Computational Studies of the Arabinofuranose Ring: Conformational Preferences of Fully Relaxed Methyl α-D-arabinofuranoside

Patrick R. McCarren, Matthew T. Gordon, Todd L. Lowary,* and Christopher M. Hadad*

Department of Chemistry, The Ohio State University, 100 West 18th Avenue, Columbus, Ohio 43210

Received: December 31, 2000; In Final Form: April 3, 2001

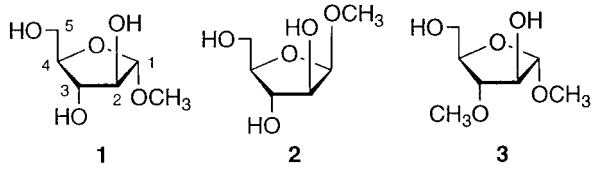
Two approaches for identifying the minimum energy conformers of methyl α-D-arabinofuranoside **1**, in the gas phase have been explored and compared. In the first approach (the constrained envelope method), 30 previously reported envelope geometries of **1** were allowed to fully optimize at the B3LYP/6-31G* level. B3LYP/6-31+G** single-point energies of these optimized structures were also determined, which led to the identification of the $^3T_4$ and $^2T_1$ ring conformers as the Northern (N) and Southern (S) minima, respectively, with the latter being the global minimum. The importance of intramolecular hydrogen bonding was probed by optimizing another set of 30 envelope geometries with initial geometries biased against the formation of these stabilizing interactions. These calculations led to the same two families of low-energy ring conformers ($^3T_4$ and $^2T_1$); however, the N, and not the S, conformer was the global minimum without hydrogen bonding. The second approach involved the identification of conformers for **1** through the use of a Monte Carlo search coupled with molecular mechanics and then further optimization of these structures at higher levels of theory (HF/6-31G* and B3LYP/6-31G*). Subsequent B3LYP/6-31+G** single-point energy calculations afforded results that are similar to the constrained envelope method, but the stochastic approach led to more low-energy conformers, and to a new global minimum. A comparison of these computational results with the experimentally determined solution conformation of **1** is also presented.

Introduction

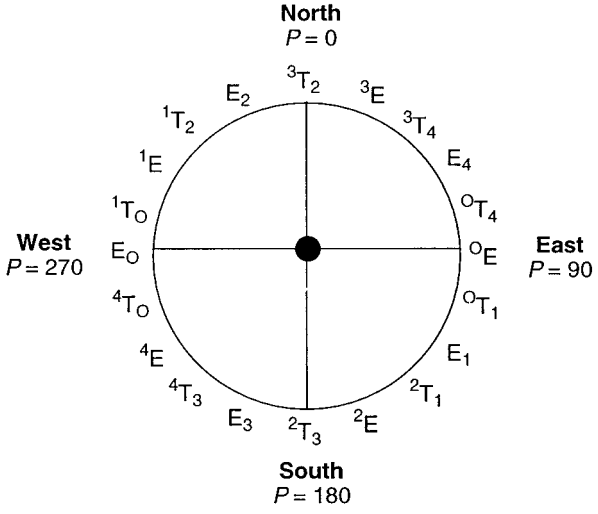
In two earlier papers, we have reported computational investigations into the potential energy surface of the D-arabinofuranose ring.[1,2] These studies were undertaken to provide structural information that would facilitate the interpretation of NMR-based conformational investigations of oligosaccharides comprised of arabinofuranose residues.[3] The glycans under investigation are fragments of two polysaccharides, an arabinogalactan (AG) and lipoarabinomannan (LAM) that are the major structural components of the cell wall complex in mycobacteria. This family of bacteria includes the human pathogens *Mycobacterium leprae* and *Mycobacterium tuberculosis*, which are the causes of leprosy and tuberculosis, respectively.[4]

In mycobacterial AG and LAM, both α-D- and β-D-arabinofuranosyl residues are present and, in our previous investigations,[1,2] we have probed the conformational preferences of methyl α-D-arabinofuranoside **1**, and methyl β-D-arabinofuranoside, **2** (Scheme 1). Using a method developed by Serianni and co-workers,[5] these studies were performed by minimization of the 10 possible envelope conformers of each ring system. Maintaining the envelope ring conformation was achieved by fixing four of the atoms in a plane, while all of the other parameters were allowed to relax to their preferred values. A number of orientations about the exocyclic bonds were also explored. In this manner, partial potential energy surfaces for both **1** and **2** were determined, and it was demonstrated that in common with other furanose rings, these systems have a local minimum in both the northern (N) and southern (S) hemisphere of the pseudorotational wheel (Scheme 2). On the basis of these investigations, it appears that these rings behave according to the traditional two-state north/south (N/S) equilibrium model

SCHEME 1



SCHEME 2



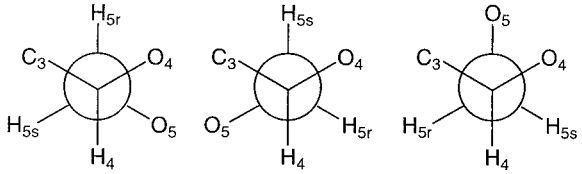
Pseudorotational Itinerary for a D-aldofuranose ring

that is generally used to describe the conformational preferences of furanose rings.[6]

From these studies, we obtained useful information, particularly the relative conformer energies and the dependence of

10.1021/jp004611p CCC: $20.00 © 2001 American Chemical Society
Published on Web 05/16/2001

SCHEME 3



various geometrical parameters (e.g., bond lengths, bond angles, and dihedral angles) as a function of ring conformation. Nevertheless, the planar restraint we employed in order to maintain the envelope conformation prohibited us not only from determining the conformational energies of the twist conformers but also from understanding the minimum energy conformers of fully relaxed α-D- and *β*-D-arabinofuranosyl rings. Furthermore, although this approach is amenable for the study of monosaccharides, the number of calculations required becomes cumbersome with larger oligomers.

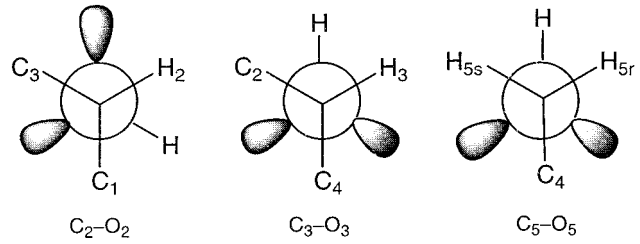
We now report studies which address these issues. In this paper, we have determined gas-phase N and S minima for **1**, which were obtained through minimization of each of our previously reported envelope geometries after removal of the planar constraint. The effect of hydrogen bonding on the potential energy surface has also been investigated. Furthermore, in the interest of developing a protocol more practical for use with oligofuranosides, we have explored the possibility of using a combination of low-level molecular mechanics and high-level ab initio and density functional theory calculations to accurately identify the minimum energy conformers. A library of conformers of **1** was generated via a Monte Carlo search, coupled with molecular mechanics optimization, and subsequently optimized at higher levels of theory (HF/6-31G* and B3LYP/6-31G*). The resulting unique structures were then subjected to single-point energy calculations with larger basis sets. These results compare very favorably with those from the optimization of the idealized envelope conformers, and this approach may be of use in computational investigations of oligosaccharides comprising furanose rings.

Methods

Ab initio molecular orbital[7] and density functional theory[8] calculations were conducted using Gaussian 98.[9] HF/6-31G* and B3LYP/6-31G* methods were used for geometry optimizations.[8c,8d] The default convergence criteria in Gaussian 98 were used. Single-point energies for unique conformers (see text) were calculated using the 6-31+G** basis set. We,[2] and others,[10] have shown that for intramolecular H-bonded molecules, this flexible basis set provides relative energies between conformers that are consistent with larger and even more flexible basis sets.

Full Optimization of Envelope Geometries. The B3LYP/6-31G* envelope structures previously calculated for methyl α-D-arabinofuranoside **1**, were used as starting points for the full optimizations. In each of these geometries, the envelope conformation was maintained by fixing four of the ring atoms in a plane. The choice of exocyclic torsion angles was arbitrary except for the C_1-O_1 and C_4-C_5 bonds (see Scheme 1 for atom numbering). The dihedral angle about the C_1-O_1 bond was initially chosen in all starting structures to maximize the exoanomeric effect; specifically, the methyl group was placed antiperiplanar to C_2. For the C_4-C_5 bond, each of the three staggered conformations, gg, gt, or tg were explored (Scheme 3). Therefore, 30 geometries of **1** were considered, all of which possessed a significant amount of intramolecular hydrogen bonding after optimization. To probe the importance of this hydrogen bonding on the potential energy surface, we constructed another series of envelope structures in which the hydroxyl hydrogens were initially oriented such that intramolecular hydrogen bonding was removed. The orientation of the hydroxyl hydrogens were as follows: OH_2 anti to C_3, OH_3 anti to C_4, and OH_5 anti to C_4 (Scheme 4). As was done previously, the C_1-O_1 torsion angle was set antiperiplanar to C_2, and each of the three possible staggered C_4-C_5 rotamers was investigated.

SCHEME 4



Vibrational frequency analyses were performed for the B3LYP/6-31G* optimized geometries to confirm each stationary point as a true minimum on the potential energy surface. The calculated zero-point vibrational energy (ZPE) corrections were scaled by 0.9806.[11]

Molecular Mechanics Calculations. The Systematic Pseudo Monte Carlo search protocol available in Macromodel Version 6.5[12] was used to generate 25 000 conformers of **1**. As part of the search, each conformer was minimized in the gas phase using the AMBER force field as modified by Homans.[13] The default convergence criteria available in MacroModel were used. This generated 177 unique converged conformers within 15 kcal/mol of the global minimum.[14] The geometrical data for these conformers were analyzed using the program, ConforMole.[15] An additional 39 conformers, which were not fully converged, were also found. These were discarded and not subjected to further calculations.

Optimization of Structures from Molecular Mechanics Calculations. All 177 conformers obtained from the Monte Carlo Search were optimized at the HF/6-31G* level of theory to provide 118 conformers.[14] These HF conformers were further refined at the B3LYP/6-31G* level to give 116 unique structures. Single-point energies of all unique HF and B3LYP geometries were, in turn, determined at the B3LYP/6-31+G** level.

