore the long-range parts of the correlation and exchange functionals work better for the OH interactions. Similar behavior was found for the weak interactions with the B3LYP functional.[18] The inclusion of the exact exchange into the functional (B3P or B3LYP methods) improves the agreement between the DFT and MP2 or composite results considerably.[13]

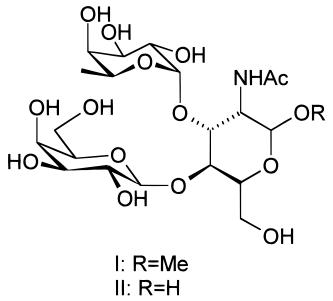

Figure 1. Schematic representation of the molecules studied in the present paper.

HF/6-31G(d) results have been used for the parametrization of MM methods for saccharides in the gas phase.[19−22] We note that carbohydrates are rather difficult tests for MM methods because they have densely packed highly polar functional groups, and the conformational energies depend on stereoelectronic effects. In this respect our earlier results show that the MM2* method provides good qualitative results for the lowest-energy rotamers of monosaccharides, but an energetically compressed conformational space with incorrectly ordered rotamers in the higher-energy region.[15,17] Damm et al.[19] showed that using all the 144 carbohydrate conformers, optimized at HF/6-31G(d) level, for the OPLS-AA torsional parameters optimization gave poor results for the low-energy structures. Thus only the 44 lowest-energy HF/6-31G(d) structures were used to fit the torsional parameters to resolve the problem. This supports our earlier observations that because of the inherent problems the MM methods are unable to yield correct energy differences for a larger energy window for saccharides. Barrows et al.[14] summarized the performance of the best MM methods for calculating the relative energies of D-glucose rotamers and they have found that HF/6-31G(d) method is clearly superior compared with any MM parametrization.

The earlier results show that the calculated HF/6-31G(d) equilibrium molecular geometries might noticeably differ from equilibrium molecular geometries calculated at the MP2 or GGA-DFT levels of theory. However, these geometry variations result in small changes in the relative energies. It was observed that reoptimizations of the HF/6-31G(d) equilibrium geometries of D-glucose rotamers with various correlation methods and basis sets cause only 0.0−0.2 kcal/mol changes in the relative energies.[13,14] However, changing the method and the basis set might provide 5−10 kcal/mol variation in the relative energies calculated with the same geometry.[13,14]

For the present conformational study we consider $2 \times 3^{15} = 28697814$ possible rotamers of Le^x, as there are 15 threefold rotating and one twofold rotating groups in the molecule. We study the conformational space of two molecules, (**I**) and (**II**) shown in Figure 1. We use the MM2*-SUMM method for the exploration of the conformational space with fixed ring puckering. The most stable $^4C_1$ pyranose ring form for the D-glucose and D-mannose and $^1C_4$ ring form for the L-fucose were used as a starting point in this study. Because the energy spectrum turned out to be quite dense we reoptimized the geometries of the lowest-energy structures obtained by the MM2*-SUMM method within a 1 kcal/mol energy window by the HF/6-31G-(d) method. We use the earlier results obtained for the conformational space of L-fucose, D-galactose, and D-mannose. The experimental structures are also analyzed.

TABLE 1: MM2* Energies (*E*), Relative Energies (Δ*E*), and C−C−O−R Torsion Angles in Most Stable Conformers of MM2*-SUMM Conformational Search for (I)

			C−C−O−R torsion angles (°) in GlcNAc[a]						C−C−O−R torsion angles (°) in Gal[a]						C−C−O−R torsion angles (°) in Fuc[a]			
no.	*E* (kJ/mol)	Δ*E* (kcal/mol)	τ1	τ2	τ3-F[b]	τ4-G[c]	τ5	τ6	τ1[d]	τ2	τ3	τ4	τ5	τ6	τ1[e]	τ2	τ3	τ4
1	−425.75	0.00	179.0	75.9	148.0	−106.9	−61.1	56.0	175.2	−174.2	−170.1	−168.0	53.8	−52.3	164.4	−96.1	170.1	167.9
2	−425.30	0.11	179.0	74.1	145.3	−101.7	−62.0	56.0	172.7	−175.4	−170.3	−165.6	**−57.7**	**−42.5**	162.1	**−89.8**	171.2	167.3
3	−425.01	0.18	179.0	76.0	148.0	−106.4	**55.4**	**−55.4**	175.0	−173.5	−170.1	−168.2	53.8	−53.8	164.5	−97.8	170.1	167.9
4	−424.86	0.21	179.0	74.1	145.1	−101.1	**54.3**	**−55.2**	172.5	−174.8	−170.3	−165.7	**−57.6**	**−42.8**	162.2	−90.6	171.1	167.3
5	−424.84	0.22	177.6	**44.0**	**−179.3**	−109.4	−60.3	54.6	175.0	−176.4	−170.9	−165.6	**−56.5**	**−44.9**	**−172.9**	**−56.2**	170.8	167.0
6	−424.37	0.33	178.4	76.0	153.5	−106.3	−60.8	55.5	175.3	−175.0	−170.2	−167.9	52.7	−55.8	165.1	**−155.9**	171.9	168.7
7	−424.31	0.34	179.1	76.1	151.3	−105.4	−60.4	55.1	170.0	**−52.4**	**−36.1**	**49.6**	**−76.9**	**155.4**	166.8	**−102.4**	**162.0**	169.7
8	−424.22	0.37	177.6	**43.9**	**−179.2**	−109.4	**56.2**	**−55.3**	175.2	−175.8	−170.9	−165.8	**−56.5**	**−45.1**	**−172.8**	**−56.2**	170.5	167.0
9	−424.14	0.38	178.4	76.0	153.4	−105.8	**55.6**	**−55.8**	175.3	−174.3	−170.2	−168.0	52.8	−57.0	164.9	**−157.0**	171.6	168.7
10	−423.45	0.55	179.0	76.1	151.2	−105.0	**54.6**	**−54.5**	170.1	**−52.7**	**−36.2**	**50.0**	**−77.1**	**155.7**	166.8	**−104.6**	**162.1**	169.6
11	−422.50	0.78	179.9	76.4	151.4	−110.8	**155.2**	**65.1**	**168.7**	**−52.7**	**−35.7**	**50.8**	**−77.5**	**154.1**	166.7	−101.6	**161.2**	169.9
12	−422.37	0.81	**172.7**	**−75.5**	149.6	−106.0	−60.8	54.7	177.0	−175.9	−170.3	−165.2	**−58.0**	**−40.8**	**147.1**	**56.3**	**−89.0**	170.4
13	−422.30	0.82	175.8	**−122.3**	**178.0**	−108.5	−59.4	54.9	177.5	−176.9	−170.7	−165.3	**−57.1**	**−44.6**	**−174.0**	**−54.5**	170.9	167.4

[a] The τ*x* torsion angles are defined as C(*x*+1)−C*x*−O−R, where R = C or H, except the τ2 torsion angle in GlcNAc is defined as C3−C2−N−H, the τ5 torsion angle is defined as O−C5−C6−O, the τ6 torsion angle is defined as C5−C6−O−H (cf. Figure 2). To facilitate the overview the torsion angles differing more than ±6° from the value in the first row are set to bold. [b] Fuc-α-1,3-GlcNAc torsion angle: τ(C1−O1[1,3]−C3−C4). [c] Gal-β-1,4-GlcNAc torsion angle: τ(C1−O1[1,4]−C4−C5). [d] Gal-β-1,4-GlcNAc torsion angle: τ(C2−C1−O1[1,4]−C4). [e] Fuc-α-1,3-GlcNAc torsion angle: τ(C2−C1−O1[1,3]−C3).

The principal aims of this work are to identify the stabilizing factors in the lowest-energy conformers of Le^x, and to provide an ab initio gas-phase energetics. These ab initio HF results may serve as a calibration for future MM developments, and they can serve as a starting point for the considerably more expensive correlation energy studies. The gas-phase energies also serve as starting points for various solvation and receptor binding energy calculations.

2. Computational Procedure

Conformational Search. The search for stable conformers in the conformational space of the selected molecules (cf. Figure 1) was carried out using the MacroModel 4.5 program package.[23] MM2* (MacroModel) is a variant of the original MM2.[24] The most important difference is in the electrostatic equation. A previous comparison of a series of MM methods has shown that the accuracy in relative conformational energies is apparently equal for MM2*, MM2(91), and MM3(92).[25]

The conformational searches were carried out with a particularly efficient systematic unbounded multiple minimum search technique (SUMM)[26] that is available in MacroModel. During the conformational searches the puckerings of the pyranose rings were not changed. The search was limited to the various rotamers of the freely rotatable bonds and to 4000 steps. Then a new conformational search was started from the global minimum limited to 2000 steps. The two resulting conformational spaces were compared, merged, and further conformational searches were started until consistent results were obtained.

