A Quantum Mechanically Derived All-Atom Force Field for Pyranose Oligosaccharides. AMBER* Parameters and Free Energy Simulations

Hanoch Senderowitz and W. Clark Still*

Department of Chemistry, Columbia University, New York, New York 10027

Received July 2, 1996[⊗]

In this paper we extend our previously reported parameterization of carbohydrates to an all-atom AMBER-like force field suitable for modeling oligosaccharides. Parameters were developed from *ab initio* calculations on small model systems having structures characteristic of 1,2-, 1,3-, and 1,4-glycosidic linkages and monosaccharides to give a complete parameter set for pyranose mono- and oligosaccharides. The accuracy of the parameter set was assessed by free energy calculations on various simple sugars and disaccharides. Solvent effects were included using the GB/SA continuum model for water. The sampling problem was solved for these systems by the recently described MC/SD and MC(JBW)/SD simulation methods that facilitate interconversion of conformational states. MC(JBW)/SD simulations were used to calculate anomeric free energies for tetrahydropyran derivatives and monosaccharides. All simulations rapidly converged to values in good agreement with the experimental results. MC/SD simulations were used to study the conformational free energies of selected disaccharides in water. In all cases, our free energy simulations generally reproduced the known conformational features of these systems, namely, the preference of the glycosidic angle (ϕ) for the *exo*-anomeric conformation and the aglyconic angle (ψ) for the *syn* conformation. All of the disaccharides spent >98% of their time in a single conformational energy well corresponding to the *syn*, *exo*-anomeric conformer, though the interresidue ϕ and ψ torsion angles fluctuated considerably in aqueous solution at rt. $^3J_{C,H}$ coupling constants across the glycosidic and aglyconic linkages were calculated using a Karplus-like equation with the distributions of the appropriate torsional angles and show reasonable agreement with the experimental data. Experimental NOE data were also semiquantitatively reproduced. X-ray structures of the disaccharides were found to correspond to populated regions of the simulated ϕ and ψ angular distribution maps.

Introduction

Carbohydrates form an exceptionally important class of molecules, and the development of computational methods for modeling their properties has received considerable attention in recent years. The need for a special molecular mechanics treatment of carbohydrates follows from their densely packed, highly polar functionalities and the dependence of their conformational behavior on stereoelectronic effects (*e.g.*, anomeric, *exo*-anomeric, and *gauche* effects). These issues have long been recognized and have resulted in the development of several parameterization schemes and their subsequent application to the conformational analysis of simple sugars and oligosaccharides, either by energy minimization techniques or by dynamics and Monte Carlo methods.[1−9] While such parameter sets have been successful in reproducing known carbohydrate structures, there are so few examples of well-established carbohydrate conformational energies that it is difficult to assess their energetic accuracy. Thus it is important that carbohydrate force fields in particular be developed on as firm a basis as possible. Like many others working in force field parameterization, we use quantum mechanics as this basis and here describe its use in developing an all-atom parameter set for carbohydrates.

We recently reported a quantum mechanically based carbohydrate parameter set for the united-atom AMBER* force field as implemented in MacroModel 5.0.[10] In a united-atom force field, hydrogens bound to carbon are not explicitly included. Instead, single superatoms are used to represent methine, methylene, and methyl groups. The AMBER* united-atom carbohydrate parameters were based on *ab initio* derived atomic partial charges and conformational energies of small model compounds having structures characteristic of simple sugars. These parameters were subsequently used in a series of MC(JBW)/SD free energy simulations to calculate equilibrium anomeric ratios of tetrahydropyran derivatives and monosaccharides in water. These anomeric ratios were found to be in generally good agreement

[⊗] Abstract published in *Advance ACS Abstracts,* February 1, 1997.

(1) (a) Melberg, S.; Rasmussen, K. *J. Mol. Struct.* **1979**, *57*, 215. (b) Melberg, S.; Rasmussen, K. *Carbohydr. Res.* **1979**, *76*, 23. (c) Melberg, S.; Rasmussen, K. *Carbohydr. Res.* **1980**, *78*, 215. (d) Rasmussen, K. *Acta Chem. Scand.* **1982**, *A36*, 323.

(2) (a) Brady, J. W. *J. Am. Chem. Soc.* **1986**, *108*, 8153. (b) Brady, J. W. *Carbohydr. Res.* **1987**, *165*, 306. (c) Brady, J. W. *J. Am. Chem. Soc.* **1989**, *111*, 5155.

(3) Jeffrey, G. A.; Taylor, R. *J. Comput. Chem.* **1980**, *1*, 99.

(4) Ha, S. N.; Giammona, A.; Field, M.; Brady, J. W. *Carbohydr. Res.* **1988**, *180*, 207.

(5) Homans, S. W. *Biochemistry* **1990**, *29*, 9110.

(6) Mardsen, A.; Robson, B.; Thompson, J. S. *J. Chem. Soc., Faraday Trans. 1* **1988**, *84*, 2519.

(7) (a) Woods, R. J.; Dwek, R. A.; Edge, C. J.; Fraser-Reid, B. *J. Phys. Chem.* **1995**, *99*, 3832. (b) Zheng, Y.-J.; Le Grand, S. M.; Merz, K. M. *J. Comput. Chem.* **1992**, *13*, 772. (c) Glennon, T. M.; Zheng, Y. J.; Le Grand, S. M.; Shutzberg, B. A.; Merz, K. M. *J. Comput. Chem.* **1994**, *15*, 1019.

