Theoretical Study of Alternative Ring Forms of α-L-Fucopyranose

Gábor I. Csonka* and Krisztina Éliás

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

István Kolossváry

Columbia University, Department of Chemistry, 410 Havemeyer Hall, New York, New York 10027

Carlos P. Sosa

Cray Research Inc., 655 E. Lone Oak Drive, Eagan, Minnesota 55123

Imre G. Csizmadia

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

Received: October 13, 1997

The ring puckering of α-L-fucopyranose was studied by the MM2*-LMOD (low mode) conformational search technique built into the Macromodel program. The molecular geometries of lowest energy (within a 10 kcal/mol energy window) were analyzed at the HF/3-21G, HF/6-31G(d), and generalized gradient approximation (GGA) levels of density functional theory (DFT). Our results show that the MM2* method tends to overestimate the number of stable conformers in the high-energy region. The MM2* method detected 33 stable conformers within the aforementioned energy window, and this number was reduced to 25 after re-minimization by the HF and GGA DFT methods. The low-energy conformers yielded by the MM2* method are in qualitative agreement with the *ab initio* results. In these conformers, intramolecular chains of hydroxyl groups are formed in counterclockwise or clockwise directions. The best calculations predict the most stable $^1C_4$ chair conformer to be lower in energy by ∼6 kcal/mol than the most stable non-$^1C_4$ conformer. The zero-point energy differences are very small. We provide a detailed study of the basis set (up to 6-311+G-(2d,p) basis set) and correlation effects (MP2 and hybrid functional study) on the relative energies and molecular geometries. The converged results are presented wherever it is possible. The largest difference between the various methods is ∼2.7 kcal/mol; however, agreement well within 1 kcal/mol is experienced for most of the conformers. We discuss the geometric consequences of the exoanomeric effect, and show how the calculated and X-ray experimental results can be brought into accordance with each other. The geometric aspects of the O···H interactions (hydrogen bonding) are described with the aid of a polar coordinate system centered at the acceptor oxygen atom. The equatorial−equatorial and the equatorial−axial type of OH interactions can be readily distinguished using the proposed geometric parameters. We present the shape of the contour surface of the Laplacian of the electron density around the oxygen atom and show how this shape influences the various types of O···H interactions. The strength and weakness of the Van den Enden's π-, sp^3-, and σ-type interactions were readily explained using the aforementioned surface.

1. Introduction

Carbohydrates play a vital role in cell−cell recognition processes. The α-L-fucose (α-L-6-deoxygalactose; 6-methyl-tetrahydro-pyran-2,3,4,5-tetraol) is an important building block of plant polysaccharides and animal glucans.[1,2] For example, L-fucose is the immunodominant monosaccharide of many blood group antigen determinants,[3] and it occurs in a tumor-associated antigen,[4] in plasma, hormonal, and serum glycoproteins,[5] and in the neural glycopeptides.[6] The serum lectin recognizes oligosaccharides on invading pathogenes and thus plays a role in the immune system. X-ray data[7,8] show the importance of the shape of presentation of the oligosaccharides to the receptor. The self-oligosaccharides are distinguished from those that are foreign using the information coded into the geometry. The Sialyl Lewis^X (SLe^X) blood group antigen tetrasaccharide is displayed 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.[9] The conformation[10] and the key structural features[11] of SLe^X, required for recognition, have been known. 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*-selectin recognition of SLe^X.

For many proteins, the crystallographic methods yield quite accurate structural information. However, the oligosaccharide groups of glycoproteins present more difficulties for the X-ray methods because the carbohydrate part exhibits greater disorder than the protein part. The conformation of carbohydrates in solution can be established by the combination of NMR spectroscopy and molecular dynamics (MD) techniques.[12] The nuclear magnetic resonance (NMR) data alone are frequently

S1089-5639(97)03320-3 CCC: $15.00
Published on Web 01/23/1998

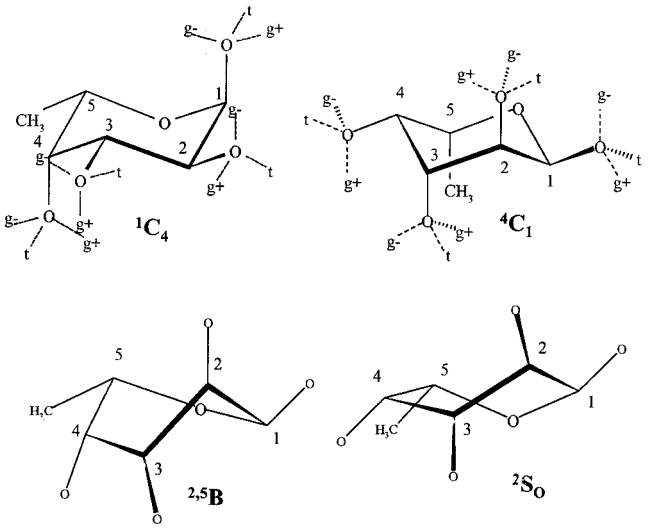

Figure 1. Schematic representations of four selected conformers of the α-L-fucose; chairs (C), boat (B), and skew (S). The idealized C(x+1)-Cx-O-H torsions of the 81 possible OH rotamers of chair conformers are denoted by *g*+, *t*, and *g*- for gauche clockwise (60°), anti (180°), and gauche counterclockwise (−60°), respectively, where the ring carbon atoms are denoted by $x = 1, 2, 3, 4$.

insufficient for the determination of the conformation of an oligosaccharide. The ambiguity arises from a paucity of nuclear Overhauser effects (NOE) between sugars. It should be noted that further problems may arise from the applied force field. Most of the force fields provide more or less reasonable geometries for simple carbohydrates; however, the relative energies are frequently in error compared with *ab initio* calculations.[13] The carbohydrates have densely packed, highly polar functional groups, and the conformational energies depend on stereoelectronic effects. This situation makes force field methods difficult to use. The force field methods frequently rely on experimental results. However, for carbohydrates, only a few of the energy differences are well established experimentally. Furthermore, it is difficult to derive the torsional parameters of the molecular mechanics from experiments because of the barely measurable torsional mode frequencies. Serious problems, arising from the inadequate quality force field parameter sets, were noticed for saccharides.[14] To remedy these problems, reparametrized force fields were proposed recently, based on comprehensive high-quality *ab initio* calculations carried out on fragments and model compounds.[15,16]

Theoretical investigation of monosaccharides is a very demanding task because recent studies have shown that to achieve converged *ab initio* results within 1 kcal/mol for relative energies of chair configurations requires basis sets of at least triple-zeta quality and correlation levels better than the MP2 method.[17] Another source of difficulties is the very large conformational space. The idealized threefold rotation model of *n* hydroxyl groups, in principle, provides 3^n different rotamers. The aldopyranosyl ring puckering (two nonequivalent chair-like,[18] $^4C_1$, and $^1C_4$, skew and skew-boat forms[19]) further increases the number of the possible conformers (cf., Figure 1). The NMR spectroscopic and crystallographic results usually identify the $^4C_1$ form of D-pyranosil rings. The NMR results show that for α-D-altropyranose, an equilibrium between the $^4C_1$ and $^1C_4$ forms is present,[20] and for α-D-idopyranose, an equilibrium between the $^4C_1$, $^1C_4$, and $^0S_2$ forms is present.[21] For L-pyranoses, the $^1C_4$ ring is the most stable (i.e., mirror image of the $^4C_1$ D-pyranosyl ring). Cremer and Pople[22] proposed three parameters to describe the puckering of six-atom rings: the magnitude of puckering, *q*, and two angles, Φ and Θ, for the relative orientation of the puckering about the ring (the important Φ angles and ring forms for α-L-fucose are printed in bold). The two ideal chairs of an L-hexapyranose are located at $\Theta = 180°$ ($^4C_1$) and at $\Theta = 0°$ ($^1C_4$) independently of the other two parameters. The ideal boat and skew forms are at $\Theta = 90°$. These forms can be readily distinguished by the Φ angle: 0° ($B_{3,0}$), 30° ($^4S_0 \equiv {}^1S_3$), 60° ($^{1,4}B$), 90° ($^1S_5 \equiv {}^4S_2$), 120° ($B^{2,5}$), 150° ($^3S_5 \equiv {}^0S_2$), 180° ($^{3,0}B$), 210° ($^3S_1$), 240° ($B^{1,4}$), 270° ($^2S_4$), 300° ($^{2,5}B$), 330° ($^2S_0$). The $^1C_4$, $^4C_1$, $^{2,5}B$, and $^2S_0$ ring forms are shown in Figure 1. This complexity makes the sugars excellent information encoders, and this conformational and configurational information is decoded by a given receptor during the biomolecular binding process. This process is entropically unfavorable due to the decrease of rotational and translational entropy.[23−25] To estimate the magnitude of this entropy decrease, the entropy of the free carbohydrates should be known in the first place.