Ab Initio Methods. The minima obtained by the MM2*-SUMM search were reoptimized at the HF/6-31G(d) level of theory using the Berny algorithm combined with redundant internal coordinates built into the GAUSSIAN 94[27] and 98 programs. Because of the differences between the MM2* and HF/6-31G(d) equilibrium geometries, single HF geometry optimizations started from MM2* geometries took about 20−30 days of supercomputer time. For this reason we limited the present study to the first 7 rotamers of **(I)**. For **(II)** it was possible to go further, so the relative energies of the first 12 rotamers originating from the MM2* conformational space were calculated. We note that using transformed ab initio geometries for similar rotamers can save considerable computing effort. The starting geometry of the higher-energy rotamers was generated from the geometry of the lower-energy rotamers. The calculations were performed on Pentium II and Silicon Graphics computers.

3. Relative Stability of Rotamers

MM2* Results. We performed conformational searches for Le^x *O*-methyl glycoside **(I)** and Le^x **(II)**. Table 1 shows the results of the MM2*-SUMM conformational search for compound **(I)**. We show here the energies, relative energies, and torsion angles necessary to identify the various rotamers. The numbering of the torsion angles is shown in Figure 2. Tables 2 and 3 show similar types of results for compound **(II)**. In Table 2 we show the results for **(II)** with fixed anti position for C2−C1−O1−H torsion angle in the GlcNAc moiety (cf. Table 2, first column for torsion angles). In Table 3 we show the results of the unconstrained conformational search for **(II)**. Releasing the first torsion angle in GlcNAc results in new and rather stable rotamers. An interaction between the O1−H group, the NAc group, and the fucose part of the molecule stabilizes considerably (up to 2 kcal/mol) these rotamers compared with the above-mentioned constrained rotamers (cf. Tables 2 and 3). The rotamers in Table 3 are not discussed further because these interactions are specific to **(II)** and these types of rotamers show only partial similarity with the rotamers of **(I)**. Moreover, the above-mentioned interactions are certainly missing from the biologically active glycopeptides. We note that the available X-ray geometry is related to **(I)**.

Our aim is to explore the ab initio HF/6-31G(d) conformational space of **(I)** and **(II)**. We attempt to model the ab initio conformational space of **(I)** using the ab initio conformational space of **(II)**. It is expected that the replacement of the O-methyl group by a hydrogen atom would not influence the relative stabilities of the various OH rotamers. However, the MM2* results contradict this expectation (cf. Figure 3). The corresponding rotamers in Tables 1 and 2 can be identified using the torsion angles. The following pairs can be formed comparing the results in the two tables: 1−1, 2−2, 3−5, 4−7, 5−6, 6−3, 7−4, 8−11, 9−8, 10−10, 11−9, 12−*x*, 13−*x*, *x*−12, where the first number is the conformer number in Table 1 and the second number is the conformer number in Table 2. If a conformer has no pair in the other table the conformer

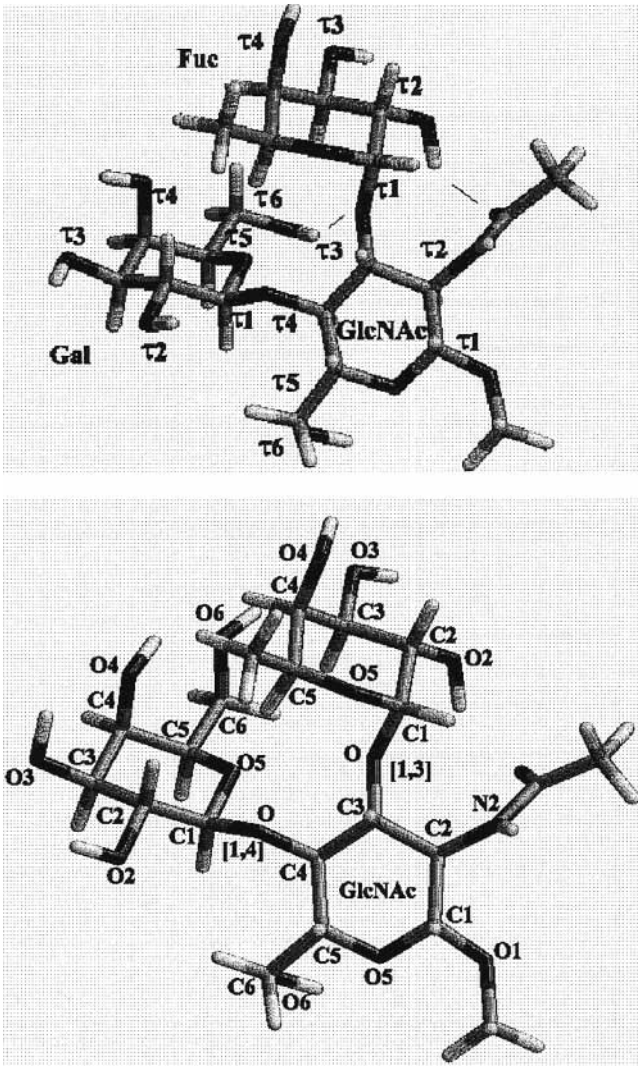

Figure 2. Three-dimensional (3D) structure of the most stable MM2* (upper) and HF/6-31G(d) (lower) conformations of (**I**). The numbering of the torsion angles is shown in the upper structure; the numbering of the atoms is shown in the lower structure. For more detailed investigation a database of 3D structures is presented at http://web.inc.bme.hu/mols/lex.

number is replaced by *x*. The order of the first two most stable rotamers agree well; however, the order of the higher-energy rotamers is quite different (cf. Figure 3). There is no correlation between the MM2* relative energies of **(I)** and **(II)** ($R^2 = 0.274$, cf. Figure 3). On the basis of the earlier experience (e.g., the conformational space of D-galactose or L-fucose),[17,19] it is not expected that the MM2* method would provide a reliable energetic order for the conformational space of Le^x or its analogues. However, such a lack of consistency in the energetic order is surprising. One reason is that the conformational space calculated by MM2* method contains small energy differences for the rotamers of **(I)** and **(II)** (cf. Tables 1 and 2).

Ab Initio Results. The HF/6-31G(d) energetic order for the rotamers of **(I)** and **(II)**, given in Tables 4 and 5, is considerably more consistent than that obtained by MM2*. The HF/6-31G(d) method provides exactly the same energetic order for the first 7 rotamers of the conformational space of the two compounds, and a very good correlation was observed for the energy differences ($R^2 = 0.994$, cf. Figure 3). This supports the view that the conformational space of the larger molecule **(I)** can be predicted from the results obtained for the smaller Le^x **(II)**. The HF/6-31G(d) energetic order in Table 4 is: 7, 1, 2, 5, 6, 3, 4, which differs considerably from the energetic order obtained by MM2*. The energy difference between the most stable and least stable rotamers is more than 4 kcal/mol according to the HF/6-31G(d) results. This is more than one order of magnitude larger than the MM2* value 0.34 kcal/mol for the corresponding 7 rotamers (cf. Tables 1 and 4). This clearly shows that the MM2* conformational space is energetically compressed in comparison with the HF/6-31G(d) conformational space. A comparison of the MM2* and HF/6-31G(d) relative energies of the first four conformers of β-D-galactose shows a difference of 3 kcal/mol between the results obtained by the two methods. Although the HF/6-31G(d) method usually provides excellent relative energies for the monosaccharides it remains to be seen whether it is also good for oligosaccharides. Actually, it is too expensive to obtain better energies by extending the basis set, incorporating electron correlation, or reoptimizing molecular geometries with correlation effects for a molecule of this size.

4. Molecular Geometries

The rotamers of Le^x are uniquely characterized by the 16 C−C−O−R-type torsion angles shown in Tables 1−5. These

TABLE 2: MM2* Energies (*E*), Relative Energies (Δ*E*), and C−C−O−R Torsion Angles in Most Stable Conformers of MM2*-SUMM Conformational Search for (II)