Nomenclature. Conformers are named by the conformation of the ring followed by the rotamer about the C_4-C_5 bond (gg, tg, or gt). Hence, $^2T_3$ gg, is a conformer in which the ring adopts a twist with C_2 above the plane and C_3 below the plane and with the C_4-C_5 orientation as the gg rotamer. For the conformers obtained by minimization of the envelope structures, a suffix is added to the name to designate whether the starting structure had intramolecular hydrogen bonding (h) or was biased against the formation of these interactions (n). In cases where more than one conformer was found with the same ring conformation and C_4-C_5 bond rotamer, but with different orientations of the other exocyclic bonds, the names are appended with numbers. These numbers correlate to the relative energy of the conformer. Lower energy conformers are appended with lower numbers. For example, the $^2T_1$ gg-1-h conformer is of lower energy than that of the $^2T_1$ gg-2-h conformer. In the following discussion, *P* refers to the Altona-Sundaralingam pseudorotational phase angle as previously defined.[6]

TABLE 1: Optimization of the 30 Envelope Geometries of 1 with the Potential for Intramolecular Hydrogen Bonding at the B3LYP/6-31G* Level of Theory

initial ring starting geometry	initial C_4-C_5 rotamer[a]	optimized conformer	relative energy[b]	P[c]	q_2[d]	exocyclic torsion angles[e]			
						C_1-O_1	C_2-O_2	C_3-O_3	C_5-O_5
	gg	$^3T_4$ gg-1-h	2.2	31	0.374	183	49	−55	51
^3E	gt	$^3T_4$ gt-h	3.0	32	0.382	183	49	−59	−54
	tg	E_4 tg-2-h	4.3	50	0.369	183	43	69	174
	gg	$^3T_4$ gg-1-h	2.2	31	0.374	183	49	−55	51
E_4	gt	$^3T_4$ gt-h	3.0	32	0.382	183	49	−59	−54
	tg	E_4 tg-2-h	4.3	50	0.369	183	43	69	174
	gg	$^3T_4$ gg-1-h	2.2	31	0.374	183	49	−55	51
^OE	gt	$^3T_4$ gt-h	3.0	32	0.382	183	49	−59	−54
	tg	E_4 tg-1-h	3.3	48	0.319	185	−57	70	175
	gg	$^2T_1$ gg-2-h	1.4	146	0.394	179	−79	−172	44
E_1	gt	$^3T_4$ gt-h	3.0	32	0.382	183	49	−59	−54
	tg	E_4 tg-2-h	4.3	50	0.369	183	43	69	174
	gg	$^2T_1$ gg-1-h	0.0	148	0.400	179	−165	−169	159
^2E	gt	^2E gt-1-h	2.6	156	0.383	177	60	−171	−54
	tg	^2E tg-h	2.6	158	0.381	177	60	−171	171
	gg	$^2T_1$ gg-1-h	0.0	148	0.400	179	−165	−169	159
E_3	gt	^2E gt-1-h	2.6	156	0.383	177	60	−171	−54
	tg	^2E tg-h	2.6	158	0.381	177	60	−171	171
	gg	$^2T_1$ gg-1-h	0.0	148	0.400	179	−165	−169	159
^4E	gt	^2E gt-2-h	6.5	153	0.311	179	62	66	52
	tg	E_4 tg-3-h	4.7	48	0.377	184	43	63	−74
	gg	$^3T_4$ gg-1-h	2.2	31	0.374	183	49	−55	51
E_O	gt	$^3T_4$ gt−h	3.0	32	0.382	183	49	−59	−54
	tg	^3E tg-h	7.5	25	0.367	183	50	−60	180
	gg	$^3T_4$ gg-2-h	2.9	45	0.395	185	−57	42	51
^1E	gt	$^3T_4$ gt-h	3.0	32	0.382	183	49	−59	−54
	tg	E_4 tg-1-h	3.3	48	0.319	185	−57	70	175
	gg	$^3T_4$ gg-1-h	2.2	31	0.374	183	49	−55	51
E_2	gt	$^3T_4$ gt-h	3.0	32	0.382	183	49	−59	−54
	tg	E_4 tg-2-h	4.3	50	0.369	183	43	69	174

[a] See Scheme 3 for definitions. [b] B3LYP/6-31+G**//B3LYP/6-31G* bottom-of-the-well energy relative to the global minimum for this table, in kcal/mol. [c] Altona−Sundaralingam pseudorotational phase angle (in degrees) as defined in ref 6. [d] Cremer−Pople puckering amplitude (in Ångstroms) as defined in ref 17. [e] Torsion angles in degrees. C_1-O_1 torsion: $C_{Me}-O_1-C_1-C_2$; C_2-O_2 torsion: $H_2-C_2-O_2-H$; C_3-O_3 torsion: $H_3-C_3-O_3-H$; C_5-O_5 torsion: $C_4-C_5-O_5-H$.

Results and Discussion

In earlier work,[1,2] we have shown that although optimizations at the B3LYP/6-31G* level provided geometries that are in close agreement with available crystal structures for these ring systems, the energetic results are not always similar to those from higher level calculations. However, the B3LYP/6-31+G**//B3LYP/6-31G* energies of these conformers do compare favorably to higher levels of theory. As has been previously reported,[2,10] this is likely due to the better treatment of intramolecular hydrogen bonding interactions when basis sets including diffuse functions are employed. Accordingly, although some subsequent tables contain the B3LYP/6-31G* optimized energies for purposes of comparison, the discussion is focused on the B3LYP/6-31+G**//B3LYP/6-31G* relative energies as we believe these are more accurate. In the discussion below, attention is given to the low-energy structures.

Conformers Found By Releasing the Planar Constraint in Envelopes. Releasing the planar constraint and full optimization (B3LYP/6-31G*) of the 30 previously reported[1] envelope geometries of **1** provided 12 unique conformers as detailed in Table 1. All conformers can be grouped into two eastern regions of the pseudorotational wheel: $^3E-E_4$ ($P = 24-50°$, N family) and $^2E-{}^2T_1$ ($P = 145-158°$, S family).

Relative Energies. Shown in Table 2 are the bottom-of-the-well energies, enthalpies, and free energies (298 K) of the 12 unique, fully optimized conformers of **1** at the B3LYP/6-31+G**//B3LYP/6-31G* level. When considering these bottom-of-the-well energies, the global minimum is the $^2T_1$ gg-1-h conformer ($P = 148°$) which is located in the southern hemisphere, whereas the northern local minimum is the $^3T_4$ gg-1-h conformer ($P = 31°$), which is 2.2 kcal/mol above that of the global minimum. All 12 conformers are within 7.5 kcal/mol of the global minimum. These energetic trends are consistent with the lowest energy envelopes identified in our previous study of **1**.[1] In that work, we demonstrated that the ^3E conformer was the local minimum in the northern hemisphere and that the ^2E conformer was the global and southern minimum. A comparison of the B3LYP/6-31G* optimized energies with the B3LYP/6-31+G** single-point energies shows that although the identities of the northern and southern minima do not change, there are substantial differences in the relative energies of some conformers. For example, the energy difference between the N and the S minima decreases from 4.4 (B3LYP/6-31G*) to 2.2 kcal/mol (B3LYP/6-31+G**//B3LYP/6-31G*).

Analysis of the ΔH_{298} and ΔG_{298} data (Table 2) indicates that the $T\Delta S_{298}$ term is greater for the northern conformers, as compared to those in the south. For example, the $T\Delta S_{298}$ term is 1.1 kcal/mol for the northern minimum, $^3T_4$ gg, but only −0.1 kcal/mol for $^2T_1$ gg, the southern minimum. Only one southern conformer (^2E gt-2-h) has a $T\Delta S_{298}$ term larger than 0.5 kcal/mol; whereas, for all northern conformers, this value is 1.0 kcal/mol or higher. These results are in contrast to other work[16] on the 3-*O*-methyl analogue of **1**, monosaccharide **3** (Scheme 1), in which entropic terms for the southern conformers were larger than for those in the north. We are unsure as to why methylation of one of the hydroxyl groups so significantly influences the thermodynamic parameters contributing to the free energy.

TABLE 2: B3LYP/6-31+G//B3LYP/6-31G* Energies[a] of the Twelve Optimized Conformers of 1 Containing Extensive Intramolecular Hydrogen Bonding[b]**

			B3LYP/6-31G* [e]	B3LYP/6-31+G**//B3LYP/6-31G* [b]				
conformer	***P*[c]**	**hemisphere[d]**	**ΔE_{BW}**	**ΔE_{BW}**	**ΔH_0**	**ΔH_{298}**	**$T\Delta S_{298}$**	**ΔG_{298}**
$^2T_1$ gg-1-h	148	S	0.0	0.0	0.0	0.0	0.0	0.0
$^2T_1$ gg-2-h	146	S	0.1	1.4	1.5	1.5	−0.1	1.6
$^3T_4$ gg-1-h	31	N	4.4	2.2	1.7	2.0	1.1	0.9
^{2}E tg-h	158	S	3.0	2.6	2.1	2.3	0.5	1.8
^{2}E gt-1-h	156	S	5.4	2.6	2.4	2.5	0.3	2.2
$^3T_4$ gg-2-h	45	N	5.0	2.9	2.4	2.7	1.0	1.7
$^3T_4$ gt-h	32	N	5.7	3.0	2.4	2.7	1.3	1.4
E_4 tg-1-h	48	N	5.7	3.3	2.2	2.6	1.2	1.4
E_4 tg-2-h	50	N	6.7	4.3	3.6	3.9	1.2	2.7
E_4 tg-3-h	48	N	6.4	4.7	4.1	4.3	1.0	3.3
^{2}E gt-2-h	153	S	7.9	6.5	5.6	6.0	1.2	4.8
^{3}E tg-h	25	N	10.3	7.5	6.4	6.9	1.8	5.1

[a] Relative energies are in kcal/mol. [b] The B3LYP/6-31+G** single-point energies were calculated using the corresponding B3LYP/6-31G* geometry as well as the scaled ZPE corrections and thermal corrections to 298 K from the B3LYP/6-31G* vibrational frequency analysis. The ΔE_{BW} energies are at the bottom-of-the-well and do not include ZPE corrections. ΔH_0 corresponds to relative enthalpies at 0 K. ΔH_{298} corresponds to relative enthalpies at 298 K. ΔG_{298} corresponds to relative free energies at 298 K. [c] Altona−Sundaralingam pseudorotational phase angle (in degrees) as defined in ref 6. [d] See Scheme 2. [e] Fully optimized B3LYP/6-31G* energies (bottom-of-the-well) for comparison.