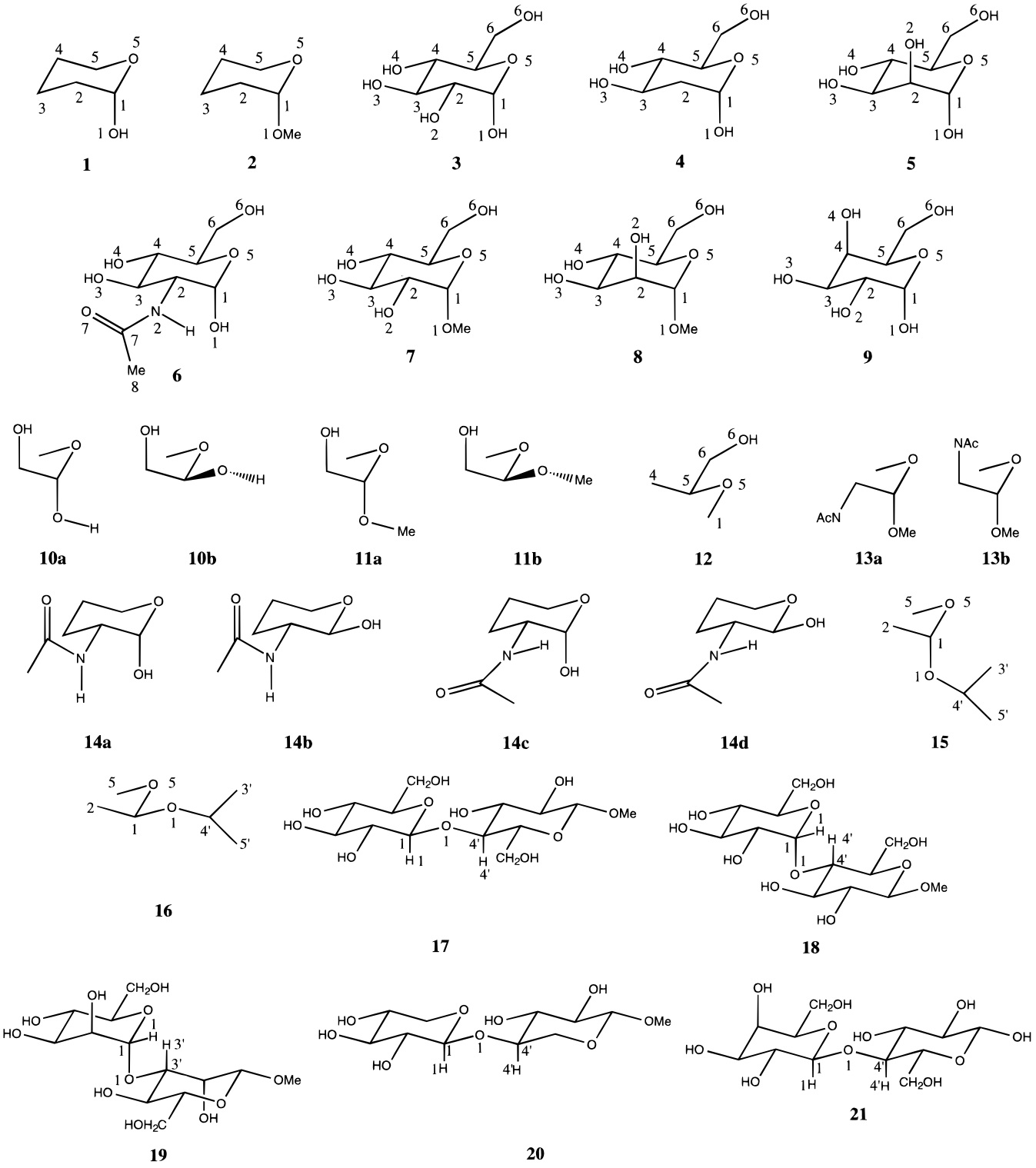
(8) (a) Grootenhuis, P. D. J.; Haasnoot, C. A. G. *Mol. Simul.* **1993**, *10*, 75. (b) Kouwijzer, M. L. C. E.; Grootenhuis, P. D. J. *J. Phys. Chem.* **1995**, *99*, 13426.

(9) Reiling, S.; Schlenkrich M.; Brickmann, J. *J. Comput. Chem.* **1996**, *17*, 450.

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

S0022-3263(96)01248-0 CCC: $14.00

Chart 1



with available experimental data. With disaccharides, however, the united-atom force field performed more poorly and predicted conformational populations that differed significantly from experiment. The problem appeared to lie in the united-atom approximation—an approximation that is good when interacting nonbonded atoms are widely separated but that is less realistic when they are in close contact. Because the most-known oligosaccharide conformations incorporate several short-range nonbonded interactions involving sugar methine or methylene groups, it appeared that explicit inclusion of all hydrogens might be necessary to provide an accurate model of such systems.

In this work we therefore extend our force field treatment to oligosaccharides by developing an all-atom parameter set for carbohydrates. We based the new parameter set on our previously reported *ab initio* calculations[10] together with new data for relevant compounds modeling the glycosidic and aglyconic linkages. To validate our work, we calculated anomeric free energies of tetrahydropyran derivatives **1**, **2** (Chart 1) and monosaccharides **3**−**9** in water for which quantitative experimental data are available. We then studied the

conformational properties of several disaccharides **17**−**21** whose preferred conformations in water have been investigated experimentally.

AMBER* Parameterization of Pyranoses. Parameterization of the all-atom AMBER* force field for carbohydrates followed the methodology described previously for producing our united-atom carbohydrate parameter set.[10] Thus atomic partial charges were derived from the *ab initio* wave functions of suitable model compounds through electrostatic fitting using the CHELPG procedure.[11] *Ab initio* torsional energy profiles were obtained by dihedral angle driving with relaxation of all other degrees of freedom. Molecular mechanics energy profiles were obtained in a similar manner using the CHELPG-derived atomic partial charges and were fitted to the *ab initio* ones by adjusting appropriate torsional parameters. Except for O−C−O bond length and angle parameters (O−C $r_{eq} = 1.425$ Å, $K_r = 350$ kcal/mol·Å; O−C−O $\theta_{eq} = 109.5°$, $K_\theta = 40$ kcal/mol·rad) and the C−O−C glycosidic bond angle ($\theta_{eq} = 114.3°$, $K_\theta = 60$ kcal/mol·rad; see below), all stretch and bend parameters were taken from the native AMBER parameters for the sugar portion of uracil.[12]

As before, we chose 2-hydroxytetrahydropyran (**1**) and 2-methoxytetrahydropyran (**2**) as model systems for parameterization of the hemiacetal and acetal fragments of sugars. For both systems, the global axial and equatorial minima are in the *exo*-anomeric conformation, the axial anomers being energetically favored by 0.7 (**1**, HF/6-311++G**//HF/6-31G**) and 0.9 (**2**, HF/6-311++G**//HF/6-31G*) kcal/mol. To provide accurate solvation energies with the GB/SA water model,[13] we obtained atomic partial charges at the PS-GVB/6-31G**//HF//6-31G** level of theory[14] for both axial and equatorial global minima of **1** and **2**. We then added the partial charges on alkyl hydrogens to the charges of attached carbons and averaged the resulting united-atom charges of both minima. Thus the charges on all alkyl hydrogens were set to zero which provides alkyl hydrogens with steric but not electrostatic properties. The only exception to this charge treatment was the anomeric hydrogen H(C1) that was assigned its conformationally averaged, *ab initio* derived charge. All atomic partial charges and other parameters developed here are given as Supporting Information.