			C−C−O−R torsion angles (°) in GlcNAc[a]						C−C−O−R torsion angles (°) in Gal[a]						C−C−O−R torsion angles (°) in Fuc[a]			
no.	*E* (kJ/mol)	Δ*E* (kcal/mol)	τ1	τ2	τ3-F[b]	τ4-G[c]	τ5	τ6	τ1[d]	τ2	τ3	τ4	τ5	τ6	τ1[e]	τ2	τ3	τ4
1	−401.18	0.00	−165.8	76.4	148.1	−106.7	−61.6	56.0	175.2	−174.1	−170.1	−168.0	53.8	−52.4	164.5	−96.9	170.2	167.9
2	−400.58	0.14	−165.7	74.6	145.3	−101.5	−62.7	56.0	172.7	−175.4	−170.3	−165.6	−57.7	−42.5	162.2	−90.2	171.2	167.2
3	−400.00	0.28	−165.9	76.6	153.5	−106.1	−61.4	55.4	175.4	−174.9	−170.2	−167.9	52.6	−55.8	165.2	−156.2	171.9	168.6
4	−399.81	0.33	−165.6	76.6	151.4	−105.1	−60.8	54.9	170.0	−52.4	−36.1	49.6	−77.0	155.1	166.8	−103.6	162.0	169.6
5	−399.58	0.38	−165.8	76.5	148.0	−106.3	55.7	−56.8	175.0	−173.5	−170.2	−168.2	53.8	−53.9	164.5	−98.5	170.2	167.9
6	−399.34	0.44	−167.8	44.9	−179.6	−109.2	−60.8	54.4	175.1	−176.4	−170.9	−165.7	−56.5	−44.9	−172.9	−56.2	170.7	167.0
7	−399.22	0.47	−165.9	74.6	145.2	−101.1	54.5	−56.5	172.5	−174.7	−170.3	−165.7	−57.7	−42.8	162.2	−91.0	171.1	167.2
8	−398.87	0.55	−165.9	76.5	153.4	−105.8	55.8	−57.2	175.3	−174.2	−170.2	−168.0	52.7	−57.2	164.9	−157.2	171.6	168.7
9	−398.85	0.56	−161.2	76.8	151.4	−110.8	155.1	65.0	168.6	−52.6	−35.7	50.9	−77.5	154.0	166.8	−102.5	161.2	169.9
10	−398.07	0.74	−165.7	76.6	151.3	−104.8	54.8	−55.9	170.0	−52.7	−36.2	50.1	−77.2	155.4	166.9	−105.6	161.2	169.6
11	−397.82	0.80	−168.0	44.9	−179.6	−109.2	56.4	−56.8	175.3	−175.7	−170.9	−165.7	−56.6	−45.1	−172.8	−56.3	170.4	167.0
12	−397.11	0.97	−156.5	76.6	148.4	−106.9	−65.7	179.3	175.4	−173.0	−170.0	−168.2	−52.3	54.0	164.5	−98.4	170.3	167.8

[a] The τ*x* torsion angles are defined as C(*x*+1)−C*x*−O−R, where R = C or H, except the τ2 torsion angle in GlcNAc is defined as C3−C2−N−H, the τ5 torsion angle is defined as O−C5−C6−O, the τ6 torsion angle is defined as C5−C6−O−H (cf. Figure 2). The C2−C1−O−H torsion angle of GlcNAc was fixed in anti position (around 180°). To facilitate the overview the torsion angles differing more than ±6° from the value in the first row are set to bold. [b] Fuc-α-1,3-GlcNAc torsion angle: τ(C1−O1[1,3]−C3−C4). [c] Gal-β-1,4-GlcNAc torsion angle: τ(C1−O1[1,4]−C4−C5). [d] Gal-β-1,4-GlcNAc torsion angle: τ(C2−C1−O1[1,4]−C4). [e] Fuc-α-1,3-GlcNAc torsion angle: τ(C2−C1−O1[1,3]−C3).

TABLE 3: MM2* Energies (*E*), Relative Energies (Δ*E*), and C−C−O−R Torsion Angles in Most Stable Conformers of MM2*-SUMM Conformational Search for (II)

			C−C−O−R torsion angles (°) in GlcNAc[a]						C−C−O−R torsion angles (°) in Gal[a]						C−C−O−R torsion angles (°) in Fuc[a]			
no.	*E* (kJ/mol)	Δ*E* (kcal/mol)	τ1	τ2	τ3-F[b]	τ4-G[c]	τ5	τ6	τ1[d]	τ2	τ3	τ4	τ5	τ6	τ1[e]	τ2	τ3	τ4
1	−409.31	0.00	83.7	−61.2	147.3	−106.6	−60.6	54.6	175.7	−176.0	−170.3	−165.2	−58.1	−41.2	149.0	56.3	−89.2	170.4
2	−408.84	0.11	83.6	−61.4	147.2	−106.0	**55.0**	**−51.0**	175.8	−175.4	−170.3	−165.3	−58.0	−41.4	148.9	56.3	−89.2	170.4
3	−408.56	0.18	**34.3**	**19.7**	**−176.5**	−109.8	−58.8	54.2	175.2	−176.7	−170.9	−165.7	−56.2	−44.9	**−173.8**	**−55.5**	**171.6**	167.2
4	−408.29	0.24	**30.5**	**20.3**	**−176.6**	−109.5	**56.6**	**−52.8**	175.4	−176.0	−170.9	−165.8	−56.2	−45.1	**−173.7**	**−55.5**	**171.6**	167.3
5	−405.37	0.94	84.0	−61.8	149.1	−111.0	**54.7**	**−59.7**	177.4	−174.7	−170.0	−167.5	**53.3**	**−50.4**	148.7	57.3	−91.5	170.9
6	−404.60	1.13	83.9	−61.9	149.0	−110.8	**55.9**	**−51.3**	177.4	−174.0	−170.0	−167.6	**53.1**	**−51.5**	148.5	57.3	−91.5	170.9
7	−404.58	1.13	83.7	−61.5	146.4	−104.3	**54.7**	**−51.1**	174.1	−175.6	−170.6	−165.8	−57.7	−41.3	147.1	50.5	**35.1**	**−73.9**
8	−404.37	1.18	**29.7**	**20.9**	**177.5**	−116.0	**57.1**	**−52.9**	−178.9	−175.5	−171.0	−168.5	**52.9**	**−72.0**	**−171.9**	59.7	**−97.3**	170.8
9	−404.16	1.23	84.0	−61.8	148.9	−110.9	**54.6**	**−59.8**	177.6	−174.9	−170.3	−168.0	54.3	−54.9	147.9	51.0	35.8	−73.6
10	−403.85	1.30	**34.2**	**20.3**	**177.5**	−115.9	−58.4	52.9	−179.4	−176.3	−170.9	−168.3	**53.2**	**−71.1**	**−172.0**	59.5	**−97.5**	170.7
11	−403.60	1.36	83.9	−62.0	148.8	−110.5	**55.8**	**−51.3**	177.6	−174.2	−170.4	−168.1	**54.2**	**−56.0**	147.7	51.1	35.8	−73.6
12	−401.45	1.88	**25.0**	**21.6**	**−175.7**	−108.8	**167.3**	**−172.1**	169.8	**157.7**	−170.2	−165.3	−56.0	−43.8	**−173.2**	**−55.1**	**174.2**	167.1
13	−401.32	1.91	85.8	−61.8	145.3	−103.3	**54.6**	**−51.3**	173.6	−174.9	−170.2	−165.8	−57.1	−43.0	154.3	**−61.3**	**170.9**	169.6

[a] The τ*x* torsion angles are defined as C(*x*+1)−C*x*−O−R, where R = C or H, except the τ2 torsion angle in GlcNAc is defined as C3−C2−N−H, the τ5 torsion angle is defined as O−C5−C6−O, the τ6 torsion angle is defined as C5−C6−O−H (cf. Figure 2). To facilitate the overview the torsion angles differing more than ±6° from the value in the first row are set to bold. [b] Fuc-α-1,3-GlcNAc torsion angle: τ(C1−O1[1,3]−C3−C4). [c] Gal-β-1,4-GlcNAc torsion angle: τ(C1−O1[1,4]−C4−C5). [d] Gal-β-1,4-GlcNAc torsion angle: τ(C2−C1−O1[1,4]−C4). [e] Fuc-α-1,3-GlcNAc torsion angle: τ(C2−C1−O1[1,3]−C3).

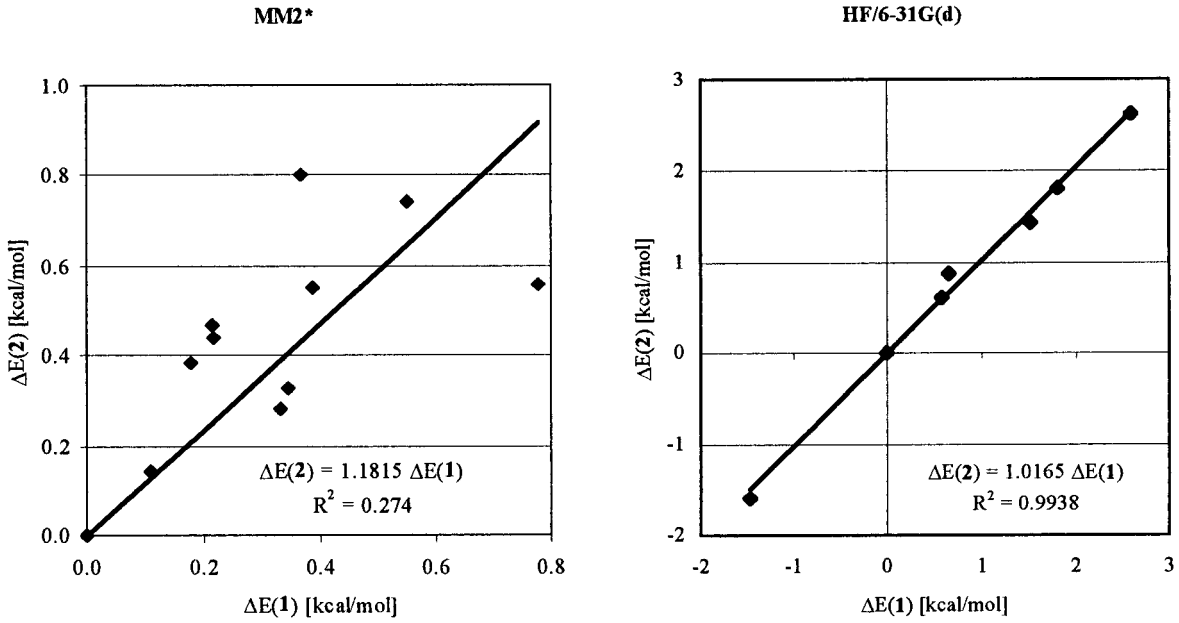

Figure 3. Correlation between the MM2* (left, cf. Tables 1 and 2) and HF/6-31G(d) (right, cf. Tables 4 and 5) energy differences for the lowest-energy rotamers of (**I**) [Δ*E*(**1**)] and (**II**) [Δ*E*(**2**)].