TABLE 3: Intramolecular Hydrogen Bonding Patterns for the Optimized Conformers of 1 at the B3LYP/6-31G* Level[a]

	$OH_3\cdots O_1$		$OH_2\cdots O_5$		$OH_5\cdots O_4$		$OH_5\cdots O_2$		$OH_3\cdots O_5$	
conformer[b]	**$H_3\cdots O_1$**	**angle**	**$H_2\cdots O_5$**	**angle**	**$H_5\cdots O_4$**	**angle**	**$H_5\cdots O_2$**	**angle**	**$H_3\cdots O_5$**	**angle**
$^2T_1$ gg-1-h	2.08	136.7	1.87	157.4	*c*	*c*	*c*	*c*	*c*	*c*
$^2T_1$ gg-2-h	2.10	135.2	2.68	87.7	2.54	107.2	2.13	126.0	*c*	*c*
$^3T_4$ gg-1-h/$^3T_4$ gg-n[d]	*c*	*c*	*c*	*c*	2.29	110.1	*c*	*c*	*c*	*c*
^{2}E gt-1-h	2.11	135.5	*c*	*c*	2.30	109.0	*c*	*c*	*c*	*c*
$^3T_4$ gg-2-h	*c*	*c*	*c*	*c*	2.32	109.3	*c*	*c*	*c*	*c*
$^3T_4$ gt-h	*c*	*c*	*c*	*c*	2.38	107.5	*c*	*c*	*c*	*c*
E_4 tg-1-h	*c*	*c*	*c*	*c*	*c*	*c*	*c*	*c*	2.16	132.0
^{2}E tg-h	*c*	*c*	*c*	*c*	*c*	*c*	*c*	*c*	2.10	136.2
E_4 tg-2-h	*c*	*c*	*c*	*c*	*c*	*c*	*c*	*c*	2.17	132.1
E_4 tg-3-h	*c*	*c*	*c*	*c*	*c*	*c*	*c*	*c*	2.24	130.5
$^2T_1$ gg-n	*c*	*c*	*c*	*c*	2.57	106.0	2.19	137.4	*c*	*c*
^{2}E gt-2-h	*c*	*c*	*c*	*c*	2.27	105.3	*c*	*c*	*c*	*c*

[a] $H_X\cdots O_Y$ distances are in Ångstroms and angles are in degrees. [b] Only structures with intramolecular hydrogen bonds are shown. [c] Absence of the specified hydrogen-bonding pattern. [d] The $^3T_4$ gg-1-h and $^3T_4$ gg-n conformers are identical.

Structural Features of Conformers: C_4-C_5 Bond Rotamer. In general, the C_4-C_5 rotamer present in the starting envelope conformer does not influence the identity of the fully optimized envelope or twist structure (Table 1). Initial envelope conformers from the northern hemisphere ($^1E-E_4$) and those from the east and west (^OE and E_O) converged to northern ring conformers. The southernmost envelope conformers (E_3 and ^2E) converged to southern conformers. Only with the E_1 and ^4E geometries was the optimized ring geometry dependent upon the C_4-C_5 rotamer. In all cases, the initial orientation about the C_4-C_5 bond is preserved in the fully optimized structure indicating that the barrier to rotation about this bond is more significant than the energetic cost of pseudorotation. Consistent with this is our earlier work[1] on the idealized envelope conformers of **1**, which identified that the barrier to pseudorotation through the eastern pathway was approximately 3 kcal/mol, a value less than the expected 5 kcal/mol barrier to rotation about the C_4-C_5 bond.

The three lowest energy conformers (Table 2, ΔE_{BW}, B3LYP/6-31+G**//B3LYP/6-31G*) are the $^2T_1$ gg-1-h, $^2T_1$ gg-2-h, and $^3T_4$ gg-1-h conformers. The gg orientation about the C_4-C_5 bond is stabilized both by a gauche interaction with the ring oxygen (O_4) as well as intramolecular hydrogen bonding. In the $^2T_1$ conformer, the formation of strong[17] $OH_2\cdots O_5$ and $OH_3\cdots O_1$ transannular H-bonds (Table 3) may be the driving force toward favoring the gg rotamer. Undoubtedly, $^2T_1$ gt and $^2T_1$ tg conformers are not present among the 12 structures shown in Table 2 because the formation of a strong transannular H-bond between O_5 and OH_2 is not possible. Although the $^2T_1$ gt conformer was not found, two closely related ^{2}E gt conformers were; however, they were of higher energy (2.6 and 6.5 kcal/mol above the global minimum).

In the $^3T_4$ conformers, strong transannular hydrogen bonds are not possible; however, the formation of a weaker $OH_5\cdots O_4$ hydrogen bond does occur. Not unexpectedly, there are smaller energy differences among the low-energy northern conformers. The $^3T_4$ gt-h conformer is 3.0 kcal/mol above the global minimum and its two gg counterparts, $^3T_4$ gg-1-h and $^3T_4$ gg-2-h, are of similar energy, 2.2 and 2.9 kcal/mol, respectively. Both gg and gt rotamers of this ring conformation can form $OH_5\cdots O_4$ hydrogen bonds, and both are also stabilized by gauche interactions between O_5 and the ring oxygen (O_4). For the conformers with energies greater than 3.0 kcal/mol above the global minimum, a distribution of tg and gt rotamers is found.

Structural Features of Conformers: Other Exocyclic Torsion Angles. In all conformers, the O_1-C_{Me} bond is antiperiplanar to the C_1-C_2 bond (Table 1). Orientation of the methyl group in this fashion would be expected to allow for maximum stabilization by the exo-anomeric effect. The orientation of the secondary OH hydrogens is strongly dependent upon the ability of the conformers to form intramolecular hydrogen bonds. In the S conformers possessing transannular hydrogen bonds ($^2T_1$ gg, ^{2}E gt), the $H_x-C_x-O_x-H_x$ torsion angles ($x = 2, 3$) approach 180°. In the other conformers, a range of staggered orientations is present.

TABLE 4: Optimization of the 30 Envelope Geometries of 1 at the B3LYP/6-31G* Level of Theory Starting from Geometries Biased against the Formation of Intramolecular Hydrogen Bonds

initial ring starting geometry	initial C_4-C_5 rotamer[a]	optimized conformer	relative energy[b]	P[c]	q_2[d]	exocyclic torsion angles[e]			
						C_1-O_1	C_2-O_2	C_3-O_3	C_5-O_5
	gg	$^3T_4$ gg-n[f]	0.0	31	0.374	183	49	−55	51
^3E	gt	$^3T_4$ gt-n	3.3	37	0.390	181	50	−60	−178
	tg	^3E tg-n	5.2	25	0.367	183	50	−61	−180
	gg	^3E gg-n	2.8	17	0.369	183	53	−57	−179
E_4	gt	$^3T_4$ gt-n	3.3	37	0.390	181	50	−60	−178
	tg	^3E tg-n	5.2	25	0.367	183	50	−61	−180
	gg	^3E gg-n	2.8	17	0.369	183	53	−57	−179
^OE	gt	$^3T_4$ gt-n	3.3	37	0.390	181	50	−60	−178
	tg	^2E tg-n	5.4	153	0.319	177	66	−57	174
	gg	^3E gg-n	2.8	17	0.369	183	53	−57	−179
E_1	gt	$^3T_4$ gt-n	3.3	37	0.390	181	50	−60	−178
	tg	^2E tg-n	5.4	153	0.319	177	66	−57	174
	gg	$^2T_1$ gg-n	3.3	134	0.350	178	64	−51	34
^2E	gt	E_1 gt-n	6.2	135	0.325	177	69	−57	−171
	tg	^2E tg-n	5.4	153	0.319	177	66	−57	174
	gg	^3E gg-n	2.8	17	0.369	183	53	−57	−179
E_3	gt	E_1 gt-n	6.2	135	0.325	177	69	−57	−171
	tg	^2E tg-n	5.4	153	0.319	177	66	−57	174
	gg	^3E gg-n	2.8	17	0.369	183	53	−57	−179
^4E	gt	E_1 gt-n	6.2	135	0.325	177	69	−57	−171
	tg	^2E tg-n	5.4	153	0.319	177	66	−57	174
	gg	^3E gg-n	2.8	17	0.369	183	53	−57	−179
E_O	gt	$^3T_4$ gt-n	3.3	37	0.390	181	50	−60	−178
	tg	^3E tg-n	5.2	25	0.367	183	50	−61	−180
	gg	^3E gg-n	2.8	17	0.369	183	53	−57	−179
^1E	gt	$^3T_4$ gt-n	3.3	37	0.390	181	50	−60	−178
	tg	^3E tg-n	5.2	25	0.367	183	50	−61	−180
	gg	^3E gg-n	2.8	17	0.369	183	53	−57	−179
E_2	gt	$^3T_4$ gt-n	3.3	37	0.390	181	50	−60	−178
	tg	^3E tg-n	5.2	25	0.367	183	50	−61	−180

[a] See Scheme 3 for definitions. [b] B3LYP/6-31+G**//B3LYP/6-31G* bottom-of-the-well energy relative to the global minimum in this table in kcal/mol. [c] Altona−Sundaralingam pseudorotational phase angle (in degrees) as defined in ref 6. [d] Cremer−Pople puckering amplitude (in Ångstroms) as defined in ref 17. [e] Torsion angles in degrees. C_1-O_1 torsion: $C_{Me}-O_1-C_1-C_2$; C_2-O_2 torsion: $H_2-C_2-O_2-H$; C_3-O_3 torsion: $H_3-C_3-O_3-H$; C_5-O_5 torsion: $C_4-C_5-O_5-H$. [f] Identical to the $^3T_4$ gg-1-h conformer (see Tables 1 and 2).