The rotational energetic profiles around the C1−O1 glycosidic bond of both axial and equatorial **1** were calculated at 60° resolution at HF/6-311++G**//HF/6-31G**. Those for **2** were available at 30° resolution at the HF/6-311++G**//HF/6-31G* level from Tvaroska and Carver.[15] The molecular mechanics C5−O5−C1−O1 and (C2,O5)−C1−O1−R (R = H, Me) torsional parameters were adjusted to reproduce the *ab initio* energy differences between the axial and equatorial anomeric minima and the rotational energy profiles. A comparison of the relative *ab initio* and AMBER* anomeric energies and rotational profiles for **1** and **2** is presented in Table 1 and indicates that the relative energies of these states are well-reproduced by the new AMBER* force field. Though *ab initio* and AMBER* disagree on the nature of the O5−C1−O1−H *trans* conformation of axial **1** (*ab initio* finds a very shallow energy minimum whereas AMBER* finds an almost featureless plane), both agree that the area around axial *trans*-**1** is energetically quite flat and lies 3.6−3.9 kcal/mol above the global minimum *gauche*$^+$-**1**. A related situation exists for the high-energy *gauche*$^-$ conformation of axial **2**. The full rotational profiles are shown in Figures 1 and 2 of the Supporting Information.

Table 1. Relative Conformational Energies (kcal/mol) of 1 and 2 as Calculated *Ab Initio* and with the Reparameterized AMBER* Force Field

conformer	*ab initio*[a]	AMBER*
	1, axial	
trans[b]	3.63[e]	3.89[e]
gauche$^-$ [c]	3.85	3.66
gauche$^+$ [d]	0.00	0.00
	1, equatorial	
trans[b]	5.36	5.32
gauche$^-$ [c]	0.69	0.68
gauche$^+$ [d]	1.46	1.35
	2, axial	
trans[b]	3.68[f]	5.35[f]
gauche$^-$ [c]	10.11[g]	8.32[g]
gauche$^+$ [d]	0.00	0.00
	2, equatorial	
trans[b]	5.48[f]	6.19[f]
gauche$^-$ [c]	0.94	0.97
gauche$^+$ [d]	3.93	4.00

[a] **1**: HF/6-311++G**//HF/6-31G**. **2**: HF/6-311++G**//HF/6-31G*. [b] Trans: O5−C1−O1−H, C ∼ 180°. [c] Gauche$^+$: O5−C1−O1−H, C ∼ 60°. [d] Gauche$^-$: O5−C1−O1−H, C ∼ −60°. [e] conformer is not a minimum on the AMBER* potential surface and only a shallow one on the *ab initio* potential surface (see Figure 1 in the Supporting Information). [f] Conformers are not minima on the potential energy surface. [g] Conformer is not a minimum on the *ab initio* potential surface and only a shallow, high-energy one on the AMBER* potential surface.

Next, we parameterized the primary and secondary hydroxy groups. Since the hydroxyls in simple sugars are closely related to those in nucleic acids, we used the same stretch, bend, and atomic partial charge parameters for nonanomeric hydroxyls as AMBER uses for uracil.[12] For hydroxyl conformational energies and barriers, however, we used HF/6-31G*//HF/6-31G* calculations on model compounds 2-methoxyethanol (for the C6 primary alcohol) and 2-propanol (for the C2−C4 secondary alcohols). The *ab initio* profiles for the C−C−O−H torsional rotations along with the corresponding all-atom AMBER* results are shown in Figures 3 and 4 of the Supporting Information.

For parameterizing the endocyclic C3−C2−C1−O1 and C5−O5−C1−O1 torsions, we used the *exo*-anomeric conformations of 2,4-dihydroxytetrahydropyran (C4 hydroxy equatorial), 2-hydroxy-6-(hydroxymethyl)tetrahydropyran, and 2-methoxy-6-(hydroxymethyl)tetrahydropyran as model compounds (in the last two systems the O5−C5−C6−O6 and C5−C6−O6−H torsions were initially set to 60° and −60°, respectively). After torsional parameterization, all-atom AMBER* calculations exactly reproduced the *ab initio* energy differences of 1.3, 0.9, and 1.1 kcal/mol (all in favor of the *trans* diastereomers) for the three model systems respectively. Small adjust-

(11) Breneman, C. M.; Wiberg, K. B. *J. Comput. Chem.* **1990**, *11*, 361.

(12) (a) Weiner, S. J.; Kollman, P. A.; Case, D. A.; Singh, U. C.; Chio, C.; Alagona, G.; Profeta, S.; Weiner, P. *J. Am. Chem. Soc.* **1984**, *106*, 765. (b) Weiner, S. J.; Kollman, P. A.; Nguyen, D. T.; Case, D. A. *J. Comput. Chem.* **1986**, *7*, 230.

(13) (a) Still, W. C.; Tempczyk, A.; Hawley, R. C.; Hendrickson, T. *J. Am. Chem. Soc.* **1990**, *112*, 6127. See also Cramer, C. J.; Truhlar, D. G. *J. Am. Chem. Soc.* **1993**, *115*, 5745. (b) Qui, D.; Shenkin, P. S.; Hollinger, F.; Still, W. C. *J. Phys. Chem.*, in press.

(14) Ringnalda, M. N.; Langlois, J. M.; Greeley, B. H.; Murphy, R. B.; Russo, T. V.; Cortis, C.; Muller, R. P.; Marten, B.; Donnelly, R. E., Jr.; Mainz, D. T.; Wright, J. R.; Pollard, W. T.; Cao, Y.; Won, Y.; Miller, G. H.; Goddard, W. A., III; Friesner, R. A. *PS-GVB Version 2.01*; Schrodinger, Inc., 1994.

(15) Tvaroska, I.; Carver, C. P., *J. Phys. Chem.* **1994**, *98*, 9477.

ments were also made in the C−O−C−C torsional parameters to reproduce the HF/6-31G*//HF/6-31G* rotational profile of methyl ethyl ether. For torsional parameters associated with the C2 hydroxyls, we fit the O2−C2−C1−O1 torsional parameters to HF/6-31G*//HF/6-31G* conformational energies of hydroxy hemiacetal **10** and hydroxy acetal **11**. The torsional parameters for rotation around the carbohydrate C5−C6 exocyclic bond were obtained by fitting HF/6-31G*//HF/6-31G* *ab initio* rotational profiles for 2-methoxy-1-propanol (**12**) with the equivalent C4−C5−O5−C1 (sugar nomenclature) torsion constrained to 60° to maintain a pyranose-like geometry and the hydroxy H initially set to be *anti*. The *ab initio* and AMBER* results are shown in Figure 5 of the Supporting Information. With all these model compounds, the agreement between quantum mechanics and molecular mechanics was very good.