TABLE 4: HF/6-31G(d) Energies (*E*), Relative Energies (Δ*E*), and C−C−O−R Torsion Angles in Most Stable Conformers of MM2*-SUMM Conformational Search for (I)

			C−C−O−R torsion angles (°) in GlcNAc[a]						C−C−O−R torsion angles (°) in Gal[a]						C−C−O−R torsion angles (°) in Fuc[a]			
no.	*E* (Hartree)	Δ*E* (kcal/mol)	τ1	τ2	τ3-F[b]	τ4-G[c]	τ5	τ6	τ1[d]	τ2	τ3	τ4	τ5	τ6	τ1[e]	τ2	τ3	τ4
1	−1994.10548	0.00	174.6	83.2	136.0	−105.5	−63.9	58.7	167.4	−169.4	−171.5	−170.3	60.0	−57.6	160.8	−79.7	165.4	169.7
2	−1994.10456	0.57	174.0	77.2	134.9	−102.6	−64.0	58.0	165.1	−171.3	−170.5	−168.5	**−56.2**	**−51.0**	158.2	−76.8	166.8	168.9
3	−1994.10259	1.81	175.0	84.6	136.4	−106.4	**58.9**	**−58.4**	168.1	−169.8	−172.0	−170.6	60.0	−60.4	160.6	−80.4	164.9	169.8
4	−1994.10134	2.60	174.3	78.4	135.3	−103.2	**58.1**	**−57.2**	165.8	−171.7	−171.0	−168.8	**−56.5**	**−51.3**	157.9	−77.1	166.5	168.8
5	−1994.10444	0.65	170.2	**27.3**	**174.8**	−106.8	−63.3	58.6	**159.3**	−172.5	−170.6	−168.5	**−57.8**	**−48.2**	**−169.0**	**−67.5**	**−170.6**	167.5
6	−1994.10305	1.52	175.0	**99.9**	**150.0**	−108.0	−63.0	58.3	167.7	−171.2	−171.8	−170.3	59.5	−68.8	**170.3**	**−157.3**	**171.7**	168.3
7	−1994.10782	−1.47	175.8	**98.2**	**142.7**	−104.3	−63.3	59.1	157.2	**−48.2**	**−41.1**	**45.8**	**−76.2**	**111.2**	**175.9**	**−96.3**	**157.2**	171.9

[a] The τ*x* torsion angles are defined as C(*x*+1)−C*x*-O−R, where R = C or H, except the τ2 torsion angle in GlcNAc is defined as C3−C2−N−H, the τ5 torsion angle is defined as O−C5−C6−O, the τ6 torsion angle is defined as C5−C6−O−H (cf. Figure 2). To facilitate the overview the torsion angles differing more than ±6° from the value in the first row are set to bold. [b] Fuc-α-1,3-GlcNAc torsion angle: τ(C1−O1[1,3]−C3−C4). [c] Gal-β-1,4-GlcNAc torsion angle: τ(C1−O1[1,4]−C4−C5). [d] Gal-β-1,4-GlcNAc torsion angle: τ(C2−C1−O1[1,4]−C4). [e] Fuc-α-1,3-GlcNAc torsion angle: τ(C2−C1−O1[1,3]−C3).

torsion angles specify the orientations of the hydroxymethylene groups (τ5 and τ6), the NAc group (τ2 in GlcNAc), the glycosidic torsion angles (τ3 and τ4 in GlcNAc and τ1 in galactose and fucose), and all the hydroxyl groups (cf. Tables 1−5).

TABLE 5: HF/6-31G(d) Energies (*E*), Relative Energies (Δ*E*), and C−C−O−R Torsion Angles in Most Stable Conformers of MM2*-SUMM Conformational Search for (II)

			C−C−O−R torsion angles (°) in GlcNAc[a]						C−C−O−R torsion angles (°) in Gal[a]						C−C−O−R torsion angles (°) in Fuc[a]			
no.	*E* (Hartree)	Δ*E* (kcal/mol)	τ1	τ2	τ3-F[b]	τ4-G[c]	τ5	τ6	τ1[d]	τ2	τ3	τ4	τ5	τ6	τ1[e]	τ2	τ3	τ4
1	−1955.07847	0.00	−173.6	84.5	136.1	−105.7	−64.2	59.1	167.5	−169.1	−171.5	−170.4	60.0	−57.8	160.9	−79.8	165.5	169.8
2	−1955.07750	0.61	−174.4	78.6	135.0	−102.7	−64.3	58.3	165.1	−171.1	−170.6	−168.6	**−56.2**	**−51.0**	158.4	−76.8	166.8	168.9
3	−1955.07618	1.44	−173.0	**100.7**	**150.4**	−108.3	−63.3	58.6	167.8	−171.0	−171.8	−170.4	59.4	**−69.7**	**170.5**	**−157.7**	**171.6**	168.3
4	−1955.08101	−1.59	−172.3	**98.3**	**142.8**	−104.4	−63.5	59.4	**157.1**	**−49.6**	**−48.3**	**45.9**	**−76.2**	**111.3**	**175.9**	**−96.4**	**157.2**	171.9
5	−1955.07559	1.81	−174.7	86.0	136.5	−106.6	**58.6**	**−59.3**	168.1	−169.6	−172.0	−170.7	60.0	−60.4	160.8	−80.5	165.0	169.8
6	−1955.07707	0.87	179.8	**28.5**	**174.7**	−106.8	−63.6	58.7	**159.3**	−172.3	−170.7	−168.6	**−57.7**	**−48.3**	**−168.9**	**−67.4**	**−170.9**	167.4
7	−1955.07430	2.61	−175.7	79.9	135.3	-103.2	**57.9**	**−58.1**	165.8	−171.5	−171.0	−168.8	**−56.5**	**−51.3**	158.0	−77.2	166.6	168.8
8	−1955.07385	2.90	−173.9	**101.7**	**151.1**	−109.5	**59.5**	**−60.3**	168.5	−171.4	−172.3	−170.7	58.7	**−72.4**	**170.4**	**−158.3**	170.9	168.6
9	−1955.07807	0.25	−169.1	**97.5**	**142.5**	**−112.0**	**147.7**	**77.1**	**160.0**	**−49.4**	**−40.3**	**48.3**	**−79.3**	**113.0**	**172.6**	**−95.0**	**155.6**	172.0
10	−1955.07709	0.87	−173.8	**99.4**	**143.3**	−105.6	**57.2**	**−58.2**	**157.2**	**−49.8**	**−41.2**	**47.0**	**−77.3**	**112.0**	**175.4**	**−98.2**	**156.9**	171.9
11	−1955.07382	2.92	177.9	**29.6**	**174.5**	−107.8	**59.0**	**−58.0**	**159.1**	−172.5	−171.1	−168.9	**−58.3**	**−48.1**	**−168.5**	**−67.8**	**−171.3**	167.6
9a	−1955.07171	4.24	−168.8	**98.5**	**143.1**	**−112.9**	**147.5**	**78.1**	**159.2**	**−49.2**	**−39.8**	**48.8**	**−79.9**	**117.6**	**171.9**	**−95.6**	**155.6**	**38.6**

[a] The τ*x* torsion angles are defined as C(*x*+1)−C*x*−O−R, where R = C or H, except the τ2 torsion angle in GlcNAc is defined as C3−C2−N−H, the τ5 torsion angle is defined as O−C5−C6−O, the τ6 torsion angle is defined as C5−C6−O−H (cf. Figure 2). To facilitate the overview the torsion angles differing more than ±6° from the value in the first row are set to bold. [b] Fuc-α-1,3-GlcNAc torsion angle: τ(C1−O1[1,3]−C3−C4). [c] Gal-β-1,4-GlcNAc torsion angle: τ(C1−O1[1,4]−C4−C5). [d] Gal-β-1,4-GlcNAc torsion angle: τ(C2−C1−O1[1,4]−C4). [e] Fuc-α-1,3-GlcNAc torsion angle: τ(C2−C1−O1[1,3]−C3).