Barrows et al.[17] have shown that the HF/6-31G(d) energetic difference of the $^4C_1$ and $^1C_4$ aldopyranose rings of D-glucose agrees well with the very expensive correlated energy differences. However, our results show that despite the good energy, the HF/6-31G(d) equilibrium molecular geometry differs considerably from the correlated equilibrium molecular geometry in several respects. Probably the most important differences between HF/6-31G(d) and correlated results are in the C—C—O—H dihedral angles.[13,26] These torsional angles are sensitive measures of the OH···O interactions. Without OH···O interaction, these angles would be close to their ideal values (−60°, +60°, or 180°). However, an interaction with a neighboring OH group usually results in a considerable deviation from these idealized angles. In extreme cases an interaction can be so strong that it might cover the energy required to turn the O—H bond into an eclipsed position with the C—C bond. The serious underestimation of the strength of hydrogen bonding at the HF/6-31G(d) level of theory results in nearly ideal C—C—O—H dihedral angles. Further increase of the basis set size at the HF/6-31G(d) level will worsen the agreement with the correlated methods,[26] because the HF/6-31G(d) error becomes more apparent and OH···O interactions will become even weaker. The importance of the electron correlation for saccharides has been demonstrated on several smaller fragments (diols, triols)[26,27] and on monosaccharides.[13,17,28,29] However, the applicability of the MP2 or CCSD methods for mono- and oligosaccharides is limited because of the associated computational expenses. The fact that these methods yield poor energies with double-zeta polarized basis sets makes such calculations fairly expensive: at least triple-zeta polarized basis sets plus further electron correlation and basis set corrections are required for sufficiently converged results. The alternative involves the use of more economical theories such as density functional theory (DFT), which requires a modest basis set for converged results in the case of saccharides and saccharide fragments. The DFT method is considerably faster and it correctly reproduces the results of the more expensive calculations. It should be noted that the reliability of the DFT methods is still under research, and for small molecules, the classical correlation methods yield more reliable results after a systematic improvement of the basis set and the correlation method. However, the size of the saccharide molecules make the latter calculations prohibitively expensive. The systematic improvement of DFT methods is less evident. However, our results[13,28,29] and Allinger's results[27] show that a carefully selected DFT method

is a viable alternative for the classical, very expensive theoretical methods in carbohydrate chemistry.

In the present paper we continue (cf., refs 13 and 27) the exploration of the conformational space of the α-L-fucose by the new MM2*-LMOD (low-mode) conformational search method.[30] We analyze the low-energy structures obtained by the MM2*-LMOD method within a 10 kcal/mol energy window by the MM2*,[31,32] HF/3-21G, HF/6-31G(d), and generalized gradient approximation (GGA) DFT levels of theory using extended basis sets up to 6-311+G(2d,p). We are seeking the computationally most effective level of the theory that provides chemically useful accuracy for ring puckering.

2. Computational Methodology

Conformational Search. The search for stable conformers in the conformational space of α-L-fucose was carried out with the MacroModel 4.5 program package.[32] In MM2* (MacroModel),[32] all force field equations are identical with those of authentic MM2[31] except several equations. The most important difference is in the electrostatic equation. A recent comparison of a series of molecular mechanics methods has shown that the accuracy in relative conformational energies is apparently equal for MM2*, MM2(91), and MM3(92).[33]

The conformational searches were carried out with the new, particularly efficient LMOD conformational search method.[30] LMOD operates as follows: an initial arbitrary minimum-energy conformer is subjected to normal-mode analysis, and the low-frequency modes are stored as an array of eigenvectors of the non-mass-weighted Hessian matrix. The number of low-frequency modes considered is determined by a user-defined frequency threshold (see later). LMOD searches the low-frequency modes systematically along the corresponding eigenvectors, which are searched in both directions. The initial structure is continuously perturbed along one of its low-mode eigenvectors in discrete steps until the increase in potential energy exceeds a user-defined threshold during a single step. The resulting structure is subsequently subjected to energy minimization. LMOD typically focuses the search to the local neighborhood of a minimum on the potential energy hypersurface. As the search progresses, an ensemble of conformers is collected, and these are used as starting structures for structural perturbation along their low-frequency modes. Furthermore, each of the low-frequency modes associated with each structure is employed for structural perturbation. The new minima found during an LMOD search become new focal points and, thus, LMOD necessarily explores the entire potential energy surface. When the systematic search directions are exhausted for a particular minimum, LMOD switches to a stochastic or Monte Carlo procedure. In the Monte Carlo mode, random directions comprised of a random mixture of the low-mode eigenvectors are searched in exactly the same manner as in the systematic mode.

LMOD generates its own search space automatically and uses only two parameters, the threshold for the low-frequency modes and the energy threshold for energy minimization. For the current study we used 250 cm^{-1} for the threshold of the low-frequency modes. The structures generated by the LMOD procedure were minimized to yield unique conformers within an energy window of 10 kcal/mol above the global minimum. Geometry optimizations were carried out with a Truncated Newton Conjugate Gradient (TNCG) technique,[34] with the maximal number of iterations set to 200 and using a convergence criterion of 0.01 for the gradient norm. This procedure resulted in 33 minima.

***Ab Initio* and Density Functional Methods.** The 33 minima obtained by the MM2*-LMOD search were re-optimized by HF and GGA DFT methods using the Berny algorithm combined with redundant internal coordinates built into the GAUSSIAN 94 program.[35]

We employed the following combinations of the GGA DFT functionals:

(i) **BP** and **BPW**: Becke 88 exchange functional[36] is combined with the correlation functionals of Perdew 86, and Perdew-Wang 91,[38] respectively.

(ii) **B3P** and **B3PW**: Hybrid methods that are linear combination of various exchange and correlation functionals in the form:

$$AE_x[\text{Exact}] + (1 - A)\,E_x[\text{S}] + B\Delta E_x[\text{B}] + E_c[\text{VWN5}] + C\Delta E_c[\text{P}] \quad (1)$$

where E_x[Exact], E_x[S], and ΔE_x[B] are the exact, Slater, and Becke[6] exchange functionals, respectively, and E_c[VWN] and ΔE_c[P] are the Vosko, Wilk, and Nussair 5[39] and Perdew[37] or Perdew-Wang[38] correlation functionals, respectively. The constants *A*, *B*, and *C* are those determined by Becke by fitting heats of formation ($A = 0.2$, $B = 0.72$, $C = 0.81$[40]). Becke used the PW correlation functional for the determination of these parameters; however, the same parameters are generally used with *P* correlation functional.

(iii) **BLYP** method, in which Becke's exchange functional[36] is combined with the correlation functional of Lee, Yang, and Parr.[41]

(iv) **B3LYP** is a hybrid method (this functional was not published before the implementation into the GAUSSIAN 92/DFT)[42] which is a logical extension of Becke's three-parameter concept using different correction functionals (e.g., LYP) in the form:

$$AE_x[\text{Exact}] + (1 - A)E_x[\text{S}] + BE_x[\text{B}] + (1 - C)E_c[\text{VWN3}] + CE_c[\text{LYP}] \quad (2)$$

The constants *A*, *B*, and *C* are selected to be equal to those determined by Becke for the B3P method.

The Gaussian 94 program[35] uses numerical quadrature to evaluate the DFT integrals. The quadrature scheme is defined by the number of points in radial and angular directions. The DFT calculations were carried out using the sg1 (pruned to ∼3000 points per atom) and fine (pruned to ∼7000 points per atom) grids. By default, we used the fine grid for the geometry optimizations throughout this paper, and we note the calculations where the sg1 grid was used. The HF and GGA DFT calculations were carried out using 3-21G, 6-31G(d), 6-31+G(d), and 6-311+G(2d,p)[43] basis sets. The calculations were performed on Silicon Graphics and Cray computers.

3. Energetics

Relative Stabilities. The MM2*, HF, and GGA DFT total energies (*E*), and energy differences (Δ*E*) of the stable conformers of α-L-fucose that are provided by the MM2*-LMOD search within a 10 kcal/mol energy window are summarized in Table 1. This search resulted in 33 stable conformers. For a given ring conformer several hydroxyl rotamers may be present in the conformational space. The orientations of the four hydroxyl groups are specified by *t*, *g*+, or *g*- corresponding to the notation in Figure 1. The meaning of *t*, *g*+, or *g*- is explained in the text of Figure 1. It should be noted that the OH···O interactions may deviate the C—C—OH dihedral angles from their ideal

TABLE 1: The Calculated MM2*, HF, and BP Total (*E*) and Relative (Δ*E*) Energies of α-L-Fucose Ring