Structural Features of Conformers: Puckering Amplitude. The puckering amplitude of the furanose ring in most structures remained relatively constant (Table 1). In all but the ^2E gt-2-h and E_4 tg-1-h conformers, the Cremer−Pople[18] puckering amplitudes are between 0.367 and 0.400 Å.

Influence of Intramolecular Hydrogen Bonding. All of the conformers discussed above had one or more intramolecular hydrogen bonds (Table 3). It is well understood[1,2,10,19] that the formation of such interactions greatly stabilizes carbohydrates in the gas phase, and we were therefore concerned that some conformations might be preferentially stabilized by intramolecular H-bonding. As intramolecular H-bonding would be expected to be diminished in water, we anticipated that comparison of these gas-phase results with experimental data on the conformation of **1** obtained from its NMR spectrum in aqueous solution would be problematic. Consequently, we constructed another family of 30 envelope structures with the hydroxyl hydrogens oriented "away" from the ring such that the formation of intramolecular H-bonds was impossible in the initial conformations. The orientation of these groups in the starting geometry is shown in Scheme 4.

Upon minimization, these 30 envelopes converged to seven unique conformers which are detailed in Table 4. One conformer ($^3T_4$ gg-n) was identical to $^3T_4$ gg-1-h. In all cases, the orientation of the hydrogens on the secondary hydroxyl groups remained in the initial position, and these groups were not involved in the formation of intramolecular hydrogen bonding in the final conformers (Table 4). In two conformers ($^3T_4$ gg-n and $^2T_1$ gg-n), minimization resulted in rotation of the OH_5 hydrogen to form a weak hydrogen bond with the ring oxygen or a moderately strong hydrogen bond to O_2 (Table 3). In the other five conformers, the OH_5 hydrogen remained in the initial position, anti to the C_4-C_5 bond.

The distribution of conformers is similar to what was observed in the structures possessing intramolecular H-bonds. All conformers converge to the same two families, one in the north ($^3E-{}^3T_4$, $P = 17-38°$) and another in the south ($E_1-{}^2E$, $P = 134-153°$). As shown in Table 5, the global minimum for these non-hydrogen-bonded structures is in the north, $^3T_4$ gg-n, and the southern minimum, which is 3.3 kcal/mol higher in energy, is $^2T_1$ gg-n. So, although the minimum energy ring conformers resulting from optimization of both series of envelope geometries are the same, the relative energies are reversed. Not surprisingly, the structures with extensive hydrogen bonding are of lower energy.[20] Comparison of the ΔH_{298} and ΔG_{298} data shows that the entropic contributions are similar for both the northern and southern non-hydrogen bonded conformers. This is in contrast to the conformers with a number of intramolecular H-bonds, in which the free energies of the northern conformers had significantly larger entropic contributions than for those in the southern hemisphere.

Similar to the conformers possessing extensive intramolecular H-bonding, in this series of structures, the rotamer about the C_4-C_5 bond does not change upon release of the planar constraint and overall, the gg rotamer is preferred. The two lowest energy structures are $^3T_4$ gg and ^3E gg. In contrast to the extensively H-bonded conformers, the C_4-C_5 rotamer in the starting structure has a more profound influence on the ring conformation of the converged structure. For example, the ^OE, E_1, E_3, and ^4E structures converge to either a northern or

TABLE 5: B3LYP/6-31+G//B3LYP/6-31G* Energies[a] of the Seven Optimized Conformers of 1 Starting from Structures Biased against the Formation of Intramolecular Hydrogen Bonds[b]**

			B3LYP/6-31G* [e]	B3LYP/6-31+G**//B3LYP/6-31G* [b]				
conformer	***P*[c]**	**hemisphere[d]**	ΔE_{BW}	ΔE_{BW}	ΔH_0	ΔH_{298}	$T\Delta S_{298}$	ΔG_{298}
$^3T_4$ gg-n	31	N	0.0	0.0	0.0	0.0	0.0	0.0
^3E gg-n	17	N	3.9	2.8	2.4	2.6	0.8	1.8
$^3T_4$ gt-n	37	N	4.4	3.3	2.8	3.0	0.5	2.5
$^2T_1$ gg-n	134	S	1.5	3.3	3.3	3.2	0.5	3.7
^3E tg-n	25	N	5.8	5.3	4.7	4.9	0.8	4.1
^2E tg-n	153	S	5.1	5.4	4.9	5.0	0.2	4.8
E_1 gt-n	135	S	6.3	6.2	5.6	5.8	0.7	5.1

[a] Relative energies are in kcal/mol. [b] The B3LYP/6-31+G** energies were calculated using the corresponding B3LYP/6-31G* geometry as well as the scaled ZPE corrections and thermal corrections to 298 K from the B3LYP/6-31G* vibrational frequency analysis. The ΔE_{BW} energies are at the bottom-of-the-well and do not include ZPE corrections. ΔH_0 corresponds to relative enthalpies at 0 K. ΔH_{298} corresponds to relative enthalpies at 298 K. ΔG_{298} corresponds to relative free energies at 298 K. [c] Altona−Sundaralingam pseudorotational phase angle (in degrees) as defined in ref 6. [d] See Scheme 2. [e] Fully optimized B3LYP/6-31G* energies (bottom-of-the-well) for comparison.

southern conformer depending upon the starting orientation about the C_4-C_5 bond (Table 4). The degree of ring puckering amplitude for the final conformers is similar regardless of the initial orientations of the hydroxyl hydrogens.

From these results, it is clear that the formation of intramolecular hydrogen bonds does preferentially stabilize some conformers of **1** relative to others. In particular, those conformers lying in the southern hemisphere (e.g., $^2T_1$ gg-1-h) are most stabilized as two strong transannular hydrogen bonds ($OH_3\cdots O_1$ and $OH_2\cdots O_5$) are present in these structures. When the starting conformers are biased such that the formation of these intramolecular H-bonds is disfavored, the northern conformers are of the lowest energy.

Comparison with Experiment. Previous NMR studies (in D_2O) have identified that **1** exists as an equilibrium mixture of $^OT_4$ (N) and $^2T_3$ (S) conformers in a 56:44 N:S ratio.[3] The distribution of conformers about the C_4-C_5 bond is 48% gg, 38% gt, and 14% tg. In the crystal structure of **1**,[21] the ring adopts the E_4 gg geometry. The gas-phase results presented here are consistent with these experimental data. First, these calculations predict that the C_4-C_5 gg rotamer will be preferred. Second, the agreement between the identity of the ring conformers is good. The northern minimum adopts a $^3T_4$ gg conformation, and other conformers in this family are in the $^3E-E_4$ portion of the pseudorotational itinerary, which is adjacent to the geometry of the crystal structure and close to the northern solution conformer. The southern minimum exists in a $^2T_1$ gg conformation which is present in a range of low energy conformers in the southeast region of the itinerary adjacent to the southern solution conformation.

However, the predicted Boltzmann distributions (Table 6) agree poorly with the experimental results. With the family of conformers having extensive intramolecular H-bonding, there is a roughly 30:70 N:S ratio, not the approximately 1:1 mixture present in aqueous solution.[3] Given that intramolecular hydrogen bonding will likely be diminished in water, the conformers with minimal intramolecular H-bonding might be expected to be a better representation of the solution conformational ensemble. However, with this family of conformers, no better agreement with experiment is seen as a 99:1 N:S ratio is predicted. It appears, therefore, that biasing the structures against the formation of hydrogen bonds is not sufficient to ensure a good agreement of conformer populations with experiment. This suggests that some intramolecular H-bonding stabilization may remain even under aqueous conditions. These results are consistent with our recent studies on glycerol,[22] which have demonstrated that the population of intramolecularly H-bonded conformers does not dramatically change between the gas and solution phases, as modeled using the MN-GSM solvation model.[23] Earlier, similar results were obtained with ethylene glycol by Cramer and Truhlar using the AMSOL solvation model.[24]

TABLE 6: Boltzmann Distribution[a] of 1 at 298 K. Conformers Were Those Obtained by Full Optimization of Envelope Geometries Following Release of Planar Constraint

conformer	**family[b]**	**percentage (h conformers)[c]**	**percentage (n conformers)[d]**
$^3T_4$ gg-1-h	N	13.3	NA
$^3T_4$ gt-h	N	5.7	NA
E_4 tg-1-h	N	5.7	NA
$^3T_4$ gg-2-h	N	3.5	NA
E_4 tg-2-h	N	0.6	NA
E_4 tg-3-h	N	0.2	NA
^3E tg-h	N	0.0	NA
$^2T_1$ gg-1-h	S	61.0	NA
$^2T_1$ gg-2-h	S	4.1	NA
^2E gt-2-h	S	4.1	NA
^2E gt-1-h	S	1.5	NA
^2E tg-h	S	0.1	NA
$^3T_4$ gg-n	N	NA	93.8
^3E gg-n	N	NA	4.5
$^3T_4$ gt-n	N	NA	1.4
^3E tg-n	N	NA	0.1
$^2T_1$ gg-n	S	NA	0.2
^2E tg-n	S	NA	0.0
E_1 tg-n	S	NA	0.0

[a] On the basis of $\Delta G°$ from B3LYP/6-31+G**//B3LYP/6-31G* single-point energies and thermal and entropic corrections from B3LYP/6-31G* vibrational frequency calculations. [b] North family: $^3E-E_4$ ($P = 9°-62°$); South family: $^2E-E_1$ ($P = 117°-170°$), where P is the Altona−Sundaralingam pseudorotational phase angle (in degrees) as defined in Reference 6. [c] Conformers listed in Table 2. [d] Conformers listed in Table 5.