Torsion Angles in GlcNAc Moiety. In the GlcNAc moiety of the Le[x] the first torsion angle τ1 is set to the t (anti) position for **(I)** and **(II)** (cf. Tables 1, 2, 4, and 5). According to the MM2* results the second torsion angle τ2 (C3−C2−N−H, related to NAc group) in the more stable rotamers falls into two ranges: 74−76° and 44−45° (cf. Tables 1 and 2 and Figure 4). The HF/6-31G(d) values for this torsion angle fall into four different ranges: 98−101°, 85−86°, 78−79°, and 27−29° (cf. Tables 4 and 5 and Figure 4). The MM2* method provides three ranges for the glycosidic torsion angle τ3-F (connecting GlcNAc and fucose): 145−148, 151−153°, and ∼180° (Tables 1 and 2 and Figure 4). The corresponding HF/6-31G(d) values fall within four ranges: 134−136°, 142−143°, 150−151°, and ∼175°(cf. Tables 4 and 5 and Figure 4). The MM2* values for the glycosidic torsion angle τ4-G (connecting GlcNAc and galactose) fall within −101° and −111° (Tables 1 and 2 and Figure 4). The corresponding HF/6-31G(d) values fall within −103° to −112° (cf. Tables 4 and 5 and Figure 4). The distribution of the 11 rotamers of Le[x] (**II**) in the τ2, τ3, and τ4 space can be followed in Figure 5. The values of τ2, τ3, and τ4 calculated from the two X-ray structures[10] are also shown in Figure 5. From these results it appears that the HF/6-31G(d) method provides larger differences between torsion angles in the various rotamers than the MM2* method. This is especially true for the τ2 angle (cf. Figure 5).

The two torsion angles of the hydroxymethylene group, τ5-(O5−C5−C6−O6) and τ6(C5−C6−O6−H), take two typical orientations: τ5 ∼60° and τ6 ∼ −60° or τ5 ∼−60° and τ6 ∼60°, respectively (cf. Tables 1, 2, 4, and 5). For the notation of the three possible values of τ5 we use capital letters: G+, G-, and T for 60°, −60°, and 180°, respectively. (Several papers use *gt*, *gg*, and *tg* notations for this single angle.[6]) For τ6 we use lower case letters; for the two typical orientations of τ5 and τ6 we use the G+ g−, and G− g+ notations. In these rotamers the hydroxymethylene group interacts with the pyranose ring oxygen of GlcNAc and turns away from the galactose moiety. The MM2* and HF results agree in that the G− g+ hydroxymethylene orientation is the most stable (cf. Tables 1, 2, 4, and 5). Only one T g+ rotamer was found within the given energy range (cf. rotamer 11 in Table 1, and rotamer 9 in Tables 2 and 5). In this rotamer the hydroxymethylene group of GlcNAc interacts with the second OH group of galactose. This rotamer is the third most stable rotamer according to the HF/6-31G(d) results.

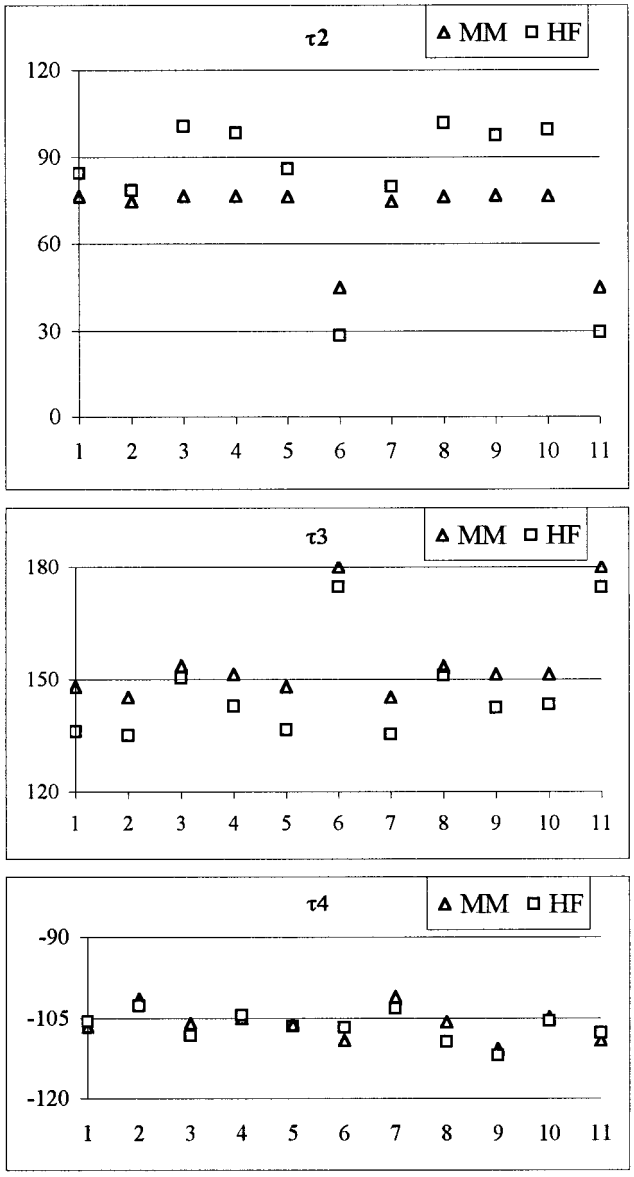

Figure 4. Comparison of the MM2* (MM) and HF/6-31G(d) (HF) τ2, τ3, and τ4 torsion angles (°) calculated for GlcNAc moiety of the 11 rotamers of Le[x](**II**) (cf. τ2, τ3, and τ4 values for GlcNAc in Tables 2 and 5).

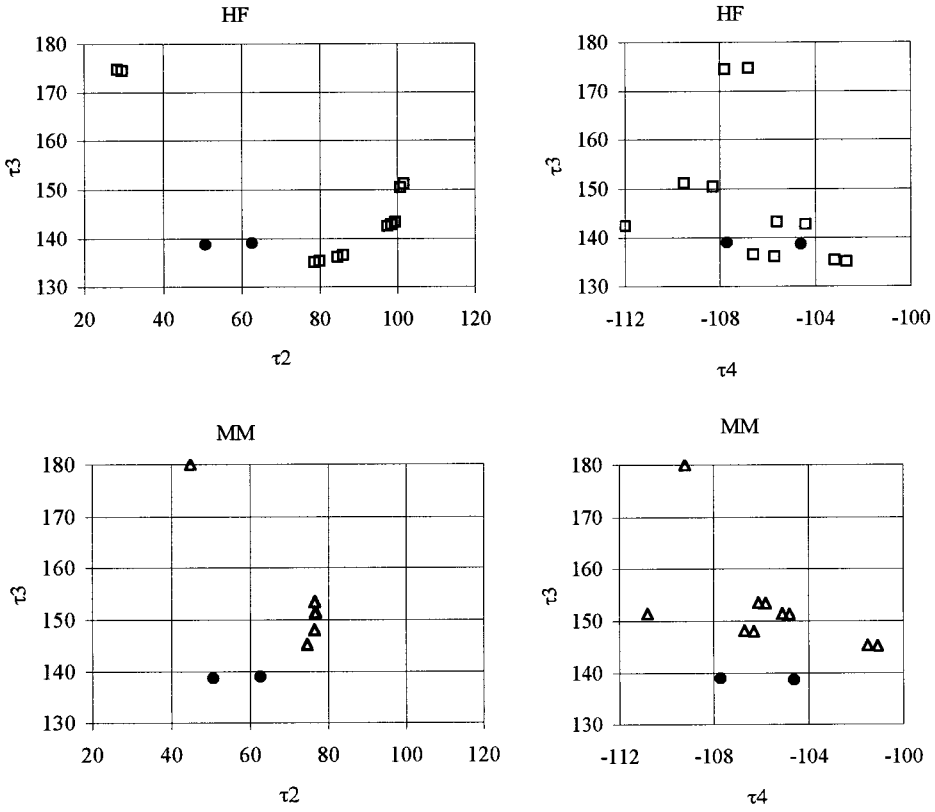

Figure 5. Distribution of the 11 rotamers of Le^x(**II**) in the $\tau 1$, $\tau 2$, and $\tau 3$ (°) conformational space of GlcNAc moiety calculated with HF/6-31G(d) (HF) and MM2* (MM) methods (cf. Tables 2 and 5). The corresponding experimental X-ray values are shown by dots (•) ($\tau 2$ ={50.7°; 62.6°}, $\tau 3$ ={138.7°; 139.0°}, and $\tau 4$ = {−104.6°; −107.7°}).

Torsion Angles in Galactose Moiety. The galactose moiety of the most stable rotamer of Le^x has the following conformation: t t t t G+ g- according to the MM2* method (cf. rotamer 1 in Tables 1 and 2). (For $\tau 5$ we use acapital letter.) In the most stable MM2* structure the orientation of the hydrogen-bonded chain of the galactose points to the counterclockwise direction. In this rotamer the hydroxymethylene group of the galactose interacts as a hydrogen bond donor with the oxygen atom (O1,3) of the Fuc-α-1,3-GlcNAc link (cf. Figure 2.). The HF/6-31G(d) method provides a quite different rotamer for the galactose moiety of Le^x, t g− g− g+ G- g++, as the most stable rotamer, where g++ stands for a large positive torsion angle ∼110° (cf. rotamer 7 in Table 4 and rotamer 4 in Table 5). In the most stable HF structure the orientation of the hydroxyl chain of galactose is in the clockwise direction and the hydroxymethylene group of galactose interacts with the third OH group of fucose (vide infra). According to the HF/6-31G-(d) results the three most stable rotamers of the β-D-galactose monomer are in order of decreasing stability, t t t t T g+, t t t t G+ g-, and g+ g- g- g+ G- g+, which lie within 1 kcal/mol energy range. Thus the most stable Tg+ hydroxymethylene orientation of the β-D-galactose monomer does not occur in Le^x. The second and third most stable rotamers do occur, however, in somewhat distorted forms due to the intermoiety interactions.