		OH rotamers				MM2*		HF/3-21G		HF/6-31G(d)		BP/6-31G(d)	
no.	conf.	1	2	3	4	E^b	ΔE^c	E^d	ΔE^c	E^d	ΔE^c	E^d	ΔE^c
1	$^1C_4$	*t*	*g*−	*t*	*t*	−97.07	0.00	−605.12709	0.00	−608.48450	0.00	−611.94819	0.00
2	$^1C_4$	*g*+	*g*+	*g*+	*g*+	−91.78	1.26	−605.12375	2.10	−608.48144	1.92	−611.94806	0.08
3	$^1C_4$	*g*+	*g*+	*g*+	*g*−	−89.90	1.71	−605.12139	3.58	−608.48017	2.71	−611.94427	2.47
4	$^1C_4$	*t* |	*t*	*t*	*t*	−89.30	1.86	−605.11973	4.62	−608.47900	3.45	−611.94168	4.09
6	$^1C_4$	*g*+	*t*	*t*	*t*	−88.30	2.10	−605.11905	5.04	−608.47877	3.59	−611.94244	3.61
5	$^1C_4$	*g*+	*g*+	*g*−	*t*	−88.50	2.05	−605.11882	5.19	−608.47852	3.75	−611.94218	3.77
8	$^1C_4$	*t* |	*g*+	*g*+	*g*+	−86.01	2.65	−605.11918	4.96	−608.47848	3.78	−611.94470	2.19
10	$^1C_4$	*t*	*g*− |	*g*+	*g*+	−81.70	3.68	−605.11949	4.77	−608.47824	3.93	−611.94438	2.40
9	$^1C_4$	*t* |	*g*+	*g*+	*g*−	−85.87	2.68	−605.11763	5.94	−608.47728	4.53	−611.94092	4.57
12	$^1C_4$	*t*	*g*− |	*g*+	*g*−	−80.89	3.87	−605.11754	5.99	−608.47727	4.54	−611.94063	4.75
11	$^1C_4$	*t*	*g*−	*t* |	*g*+	−81.20	3.80	−605.11757	5.97	−608.47703	4.69	−611.94287	3.34
7	$^1C_4$	*t* |	*g*+	*g*−	*t*	−86.88	2.44	−605.11700	6.33	−608.47691	4.76	−611.94010	5.08
14	$^1C_4$	*g*+	*t*	*t*	*g*+ |	−74.89	5.31	−605.11165	9.69	−608.47232	7.64	−611.93901	5.76
17	$^1C_4$	*t* |	*t*	*t*	*g*+ |	−72.26	5.93	−605.10892	11.40	−608.47125	8.31	−611.93607	7.61
21	$^4C_1$	*t*	*t*	*g*−	*t*	−67.06	7.18	−605.11827	5.53	−608.47458	6.22	−611.93733	6.82
23	$^4C_1$	*t*	*t*	*g*−	*g*−	−66.78	7.25	−605.11887	5.16	−608.47431	6.39	−611.93610	7.59
26	$^4C_1$	*t*	*t*	*t*	*g*+	−65.18	7.63	−605.11702	6.32	−608.47410	6.52	−611.93682	7.14
20	$^{2,5}B$	*t*	*t*	*g*−	*g*+	−67.68	7.03	−605.11977	4.59	−608.47315	7.12	−611.94149	4.21
27	$^2S_0$	*t*	*t*	*t*	*g*+	−64.80	7.72	−605.11787	5.79	*−608.47315*	*7.12*	*−611.94149*	*4.21*
19	$^2S_0$	*t*	*t*	*g*−	*g*−	−68.33	6.87	−605.12012	4.37	−608.47263	7.45	−611.93675	7.18
24	$^2S_0$	*t*	t	*g*−	*t*	−65.99	7.43	−605.11966	4.66	−608.47224	7.69	−611.93786	6.48
25	$^2S_4$	*t*	*t*	*t*	*t*	−65.23	7.62	−605.11418	8.10	−608.47120	8.34	−611.93512	8.21
29	$^4C_1$	*g*−	*g*+	*g*−	*g*−	−60.68	8.71	−605.11262	9.08	−608.47114	8.38	−611.93451	8.59
22	$^3S_1$	*t*	*t*	*t*	*t*	−66.98	7.20	−605.11367	8.42	−608.47112	8.40	−611.93447	8.61
31	$^3S_5$	*t*	*t*	*t*	*t*	−59.32	9.03	−605.11692	6.38	−608.46862	9.96	−611.93498	8.29
33	$^3S_1$	*g*−	*g*+	*g*−	*t*	−56.97	9.59	−605.10686	12.69			−611.93258	9.80

[a] The hydroxyl groups are numbered according to Figure 1, and the *t*, *g*+, and *g*− symbols are used to identify the rotamers. The first 17 $^1C_4$ rotamers were analyzed in a more detailed manner in ref 13. The vertical bars between the hydroxyl groups denotes the lack of interaction between hydrogen in OH groups and the lone pair of electrons with neighboring oxygen atoms (OH···O). The conformers are ordered according to the HF/6-31G(d) energetic order, except the two highest energy $^1C_4$ rotamers (14 and 17). Conformer 27 is shown in italic, because the HF and BP/6-31G(d) calculations show that this conformer is a transition state. This conformer was transformed on conformer 20 during the geometry optimization. [b] kJ/mol. [c] kcal/mol. [d] Hartree.

values; thus, the notation used in Table 1 is only qualitative. The geometries of the first 17 $^1C_4$ rotamers agree with our previous MM2*-SUMM[44] conformational search results.[13] These rotamers are ordered in increasing HF/6-31G(d) relative energy because previous studies have shown that the HF/6-31G-(d) energetic differences are surprisingly close to the very high level, MP2/cc-pVTZ basis set and correlation-corrected results.[17,28] A new rotamer, the 18th (*t*|*g*+|*t*|*g*-) $^1C_4$ rotamer, has also been found (we put a vertical bar between the noninteracting OH groups); however, the *ab initio* calculations show that this rotamer is not stable and it was smoothly transformed by the geometry optimization to conformer number 7 (*t*|*g*+*g*- *t*) in Table 1. This result is in agreement with our earlier experience that some of the high-energy $^1C_4$ rotamers (e.g., 13, 15, 16) provided by the MM2*-SUMM search are not stable according to the *ab initio* calculations.[13] Our accumulated experience shows that the MM2* method tends to overestimate the number of the stable conformers in the high-energy region. However, the low-energy conformers are in qualitative agreement with the *ab initio* results. The results in Table 1 show that there is a considerable agreement between the methods with respect the most stable $^1C_4$ rotamers. In these rotamers, intramolecular chains of hydroxyl groups are formed in counterclockwise or clockwise directions. These unidirected chains of OH groups are more stable than the non-unidirected chains (cf., rotamers 1, 2, 5, and 6 in Table 1). Similar unidirected patterns were found earlier by French et al.[45a] and Tran and Brady[45b] using molecular modeling techniques, by Cramer and Truhlar using AM1 and PM3 methods,[46] and by Polavarapu et al.[47] at the HF/4-31G level of theory for D-glucose. In most of the $^4C_1$, S, and B conformations, the two pairs of the OH groups are sterically separated, thus the OH···O chain formation is less frequently found.

Sources of Error. Recent investigations have shown that the AM1 and PM3 semiempirical methods fail for 1,2-ethanediol[48] and for the energetic order of the rotamers of $^1C_4$ α-L-fucose[13] and α-D-glucose.[17,46] The source of problems with the PM3 method for the OH···O interactions has been identified recently.[49]

The MM2* results agree considerably better with the *ab initio* results than the AM1 and PM3 semiempirical results; however, MM2* underestimates the energy differences and yields wrong ordering and several extra conformers. The source of the underestimation of energetic differences and the wrong ordering is probably the simple Coulomb interaction approximation used in the MM2* method. In the aldohexapyranoses, the highly polar hydroxyl groups interact strongly, and it is assumed that the Coulomb approximation cannot differentiate correctly between the various interactions. The source of the additional high-energy rotamers is the following: The MM2* method erroneously stabilizes two hydroxyl groups in the anti position.[26,50] A lone pair−lone pair repulsive interaction destabilizes these rotamers and the *ab initio* methods rotate these hydroxyl groups to an energetically more advantageous position.[26] For a more convenient overview, Figure 2 shows the relative energies of the various conformers. The conformers are divided into two groups, the first 17 $^1C_4$ rotamers are on the left-hand side and the other conformers are on the right-hand side. The conformers are ordered in increasing HF/6-31G(d) energy, similar to the order shown in Table 1.

The problems with the HF/3-21G relative energies are clearly visible in Figure 2. This method provides very polar hydroxyl groups and thus seriously overestimates the energetic differences between the $^1C_4$ rotamers and erroneously stabilizes the $^4C_1$, S, and B conformations where a strong interaction of the hydroxyl groups is possible. This situation is in agreement with the results

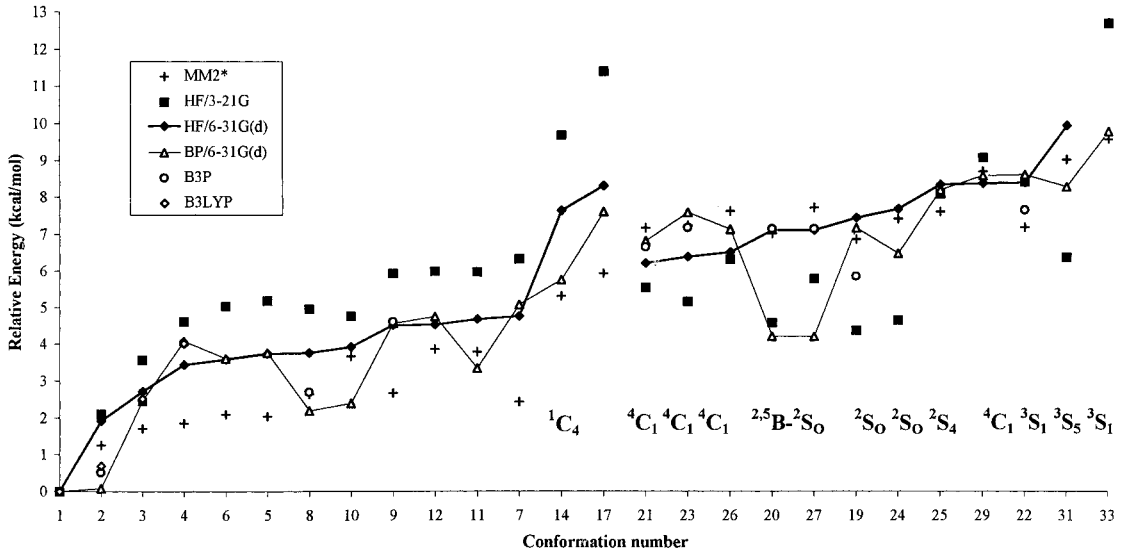

Figure 2. The relative energies of the α-L-fucose conformers within 10 kcal/mol energy window. The numbering of the conformers follows the order yielded by the MM2* method. The missing conformers (13, 15, 16, 18, 28, 30, 32) are present only in the MM2* conformational space and they are missing from the *ab initio* conformational space. The first 17 conformers are $^1C_4$ rotamers (cf., ref 13). The conformers are reordered according to HF/6-31G(d) energetic order. The conformer 27 is missing from the HF and BP/6-31G(d) conformational space, the $^2S_0$ conformer is transformed to $^{2,5}B$ conformer.