For *N*-acetylglucosamine we used the same parameters for the *N*-acetyl part that are used in AMBER for peptide acetamides.[12] The O1−C1−C2−N2 torsional parameters were then adjusted to reproduce the HF/6-31G*//HF/6-31G* energy difference between the two diagrammed conformers of acetamido acetal **13** (*ab initio* and AMBER* calculations gave 0.19 and 0.22 kcal/mol in favor of **13a**), and the O1−C1−C2−C3, C3−C2−N2−C7, and C1−C2−N2−C7 torsional parameters were adjusted to reproduce HF/6-31G*//HF/6-31G* *ab initio* calculations of *cis*- and *trans*-3_{eq}-acetamido-2-hydroxytetrahydropyran (**14**). (*Ab initio*: **14a**, 0.0; **14b**, 3.4; **14c**, 7.6; **14d**, 3.8 kcal/mol. AMBER*: **14a**, 0.0; **14b**, 3.6; **14c**, 9.1; **14d**, 3.8 kcal/mol.) Several torsional parameters for the *N*-acetyl part were taken from the McDonald−Still peptide backbone parameters (see Supporting Information).[16]

The final AMBER* parameter set was used to calculate the energy difference between two particular conformations of α- and β-glucose *in vacuo*. The result, 1.28 kcal/mol in favor of the former, is in good agreement with the HF/6-31G*//HF/6-31G* estimate of 1.12 kcal/mol.

Finally, parameters for the glycosidic and aglyconic linkages were developed from *ab initio* calculations on pseudoaxial and pseudoequatorial acetals **15** and **16**. During the parameterization procedure, the C5−O5−C1−C2 torsion was constrained to −60° to maintain a pyranose-like geometry. As evident from the structures of model compounds **15** and **16**, the new parameters are adequate for all types of disaccharides except 1,1- and 1,6-linked ones. For simplicity, we will use a notation appropriate for a 1,4-linkage in the following. First, a θ_0 value for the C1−O1−C4′ bond angle and parameters for the C5−O5−C1−O1 and C5−O5−C1−H(C1) torsions were obtained by fitting the HF/6-31G*//HF/6-31G* geometries and relative energies of the two model systems. The final reparameterized AMBER* results are in good agreement with the *ab initio* ones. (*Ab initio*: **15** E_{rel} = 0.0 kcal/mol, C1−O1−C4′ = 116.7°, **16** E_{rel} = 1.90 kcal/mol, C1−O1−C4′ = 117.4°. AMBER*: **15** E_{rel} = 0.0 kcal/mol, C1−O1−C4′ = 117.0°, **16** E_{rel} = 1.94 kcal/mol, C1−O1−C4′ = 117.1°.) Next, torsional parameters around the C1−O1 and O1−C4′ bonds were developed by fitting the appropriate HF/6-31G*//HF/6-31G* rotational profiles of **15** and **16**. A comparison between *ab initio* and AMBER* results is provided in Figures 1 and 2 and shows good agreement between the two methods. Parameters for C−O−C−C−O−C sequences in disaccharides were taken directly from Kollman *et al.*[17] A complete set of the new AMBER* parameters and partial charges for all the systems described above is provided as Supporting Information.

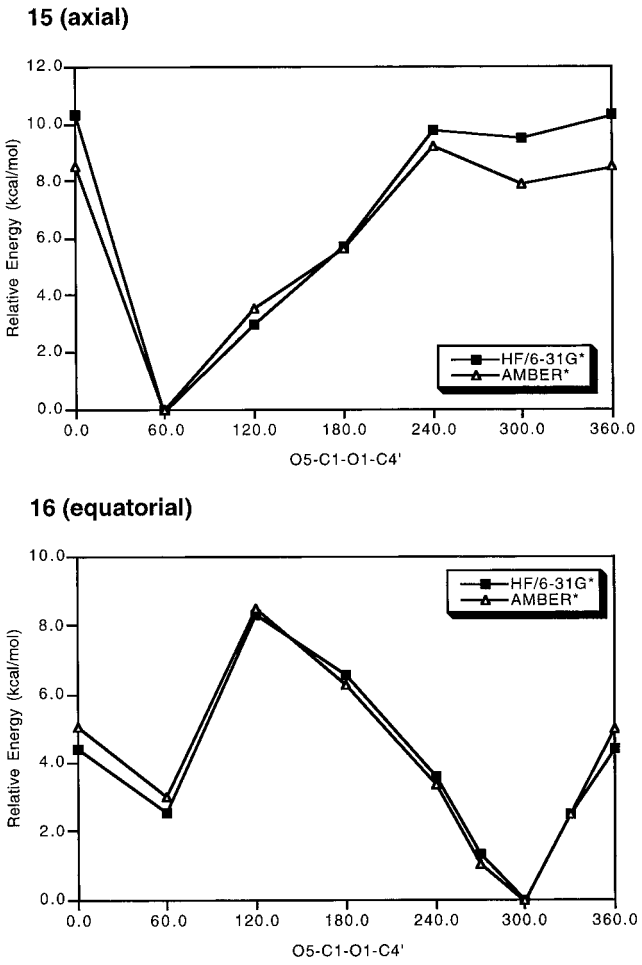


Figure 1. Rotational profiles around the C1−O1 (glycosidic) torsion in acetals **15** and **16** as calculated *ab initio* (HF/6-31G*//HF/6-31G*) and with the reparameterized AMBER* force field.

Anomeric Free Energies of Pyranoses. The strategy for calculating anomeric free energies using the MC(JBW)/SD simulation technique[18] has been described in detail elsewhere and is discussed only briefly here.[10] This simulation method allows for rapid interconversion between anomeric and conformational isomers, and the anomeric ratio is determined by counting the number of structures that are α and the number that are β in the ensemble of states generated. All the calculations were performed with the above-described all-atom AMBER* force field as implemented in MacroModel 5.5.[19] Input conformations were obtained from 5000 (tetrahydropyran derivatives) and 30 000 (monosaccharides) step conformational searches using the internal coordinate SUMM method.[20] For the tetrahydropyran derivatives, all conformations within ∼10 kcal/mol of the global minimum were used as input for the MC(JBW)/SD algorithm. For

(16) McDonald, D. Q.; Still, W. C. *Tetrahedron. Lett.* **1992**, *33*, 7743.

(17) Billeter, M.; Howard, A. E.; Kuntz, I. D.; Kollman, P. A. *J. Am. Chem. Soc.* **1988**, *110*, 8385.