Torsion Angles in Fucose Moiety. The fucose moiety of the most stable rotamer of Le^x has the following conformation: t g− t t according to the MM2* method (cf. rotamer 1 in Tables 1 and 2). A similar rotamer occurs among the most stable HF/6-31G(d) conformers (cf. rotamer 7 in Table 4 and rotamer 4 in Table 5). This clockwise direction of the hydrogen bonds corresponds to the pattern found in the most stable $^1C_4$ L-fucose monomer.[15,17] Because of the strong donor interaction with the oxygen atom of the NAc group (cf. Figure 2), the $\tau 2$ torsion angle for the second OH group of the fucose moiety deviates considerably from −60° (cf. Tables 1, 2, 4, and 5). It is interesting to note that in those rotamers, where the oxygen atom of the NAc group is engaged by the other OH group, the $\tau 2$ torsion angle of the fucose moiety is close to ±60° (cf. Table 3). We calculated a slightly changed rotamer 9 of **(II)**. This rotamer is noted as 9a in Table 5. The main difference between the rotamer 9 and 9a of **(II)** is in the $\tau 4$ of fucose. The Fuc-(O4H)→(O3)Fuc interaction in rotamer 9 was broken and replaced by the Fuc(O4H)→(O5)Fuc interaction in rotamer 9a (cf. Table 5). The HF/6-31G(d) method shows that rotamer 9a is 4 kcal/mol less stable than rotamer 9 of **(II)** (cf. Table 5).

Interresidue Connections. In the most stable rotamers for the galactose and fucose moieties of Le^x, long intramolecular chains of hydroxyl groups are formed in counterclockwise or clockwise directions as viewed from the β face of the anomeric carbon atom (the directed chains are more stable than the nondirected chains). Similar directed patterns were found for monosaccharides earlier by French et al., Tran and Brady[28] using molecular modeling techniques, by Cramer and Truhlar using AM1 and PM3 methods,[29] and by Polavarapu and Ewig[30] at the HF/4-31G level of theory for D-glucose. It is interesting to observe how the previous experience gained on the monosaccharides might be applied on Le^x. Because of the substituents in 1, 2, 3, and 4 positions of GlcNAc the experience gained on glucose can only be applied on the hydroxymethylene group of GlcNAc. It was observed that the different orientations of this group change only slightly the relative energy,[14,16] and this remains true in Le^x. In fucose and in galactose only the anomeric OH group is substituted; thus the remaining OH groups show the typical orientations characteristic for the most stable monomers. Certain differences occur because these chains of OH groups might connect the galactose, fucose, and GlcNAc moieties in several ways (cf. Figure 2). One typical intermoiety

interaction is the Gal(O6H)→(O1)Fuc interaction favored by the MM2* method (cf. Figure 2, MM2* result). Another frequently occurring interaction is the Fuc(O2H)→(O═C)-GlcNAc interaction. A long chain of seven directed OH interactions including a Gal(O6H)→(O3)Fuc interaction (cf. Figure 2) can be observed in the most stable HF/6-31G(d) structure. The longest continuous chain of eight OH interactions occurs in rotamer 11 of **(I)** (cf. Tables 1 and 4) and in rotamer 9 of **(II)** (cf. Tables 2 and 5). The GlcNAc(O6H)→(O2)Gal, the Gal(O6H)→(O3)Fuc, and the Fuc(O2H)→(O═C)GlcNAc donor→acceptor interresidue interactions stabilize this rotamer. According to MM2* this rotamer is not particularly stable; however, according to HF/6-31G(d) this is the third most stable rotamer. The geometry difference between the most stable and the third most stable HF/6-31G(d) rotamers is in the orientation of the hydroxymethylene group of GlcNAc. Replacing the GlcNAc(O6H)→(O5)GlcNAc intraresidue interaction in rotamer 4 of **(II)** by GlcNAc(O6H)→(O2)Gal interresidue interaction in rotamer 9 of **(II)** destabilizes the molecule by 1.8 kcal/mol according to the HF/6-31G(d) method (cf. Table 5).

Comparison of Experimental Results with Calculated Geometries. The glycosidic torsion angles play an essential role in the molecular shape; thus we analyze these angles first. The NMR NOE experiments provide H−C−O−C-type torsion angles. Following Poppe et al.[6] we define $\Phi 3 = \tau$(H1−C1−O1[1,3]-C3), $\Psi 3 = \tau$(C1−O1[1,3]-C3−H3) torsion angles for the Fuc-α-1,3-GlcNAc glycosidic bond and $\Phi 4 = \tau$(H1−C1−O1[1,4]-C4), $\Psi 4 = \tau$(C1−O1[1,4]-C4−H4) for the Gal-β-1,4-GlcNAc glycosidic bond. (For Ψ3 and Ψ4 the following approximations are valid within few degrees: $\Psi 3 \cong \tau 3 - 120°$ and $\Psi 4 \cong \tau 4 + 120°$, where $\tau 3$ and $\tau 4$ are the glycosidic torsion angles in GlcNAc. Similarly, Φ3 can be derived from $\tau 1$ of fucose and Φ4 can be derived from $\tau 1$ of galactose). Table 6 shows the values and the error bars for these angles in sLe^x measured by NOE NMR in the water solution and in the bound states for three selectins (E, P, and L).[6] In Table 6 we also present Φ3, Ψ3, Φ4, and Ψ4 torsion angles obtained from the X-ray experimental coordinates.[10] Comparison of the two different structures found in the asymmetric unit of crystal shows that the important glycosidic torsion angles are similar in both structures (cf. Figure 6). Although there are somewhat larger differences (5−8°) in Φ torsion angles, the Ψ angles agree with each other within 1° (cf. X-ray values in Table 6). We note that the positions of hydrogen atoms are not very precise in the X-ray experiment.

Comparisons of the molecular structures in solution (NMR), solid phase (X-ray), and gas phase (calculated r_e structure) should be done with care. For example, the solvent effect influences the OH torsion angles (the intramolecular hydrogen bonding is competing with intermolecular solvent−solute interactions). In the solid phase the crystal field effects and strong intermolecular interactions distort the molecular shape. Another important feature is that whereas the calculations can characterize the various rotamers of the conformational space, the NMR experiment sees only the average of the rotamers in the solution. Thus it should be kept in mind that these data should not agree with each other. If some agreement is found, then that signals the rigidity of the structure.

A rather interesting feature of Le^x is that the galactose and fucopyranose rings take approximately parallel positions. In this stacked arrangement the hydrophobic side of fucose turns toward the galactose. Earlier MM studies showed that the Fuc-α-1,3-GlcNAc glycosidic torsion angles correspond to one of the two minima for the disaccharide.[31] However, the fucose−galactose interaction forces the Gal-β-1,4-GlcNAc glycosidic torsion angles out of the minima related to the corresponding disaccharide.[31] The most stable conformation predicted by the MM studies are close to that obtained by NMR experiments for the Le^x.[32]

The HF/6-31G(d) and MM2* results for the Φ3, Ψ3, Φ4, and Ψ4 torsion angles are shown in Table 6 for the first 10 rotamers of **(I)** and **(II)**. There is a very good agreement between the HF/6-31G(d) values for **(I)** and **(II):** the difference is less than 0.2°. This supports the use of the computationally less demanding **(II)** for these studies. The same is true for the MM2*

TABLE 6: Experimental (NMR,and X-ray), HF/6-31G(d) (Noted as HF), and MM2* (Noted as MM) H1−C1−O1,*x*−C*x* (Φ*x*) and C1−O1,*x*−C*x*−H*x* (Ψ*x*) Torsion Angles for (I) (Noted as Me) and (II) (Noted as H), Where *x* = 4 for Gal and *x* = 3 for Fuc