published by Barrows et al.,[17] who found that the HF/6-31G(d) method shows quite good performance for the ring-puckering energetics; this good performance is why we reordered the conformations according to HF/6-31G(d) energy. However, we should not forget that the Coulomb correlation between electrons of parallel spins is not treated correctly by the HF method. These electron correlation effects play an important role in the correct description of OH···O interactions. The deficiencies of the HF method are quite apparent if a sufficiently large basis set is used. The deficiencies of the HF/6-31G(d) basis set manifest themselves in correlated calculations. The MP2/6-31G(d) calculations show that this basis set is far from being converged. Even the very demanding MP2/triple-zeta calculations require further basis set and correlation corrections. In this paper we model the electron correlation effects by the significantly less expensive DFT functionals. We note that the computational expense of the MP2 method is formally $O(N^5)$, where N is the number of basis functions in the molecule. The cost of the DFT methods is formally $O(N^3)$, which may be further reduced by efficient implementations.[51]

We have already mentioned that the application of DFT is not free of problems; thus, we have selected the generally well-performing GGA and hybrid DFT functionals for this study. The B3P, B3PW, or B3LYP type hybrid functionals are known to provide better energies than the pure GGA methods (e.g., BP, BPW, or BLYP).[39] Furthermore, the addition of diffuse functions should considerably improve the DFT energetic order because the diffuse functions provide sufficient space for the electrons far from the nuclei, thus the long-range part of the correlation and exchange functionals work better for the OH···O interactions. Similar behavior was experienced by Del Bene et al.[52] and by Novoa and Sosa[53] for the weak interactions with B3LYP functional.

For several $^1C_4$ rotamers, the BP, BLYP, and B3P ΔE values follow quite closely the HF/6-31G(d) ΔE values (cf., rotamers 3, 5, 6, 7, 9, and 12 in Figure 1). However, for other rotamers (e.g., 2, 8, 10, 11, 14, and 17), the DFT methods with the 6-31G-(d) basis set provide a considerable (∼2 kcal/mol) stabilization effect. In these rotamers, the 4th $g+(a)$ hydroxyl group interacts with the ring oxygen; thus, the inclusion of the electron

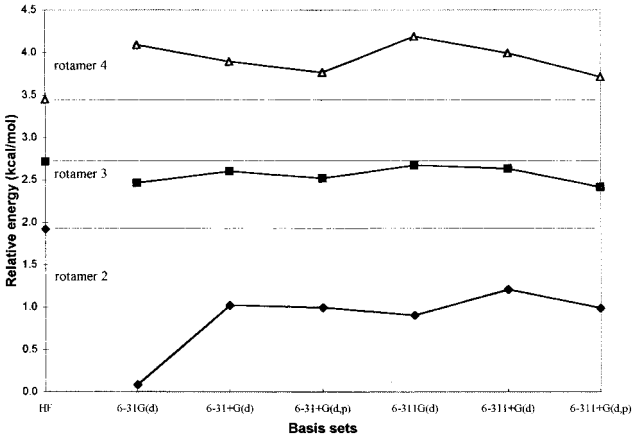

Figure 3. The basis set dependence of the relative energies of the first four $^1C_4$ α-L-fucose rotamers. The BP functional and the BP/6-31G(d) equilibrium geometries were used. The HF/6-31G(d) relative energies are represented by horizontal lines.

correlation is necessary to recover this effect. We performed a basis set study for the first four rotamers. We selected these rotamers because rotamers 2 and 4 show negative and positive correlation energetic effects, respectively, whereas for rotamer 3, the HF, BP, B3P, and B3LYP results agree quite well.[13] Inspection of Figure 3 reveals that for rotamer 2, most of the 2 kcal/mol stabilization effect provided by the BP functional relative to the HF result is in fact a basis set effect. Inclusion of diffuse functions into the basis set (6-31+G(d)) provides less stabilization, only 1 kcal/mol, and further increases in the quality of the basis set (up to 6-311+G(2d,p)) support this 1 kcal/mol value (within 0.2 kcal/mol). For rotamer 3, the relative energies do not change with the increase of the basis set above 6-31G-(d) (cf., Figure 3). For rotamer 4, slight (0.3 kcal/mol) convergence toward the HF value can be observed (cf., Figure 3 and ref 13), and the BP method seems to be converged even at double-zeta level. We also performed a DFT method dependence study. For rotamer 2, the relative energies are 0.21 and 0.69 kcal/mol calculated by BPW and B3LYP functionals, respectively, using the 6-31G(d) basis set. For rotamers 3 and 4, there is no noticeable functional dependence.[13]

TABLE 2: The Method and Basis Set Dependence of the Relative Energies of Several Selected Conformers of α-L-Fucose[a]

		OH rotamers				BPW/6-31G(d)		B3P/6-31+G(d)		BP3P/6-31+G(d)//		BP3PW/6-311+G(2d,p)//		B3LYP/6-311+G(2d,p)//	
no.	conf.	1	2	3	4	E^b	ΔE^c	E^b	ΔE^c	E^b	ΔE^c	E^b	ΔE^c	E^b	ΔE^c
1	$^1C_4$	*t*	*g*−	*t*	*t*	−611.87357	0.00	−613.65619	0.00	−613.56197	0.00	−611.94247	0.00	−612.17770	0.00
2	$^1C_4$	*g*+	*g*+	*g*+	*g*+	−611.87323	0.21								
19	$^2S_0$	*t*	*t*	*g*−	*g*−	−611.86184	7.36	−613.55587	5.85	−613.55205	6.23	−611.93071	7.38	−612.16579	7.47
20	$^{2,5}B$	*t*	*t*	*g*−	*g*+	−611.86636	4.52	−613.55380	7.14	−613.55048	7.21	−611.93268	6.14	−612.16724	6.56
21	$^4C_1$	*t*	*t*	*g*−	*t*	−611.86244	6.98	−613.55457	6.66	−613.55120	6.76	−611.93183	6.68	−612.16720	6.59
22	$^3S_1$	*t*	*t*	*t*	*t*			−613.55298	7.66	−613.54966	7.73	−611.93024	7.68	−612.16543	7.70
23	$^4C_1$	*t*	*t*	*t*	*g*+			−613.55374	7.18			−611.93101	7.19	−612.16646	7.05

[a] The hydroxyl groups are numbered according to Figure 1, and the *t*, *g*+, and *g*− symbols are used to identify the rotamers. The vertical bars between the hydroxyl groups denote the that there is no OH···O interaction. For the B3P/6-31+G(d)// calculation, the BP/6-31G(d) geometry was used. For the B3PW or B3LYP/6-311+G(2d,p)// calculations, the B3P/6-31+G(d) geometry was used. The sg1 grid was used with the diffuse basis sets. [b] kcal/mol. [c] Hartree.

Further inspection of Figure 2 reveals that the correct description of the energetics of the ring puckering is a more difficult task for theoretical methods than the description of the energetics of the hydroxyl rotamers. The results of a basis set and method dependence study for conformers 19−23 are shown in Table 2. The more recent PW (1991) functional provides very similar energy differences compared with the P (1986) functional for the relative energies of hexapyranoses. The PW functional slightly increases the energy differences (cf., Table 2).

Considerable discrepancies between the HF and BP/6-31G-(d) relative energies have been found for conformer 20 ($^{2,5}B$). Although the BP and BPW results agree with each other for the relative stability of this conformer, the inclusion of the exact exchange (B3P) and addition of diffuse functions (+) to the basis set, destabilizes it relative to conformer 19 (thus, supporting the HF result). However, further increase of the basis set (B3PW/6-311+G(2d,p)//B3P/6-31+G(d) calculations using sg1 grid) stabilizes this conformer (cf., Table 2; 1 kcal/mol stabilization). For conformer 19, the HF, BP, and BPW results with the 6-31G(d) basis set agree quite well; however, the B3P/6-31+G(d) method shows a considerable stabilization effect, thus supporting the MM2* results. Further improvement of the basis set to 6-311+G(2d,p) supports the HF, BP, and BPW results with the 6-31G(d) basis set (cf., Table 2). For conformer 21, considerable agreement was found between the results. For conformer 22, the agreement is less perfect, however it is within 1 kcal/mol. The energetic effect of the geometry reoptimization is <0.1 kcal/mol; this value justifies the utilization of the BP/6-31G(d) geometries. For conformer 19, this effect is larger, nearly 0.4 kcal/mol, but this value is well within the precision expected from DFT methods. However, this behavior warns us that if the basis set extension provides a large energy change, then geometry reoptimization is advisable for greater precision, otherwise it seems unnecessary. The B3PW/6-311+G(2d,p)//B3P/6-31+G(d) results in Table 2 suggest that the BP or BPW/6-31G(d) calculations are converged within 1 kcal/mol for conformers 19, 21, 22, and 23. The general agreement between the HF/6-31G(d) and B3PW/6-311+G(2d,p)//B3P/6-31+G(d) calculations in Table 2 supports the use of HF/6-31G(d) method for the energetic studies of the pyranose ring puckering.