(18) Senderowitz, H.; Guarnieri, F.; Still, W. C. *J. Am. Chem. Soc.* **1995**, *117*, 8211.

(19) Mohamadi, F.; Richards, N. G. J.; Guida, W. C.; Liskamp, R.; Lipton, M.; Caufield, C.; Chang, G.; Hendrickson, T.; Still, W. C. *J. Comput. Chem.* **1990**, *11*, 440.

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

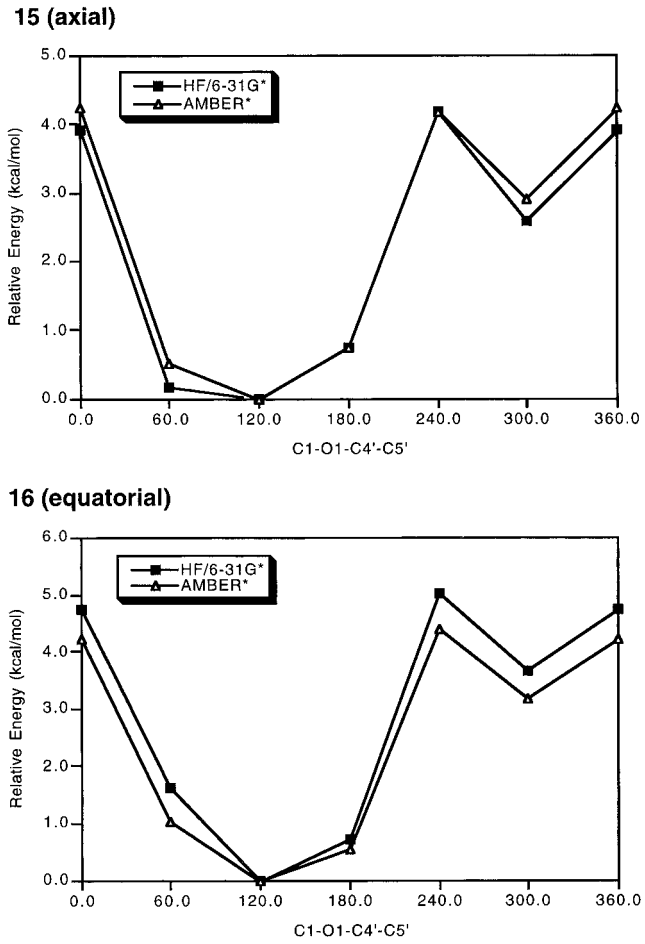


Figure 2. Rotational profiles around the O1−C4′ (aglyconic) torsion in acetals **15** and **16** as calculated *ab initio* (HF/6-31G*//HF/6-31G*) and with the reparameterized AMBER* force field.

the more conformationally rich monosaccharides, only the 100 lowest energy structures were used (typically 2.4−3.4 kcal/mol of the global minimum). MC(JBW)/SD simulations were run each for 10 ns in vacuum and GB/SA chloroform and/or GB/SA water as appropriate for the system under study. In all cases, interconversions between different conformers or anomers occurred at least once every 0.6 ps on average, assuring good convergence of the anomeric ratios. Free energy differences were determined from the distribution of the C3−C2−C1−O1 torsion for tetrahydropyran derivatives (−60 ± 30° and 180 ± 30° for the axial and equatorial conformers, respectively; see **1**) and the improper torsion C2−C1−O5−O1 for monosaccharides (−120 ± 30° and 120 ± 30° for the α and β anomers respectively; see **3**). Convergence was assessed by monitoring these distributions in 1 ns blocks with standard deviations being calculated by the method of block averages.

As before, we tested our new parameter set and sampling methodology by anomeric free energy calculations on 2-hydroxytetrahydropyran (**1**) and 2-methoxytetrahydropyran (**2**). Both molecules have been studied extensively by experiment and calculation.[21−23] The conformational properties of these molecules are partly governed by the anomeric and *exo*-anomeric effects.[24] Both show preference for the axial conformation in the gas phase and nonpolar solvents and for the equatorial conformation in water. The results of our 10 ns MC(JBW)/SD simulations of **1** in vacuum and GB/SA water and **2** in vacuum, GB/SA water, and GB/SA chloroform are summarized in the first five entries of Table 2. In all cases the free energies after 1 and 10 ns differ by no more than 0.1 kcal/mol. The final results reproduce the experimentally observed trend on going from nonpolar to polar solvents and are in good agreement with available experimental data.

Several free energy calculations involving the anomeric behavior of glucose are found in the literature. Merz *et al.* used a reparameterized AMBER force field and SPC/E water to obtain an α/β free energy difference of 3.03 ± 0.04 kcal/mol,[7c] Karplus *et al.* used the CHARMM force field and TIP3P water to calculate an α/β free energy difference of 0.31 ± 0.43 kcal/mol,[25a] and Van Eijck *et al.* used the GROMOS force field and SPC/E water to obtain an α/β free energy difference of −0.86 ± 0.15 kcal/mol.[25b] The two latter values are in reasonable agreement with the experimental measurement of −0.34 kcal/mol (in favor of β).[26] Our previously reported estimate of this energy difference using the united-atom AMBER* force field and GB/SA water is −0.22 ± 0.02 kcal/mol.[10]

Our current simulation of all-atom glucose began with a conformational search of the α and β anomers using the new all-atom AMBER* force field and GB/SA water. The lowest 100 structures spanned an energy window of 3.1 kcal/mol of the global minimum and were used as input for a 10 ns MC(JBW)/SD free energy simulation. The α/β ratios were obtained by monitoring the C2−C1−O5−O1 torsion as noted above. The results of our simulation (Table 2, entry 6) favor the β anomer by 0.12 kcal/mol, in good agreement with experiment.

The anomeric ratios in pyranoses are known to depend on the substitution pattern at C2 of the pyranose ring. Thus, the α/β free energy differences for glucose (**3**, equatorial OH at C2), 2-deoxyglucose (**4**, no substituent at C2), mannose (**5**, axial OH at C2), and *N*-acetylglucosamine (**6**, equatorial acetamide at C2) have been experimentally measured in water and found to be −0.34,[26] −0.05,[27] 0.34−0.45,[26] and 0.51[28] kcal/mol, respectively. Our results, obtained from 10 ns MC(JBW)/SD simulations in GB/SA water, are −0.12, −0.42, 0.24, and 0.56 kcal/mol for the four systems, respectively, all in reasonable agreement with experiment (Table 2, entries 6−9).