Conformers Found Via Monte Carlo Approach. To obtain a protocol more amenable to the study of oligofuranose-containing molecules than the one described above, we explored a stochastic method for determining the preferred geometries of **1**. At the outset, the major question we wanted to answer was whether this Monte Carlo-based approach (MC) would generate a family of conformers similar to those obtained from the constrained envelope method (CE) described above. We expected that many additional conformers would also be generated by the MC method. Moreover, we envisioned that if we had a large diversity of conformers for **1** at our disposal, then higher-level calculations, especially with consideration of solvation via either continuum models or via the inclusion of specific water molecules, would provide better agreement with experiment. Although these conformers could be generated manually, as has been reported for other furanose rings,[25] we

TABLE 7: Summary of the AMBER, HF/6-31G*, and B3LYP/6-31G* Optimizations Depicting the Conformers within the Lowest 5.0 kcal/mol[a]

AMBER optimized				HF optimized				B3LYP optimized			
conformer	relative energy[b]	*P*[c]	q_2[d]	conformer	relative energy[e]	*P*[c]	q_2[d]	conformer	relative energy[f]	*P*[c]	q_2[d]
$^2T_1$ gg	0.0	138	0.422	$^2T_1$ gg	0.0	143	0.381	$^2T_1$ gg	0.0	153	0.395
$^2T_1$ gg	0.7	138	0.418	$^2T_1$ gg	0.7	142	0.385	$^2T_1$ gg	0.6	148	0.400
$^2T_1$ gg	1.2	138	0.427	$^2T_1$ gg	1.8	137	0.382	$^2T_1$ gg	2.0	146	0.394
$^2T_1$ gg	1.2	138	0.427	^3E gg	2.7	25	0.377	$^3T_4$ gg	2.9	31	0.374
E_1 gg	1.5	135	0.394	$^2T_1$ gt	2.8	147	0.370	^2E gt	3.0	157	0.383
E_1 gg	1.5	135	0.394	$^2T_1$ gt	3.1	147	0.371	^2E gt	3.2	156	0.383
$^3T_4$ gt	1.6	37	0.427	E_1 gg	3.3	135	0.384	$^2T_1$ gg	3.3	144	0.397
E_4 gt	1.9	57	0.424	$^3T_4$ gg	3.4	44	0.396	$^2T_1$ gg	3.4	150	0.398
E_4 gg	1.9	58	0.432	$^3T_4$ gt	3.5	28	0.383	$^3T_4$ gg	3.6	45	0.395
E_4 gg	1.9	57	0.432	$^2T_1$ gg	3.5	143	0.384	$^3T_4$ gt	3.6	32	0.382
$^2T_1$ gt	2.1	137	0.409	E_4 tg	3.8	46	0.376	E_4 tg	3.9	48	0.388
$^2T_1$ gt	2.1	137	0.409	$^2T_1$ gg	4.0	140	0.339	$^2T_1$ gg	4.0	147	0.347
$^1T_2$ gg	2.1	320	0.427	E_4 gg	4.2	58	0.364	E_4 gg	4.4	54	0.368
$^OT_4$ gt	2.2	71	0.416	^2E tg	4.3	153	0.366	^2E tg	4.4	160	0.381
$^OT_4$ gg	2.4	75	0.414	$^3T_4$ gt	4.7	45	0.387	^2E tg	4.8	160	0.379
E_4 tg	2.4	55	0.435	$^2T_1$ tg	4.7	153	0.365	E_4 gt	4.8	47	0.390
E_1 gg	2.5	135	0.398	E_4 tg	4.7	47	0.358	E_4 tg	4.9	50	0.369
$^2T_1$ gt	3.0	139	0.414	$^2T_1$ gg	4.9	138	0.382	$^2T_1$ gg	4.9	149	0.342
^4T_O gg	3.0	244	0.314	$^2T_1$ tg	4.9	151	0.368	^2E tg	5.0	158	0.382
^4E gg	3.2	241	0.334								
$^3T_2$ gt	3.4	6	0.420								
$^3T_2$ gt	3.4	6	0.420								
$^3T_2$ gt	3.4	7	0.420								
E_4 tg	3.5	60	0.445								
E_4 tg	3.5	60	0.445								
E_1 gg	3.5	133	0.404								
$^OT_4$ gt	3.6	64	0.425								
E_4 gt	3.6	59	0.430								
E_4 gg	3.6	49	0.430								
E_4 gg	3.6	49	0.429								
$^OT_4$ gt	3.7	71	0.425								
E_4 gg	3.7	61	0.421								
E_1 gt	3.7	129	0.455								
$^1T_2$ gg	3.8	320	0.411								
^4E gg	3.8	241	0.308								
^4E gg	3.9	238	0.348								
$^3T_2$ gg	4.0	5	0.421								
$^1T_2$ gg	4.0	321	0.421								
$^2T_1$ gg	4.1	138	0.424								
$^2T_1$ gg	4.1	138	0.425								
$^2T_1$ gg	4.1	138	0.424								
$^3T_2$ gg	4.4	9	0.409								
$^OT_4$ tg	4.6	69	0.427								
$^3T_4$ gg	4.6	28	0.420								
$^3T_4$ gt	4.6	28	0.410								
$^3T_4$ gt	4.6	28	0.410								
E_1 gg	4.7	131	0.406								
$^OT_4$ gg	4.7	68	0.418								
^3E tg	4.8	22	0.442								
E_4 gg	4.9	56	0.424								
^4E gg	4.9	238	0.321								
$^2T_1$ gt	5.0	139	0.414								

[a] See text for details on the protocol. [b] AMBER//AMBER energy in kcal/mol. [c] Altona−Sundaralingam pseudorotational phase angle (in degrees) as defined in ref 6. [d] Cremer−Pople puckering amplitude (in Ångstroms) as defined in ref 17. [e] B3LYP/6-31+G**//HF/6-31G* bottom-of-the-well energy in kcal/mol relative to the global minimum at this level of theory. [f] B3LYP/6-31+G**//B3LYP/6-31G* bottom-of-the-well energy in kcal/mol relative to the global minimum at this level of theory.

envisioned that an approach unbiased by intuition would be preferable. The use of molecular mechanics algorithms[26] to generate and minimize a large number of conformers, a portion of which are then studied at higher levels of theory has been previously reported for pyranose ring systems, including di- and trisaccharides.[27] However, these methods have not been widely used with furanose-containing molecules.[28]

A Systematic Pseudo Monte Carlo search, using the protocol available in Macromodel 6.5,[12] coupled with the AMBER force field as parametrized by Homans[13] was carried out for **1**. A total of 25 000 structures were generated in 5 CPU hours, which led to a total of 177 fully converged conformers. All conformers were within 15 kcal/mol of the molecular mechanics global minimum.

Each of these conformers was optimized at the HF/6-31G* level of theory and, in turn, these unique structures were again further refined at the B3LYP/6-31G* level. B3LYP/6-31+G** single-point energies of the HF/6-31G* and B3LYP/6-31G* conformers were also calculated and, because we believe these are the most accurate energies, are used in the following discussion. Listed in Table 7 are the conformers obtained from the AMBER optimizations within 5 kcal/mol of the global

TABLE 8: Conformer Distribution from Monte Carlo Approach

		AMBER	HF	B3LYP
all conformers	total number	177	118	116
	number in N/S family regions[a]	95	89	93
	number of eastern conformers outside N/S family regions[a]	21	5	2
	number of western conformers[b]	61	24	21
conformers between 0 and 5 kcal/mol of global minimum	total number	52	19	19
	number of eastern conformers outside N/S family regions[a]	6	0	0
	number of western conformers[b]	13	0	0

[a] North family region: $^3E-E_4$ ($P = 9°-62°$); South family region: $^2E-E_1$ ($P = 117°-170°$), where P is the Altona−Sundaralingam pseudorotational phase angle (in degrees) as defined in ref 6. [b] E_3-E_2 regions of the pseudorotational itinerary; $P = 190°-350°$.

minimum. Also presented in Table 7 are the HF/6-31G* and B3LYP/6-31G* optimized conformers with B3LYP/6-31+G** single-point energies between 0 and 5 kcal/mol. The discussion below focuses on the conformers within this window as they would be expected to be the major contributors to the Boltzmann distribution. Shown in Table 8 is a summary of conformer numbers, location in the pseudorotational itinerary, and energy at the three levels of theory.

Conformers Found after AMBER Optimization. The 177 AMBER conformers were distributed over the entire pseudorotational wheel. The majority of structures were found in the eastern portion of the itinerary and were localized between ^3E and E_4 (the "N family region") and between E_1 and ^2E (the "S family region"). Unexpected, however, was that 61 of the 177 total conformers (34%, see Table 8) were present in the western hemisphere (E_3-E_2, $P = 190-350°$). The global minimum is a $^2T_1$ gg conformer, in the south. When considering the lowest energy structures, 52 of the 177 conformers (29%) were within 5.0 kcal/mol of the global minimum. Although the majority of the structures in this window were eastern conformers, 13 from the western hemisphere were also identified at the AMBER level. The preferred orientations about the C_4-C_5 bond are gg and gt. The gg:gt:tg ratio of the 52 lowest-energy structures is 60%:30%:10%.

As is apparent from the data in Table 7, a number of the AMBER conformers were structurally very similar and nearly degenerate in energy. In most cases, the only differences were slight deviations in the pseudorotational phase angle (<1°), ring puckering amplitude, and orientation of the exocyclic groups (Table 7 and Table S1 in the Supporting Information).

No molecular mechanics force field has been explicitly parametrized for furanose rings, with the exception of those found in nucleic acids.[29] Accordingly, we were concerned that concentrating only on the low-energy AMBER conformers might result in the loss of structures that would be important at higher levels of theory. We therefore decided to carry out further calculations on all 177 AMBER conformers. The approach we took was to first perform HF/6-31G* optimizations, which we hoped would significantly reduce the number of conformers that needed to be studied using the more expensive B3LYP/6-31G* method. As discussed below, this approach did indeed work, and it is the approach that we recommend. Nevertheless, we were initially concerned that a loss of conformer diversity might occur upon doing the HF/6-31G* optimizations. Consequently, we also carried out B3LYP/6-31G* minimizations on all 177 AMBER conformers, and obtained the same B3LYP/6-31G* final conformers as when the HF/6-31G* calculations were used first to reduce the number of conformers. The relative computational cost between both approaches was very similar (within 10%). Table S4 in the Supporting Information shows the fate of all 177 AMBER conformers upon optimization at increasingly higher levels of theory.