Conformers 28, 30, and 32 afforded by the MM2* conformational search were unstable and were transformed by the HF or BP geometry optimizations to conformers 22, 29, and 20, respectively. Conformer 27 ($^2S_0$) requires special attention. This conformer was stable by the HF/3-21G method, however, the HF/6-31G(d) and BP/6-31G(d) methods transformed this conformer to conformer 20 ($^{2,5}B$). If we consider the latter results, the final number of the stable conformers, within the 10 kcal/mol energy window, is 25.

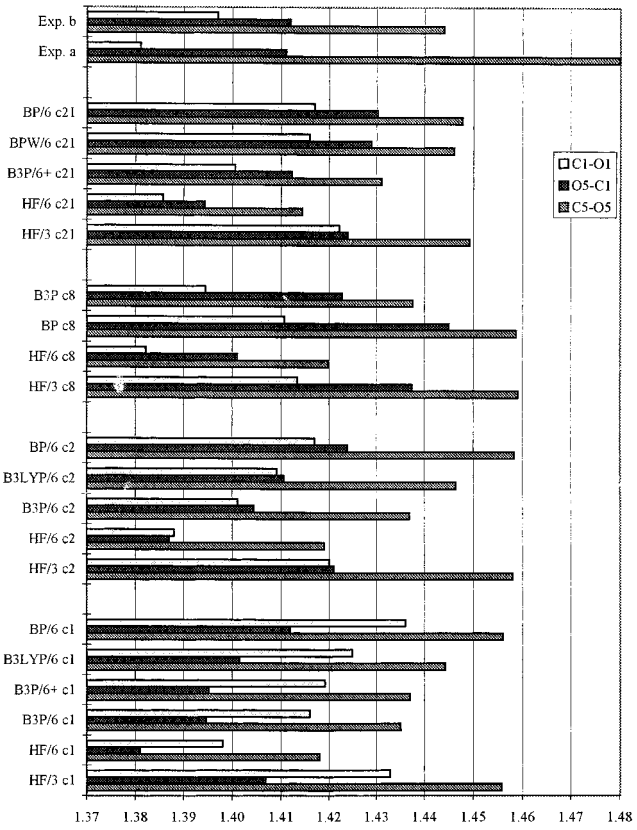

Figure 4. The method and basis set dependence of the C—O bond lengths (Å) in the conformers 1 (c1), 2 (c2), 8 (c8), and 21 (c21), where /3, /6, /6+ denotes 3-21G, 6-31G(d), and 6-31+G(d), respectively. Experimental results: Exp. a (α-L-fucose) and b (methyl α-L-fucoside), are from refs 57 and 58, respectively.

We performed a frequency analysis for the first four $^1C_4$ and for the first three non-$^1C_4$ conformers. The results show that the zero point energy differences between the conformers are rather small, <0.2 kcal/mol; thus, this effect is well below the error bars of the present GGA DFT methods and can be neglected.

4. Molecular Geometries

Bond Lengths. Figure 4 shows the calculated and experimental C1—O1, O5—C5, and C5-O5 bond lengths. Inspection of this figure reveals the typical HF/6-31G(d) error. This method provides too short lengths for the C—O bonds. Improvements of the basis set would provide further bond shortening because of the incorrect treatment of the Coulomb correlation in the HF method. The introduction of the Coulomb electron correlation results in a considerable bond lengthening

effect. The analysis of the equilibrium geometries show that the bond lengthening effects are largest in the BP/6-31G(d) method. Adding diffuse functions to the heavy atoms provides a very small further bond lengthening (∼0.002 Å). Our previous results show that the bond lengthening effect of the BP and BLYP methods is larger than that of the MP2 method for D-glucose (cf., ref 13). The introduction of the exact exchange in the hybrid forms (e.g., B3P, B3PW, or B3LYP) decreases the bond lengthening effects. All these geometry changes are governed by the changes in the spatial electron density distribution, $\rho(\boldsymbol{r})$. It can be observed that two bonded atoms are connected by a maximum electron density (MED) path that terminates near the neighboring nuclei. In this sense, a molecular graph is a network of bond paths. Along a bond path there exists a critical point where the $\nabla\rho(\boldsymbol{r}) = 0$. This critical point is called the bond critical point (BCP) and it is a minimum of $\rho(\boldsymbol{r})$ along the MED path and a maximum of $\rho(\boldsymbol{r})$ in all directions perpendicular to the MED path. The BCP corresponds to a multiple saddle point of $\rho(\boldsymbol{r})$ in three dimensions. Our analysis of the electron density shows that the aforementioned bond shortening effect can be attributed to the fact that the inclusion of the exact exchange increases the electron density in the BCP of the covalent bonds.[54]

The results in Figure 4 show that the HF/3-21G C—O bond lengths agree quite well with the GGA DFT results. The following relations can be established for the C—C and C—O bond lengths: r(HF/6) < r(B3P/6) < r(B3P/6+) < r(MP2/6) < r(B3LYP/6) < r(BPW/6) ≅ r(BP/6), where 6 and + denotes 6-31G(d) and 6-31+G(d) basis sets, respectively. This observation is in perfect agreement with our earlier observations for C—O and O—H bond lengths with double- and triple-zeta quality basis sets.[26]

The X-ray experimental results in Figure 4 show the usual trend in the pyranosidic compounds: the C1—O1 bond is shorter than the O5—C1 bond in α-L-fucose[55] and in methyl α-L-fucoside.[56] This difference is attributed to the exoanomeric effect. There is a considerable difference between the two experiments (cf., Figure 4), and the calculated equilibrium geometries do not show a clear agreement with the experimental results in this respect. Before further discussion it should be noted that the experimental errors for these bond lengths are 0.009 and 0.004 Å for α-L-fucose and methyl α-L-fucoside, respectively. We note that the crystal packing forces and intermolecular interactions are dominant in the solid phase, thus the molecular structure and conformation differs between gas phase and solid phase. Moreover, it is known that different crystal forms result in different X-ray structures for sugars.

For the most stable conformer (c1 in Figure 4), all the methods are consistent in that the C1—O1 bond is longer by ∼0.024 Å than the O5—C1 bond, thus contradicting the X-ray results. It can be assumed that in the exoanomeric effect the exocyclic O1 lone pair delocalizes into the endocyclic σ^*-(C1—O5) bond orbital. This may occur for g-(a) or t(a) orientations of the anomeric hydroxyl group, whereas g+(a) orientation does not permit exoanomeric stabilization according to this model.

Let us next consider the conformer number 8. In this conformer the C1—O1 bond is shorter by 0.034 Å than the O5—C1 bond, which is in agreement with X-ray results (cf., Figure 4 and Table 3). The exoanomeric effect is also present in this conformer because of the t(a) position of the anomeric hydroxyl group. By systematic comparisons of the C1—O1 and O5—C1 bond lengths in the various rotamers, one can understand how the positions of the hydroxyl groups influence the aforementioned two bond lengths. The corresponding BP/6-31G(d) results are shown in Table 3. Turning the anomeric hydroxyl group from t(a) to g+(a) position shortens the O5—C1 bond by 0.021 Å, whereas it lengthens the C1—O1 bond length slightly by 0.006 Å (cf., 8 → 2 transition in Table 3). This example is very interesting because the exoanomeric effect is switched off in the g+(a) position, thus the geometric consequences of the exoanomeric effect are clearly visible. The exoanomeric effect lengthens the O5—C1 and shortens C1—O1 bond lengths as expected. The energetic effect is overshadowed by a large stabilization effect of the formation of a continuous hydroxyl chain in conformer 2. The transition 8 → 9 rotates the 4th OH group away from the O5, and thus shortens the O5—C1 bond by 0.011 Å.

TABLE 3: The O5—C1 and C1—O1 Bond Lengths, Their Difference, and the Relative Bond Length Changes Due to the Rotation of Various Hydroxyl Groups Compared with Conformer 8 Calculated with the BP/6-31G(d) Method for $^1C_4$ α-L-Fucose

		bond length[a]		
transition	rotated hydroxyl	O5—C1	C1—O1	difference[a]
Conformer 8		1.445	1.411	−0.034
8 → 2	1:t → g+	−0.021	0.006	0.027
8 → 9	4:g+ → g−	−0.011	0.006	0.017
8 → 10	2:g+ → g−	−0.016	0.017	0.033
8 → 10 → 11	2. 3:g+ → t	−0.019	0.018	0.037
8 → 10 → 11 → 1	2. 3. 4:g+ → t	−0.033	0.025	0.058

[a] Å.