(21) (a) Wiberg, K. B.; Murcko, M. A. *J. Am. Chem. Soc.* **1989**, *111*, 4821. (b) Zheng, Y-J.; Le Grand, S. M.; Merz, K. M., Jr. *J. Comput. Chem.* **1992**, *13*, 772. (c) Salzner, U.; Schleyer, P. v. R. *J. Org. Chem.* **1994**, *59*, 2138. (d) Jorgensen, W. L.; Morales de Tirado, P. I.; Severance, D. L. *J. Am. Chem. Soc.* **1994**, *116*, 2199.

(22) Lemieux, R. U.; Pavia, A. A.; Martin, J. C.; Watanabe, K. A. *Can. J. Chem.* **1969**, *47*, 4427.

(23) (a) de Hoog, A. J.; Buys, H. R.; Altona, C.; Havinga, E. *Tetrahedron* **1969**, *25*, 3365. (b) El-Kafrawy, A.; Perraud, R. *C. R. Acad. Sci. Paris, Ser. C* **1975**, *280*, 1219. (c) Praly, J-P.; Lemieux, R. U. *Can. J. Chem.* **1987**, *65*, 213. (d) Wiberg, K. B.; Marquez, M. *J. Am. Chem. Soc.* **1994**, *116*, 2197.

(24) Kirby, A. J. *The Anomeric Effect and Related Stereoelectronic Effects on Oxygen*; Springer Verlag: Berlin, 1983 and references cited therein.

(25) (a) Ha, S.; Gao, J.; Tidor, B.; Brady, J. W.; Karplus, M. *J. Am. Chem. Soc.* **1991**, *113*, 1553. (b) Van Eijck, B. P.; Hooft, R. W. W.; Kroon, J. *J. Phys. Chem.* **1993**, *97*, 12093. (c) Schmidt, R. K.; Karplus, M.; Brady, J. W. *J. Am. Chem. Soc.* **1996**, *118*, 541.

(26) (a) Rudrum, M.; Shaw, D. F. *J. Chem. Soc.* **1965**, 52. (b) Reference 24, p 7. (c) Stoddart, J. F. *Stereochemistry of Carbohydrates*; Wiley Interscience: New York, 1971; p 92.

(27) (a) Angyal, S. J. *Aust. J. Chem.* **1968**, *21*, 2737. (b) Angyal, S. J. *Angew. Chem., Int. Ed. Engl.* **1969**, *8*, 157. (c) Pfeffer, P. E.; Parrish, W.; Unruh, J. *Carbohydr. Res.* **1980**, *84*, 13.

(28) (a) Horton, D.; Jewell, J. S.; Philips, K. D. *J. Org. Chem.* **1966**, *31*, 4022. (b) Okumura, H.; Azuma, I.; Kiso, M.; Hasegawa, A. *Carbohydr. Res.* **1983**, *117*, 298.

Table 2. Calculated and Experimental Anomeric Free Energies (kcal/mol) for the α, β Equilibrium of the Tetrahydropyran Derivatives and Monosaccharides (Positive Values Favor the α Conformer and Negative, the β Conformer)

	ΔG (kcal/mol)		
	AMBER* calculated		
system	Boltzmann[f]	MC(JBW)/SD[g]	experiment/ ab initio
1 (vacuum)	0.83	0.17 ± 0.02	−0.07,[a] 0.69[b]
1 (H_2O)	−0.92	−1.20 ± 0.05	−0.95[23c]
2 (vacuum)	0.99	0.73 ± 0.03	0.90,[c] 0.94[d]
2 ($CHCl_3$)	0.47	0.24 ± 0.04	0.64[24]
2 (H_2O)	−0.50	−0.70 ± 0.03	0.1 −(−0.7)[23a−d]
glucose (H_2O)	0.63	−0.12 ± 0.01	−0.34[26]
2-deoxyglucose (H_2O)	0.03	−0.42 ± 0.01	−0.05[27]
mannose (H_2O)	0.06	0.24 ± 0.01	0.34−0.45[26]
N-acetylglucosamine (H_2O)	0.80	0.56 ± 0.01	0.51[28]
methyl glycoside (H_2O)	0.90	0.64 ± 0.01	0.42 (MeOH)[e]
methyl mannoside (H_2O)	0.76	0.79 ± 0.02	1.70 (MeOH)[e]
galactose (H_2O)	0.69	0.02 ± 0.02	−0.37[26]

[a] In CCl_4.[23c] [b] The energy difference between the lowest axial (C5−O5−C1−O1 = 60°, O5−C1−O1−H = 60°) and lowest equatorial (C5−O5−C1−O1 = 180°, O5−C1−O1−H = −60°) conformations from HF/6-311++G**//HF/6-31G** *ab initio* calculations in this work. [c] In CCl_4.[24] [d] The energy difference between the lowest axial (C5−O5−C1−O1 = 60°, O5−C1−O1−Me = 60°) and lowest equatorial (C5−O5−C1−O1 = 180°, O5−C1−O1−Me = −60°) conformations from HF/6-311++G**//HF/6-31G* *ab initio* calculations in ref 15. [e] The experimental value for equilibration in 1% methanolic HCl at 35 °C.[29] [f] Results based on Boltzmann-weighted average of minimum energy conformers and anomers using the new AMBER* force field. [g] Results based on 10 ns MC(JBW)/SD free energy simulation at 300 K using the new AMBER* force field. Values include statistical uncertainty (1σ) in the result and were computed from five block averages.

The experimental anomeric ratios for methyl glucoside (**7**) and methyl mannoside (**8**) were measured in methanol solution and found to favor the α anomer by 0.41[29] and 1.70[29] kcal/mol, respectively. Our 10 ns MC(JBW)/SD simulations in GB/SA water gave 0.64 and 0.79 kcal/mol in favor of the α anomer for these two systems. These anomeric free energies are similar to, but not strictly comparable with, the experimental ones because of the differing solvents (Table 2, entries 10−11). Finally, the anomeric free energy of galactose (**9**) is small and in reasonable accord with experiment (0.37[26] kcal/mol in favor of the β anomer by experiment and 0.02 kcal/mol in favor of the α anomer by calculation; Table 2, entry 12).

As the results in Table 2 indicate, for all the molecular systems studied in this work, the calculated and experimental free energies are in qualitative agreement with an average unsigned error for all solvated systems (excluding aqueous **2** for which the experimental value is not well-defined) of 0.32 kcal/mol and a maximum error of 0.91 kcal/mol. Furthermore, in all cases, the standard deviations of our free energy results, as calculated by the method of block averages, are equal to or less than 0.05 kcal/mol and convergence was obtained within ∼1 ns of simulation time. Also given in Table 2 are the corresponding α/β free energy differences calculated by simple Boltzmann averaging of the steric energies of all minima within 50 kJ/mol of the global minimum. The average unsigned error for the solvated systems based on these energies is somewhat larger, at 0.48 kcal/mol with a maximum error of 1.06 kcal/mol.