	Gal-β-1,4-GlcNAc		Fuc-α-1,3-GlcNAc	
	Φ4 (°)	Ψ4 (°)	Φ3 (°)	Ψ3 (°)
solution[a]	46 ± 1	18 ± 2	48 ± 2	24 ± 1
bound E[a]	24 ± 5	34 ± 3	71 ± 3	14 ± 2
bound P[a]	45 ± 4	18 ± 4	61 ± 10	26 ± 6
bound L[a]	33 ± 10	26 ± 12	42 ± 11	16 ± 3
X-ray 1[b]	36.3	14.8	40.5	20.2
X-ray 2[b]	44.3	15.5	35.4	20.7
HF Me 1[c]	46.6	14.7	41.4	17.0
MM Me 1[c]	49.8	14.1	41.5	26.9
HF Me 2	44.4	18.1	39.0	16.0
MM Me 2	47.6	20.2	39.3	24.2
HF Me 3	48.1	14.4	41.2	17.8
MM Me 3	49.6	13.9	41.7	27.0
HF Me 4	46.0	18.1	38.6	16.7
MM Me 4	47.3	20.1	39.3	24.2
HF Me 5	38.6	13.5	72.7	59.1
MM Me 5	50.0	11.7	67.0	64.1
HF Me 6	47.1	11.9	50.5	31.5
MM Me 6	50.2	14.6	42.5	33.7
HF Me 7	36.1	16.1	55.8	23.5
MM Me 7	44.9	15.8	44.0	30.1
HF H 8	48.6	10.8	50.6	32.9
MM H 8	50.1	14.2	42.3	33.8
HF H 9	39.4	8.0	52.3	23.1
MM H 9	45.1	9.7	43.9	30.2
HF H 10	36.9	15.4	55.3	24.4
MM H 10	44.9	15.5	44.1	30.5

[a] Ref 6, sLe^x in water, bound to E-, P-, and L-selectin. [b] Reference 10, two conformers of Le^x. [c] HF Me 1 is the first HF/6-31G(d) conformer of **(I)**, cf. Table 4. [d] MM Me 1 is the first MM2* conformer of **(I)**, cf. Table 1. The lowest-energy conformers are set to italic.

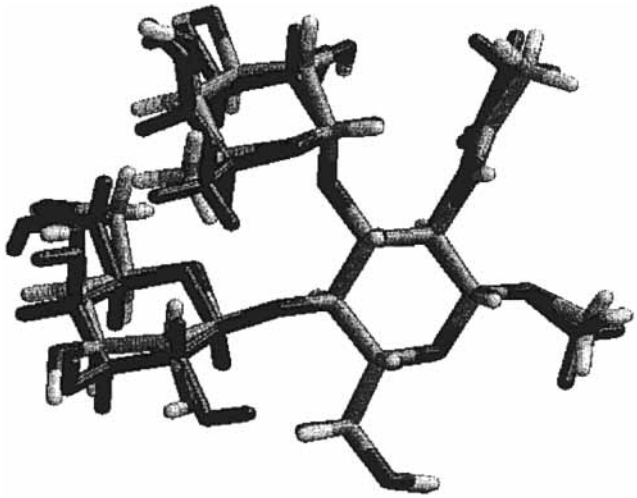

Figure 6. Overlap of the two (black and gray) experimental X-ray structures (**I**).[10]

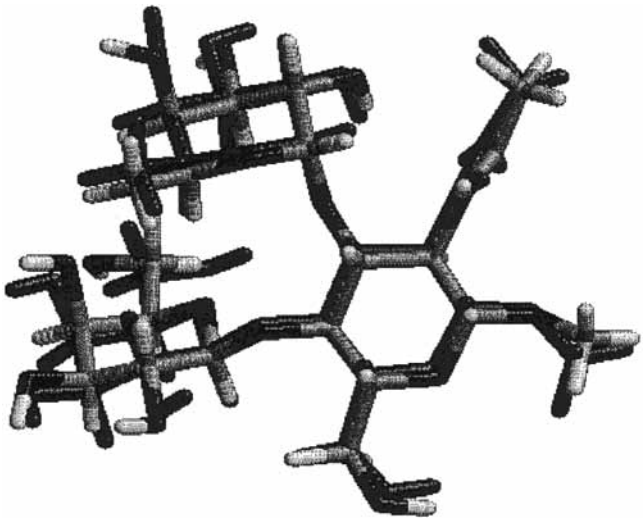

Figure 7. Overlap of the most similar HF/6-31G(d) (black, rotamer 3 in Table 4) and one of the X-ray (variable gray) structures (**I**).

values. For brevity we show only the nonredundant torsion angles in Table 6 [the first seven rotamers of (**I**) and the next four rotamers of (**II**)]. From the energetic point of view the most important rotamers are 1, 2, 5, and 7; the latter is the most stable for (**I**) according to the HF results. It can be seen that large differences (3−11°) occur between the HF/6-31G(d) and MM2* values for Φ4. In the most stable four rotamers the HF values for Φ4 are in the range of 36−47° and the MM2* values are in the range of 45−50°. The experimental Φ4 angles according to the HF/6-31G(d) results, fall in these ranges with only one exception (cf. Table 6, bound E). This means that the Φ4 angles mostly keep their values in the different environments (solid, liquid, and gas phase, and bound). There are no large differences between the HF and MM2* values for Ψ4, and the values for this angle fall within a rather narrow range of 14−20°. This observation implies that Ψ4 is more rigid, which is also supported by the experimental range of 14−18°, (cf. Table 6). Probably large intermolecular forces distort this angle in the bound sLe^x to 34° and 26° (cf. Table 6, and MM3 energy maps in ref 32).

The HF values for Φ3 fall in the range of 39−56°; the MM2* values are in the range of 39−44° for rotamers 1, 2, and 7. The solution and X-ray torsion angles are within these ranges with the exception of the sLe^x bound to P- and E-selectin (cf. Table 6). These relatively large values (61−71°) for Φ3 occur in the rather stable rotamer 5 (cf. Table 6). However, the large Ψ3 values (59−64°) calculated for rotamer 5 cannot be found in the experiments. Otherwise the experimental results agree relatively well with the calculated Ψ3 (16−30°, cf. Table 6). It can be concluded that the rigidity of Le^x is supported by the current ab initio results. Figure 7 illustrates that there is a rather good overlap between the GlcNAc moiety of one of the X-ray structures and the most similar HF/6-31G(d) structure. Even the glycosidic angles are similar. The stacking of the fucose and the galactose moieties is also similar. Naturally the orientations of the hydroxyl groups are mostly different, because the intramolecular hydrogen bonds dominate in the HF structure. In the X-ray structure the hydroxyl groups mostly provide intermolecular hydrogen bonds and are oriented accordingly (cf. Figure 7). As expected, neither of the calculated structures agrees with the experimental structures. However, there is a fair agreement for the glycosidic angles and for the bulk shape of the molecule.

5. Conclusions

The following conclusions can be drawn from our analysis:

The unconstrained MM2*-SUMM conformational search in the conformational space of torsion angles for **(I)** provides 13 rotamers and the constrained conformational search for **(II)** provides 12 rotamers within 1 kcal/mol energy range. The only constraint for (**II**) was to keep the τ1(C2−C1−O−H) torsion angle in the GlcNAc moiety in the anti position for the purpose of conserving the similarity between the two conformational spaces. No correlation was found between the MM2* relative energies of **(I)** and **(II)** ($R^2 = 0.274$). The MM2* method provides very small differences between the rotamers; for example, the energy range for the first 7 rotamers is 0.3−0.5 kcal/mol.

Comparison of the 7 rotamers of **(I)** and **(II)** obtained from the HF/6-31G(d) method shows the same energetic order for rotamers of **(I)** and **(II)** and a good correlation for the energy differences ($R^2 = 0.994$). The HF/6-31G(d) energetic order for **(I)** is 7, 1, 2, 5, 6, 3, 4, which differs considerably from the energetic order obtained by the MM2* method. The HF/6-31G-(d) method yields more than 4.5 kcal/mol energy difference for the first 11 most stable rotamers of (**II**).

Comparison of MM2* and HF/6-31G(d) results indicates that the HF method provides finer classification and distinguishes better among the various rotamers. The HF/6-31G(d) method provides larger differences among the torsion angles in the various rotamers than the MM2* method. This is especially true for the τ2 angle of GlcNAc. The τ5 and τ6 torsion angles of the hydroxymethylene group of GlcNAc take the G+ g− and G− g+ orientations, which interact with the pyranose ring oxygen (O5) of GlcNAc. The MM2* and HF/6-31G(d) results agree that the G− g+ orientation is the more stable. The energy of Le^x increases by 4 kcal/mol as the Fuc(O4H)→(O3)Fuc interaction is replaced by the Fuc(O4H)→(O5)Fuc interaction according to the HF/6-31G(d) results.

In the most stable MM2* Le^x structures the hydrogen-bonded chains of galactose (with a counterclockwise direction in the t t t G+ g− rotamer) and fucose (with a clockwise direction in the t g− t t rotamer) moieties are not connected directly. The Gal(O6H) is a hydrogen bond donor (in clockwise direction) to the O3 glycosidic oxygen of GlcNAc; thus the hydrogen-bonded chain is broken in the galactose moiety. The Fuc-(O2H)→(O═C)GlcNAc interaction connects the fucose and GlcNAc moieties. On the other hand the most stable HF/6-31G-(d) structure has a long chain of seven ordered hydrogen bonds including a Gal(O6H)→(O3)Fuc interaction observed (with clockwise hydrogen-bonded chain in galactose−t g− g− g+ G− g++ rotamer−and fucose units). The longest continuous chain of eight hydrogen bonds [with the added GlcNAc(O6H)→-(O2)Gal hydrogen bond] is the third most stable rotamer, which is 1.8 kcal/mol less stable than the most stable rotamer according to the HF/6-31G(d) method. But according to the MM2* method this rotamer is not particularly stable.