The most stable conformer can be reached starting form conformer 8 by three consecutive rotations of the 2nd, 3rd, and 4th hydroxyl groups. This is the 8 (t | g+ g+ g+) → 10 (t g- | g+ g+) → 11 (t g- t | g+) → 1 (t g- t t) transition in Table 3. The first step (8 → 10 transition that forms an O1···H—O2 hydrogen bond) increases the C1—O1 bond length by 0.017 Å and shortens the O5—C1 bond by 0.016 Å. This latter effect can be attributed to the partial occupation of the O1 lone pair that weakens the $n \rightarrow \sigma^*$ interaction. The third step (11 → 1 transition) has a rather strong synergetic effect with the first step. Turning the 4th OH group away from the O5 atom, so as to form a continuous chain of OH groups, lengthens the C1—O1 bond and shortens the O5—C1 bond (cf., transition 8 → 9 in Table 3). The effects occurring in the first and third steps (forming O1···H—O2 and breaking O5···H—O4 hydrogen bonds) clearly override the exoanomeric effect and lead to longer C1—O5 and shorter O5—C1 bonds in conformer 1.

This analysis shows that three conditions have to be fulfilled to reproduce the X-ray results: the first hydroxyl group must be in the anti position, the 2nd hydroxyl should not interact with the first, and the O5 should be bound by a hydrogen bond. These conditions are fulfilled in the solid phase (the O5 is bound by an intermolecular hydrogen bond), thus the X-ray results can be explained by the present theoretical results.

Dihedral Angles. The C—C—O—H dihedral angles in Table 4 frequently show large deviation (up to 60°) from ideal values (−60°, +60°, and 180°). This deviation is a consequence of the OH···O interactions. The inclusion of the electron correlation turns these dihedral angles further away from their ideal values so that the hydrogen bonding becomes more perfect. The GGA DFT results presented in Table 4 account for the expected correlation effects.

Earlier characterization of the O···H interactions used the O···H distance and one only angle (e.g. O···H—O or R—O—C···H out of plane angle[57]). Although the O···H distance is certainly a very important parameter to characterize

TABLE 4: Geometrical Parameters of Some Selected α-L-Fucose Conformations Calculated by HF and Different GGA DFT Methods Supplemented with 6-31G(d) Basis Set[a]

		C—C—O—H				OH				OH					OH					OH					
no.	method	1	2	3	4	x	y	O···H	C—O···H	δ	x	y	O···H	C—O···H	δ	x	y	O···H	C—O···H	δ	x	y	O···H	C—O···H	δ
1	HF/3	178.9	−79.1	163.6	168.0	1	2	2.203	82.5	145.7	2	3	2.389	75.7	−115.7	3	4	2.166	82.5	133.4					
1	HF	175.1	−78.6	165.8	169.6			2.256	81.8	146.1			2.422	74.8	−113.0			2.253	81.1	137.9					
1	BP	171.4	−78.1	159.0	161.2			2.238	82.0	142.0			2.316	74.9	−113.6			2.133	80.6	129.0					
1	B3P	172.0	−78.3	160.4	163.0			2.220	82.1	142.5			2.329	74.9	−113.9			2.145	80.9	130.8					
1	B3P/+	170.4	−79.0	163.2	166.4			2.228	82.4	141.1			2.386	75.1	−113.9			2.211	81.1	134.3					
2	HF/3	43.6	49.9	37.6	28.5	2	1	2.135	82.3	137.4	3	2	2.491	74.2	−122.5	4	3	2.251	79.5	−124.8	5	4	2.286	76.9	−105.5
2	HF	44.1	47.7	35.5	33.6			2.222	81.2	140.8			2.468	74.3	−119.0			2.282	79.0	−117.7			2.429	73.6	−103.0
2	BP	31.4	42.2	42.3	37.0			2.089	80.1	131.6			2.401	73.6	−118.9			2.269	78.7	−113.4			2.330	75.8	−101.5
2	BP	31.4	42.2	42.3	37.0			2.089	80.1	131.6			2.401	73.6	−118.9			2.269	78.7	−113.4			2.330	75.8	−101.5
2	B3P	34.3	43.2	41.4	35.9			2.101	80.5	133.0			2.400	73.8	−119.2			2.256	78.8	−114.8			2.333	75.5	−102.0
2	B3LYP-sg	35.2	44.1	41.9	36.4			2.124	80.8	134.9			2.430	73.7	−118.5			2.279	78.8	−114.4			2.365	75.2	−101.7
3	HF/3	46.7	48.7	40.2	−65.2	2	1	2.175	81.6	134.3	3	2	2.414	75.7	−118.0	4	3	2.139	84.1	143.6					
3	HF	45.6	48.1	42.8	−69.9			2.234	81.0	139.6			2.437	75.1	−114.8			2.219	82.3	140.5					
3	BP	34.5	41.8	43.2	−73.9																				
3	B3P	37.1	43.3	42.3	−71.6			2.126	80.1	131.3			2.341	75.1	−116.9			2.164	82.6	138.8					
3	B3LYP-sg	38.0	43.4	43.6	−73.0			2.148	80.4	132.3			2.366	75.1	−115.5			2.204	82.3	137.4					
4	HF/3	181.7	173.9	166.9	166.2						2	3	2.279	80.7	139.9	3	4	2.108	83.6	137.6					
4	HF	179.6	171.5	170.8	169.2								2.361	78.5	138.6			2.226	81.8	143.4					
4	BP	180.3	165.0	159.6	160.3								2.219	78.6	129.2			2.084	81.7	130.4					
4	B3P	179.9	166.8	162.3	162.2								2.236	78.8	131.8			2.101	81.9	133.6					
5	HF/3	50.5	57.2	−82.2	162.1	2	1	2.206	81.9	142.9	3	2	2.449	77.6	123.8	3	4	2.201	79.9	−111.5					
5	HF	48.9	56.2	−81.6	164.9			2.253	81.3	148.1			2.465	76.8	124.5			2.276	79.0	−108.4					
5	BP	37.1	46.7	−76.6	157.9			2.121	80.4	134.1			2.371	76.0	125.4			2.166	78.7	−104.5					
6	HF/3	42.9	158.1	172.3	170.4	2	1	2.237	78.9	−115.8	2	3	2.427	77.5	121.8	3	4	2.182	82.7	142.8					
6	HF/6	41.8	156.0	174.2	172.1			2.274	78.6	−112.0			2.450	76.5	122.1			2.264	81.5	147.2					
6	BP	31.7	163.6	164.9	163.8			2.159	77.9	−106.9			2.361	76.0	125.7			2.148	81.1	135.3					
19	HF/3	−172.4	−157.4	−28.0	−77.0	1	2	2.003	89.0	−151.1						4	3	1.926	89.1	−174.1					
19	BP	−170.4	−153.4	−22.2	−46.5			1.961	88.3	−151.4								1.888	88.2	−147.1					
19	BPW	−170.9	−154.0	−22.3	−46.9			1.976	88.2	−152.1								1.900	88.0	−147.6					
19	B3P/+	179.2	−160.2	−26.0	−58.4			2.043	88.8	−164.5								1.939	89.1	−161.8					
20	HF/3	−171.3	−161.7	−28.6	38.5	1	2	2.029	88.7	−150.2						4	3	1.955	88.7	−89.2	5	4	2.236	77.6	−111.8
20	HF	−175.9	−162.1	−57.8	52.8			2.105	87.6	−159.1								2.281	85.5	−72.3			2.695	68.4	−110.0
20	BP	−166.4	−157.8	0.5	36.8			1.973	87.0	−144.0								1.833	85.9	−96.7			2.107	79.7	−104.9
20	BPW	−166.4	−157.8	0.4	36.9			1.979	86.9	−144.2								1.850	85.8	−96.7			2.124	79.4	−105.0
20	B3P/+	−174.8	−164.1	4.2	32.9			2.048	87.4	−153.6								1.895	85.9	−103.1			2.174	78.8	−102.6
21	HF/3	−156.8	−156.5	−46.2	−143.9	1	2	2.120	82.5	−123.4						4	3	2.255	81.2	122.5					
21	HF	−164.8	−159.6	−44.5	−151.4			2.194	81.6	−134.4								2.292	79.6	111.2					
21	BP	−159.8	−152.6	−36.0	−161.5			2.064	82.1	−128.0								2.130	80.4	103.0					
21	BPW	−160.1	−152.7	−36.3	−161.1			2.078	81.8	−128.3								2.147	80.1	103.1					
21	BP3P/+	−163.4	−156.8	−41.0	−151.9			2.130	82.6	−133.0								2.213	80.5	110.8					
22	HF/3	−150.9	−137.2	175.6	175.4	1	2	1.915	89.8	−138.8	2	3				3	4	2.378	78.1	139.8					
22	HF	−166.4	−154.1	178.6	174.6			2.074	87.6	−152.2								2.365	78.4	146.3					
22	BP	−153.6	−143.8	170.9	166.2			1.935	87.0	−134.0			2.887	62.0	169.4			2.203	79.4	136.2					
22	B3P/+	−164.2	−148.9	177.4	172.3			2.009	88.1	−147.6			2.818	66.0	167.7			2.336	78.6	144.0					
23	HF/3	−157.1	−156.9	−51.9	−76.0	1	2	2.123	82.4	−123.7						4	3	2.246	83.8	−169.2					
23	HF	−164.9	−160.1	−52.1	−72.2			2.190	81.7	−134.5								2.296	82.2	−170.2					
23	BP	−160.8	−153.1	−44.1	−59.0			2.063	82.2	−129.2								2.178	82.7	−153.4					
23	B3P/+	−165.0	−157.3	−50.1	−74.6			2.129	82.7	−134.8								2.250	83.0	−172.1					
24	HF/3	−172.3	−156.6	−15.1	−153.2	1	2	1.996	89.3	−152.1						4	3	1.962	84.9	112.3					
24	HF	−177.3	−160.8	−10.7	−160.9			2.103	87.9	−162.1								2.044	83.3	101.1					