Conformational Free Energies of Disaccharides. Disaccharides are an important class of carbohydrates, both from the standpoint of their chemical and biological importance and also as the transition between mono- and polysaccharides. Thus their conformations, defined by the glycosidic (ϕ = H1−C1−O1−C4′) and aglyconic (ψ = C1−O1−C4′−H4′) torsions, have been studied extensively by experimental and theoretical methods.[30−32]

Experimentally, solution conformations of disaccharides are deduced primarily from NMR studies, in particular by interglycosidic NOE and $^3J_{^{13}C,^1H}$ coupling constant measurements. The derivation of interatomic distances from NOE data is subject to many possible sources of error and is also known to depend on experimental conditions.[33−35] Furthermore, due to the r^{-6} dependence of the NOE signals on interproton distances, significant NOEs may appear between protons which spend only a small fraction of the time in close proximity. Thus, relative proton−proton distances from NOE experiments are usually used as qualitative "molecular rulers" to constrain molecular simulations to experimentally accessible regions of the potential energy surface. Quantitative (though not unambiguous) torsional information can be obtained from measurements of time-averaged $^3J_{C,H}$ coupling constants across the glycosidic and aglyconic bonds. These have been assumed to obey a Karplus-like equation and fitted to a set of X-ray structures to obtain the appropriate coefficients.[36−38] However, on the basis of the rather limited set of data used in the fitting procedures and the large error bars, the resulting equations appear accurate to only ±1 Hz.[36]

Published computational work on disaccharides consists mainly of molecular mechanics based calculations both with and without consideration of solvent. The energetic procedures typically include grid search/energy minimization methods to generate nonrelaxed or relaxed

(29) Smirnyagin, V.; Bishop, C. T. *Can. J. Chem.* **1968**, *46*, 3085.

(30) *Computer Modeling of Carbohydrate Molecules*; French, A. D., Brady, J. W., Eds.; ACS Symposium Series 430; American Chemical Society: Washington, DC, 1990.

(31) Homans, S. W. In *Molecular Glycobiology*; Fukuda, M., Hindsgaul, O., Eds.; Oxford University Press, Oxford, 1994.

(32) Lemieux, R. U.; Koto, S. *Tetrahedron* **1974**, *30*, 1933.

(33) Dabrowski, J.; Kozár, T.; Grosskurth, H.; Nifant'ev, N. E. *J. Am. Chem. Soc.* **1995**, *117*, 5534.

(34) Cumming, D. A.; Carver, J. P. *Biochemistry* **1987**, *26*, 6664.

(35) Schirmer, R. E.; Noggle, J. H.; Davis, J. P.; Hart, P. A. *J. Am. Chem. Soc.* **1970**, *92*, 3266.

(36) Mulloy, B.; Frenkiel, T. A.; Davies, D. B. *Carbohydr. Res.* **1988**, *184*, 39.

(37) Tvaroska, I.; Hricovini, M.; Pertáková, E. *Carbohydr. Res.* **1989**, *189*, 359.

(38) Hamer, G. K.; Balza, F.; Cyr, N.; Perlin, A. *Can. J. Chem.* **1978**, *56*, 3109.

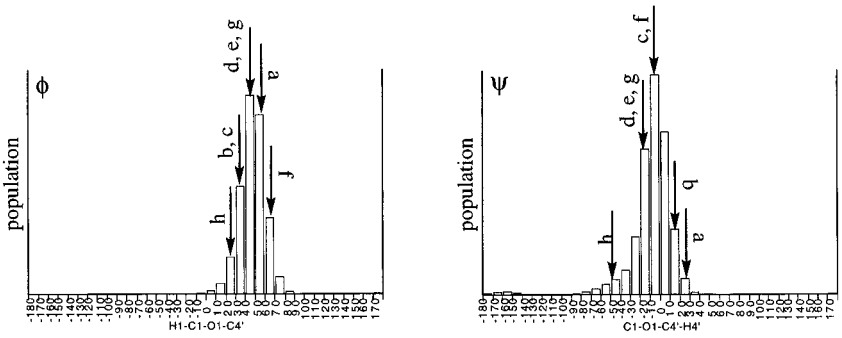


Figure 3. Distribution of the glycosidic (ϕ = H1−C1−O1−C4′) and aglyconic (ψ = C1−O1−C4′−H4′) torsions of methyl β-cellobioside (**17**) at the end of a 10 ns MC/SD simulation. Arrows point to the positions of ϕ and ψ of selected X-ray structures with the cellobiose core, retrieved from the Cambridge Structural Database. The following X-ray structures are included (ϕ and ψ are given in parentheses): β-D-acetylcellobiose (50.6°, 21.4°);[70a] 6′-*O*-trityl-α-cellobiose heptaacetate (35.3°, 13.2°);[70b] β-cellobiose (38.9°, −9.0°);[70c] β-cellobiose (44.4°, −11.8°);[70d] β-cellobiose (42.3°, −17.9°);[70e] methylhepta-*O*-nitro-β-cellobiose (67.2°, −3.7°);[70f] methyl-6,6′-dinitrato-β-cellobiose (48.0°, −16.4°);[70g] methyl-β-cellobiose (24.7°, −47.7°; the CSD structure corresponds to L-sugars. Thus, the values of ϕ and ψ of were inverted).[70h]