Conformers Found after HF/6-31G* Optimization. Subjecting the 177 conformers identified by the AMBER search to HF/6-31G* optimization resulted in a number of them converging to similar structures and eventually afforded 118 unique structures. The B3LYP/6-31+G** single-point energies of these conformers were determined, and those within 5.0 kcal/mol of the global minimum are shown in Table 7. A total of 19 conformers were found within this window. Upon HF/6-31G* minimization, the majority of the AMBER conformers in the western hemisphere converged to structures in the eastern portion of the pseudorotational itinerary. Although many of the low-energy AMBER conformers optimized to structures that had B3LYP/6-31+G**//HF/6-31G* energies near the global minimum, others were more than 3 kcal/mol higher in energy. Some structures with high AMBER energies (>5.0 kcal/mol) minimized to conformers that were within the 5.0 kcal/mol window on the HF potential energy surface (Table S4 in the Supporting Information).

Of the 118 unique, HF-optimized conformers, 24 were present in the west, but none of these had a B3LYP/6-31+G**//HF/6-31G* energy within the 0−5 kcal/mol window (Table 8). The low-energy structures were distributed between the $^3E-E_4$ and E_1-^2E regions, that is, the "N family" and "S family" regions. In the conformers within 5.0 kcal/mol of the global minimum, the preferred rotation about the C_4-C_5 bond is gg. However, in contrast to the AMBER results, there is an approximately equal number of conformers with the gt and tg orientation (gg:gt:tg = 53%:21%:26%).

In cases where the ring form did not change significantly upon HF minimization, the orientation about the exocyclic groups remained close to their positions in the AMBER conformers. However, where there were gross changes in ring conformation upon optimization, e.g., a western conformer moving to a southern conformation, the orientation of the exocyclic groups generally changed significantly. However, the orientation about the C_4-C_5 bond was usually conserved.

The southern minimum after HF/6-31G* optimization was $^2T_1$ gg, as was the case for the AMBER optimized conformers. However, the exact values of P and ring puckering amplitude of these two $^2T_1$ gg conformers were slightly different (Table 7) as was the orientation of the exocyclic groups (see Tables S1 and S2 in the Supporting Information). The northern minimum was ^3E gg, not $^3T_4$ gt, as found in the AMBER family of conformers. Although these two levels of theory do not provide exact agreement for the structure of the northern minimum, both identify a conformer in the $^3E-E_4$ region of the itinerary. The HF energy difference between the N and S conformers is 2.7 kcal/mol, which is larger than the energy gap determined from the AMBER optimizations (1.6 kcal/mol), but consistent with the constrained envelope approach (2.2 kcal/mol).

Conformers Found after B3LYP/6-31G* Optimization. When the HF/6-31G* optimized conformers were further refined

TABLE 9: Correlation of B3LYP/6-31G*-optimized, HF/6-31G*-optimized AMBER Conformers with the Results of Fully Optimizing 60 Envelope Conformations of 1 at the B3LYP/6-31G* Level of Theory

MC conformer	AMBER conformer no.	energy[a]	*P*[b]	q_2[c]	matching CE conformer[d]	MC conformer	AMBER conformer no.	energy[a]	*P*[b]	q_2[c]	matching CE conformer[d]
$^2T_1$ gg	2	0.0	153	0.395		$^2T_1$ gg	30	6.5	139	0.346	
$^2T_1$ gg	1	0.6	148	0.400	$^2T_1$ gg-1-h	^2E gt	182	6.6	158	0.384	
$^2T_1$ gg	3	2.0	146	0.394	$^2T_1$ gg-2-h	$^3T_2$ gt	25	6.6	6	0.376	
$^3T_4$ gg	38	2.9	31	0.374	$^3T_4$ gg-1-h/$^3T_4$ gg-n	$^2T_1$ gg	158	6.6	145	0.396	
^2E gt	11	3.0	157	0.383		$^3T_4$ gg	69	6.7	35	0.391	
^2E gt	20	3.2	156	0.383	^2E gt-1-h	$^3T_4$ tg	81	7.0	38	0.384	
$^2T_1$ gg	144	3.3	144	0.397		^3E gg	55	7.1	20	0.390	
$^2T_1$ gg	48	3.4	150	0.398		^3E tg	162	7.3	20	0.378	
$^3T_4$ gg	9	3.6	45	0.395	$^3T_4$ gg-2-h	$^3T_4$ tg	192	7.4	39	0.390	
$^3T_4$ gt	7	3.6	32	0.382	$^3T_4$ gt-h	$^2T_1$ gg	105	7.5	148	0.345	
E_4 tg	16	3.9	48	0.388	E_4 tg-1-h	$^3T_4$ gg	146	7.5	44	0.362	
$^2T_1$ gg	5	4.0	147	0.347		E_4 gt	32	7.6	50	0.396	
E_4 gg	15	4.4	54	0.368		$^3T_4$ gg	93	7.6	40	0.362	
^2E tg	89	4.4	160	0.381		E_4 gg	59	7.6	50	0.404	
^2E tg	134	4.8	160	0.379		E_3 gg	43	7.7	207	0.248	
E_4 gt	8	4.8	47	0.390		E_4 gg	62	7.7	51	0.390	
E_4 tg	121	4.9	50	0.369	E_4 tg-2-h	E_4 gt	36	7.8	58	0.396	
$^2T_1$ gg	17	4.9	149	0.342		^3E tg	61	7.8	16	0.399	
^2E tg	115	5.0	158	0.382		^2E tg	185	7.8	162	0.382	
$^2T_1$ gg	73	5.2	147	0.394		$^3T_4$ gt	65	7.9	37	0.386	
^2E tg	161	5.2	157	0.384		^3E tg	82	8.1	24	0.366	^3E tg-h/^3E tg-n
E_4 gt	14	5.3	54	0.388		$^3T_4$ gg	100	8.2	45	0.378	
$^2T_1$ tg	74	5.6	143	0.422		^3E tg	164	8.3	14	0.365	
^3E gg	13	5.7	17	0.369	^3E gg-n	E_1 tg	91	8.3	130	0.384	
$^2T_1$ gt	39	5.8	138	0.433		E_2 gg	174	8.4	335	0.359	
$^3T_2$ gg	45	5.9	8	0.376		E_1 gg	98	8.4	134	0.369	
^2E gt	64	6.0	155	0.386		^2E tg	203	8.4	160	0.384	
^2E gt	84	6.1	157	0.386		^3E gg	80	8.5	12	0.371	
^2E gt	99	6.1	159	0.383		$^2T_1$ tg	178	8.5	145	0.419	
$^3T_4$ gg	101	6.1	33	0.401		E_2 gg	46	8.6	344	0.375	
$^3T_4$ gt	31	6.1	38	0.390	$^3T_4$ gt-n	E_4 gg	128	8.8	50	0.373	
$^OT_4$ tg	54	6.1	64	0.384		^2E tg	90	8.8	161	0.323	
E_4 tg	28	6.2	52	0.407		E_4 tg	209	8.9	62	0.365	
^2E gt	97	6.2	157	0.387		$^4T_3$ gg	67	8.9	225	0.310	
$^2T_1$ gt	87	6.3	143	0.419		$^3T_2$ gg	107	9.0	5	0.372	
$^2T_1$ gg	104	6.3	149	0.341		$^2T_1$ gt	143	9.2	141	0.430	
$^3T_4$ gg	35	6.4	44	0.410							

[a] B3LYP/6-31+G** energy of the AMBER conformers that were optimized first at the HF/6-31G* level of theory and then at the B3LYP/6-31G* level of theory (B3LYP/6-31+G**//B3LYP/6-31G*/HF/6-31G*/AMBER) in kcal/mol. [b] Altona−Sundaralingam pseudorotational phase angle (in degrees) as defined in ref 6. [c] Cremer−Pople puckering amplitude (in Ångstroms) as defined in ref 17. [d] Related B3LYP/6-31G*-optimized constrained envelope conformer, as determined by an exact match of absolute energies and geometry. One conformer found by the CE method (^2E tg-h) was not identified by the MC method.

at the B3LYP/6-31G* level, 116 unique structures resulted. As was done with the HF/6-31G* conformers, B3LYP/6-31+G** single-point energies were calculated for all 116 structures; 19 conformers are within 5.0 kcal/mol of the global minimum (Table 7 and Figure 1). Similar to the HF-optimized geometries, the overwhelming majority of these conformers are located in the $^3E-E_4$ and E_1-^2E regions of the pseudorotational itinerary. As indicated in Table 8, only 23 of the 116 conformers are found outside these regions, and none of the 23 are included in the 19 lowest-energy conformers shown in Table 7. Further refinement of the HF-optimized geometries at the B3LYP/6-31G* level does not result in large structural changes, with only isolated conformers moving from the northern to the southern family. Similarly, the orientation of the exocyclic bonds was also largely unchanged, and the ratio of C_4-C_5 rotamers was 48%:26%:26% gg:gt:tg. The major changes observed are more reasonable bond lengths due to the inclusion of electron correlation, which is missing at the HF level of theory.[7]

The B3LYP-optimized potential energy surface is very similar to that obtained from the AMBER and HF-optimized family of conformers. The global minimum is a southern conformer, $^2T_1$ gg, and the northern minimum is $^3T_4$ gg. The energy difference between the two is 2.9 kcal/mol. The exact values of *P* and the puckering amplitude differ slightly from the structures on the HF/6-31G* potential energy surface. Table S4 in the Supporting Information contains the orientation of all the exocyclic groups in these conformers.

Comparison of Constrained Envelope (CE) and Monte Carlo (MC) Approaches. It was expected that the B3LYP/6-31+G**//B3LYP/6-31G* conformer distribution and energetic trends from the MC approach would be similar to those obtained from the CE method discussed previously. This is indeed the case. In general, all of the major conformers that were found to be major contributors to the Boltzmann population by optimization of the envelope geometries of **1** were also identified by the MC method. *However, many additional conformers were found by the stochastic method.*

We first compared the MC conformer distribution with that obtained by the optimization of the 30 envelope conformers of **1** that possessed intramolecular hydrogen bonds. Minimization of all 30 envelopes had identified 12 conformers (Table 2), eight of which contributed more than 1% to the Boltzmann population at room temperature (Table 6). As illustrated in Table 9, the Monte Carlo search identified nine of these 12 conformers, including seven of the eight conformers that are appreciable contributors to the CE Boltzmann population. The only significant contributor (4.1%) that was found by the CE approach, but not by the MC method, was ^2E tg-h.