It has been found that the Φ3 = τ(H1−C1−O1[1,3]-C3), Ψ3 = τ(C1−O1[1,3]-C3−H3) torsion angles for the Fuc-α-1,3-GlcNAc glycosidic bond and the Φ4 = τ(H1−C1−O1[1,4]-C4), Ψ4 = τ(C1−O1[1,4]-C4−H4) torsion angles for the Gal-β-1,4-GlcNAc glycosidic bond mostly keep their values in the different environments (solid, liquid, gas phase). For example there is a rather good overlap between the GlcNAc moiety of one of the X-ray structures and the most similar HF/6-31G(d) structure. The stacking of the fucose and galactose moieties is similar. Naturally the orientations of the hydroxyl groups are mostly different as a result of intramolecular hydrogen bonding

in the gas-phase HF structure versus intermolecular hydrogen bonding in the solid-phase X-ray structure.

Acknowledgment. We thank G. Schubert for assistance. We also thank the Corporate Computing Network Department at Cray Research/Silicon Graphics Inc. for providing computer time to carry out some of these calculations. The financial support of the Hungarian Research Foundation (OTKA T16328) is acknowledged. The continuous financial support of the Natural Sciences and Engineering Research Council (NSERC) is gratefully acknowledged.

References and Notes

(1) (a) Phillips, M. L.; Nudelman, E.; Gaeta, F. C. A.; Perez, M.; Singhal, A. K.; Hakomori, S.; Paulson, J. C. *Science* **1990**, *250*, 1130. (b) Lowe, J. B.; Stoolman, L. M.; Nair, R. P.; Larsen, R. D.; Berhend, T. L.; Marks, R. M. *Cell* **1990**, *63*, 475. (c) Springer, T. A.; Lasky, L. A. *Nature* **1991**, *349*, 196. (d) Lasky, L. A. *Science* **1992**, *258*, 964.

(2) (a) 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. (b) Lin, Y.-C.; Hummel, C. W.; Huang, D.-H.; Ichikawa, Y.; Nikolau, K. C.; Wong, C.-H. *J. Am. Chem. Soc*. **1992**, *114*, 5452.

(3) Ichikawa, Y.; Lin, Y.-C.; 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.

(4) (a) Tyrell, D.; James, P.; Rao, B. N. N.; Foxall, C.; Abbas, S.; Dasgupta, F.; Nashed, M.; Hasegawa, A.; Kiso, M.; Asa, D.; Kidd, J.; Bradley, B. K. *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*, 10732. (b) Bradley, B. K.; Kiso, M.; Abbas, S.; Nikrad, P.; Srivastava, O.; Foxall, C.; Oda, Y.; Hasegawa, A. *Glycobiology* **1993**, *3*, 633. (c) DeFrees, S. A.; Gaeta, F. C. A.; Lin, Y.-C.; Ichikawa, Y.; Wong, C.-H. *J. Am. Chem. Soc.* **1993**, *115*, 7549. (d) Rampal, J. Y.;. Zheng, Z.-L.; Perez, C.; Walker, L. E.; DeFrees, S. A.; Gaeta, F. C. A. *J. Med. Chem.* **1994**, *37*, 3459. (e) Giannis, A. *Angew. Chem., Int. Ed. Engl.* **1994**, *33*, 178.

(5) Simanek, E. E.; McGarvey, G. J.; Jablonowski, J. A.; Wong, C.-H. *Chem. Rev.* **1998,** *98*, 833.

(6) Poppe, L.; Brown, G. S.; Philo, J. S.; Nikrad, P. V.; Shah, B. H. *J. Am. Chem. Soc.* **1997**, *119*, 1727.

(7) Ascensio, J. L.; Canada, J.; Martin-Pastor, M.; Espinosa, J. F.; Jimenez-Barbero, J. *J. Am. Chem. Soc.* **1996**, *118*, 10862.

(8) Poveda, A.; Ascensio, J. L.; Espinosa, J. F.; Martin-Pastor, M.; Canada, J.; Jimenez-Barbero, J.; *J. Mol. Graphics* **1997**, *15*, 9.

(9) Miller, K. E.; Mukhopadhyay, C.; Cagas, P.; Bush, C. A. *Biochemistry* **1992**, *31*, 6703.

(10) Perez, S.; Mouhous-Riou, N.; Nifant'ev, N. E.; Tsvetkov, Y. E.; Bachet, B.; Imberty, A. *Glycobiology* **1996**, *6*, 537.

(11) Kojima, N.; Fenderson, B. A.; Stroud, M. R.; Goldberg, R. I.; Habermann, R.; Toyokuni, T.; Hakomori, S. *Glycoconjugate J.* **1994**, *11*, 238.

(12) Barrows, S. E.; Dulles, F. J.; Cramer, C. J.; French, A. D.; Truhlar, D. G. *Carbohyd. Res.* **1995**, *276*, 219.

(13) Csonka, G. I.; Éliás, K.; Csizmadia, I. G. *Chem. Phys. Lett.* **1996**, *257*, 49.

(14) Barrows, S. E.; Strorer, J. W.; Cramer, C. J.; French, A. D.; Truhlar, D. G. *J. Comput. Chem.* **1998**, *19*, 1111.

(15) Csonka, G. I.; Éliás, K.; Kolossváry, I.; Sosa, C. P.; Csizmadia, I. G. J. *Phys. Chem. A* **1998**, *102*, 1219.

(16) Csonka, G. I.; Kolossváry, I.; Császár, P.; Éliás, K.; Csizmadia, I. G. *J. Mol. Struct.: THEOCHEM* **1997**, *395*−*396*, 29.

(17) Csonka, G. I.; Éliás, K.; Csizmadia, I. G. *J. Comput. Chem.* **1997**, *18*, 330.

(18) Del Bene, J. E.; Person, W. B.; Szczepaniak, K. *J. Chem. Phys*. **1995**, *99*, 10705.

(19) Damm, W.; Frontera, A.; Tirado-Rives, J.; Jorgensen, W. L. *J. Comput. Chem*. **1997**, *18*, 1955.

(20) Reilig, S.; Schlenkrich, M.; Brinkmanmn, J. *J. Comput. Chem.* **1996**, *17*, 450.

(21) Senderowitz, H.; Parish, C.; Still, W. C. *J. Am. Chem. Soc.* **1996**, *118*, 2078.

(22) Kneisler, J. R.; Allinger, N. L. *J. Comput. Chem.* **1996**, *17*, 757.

(23) Still, W. C. MacroModel 4.5, Columbia University, New York; http://www.columbia.edu/cu/chemistry/mmod/mmod.html.

(24) (a) Allinger, N. L. *J. Am. Chem. Soc*. **1977**, *99*, 8127. (b) Different versions are available from Quantum Chemistry Program Exchange, University of Indiana, Bloomington, IN 47405, http://ccl.osc.edu/ccl/qcpe/QCPE/catalog.html.

(25) Gundertofte, K.; Liljefors, T.; Norrby, P.-O.; Pettersson, I. *J. Comput. Chem*. **1996**, *17*, 429.

(26) Goodman, J. M.; Still, W. C. *J. Comput. Chem*. **1991**, *12*, 1110.

(27) Frisch, M. J.; Trucks, G. W.; Head-Gordon, M.; Gill, P. M. W.; Wong, M. W.; Foresman, J. B.; Johnson, B. G.; Schlegel, H. B.; Robb, M. A.; Replogle, E. S.; Gomperts, R.; Andres, J. L.; Raghavachari, K.; Binkley, J. S.; Gonzalez, C.; Martin, R. L.; Fox, D. J.; DeFrees, D. J.; Baker, J.; Stewart, J. J. P.; Pople, J. A. GAUSSIAN 94, Revision B, Gaussian, Inc., Pittsburgh, PA, 1995.

(28) (a) French, A. D.; Rowland, R. S.; Allinger, N. L. In *Computer Modeling of Carbohydrate Molecules*; French, A. D., Brady, J. W. Eds.; ACS Symposium Series 430; American Chemical Society: Washington, DC, 1990; p 120. (b) Tran, V.; Brady, J. W. In *Computer Modeling of Carbohydrate Molecules*; French, A. D., Brady, J. W. Eds.; ACS Symposium Series 430; American Chemical Society: Washington, DC, 1990; p 213.

(29) Cramer, C. J.; Truhlar, D. E. *J. Am. Chem. Soc.* **1993**, *115*, 5745.

(30) Polavarapu, P. L.; Ewig, C. S. *J. Comput. Chem*. **1992**, *13*, 1255.

(31) Imberty, A.; Micros, E.; Koca, J.; Mollicone, R.; Oriol, R.; Perez, S. *Glycoconjugate J.* **1995**, *12*, 331.

(32) Imberty, A. *Curr. Opin. Struct. Biol.* **1997**, *7*, 617, and references therein.