TABLE 4 (Continued)

24	BP	−168.7	−152.2	−6.3	−164.5			1.954	88.4	−149.9					1.903	84.2	95.3
25	HF/3	−157.0	−160.7	−168.5	159.6	1	2	2.111	82.1	−121.2			3	4	1.949	89.4	171.1
25	HF	−163.0	−161.7	−169.9	165.3			2.173	81.5	−131.2					2.095	87.2	172.3
25	BP	−152.4	−155.8	168.9	154.7			2.020	82.0	−119.5					1.952	87.0	146.5
26	HF/3	−154.8	−155.0	−160.4	84.4	1	2	2.133	81.9	−120.5			3	4	2.165	81.7	−128.7
26	HF	−162.2	−158.0	−160.2	83.5			2.190	81.4	−131.2					2.242	80.7	−131.1
26	BP	−157.1	−151.3	−158.9	87.2			2.059	81.8	−124.9					2.120	81.4	−128.8
27	HF/3	−168.7	−154.2	−144.5	85.7	1	2	1.974	89.4	−147.7			3	4	1.961	87.3	−126.1
29	HF/3	−56.6	84.7	−48.3	−68.4	2	1	2.214	83.6	−129.3			4	3	2.205	84.2	−160.9
29	BP	−43.7	86.7	−39.9	−51.2			2.068	83.5	−125.0					2.135	82.7	−144.3
31	HF/3	−151.3	−88.3	156.0	162.0	1	2	1.960	89.3	−172.8	2	3	2.193	80.1	−119.4	3	4	2.044	84.5	128.8
31	BP	−154.8	−82.5	150.3	155.7			2.036	88.4	−170.8			2.134	78.7	−112.5			2.001	83.0	124.2
33	HF/3	−33.1	70.5	−81.6	171.6	2	1	1.954	88.9	−159.4	3	2	2.819	70.1	120.6	3	4	2.404	76.3	−115.0
33	BP	−33.0	97.0	−72.4	162.4			1.939	85.9	−125.2			3.180	65.9	132.5			2.200	77.7	−103.9

[a] The dihedral angles (C—C—O—H, and $\delta \equiv$ H···C—O—H or H···C—O—C for 4···5 interactions), and angles (C—O···H) are in degree distances (O···H) are in Å; the hydrogen bonds are identified by the numbers referring to the interacting Ox and Hy atoms; sg1 notates the use of sg1 grid for the DFT integrals; + notates the addition of diffuse functions to the 6-31G(d) basis set.

O···H interaction, one additional angle is clearly not enough to describe a relative position in the space. Moreover, the O···H—O angle does not reflect the relation between the position of the lone pair of the oxygen atoms and the position of the hydrogen atom. The relative position of an interacting hydrogen atom can be unambiguously given in a polar coordinate system centered on the O atom. The three coordinates that provide the exact position of the H atom are the O···H distance, the C—O···H angle, and the R—O—C···H dihedral angle. Bader et al.[58] showed that two minima exist for the Laplace concentrations of electron density around the O atoms. These minima can be taken as rigorous and exact mathematical representations of the centers of the lone pairs of the oxygen atom. The position of these concentrations can be characterized the same way in the polar coordinate system as was shown for the H atoms. Our detailed analysis of water and 1,2-ethanediol shows that the angular positions of the minima of the negative Laplacian concentrations (LC) are at 99−100° and ±104−105° for the C—O—LC and R—O—C···LC angles, respectively. Further analysis shows that the negative Laplacian concentration remains considerably large between the two minima (large torsion angles) and a sharp cut off experienced at small angles (<90°). Our earlier results show that the shape of the negative Laplacian concentration around the minima plays a role in the O···H interactions.[59]

The results in Table 4 show that the O*x*···H*y* interactions can readily be differentiated by the proposed geometric parameters, where *x* refers to the serial number of the acceptor O (e.g., O1 is the anomeric oxygen) and *y* refers to the serial number of the donor OH group. It is apparent that in the $^1C_4$ conformers the O1···H2 and O3···H4 axial−equatorial O···H interactions are quite different from the O2···H3 equatorial−equatorial interactions (cf., Table 4). This latter interaction can be characterized by a considerably longer O···H distance, and smaller C—O···H angle. For the $^4C_1$ conformers, the O···H distances are frequently considerably smaller (1.85−2.25 Å) and the C—O···H angles are larger (88−89°) than in the $^1C_4$ conformers. This result signals a stronger interaction. This stronger interaction is reflected in the C—C—O—H dihedral angles, especially for conformer 20 in which the third C—C—O—H dihedral angle is about zero according to the GGA DFT methods (cf., Table 4). Thus the O···H interaction counterbalances the C—C O—H eclipsed repulsion. Most of these interactions can be characterized by a large R—O—C···H dihedral angle (150−160°). Van den Enden et al.[57] classified the O···H nonbonded interactions into three major groups: π-, sp^3-, and σ-type interactions. The ideal values for the out-of-plane angle between an Ox···Hy vector and the Cx-Ox-Hx plane would be 90°, 125°, and 180°, respectively. The sp^3 type of interaction seems to be energetically favorable, and hydrogen atoms are trying to keep this angle as much as possible. Although the small angles for C—O···H or R—O—C···H provide weaker, π-type interaction, the large R—O—C···H torsion angles (>120°), the σ-type interactions are not so disadvantageous. This result supports our proposition that the shape of the negative Laplacian concentration around the O atom plays a role in the interaction.

Finally, we note that according to the Bader's electron density analysis, the necessary condition for a bond is the existence of a MED (bond) path between two nuclei.[60,61] Our results for 1,2-ethanediol[59] and for various monosaccharides[29] showed that in most of the cases, no such bond path exists between the interacting OH groups. Thus, these interactions cannot be called bonds in Bader's sense.

5. Conclusions

The following conclusions can be drawn from our analysis:

1. The MM2*-LMOD search provided 33 conformers for the α-L-fucose within a 10 kcal/mol energy window. The HF/3-21G method reduced this number to 26. The HF/6-31G(d) and BP/6-31G(d) methods eliminated one more.

2. The BP, B3P, and B3LYP methods provided very similar and consistent results for the energetic order of the various rotamers for $^1C_4$ α-L-fucose. The GGA DFT methods show a relative stabilization effect compared with the HF method for the structures in which the fourth hydroxyl group interacts with the oxygen atom of the pyranose ring. The basis set extension and DFT functional studies show that these correlation effects are exaggerated by the BP/6-31G(d) method, and the converged relative energies are between the HF/6-31G(d) and BP/6-31G(d) values. The most stable $^1C_4$ α-L-fucose conformers are considerably (by 6 kcal/mol) more stable than the other ring conformations. The zero-point energy difference between the most stable $^1C_4$ and non-$^1C_4$ conformers is <0.2 kcal/mol.

3. The orientations of the hydroxyl groups are not independent of each other. In the most stable rotamers, the number of possible OH···O interactions is maximal, leading to the formation of an intramolecular chain of hydroxyl groups. The formation of these chains leads to counterclockwise or clockwise unidirected, or concentrated, non-unidirected patterns. The concentration of the interactions results in somewhat less stable conformations. These interactions dramatically reduce the number of possible rotamers and the rotational entropy of the hydroxyl groups, thereby making sugars effective information encoders.

4. The following relations were established for the C—C and C—O bond lengths: r(HF/6) < r(B3P/6) < r(B3P/6+) < r(MP2/6) < r(B3LYP/6) < r(BPW/6) ≅ r(BP/6), where 6 and + denotes 6-31G(d) and 6-31+G(d) basis sets, respectively. The incorrect treatment of the electron correlation effects in the HF method leads to bonds that are too short. The GGA DFT methods correct this problem and provide similar results to that of the MP2 method.

5. The X-ray experimental results show that the C1—O1 bond is shorter than the O5—C1 bond in the α-L-fucose and in the methyl α-L-fucoside. For the most stable conformer, all the theoretical methods are consistent in that the C1—O1 bond is longer by ∼0.024 Å than the O5—C1 bond, thus contradicting the X-ray results. We have shown that three conditions have to be fulfilled to reproduce the X-ray results: the first hydroxyl group must be in the anti position, the 2nd hydroxyl should not interact with the first, and the O5 should be bound by a hydrogen bond. These conditions are fulfilled in the solid phase; thus, the X-ray results can be explained by the present theoretical results.

6. We propose using a polar coordinate system centered on the acceptor oxygen atom to characterize the O···H interactions. The three coordinates that provide the exact position of the H atom are the O···H distance, the C—O···H angle, and the R—O—C···H dihedral angle. The axial−equatorial and the equatorial−equatorial hydroxyl interactions were readily distinguished using these coordinates.