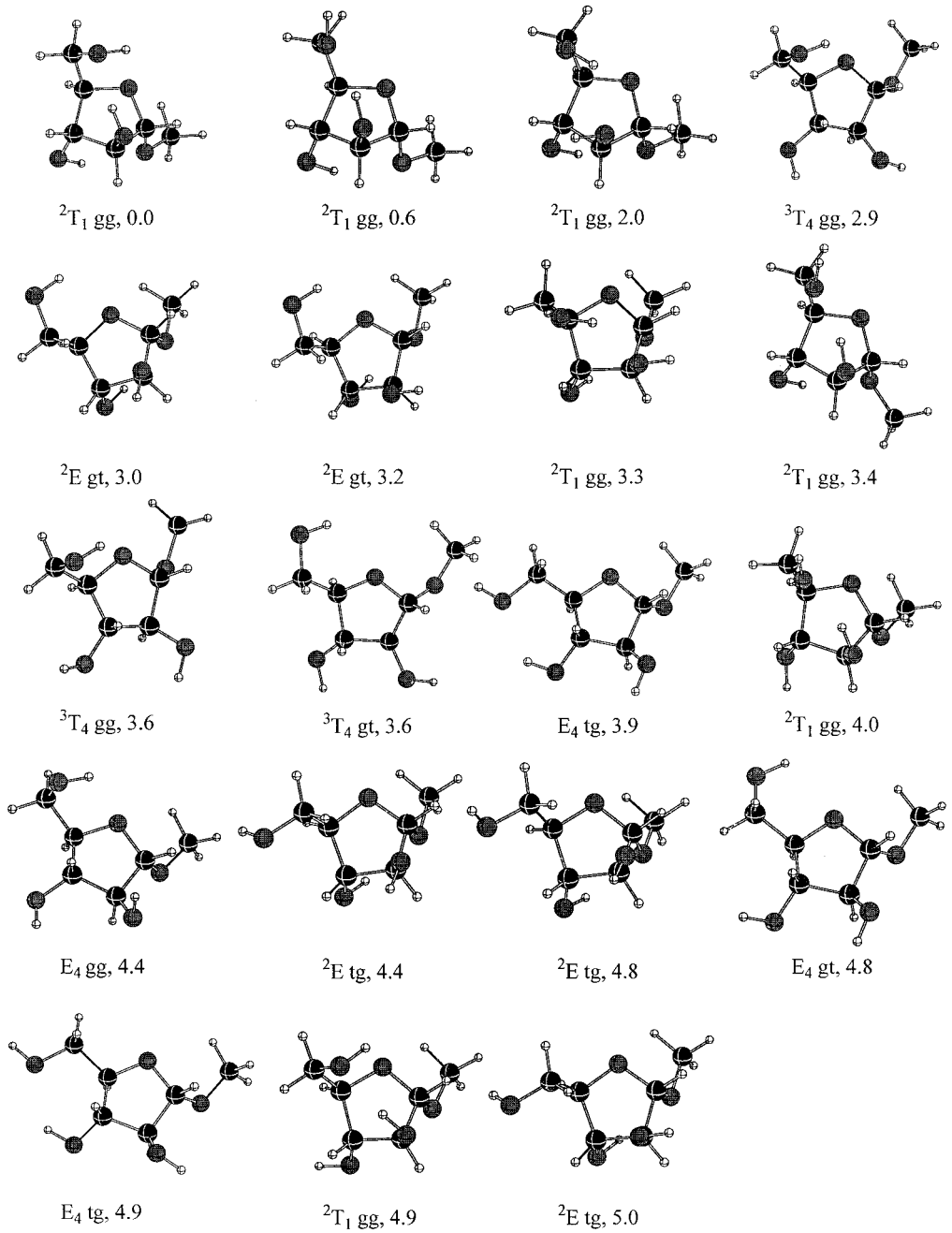


Figure 1. B3LY/6-31G* conformers of **1** obtained by the MC method. Only those conformers within 5.0 kcal/mol of the global minimum are shown, see Table 7. Exocyclic torsion angles can be found in Table S3 in the Supporting Information.

Minimization of the 30 envelope conformers of **1** with the hydroxyl hydrogens oriented such that intramolecular H-bonding was unfavorable provided seven conformers (Table 5). The MC method identified four of these, including all of those that contribute more than 1% to the CE Boltzmann population. The three conformers which were not found each contribute less than 0.2%.

The most notable difference between the MC and CE potential energy surfaces is that the global minimum from each method is not the same. The lowest energy conformer obtained from the MC method, a $^2T_1$ gg structure, was not found by the CE approach. The energy difference (ΔE_{BW}, B3LYP/6-31+G**//B3LYP/6-31G*) between the MC and CE global minima is 0.6 kcal/mol. However, the MC and CE minima are very similar, differing only slightly in terms of *P*, puckering amplitude and exocyclic torsion angles. The most significant difference between the two conformers is that in the MC minimum, the O_5-H group has rotated to form an intramolecular H-bond with the ring oxygen. The global minimum obtained by the CE method is the second-lowest energy conformer on the MC potential energy surface (Table 9). Subjecting the low-energy conformers found by the MC approach to vibrational frequency calculations and determination of free energies would likely produce a Boltzmann distribution even more biased toward the southern conformers than that obtained from the "h-conformers" listed in Table 6. So, whereas we feel that the stochastic approach is the best way to generate conformers, treatment of solvation on the structures generated will still be necessary in order to achieve results in agreement with experiment.

A final issue concerns the number of HF/6-31G* and B3LYP/6-31G* optimizations of the AMBER conformers that is necessary to avoid "missing" structures that are contributors to

the conformational ensemble. This information can be gleaned from Table 9, in which the final conformers obtained from the MC method are correlated with the AMBER conformer number from which they were obtained. Had only the AMBER conformers with a relative energy within 5.0 kcal/mol of the global minimum been carried on (AMBER conformer numbers 1−52), 5 of the 19 final conformers would not have been found. However, these are all relatively high-energy species that would be expected to be only minor contributors to the Boltzmann distribution. Consequently, although in these studies we carried through all 177 of the fully converged AMBER conformers for further investigation, it is clear that continuing with only the 50 lowest-energy conformers would have been sufficient.

Conclusions

In this paper, we report a computational study of the conformational preferences of unconstrained methyl α-D-arabinofuranoside **1**, by two approaches. The first was by release of the planar constraint in a series of envelope conformers (the CE method), the second approach (the MC method) involved the generation of a family of conformers with a Monte Carlo search followed by optimization at higher levels of theory. Our conclusions are as follows:

1. Starting from envelope conformers possessing intramolecular hydrogen-bonding, the CE method identified a family of 12 conformers of **1**. Calculation of free energies of each conformer provided a Boltzmann distribution that contained, as the major conformers, structures in reasonable agreement with previous experimental results. However, the N:S ratio was biased toward the southern conformers, which is not in accordance with experiment.

2. In an attempt to obtain better agreement with experiment, we optimized a set of envelope conformers of **1** biased against the formation of intramolecular hydrogen bonding. A total of seven conformers resulted. Following calculation of their free energies, the Boltzmann distribution was strongly biased toward northern conformers. These results agreed no better with experiment than did those obtained from the intramolecular hydrogen-bonded conformers. This suggests that simply preventing intramolecular hydrogen bonding in the conformers of **1** is not sufficient to accurately approximate the solution conformation. It is likely, therefore, that the aqueous solution population is a mixture of intramolecular H-bonded and non-H-bonded conformers.

3. The MC method provided a family of conformers in good agreement with those obtained from the CE method and from experiment. As expected, the MC method identified many other conformers as well, and most notably, the global minimum found by the MC approach was not identified by the CE method. Although the AMBER relative energies for some conformers were similar to those obtained at higher levels of theory, others were significantly different. For these optimizations, there was a strong preference to conserve the C_4-C_5 rotamer orientation from one level of theory to another.

4. In our opinion, the MC approach is a better method for identifying the low energy conformers of **1** and other furanose containing structures in that it does not rely on intuition to generate any conformers and the method is very effective at sampling the entire conformational space. With a diverse family of conformers in hand, higher level calculations can be carried out, including determination of the relative free energies and consideration of solvation effects. The use of intermediate level (HF/6-31G*) optimizations to refine the litany of necessary B3LYP/6-31G* optimizations yielded the same final results as direct B3LYP/6-31G* minimization of the AMBER conformers. Indeed, the B3LYP and HF surfaces are quite similar for **1**. Studies directed toward addressing the effect of solvation on the potential energy surface of **1** are currently underway.

Acknowledgment. This research has been supported by the Ohio State University and the National Science Foundation (T.L.L.: CHE-9875163 and C.M.H.: CHE-9733457). We acknowledge support from the Ohio Supercomputer Center where some of these calculations were performed.

Supporting Information Available: Cartesian coordinates for B3LYP/6-31G* optimized structures listed in Table 7, and tables of geometrical parameters for conformers optimized at each level of theory. This material is available free of charge via the Internet at http://pubs.acs.org.

References and Notes

(1) Gordon, M. T.; Lowary, T. L.; Hadad, C. M. *J. Am. Chem. Soc.* **1999**, *121*, 9682.

(2) Gordon, M. T.; Lowary, T. L.; Hadad, C. M. *J. Org. Chem.* **2000**, *65*, 4954.

(3) D′Souza, F. W.; Ayers. J. D.; McCarren, P. R.; Lowary, T. L. *J. Am. Chem. Soc.* **2000**, *122*, 1251.

(4) Brennan, P. J.; Nikaido, H. *Annu. Rev. Biochem.* **1995**, *64*, 29.