7. The GGA DFT methods reflect correctly the expected structural changes due to the inclusion of the electron correlation. The so-called sp^3 type of interaction between the hydroxyl groups is energetically favorable; and, the interacting hydrogen atoms are trying to keep 105−120° for the Hx—Ox—Cx···Hy dihedral angle. The deficiencies of the HF method result in the wrong C—O—H equilibrium bond and C—C—O—H dihedral angles.

8. The stronger hydroxyl−hydroxyl interactions in $^4C_1$ conformers are reflected in the C—C—O—H dihedral angles, especially for the conformer 20 in which the third C—C—O—H dihedral angle is about zero according to the GGA DFT methods. Thus the O···H interactions are able to counterbalance the C—C O—H eclipsed repulsion.

Acknowledgment. G.I.C. acknowledges the PAST professorship to professor Jean-Louis Rivail and the French Government. We 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) Lamba, D.; Segre, A. L.; Fabrizi, G.; Matsuhiro, B. *Carbohydrate Res.* **1993**, *243*, 217, and references cited therein.
(2) Flowers, H. M. *Adv. Carbohydr. Chem. Biochem.* **1981**, *39*, 279.
(3) Lemieux, R. U. *Chem. Soc. Rev.* **1978**, 423.
(4) Diakun, K. R.; Matta, K. L. *J. Immunol.* **1989**, *142*, 2037.
(5) Kennedy, J. F.; White, C. A. *Bioactive Carbohydrates in Chemistry, Biochemistry and Biology*; Ellis Horwood: Chichester, 1983; p 182.
(6) Santer, U. V.; Glick, M. C.; Van Halbeek, H.; Vliegenthart, J. F. G. *Carbohydr. Res.* **1983**, *120*, 197.
(7) Weis, W. I.; Dirkamer, K. *Curr. Biol.* **1994**, *91*, 728.
(8) Sherrif, S.; Chang, C. Y.; Ezekowitz, R. A. B. *Nat. Struct. Biol.* **1994**, *11*, 789.
(9) (a) Phillips, M. L.; Nudelman, E.; Gaeta, F. C. A.; Perez, M.; Singhal, A. K.; Hakomori, S.; Paulson, J. C. *Science* **1994**, *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.
(10) (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; (c) 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.
(11) (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. USA* **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., Intl. Ed. Engl.* **1994**, *33*, 178.
(12) Brady, J. W.; Schmidt, R. K. *J. Phys. Chem.* **1993**, *97*, 958.
(13) Csonka, G. I.; Éliás, K.; Csizmadia, I. G. *J. Comput. Chem.* **1997**, *18*, 330.
(14) Asensio, J. L.; Jimenez-Barbero, J. *Biopolymers* **1995**, *35*, 55, and references cited therein.
(15) Reilig, S.; Schlenkrich, M.; Brinkmanmn, J. *J. Comput. Chem.* **1996**, *17*, 450.
(16) Senderowitz, H.; Parish, C.; Still, W. C. *J. Am. Chem. Soc.* **1996**, *118*, 2078.
(17) Barrows, S. E.; Dulles, F. J.; Cramer, C. J.; French, A.D.; Truhlar, D.G. *Carbohyd. Res.* **1995**, *276*, 219.
(18) Reeves, R. E. *J. Am. Chem. Soc.* **1949**, *71*, 215.
(19) Dowd, M. K.; French, A. D.; Reilly, P. J. *Carbohydr. Res.* **1994**, *264*, 1.
(20) Angyal, S. J.; Pickles, V. A. *Aust. J. Chem.* **1972**, *25*, 1695.
(21) (a) Snyder, J. R.; Serianni, A. S. *J. Org. Chem.* **1986**, *51*, 2694; (b) Forster, M. J.; Mulloy, B. *Biopolymers* **1993**, *33*, 575.
(22) Cremer, D.; Pople, J. A. *J. Am. Chem. Soc.* **1975**, *97*, 1345.
(23) Cumming, D. A.; Carver, J. P. *Biochemistry* **1987**, *26*, 6664.
(24) Carver, J. P.; Michnik, S. W.; Imberty, A.; Cumming, D. A. *Carbohydrate Recognition in Cellular Function*, CIBA Foundation Symposium 158; Wiley: New York, 1989; p 231.
(25) Carver, J. P. *Pure Appl. Chem.* **1993**, *65*, 763.

(26) Csonka, G. I.; Csizmadia, I. G. *Chem. Phys. Lett.* **1995**, *243*, 419.
(27) Kneisler, J. R.; Allinger, N. L. *J. Comput. Chem.* **1996**, *17*, 757 .
(28) Csonka, G. I.; Éliás, K.; Csizmadia, I. G. *Chem. Phys. Lett.* **1996**, *257*, 49.
(29) Csonka, G. I.; Kolossváry, I.; Császár, P.; Éliás, K.; Csizmadia, I. G. *J. Mol. Struct. (THEOCHEM)* **1997**, *395*−*396*, 29.
(30) Kolossváry, I.; Guida, W. C. *J. Am. Chem. Soc.* **1996**, *118*, 501. The LMOD procedure is now available in the 6.0 release of MacroModel.
(31) (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.
(32) Still, W. C. MacroModel 4.5, Columbia University, New York, NY, http://www.columbia.edu/cu/chemistry/mmod/mmod.html.
(33) Gundertofte, K.; Liljefors, T.; Norrby, P.-O.; Pettersson, I. *J. Comput. Chem.* **1996**, *17*, 429.
(34) (a) Ponder, J. W.; Richards, F. M. *J. Comput. Chem.* **1987**, *8*, 1016. (b) Schlick, T.; Overton, M. *J. Comput. Chem.* **1987**, *8*, 1025.
(35) 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: Pittsburgh PA, 1995.
(36) Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098.
(37) Perdew, J. P. *Phys. Rev. B* **1986**, *33*, 8822.
(38) Perdew, J. P. In *Electronic Structure of Solids '91*; Ziesche, P., Eschrig, H., Eds.; Akademie Verlag: Berlin, 1991; p 11.
(39) Vosko, S. H.; Wilk, L.; Nussair, M. *Can. J. Phys.* **1980**, *58*, 1200.
(40) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648.
(41) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785.
(42) 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 92/DFT, Revision F*; Gaussian: Pittsburgh PA, 1993.
(43) Hehre, W. J.; Radom, L.; L.; Schleyer, P. v. R.; Pople, J. A. *Ab Initio Molecular Orbital Theory*; John Wiley & Sons: New York, 1986, and references therein.
(44) Goodman, J. M.; Still, W. C. *J. Comput. Chem.* **1991**, *12*, 1110.
(45) (a) French, A. D.; Rowland, R. S.; Allinger, N. L. In *Computer Modeling of Carbohydrate Molecules, ACS Symposium Series, No. 430*; French, A. D., Brady, J. W., Eds.; American Chemical Society: Washington, D.C., 1990; p 120; (b) Tran, V.; Brady, J. W. In *Computer Modeling of Carbohydrate Molecules, ACS Symposium Series, No. 430*; French, A. D., Brady, J. W., Eds.; American Chemical Society: Washington, D.C., 1990; p 213.
(46) Cramer, C. J.; Truhlar, D. E. *J. Am. Chem. Soc.* **1993**, *115*, 5745.
(47) Polavarapu, P. L.; Ewig, C. S. *J. Comput. Chem.* **1992**, *13*, 1255.
(48) Cramer, C. J.; Truhlar, D. E. *J. Am. Chem. Soc.* **1994**, *116*, 3892.
(49) Csonka, G. I.; Ángyán, J. G. *J. Mol. Struct. (THEOCHEM)* **1997**, *393*, 31.
(50) Treppen, B. J.; Cao, M.; Frey, R. F.; Van Alsenoy, C.; Miller, D. M.; Schäfer, L. *J. Mol. Struct. (THEOCHEM)* **1994**, *314*, 169.
(51) Gill, P. M. W.; Johnson, B. G.; Pople, J. A. *Chem. Phys. Lett.* **1994**, *217*, 65.
(52) Del Bene, J. E.; Person, W. B.; Szczepaniak, K. *J. Phys. Chem.* **1995**, *99*, 10705.
(53) Novoa, J. J.; Sosa, C. *J. Phys. Chem.* **1995**, *99*, 15837.
(54) Csonka, G. I.; Nguyen, N. A.; Kolossváry, I. *J. Comput. Chem.*, in press.
(55) Longchambon, F.; Giller-Pandraud, H. *Acta Crystallogr., B* **1977**, *33*, 2094.
(56) Lamba, D.; Segre, A. L.; Fabrizi, G.; Matsuhiro, B. *Carbohydrate Res.* **1993**, *243*, 217, and references therein.
(57) Van den Enden, L.; Van Alsenoy, C.; Scarsdale, J. N.; Shäfer, L. *J. Mol. Struct. (THEOCHEM)* **1983**, *104*, 471.
(58) Bader, R. F. W.; MacDougall, P. J.; Lau, C. D. H. *J. Am. Chem. Soc.* **1984**, *106*, 1594.
(59) Csonka, G. I.; Anh, N.; Ángyán, J.; Csizmadia, I. G. *Chem. Phys. Lett.* **1995**, *245*, 129.
(60) Bader, R. F. W. *Chem. Rev.* **1991**, *91*, 993.
(61) Bader, R. F. W. *Atoms in Molecules − A Quantum Theory*; University of Oxford Press: Oxford, 1990.