(5) (a) Garrett, E. C.; Serianni, A. S. *In Computer Modeling of Carbohydrate Molecules*; French, A. D., Brady, J. W., Eds.; ACS Symposium Series 430. American Chemical Society: Washington, DC, 1990; pp 91−119. (b) Serianni, A. S.; Chipman, D. M. *J. Am. Chem. Soc.* **1987**, *109*, 5297. (c) Church, T. J.; Carmichael, I.; Serianni, A. S. *J. Am. Chem. Soc.* **1997**, *119*, 8946.

(6) Altona, C.; Sundaralingam, M. *J. Am. Chem. Soc.* **1972**, *94*, 8205.

(7) Hehre, W. J.; Radom, L.; Schleyer, P. v. R.; Pople, J. A. *Ab Initio Molecular Orbital Theory*; John Wiley & Sons: New York, 1986.

(8) (a) Labanowski, J. W.; Andzelm, J. *Density Functional Methods in Chemistry*; Springer: New York, 1991. (b) Parr, R. G.; Yang, W. *Density Functional Theory in Atoms and Molecules*; Oxford University Press: New York, 1989. (c) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648. (d) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785.

(9) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Zakrzewski, V. G.; Montgomery, J. A., Jr.; Stratmann, R. E.; Burant, J. C.; Dapprich, S.; Millam, J. M.; Daniels, A. D.; Kudin, K. N.; Strain, M. C.; Farkas, O.; Tomasi, J.; Barone, V.; Cossi, M.; Cammi, R.; Mennucci, B.; Pomelli, C.; Adamo, C.; Clifford, S.; Ochterski, J.; Petersson, G. A.; Ayala, P. Y.; Cui, Q.; Morokuma, K.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Cioslowski, J.; Ortiz, J. V.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Gonzalez, C.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Andres, J. L.; Gonzalez, C.; Head-Gordon, M.; Replogle, E. S.; Pople, J. A. Gaussian 98, Revision A.7, Gaussian, Inc.; Pittsburgh, PA, 1998.

(10) (a) Lii, J. H.; Ma, B.; Allinger, N. L. *J. Comput. Chem.* **1999**, *15*, 1593. (b) Csonka, G.; Éliás, K.; Csizmadia, I. G. *J. Comput. Chem.* **1996**, *18*, 330. (c) Csonka, G. I.; Éliás, K.; Kolossváry, I.; Sosa, C. P.; Csizmadia, I. G. *J. Phys. Chem. A* **1998**, *102*, 1219. (d) Csonka, G. I.; Kolossváry, I.; Császár, P.; Éliás, K.; Csizmadia, I. G. *J. Phys. Chem. A* **1998**, *102*, 1219. (e) Del Bene, J. E.; Person, W. B.; Szczepaniak, K. *J. Phys. Chem.* **1995**, *99*, 10 705.

(11) Scott, A. P.; Radom, L. *J. Phys. Chem*. **1996**, *100*, 16 502.

(12) (a) MacroModel V6.5: Mohamadi, F.; Richards, N. G. J.; Guida, W. C.; Liskamp, R.; Lipton, M.; Caulfield, C.; Chang, G.; Hendrickson, T.; Still, W. C. *J. Comput. Chem.* **1990**, *11*, 440. (b) Goodman, J.; Still, W. C. *J. Comput. Chem.* **1991**, *12*, 1110.

(13) Homans, S. W. *Biochemistry* **1990**, *29*, 9112.

(14) **AMBER:** MacroModel determined the uniqueness of conformations by first superimposing two conformations if they were within 1 kcal/mol of each other and then comparing the position of identical heavy atoms in each conformation (e.g., C_1 in both conformations). If any two pairs of equivalent atoms were separated by more than 0.25 Å, the structures were determined unique. **Gaussian:** The uniqueness of conformations found by optimization with Gaussian was determined by comparing the puckering parameters (P and q_2), the orientation about the C_1-O_1, C_2-O_2, C_3-O_3, C_5-O_5, C_4-O_5 bonds, as well as the conformation's energy. Two conformations were the same if their parameters were within the following criteria: P and exocyclic angles $= \pm 1°$, $q_2 = \pm 0.005$ Å, and $E = \pm 0.1$ kcal/mol. A similar comparison of the AMBER conformations decreased the number of unique conformations to 159.

(15) ConforMole, P. R. McCarren, The Ohio State University. This program is available upon request.

(16) Houseknecht, J. B.; McCarren, P. R.; Lowary, T. L.; Hadad, C. M. *J. Am. Chem. Soc.*, submitted.

(17) Jeffrey, G. A. *An Introduction to Hydrogen Bonding*; Oxford Press: New York, 1997; pp 11−32.

(18) Cremer, D.; Pople, J. A. *J. Am. Chem. Soc.* **1975**, *97*, 1354.

(19) (a) Jebber, K. A.; Zhang, K.; Cassady, C. J.; Chung-Phillips, A. *J. Am. Chem. Soc.* **1996**, *118*, 10 515. (b) Brown, J. W.; Wladkowski, B. D. *J. Am. Chem. Soc.* **1996**, *118*, 1190. (c) Barrows, S. E.; Dulles, F. J.; Cramer, C. J.; French, A. D.; Truhlar, D. G. *Carbohydr. Res.* **1995**, *276*, 219. (d) Cramer, C. J.; Truhlar, D. G. *J. Am. Chem. Soc.* **1993**, *115*, 5745. (e) Polavarapu, P. L.; Ewig, C. S. *J. Comput. Chem.* **1992**, *13*, 1255. (f) Barrows, S. E.; Storer, J. W.; Cramer, C. J.; French, A. D.; Truhlar, D. G. *J. Comput. Chem.* **1998**, *19*, 1111. (g) Csonka, G. I.; Csizmadia, I. G. *Chem. Phys. Lett.* **1995**, *243*, 419. (g) Ma, B.; Schaefer, H. F., III.; Allinger, N. L. *J. Am. Chem. Soc.* **1998**, *120*, 3411.

(20) For the non-hydrogen bonded envelopes as initial structures, the global minimum conformer is $^3T_4$ gg−n, which is identical to $^3T_4$ gg-1-h. On the potential energy surface obtained from the hydrogen−bonded envelopes, this conformer has an energy that is 2.2 kcal/mol above the global minimum.

(21) Evdokimov, A. G.; Kalb (Gilboa), A. J.; Koetzle, T. F.; Klooster, W. T.; Martin, J. M. L. *J. Phys. Chem. A* **1999**, *103*, 744.

(22) Callam, C. S.; Singer, S. J.; Lowary, T. L.; Hadad, C. M., manuscript in preparation.

(23) (a) Xidos, J.D.; Li, J.; Hawkins, G. D.; Liotard, D. A.; Cramer, C. J.; Truhlar, D. G.; Frisch, M. J. MN-GSM, version 99.2, University of Minnesota, Minneapolis, MN 55455. (b) Li, J.; Zhu, T.; Cramer, C. J.; Truhlar, D. G. *J. Phys. Chem. A* **1998**, *102*, 1820. (c) Li, J.; Hawkins, G. D.; Cramer, C. J.; Truhlar, D. G. *Chem. Phys. Lett.* **1998**, *288*, 293. (d) Zhu, T.; Li, J.; Hawkins, G. D.; Cramer, C. J.; Truhlar, D. G. *J. Chem. Phys.* **1998**, *109*, 9117. (e) Li, J.; Zhu, T.; Hawkins, G. D.; Winget, P.; Liotard, D. A.; Cramer, C. J.; Truhlar, D. G. *Theor. Chem. Acc.* **1999**, *103*, 9.

(24) Cramer, C. J.; Truhlar, D. G. *J. Am. Chem. Soc.* **1994**, *116*, 3892.

(25) (a) Evdokimov, A. G.; Martin, J. M. L.; Kalb (Gilboa), A. J. *J. Phys. Chem. A* **2000**, *104*, 5291. (b) French, A. D.; Dowd, M. K. *J. Comput. Chem.* **1994**, *15*, 561. (c) French, A. D.; Dowd, M. K.; Reilly, P. J. *J. Mol. Struct. (Theochem.)* **1997**, *395*−*396*, 271. (d) Cassett, F.; Imberty, A.; Hérve du Penhoat, C.; Koca, J.; Perez, S. *J. Mol. Struct. (THEOCHEM)* **1997**, *395*−*396*, 211. (e) Cros, S.; Hérve du Penhoat, C. H.; Pérez, S.; Imberty, A. *Carbohydr. Res.* **1993**, *248*, 81. (f) French, A. D.; Tran, V. *Biopolymers* **1990**, *29*, 1599. (f) Mazeau, K.; Perez, S. *Carbohydr. Res.* **1998**, *311*, 203.

(26) Saunders, M.; Houk, K. N.; Wu, Y. D.; Still, W. C.; Lipton, M.; Chang, G.; Guida, W. C. *J. Am. Chem. Soc.* **1990**, *112*, 1419.

(27) (d) Csonka, G. I.; Sosa, C. P. *J. Phys. Chem. A* **2000**, *104*, 7113. (e) Csonka, G. I.; Sosa, C. P.; Csizmadia, I. G. *J. Phys. Chem. A* **2000**, *104*, 3381.

(28) (a) Gohlke, H.; Immel, S.; Lichtenthaler, F. W. *Carbohydr. Res.* **1999**, *321*, 96. (b) Immel, S.; Schmitt, G. E.; Lichtenthaler, F. W. *Carbohydr. Res.* **1998**, *313*, 91. (c) Gruza, J.; Koca, J.; Perez, S.; Imberty, A. *J. Mol. Struct. (THEOCHEM)* **1998**, *424*, 269.

(29) Pérez, S.; Imberty, A.; Engelsen, S. B.; Gruza, J.; Mazeau, K.; Jimenez-Barbero, J.; Poveda, A.; Espinosa, J.-F.; van Eyck, B. P.; Johnson, G.; French, A. D.; Kouwijzer, M. L. C. E.; Grootenuis, P. D. J.; Bernardi, A.; Raimondi, L.; Senderowitz, H.; Durier, V.; Vergoten, G.; Rasmussen, K. *Carbohydr. Res.* **1998**, *314*, 141.