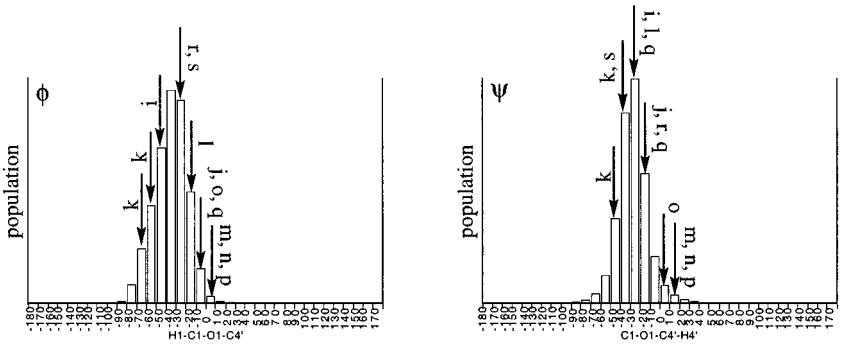


Figure 4. Distribution of the glycosidic (ϕ = H1−C1−O1−C4′) and aglyconic (ψ = C1−O1−C4′−H4′) torsions of methyl β-maltoside (**18**) at the end of a 10 ns MC/SD simulation. Arrows point to the positions of ϕ and ψ of selected X-ray structures with the maltose core, retrieved from the Cambridge Structural Database. The following X-ray structures are included (ϕ and ψ are given in parentheses): methyl α-maltotrioside tetrahydrate (−43.6°, −28.9°);[70i] erlose monohydrate trisaccharide (−9.5°, −12.9°);[70j] erlose trihydrate (−68.8° and −50.3°, −42.4° and −32.3°);[70k] α,β-panose (−16.8°, −21.6°);[70l] β-maltose monohydrate (3.9°, 12.2°);[70m] β-maltose monohydrate (neutron diffraction, 4.8°, 13.3°);[70n] α-maltose (−3.6°, 4.1°);[70o] methyl β-maltoside monohydrate (5.6°, 15.9°);[70p] phenyl α-maltoside (−7.6° and −8.1°, −15.8° and −25.5°);[70q] α-panose (−23.9°, −12.8°);[70r] β-maltose octaacetate (−28.6°, −35.6°).[70s]

potential energy surfaces and to a lesser extent molecular dynamics or Monte Carlo simulations. Energy minimization techniques generate a static picture of the system at low temperature and do not account for entropic effects. Dynamic methods are much more suitable for calculating ensemble-averaged properties, but flexible, multiconformational systems such as oligosaccharides are haunted by difficulties in crossing conformational energy barriers. In spite of these difficulties, available experimental and theoretical studies present the following qualitative picture of disaccharide conformation.[30−32,39] In solution, all pyranose rings are found in the chair conformation. The glycosidic torsion prefers the *exo*-anomeric conformation (ϕ(H1−C1−O1−C4′) = −60° ± 60° (α-D sugars), 60° ± 60° (β-D sugars)) while the aglyconic torsion prefers a *syn* conformation (ψ(C1−O1−C4′−H4′) = 0° ± 90°), both with considerable conformational flexibility.

In this work, we carried out free energy simulations for five selected disaccharides, **17**−**21**. Conformational freedom in these systems follows from rotations around the glycosidic and aglyconic interresidue bonds and around the primary and secondary hydroxyl groups. These degrees of freedom give rise to a densely populated conformational space that can be efficiently sampled by simple Metropolis Monte Carlo methods. We have therefore carried out a series of 10 ns MC/SD simulations[40] using our all-atom AMBER* force field described above as implemented in MacroModel 5.5.[19] Solvent was modeled by GB/SA continuum water.[13] The distribution of the glycosidic (ϕ) and aglyconic (ψ) torsions were monitored over 1 ns intervals. Structures sampled over the course of the simulation (Figures 3−7) were used to calculate averaged $^3J_{C,H}$ coupling constants across the H1−C1−O1−C4′ (glycosidic) and C1−O1−C4′−H4′ (aglyconic) sequences *via* the Karplus-like equation of Mulloy *et al.*[36] We did not attempt quantitative NOE analyses of our simulation results, but we did monitor distances (r) between protons for which significant NOE signals were observed to see if a qualitative relationship between proton−proton proximity and NOE signals could be found. Thus we accumulated $1/r^6$ averages ($\langle r^{-6}\rangle$) for relevant proton−proton pairs over the course of our simulations for comparison with experimental relative NOE intensities.

Methyl β-cellobioside (**17**) and methyl β-maltoside (**18**) are important disaccharides and are also building blocks

(39) Wang, Y.; Goekjian, G.; Ryckman, D. M.; Miller, W. H.; Babirad, S. A.; Kishi, Y. *J. Org. Chem.* **1992**, *57*, 482.
(40) Guarnieri, F.; Still, W. C. *J. Comput. Chem.* **1994**, *15* 1302.
(41) Melberg, S.; Rasmussen, K. *Carbohydr. Res.* **1979**, *71*, 25.
(42) Tvaroska, I. *Biopolymers* **1984**, *23*, 1951.

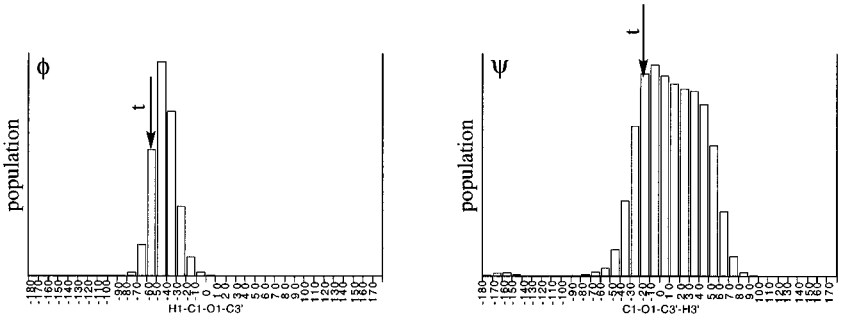


Figure 5. Distribution of the glycosidic (ϕ = H1−C1−O1−C3′) and aglyconic (ψ = C1−O1−C3′−H3′) torsions of mannosyl-(α-1,3)-mannose-β-OMe (**19**) at the end of a 10 ns MC/SD simulation. Arrows point to the positions of ϕ and ψ of selected X-ray structures with the dimannose core, retrieved from the Cambridge Structural Database. The following X-ray structures are included (ϕ and ψ are given in parentheses): *O*-α-D-mannopyranosyl-(1−3)-*O*-β-D-mannopyranosyl-(1−4)-2-acetamido-2-deoxy-α-D-glucopyranose (−57.6°, −19.4°).[70t]

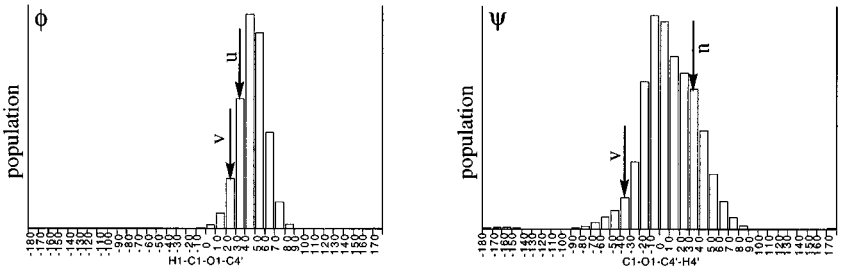


Figure 6. Distribution of the glycosidic (ϕ = H1−C1−O1−C4′) and aglyconic (ψ = C1−O1−C4′−H4′) torsions of methyl β-xylobioside (**20**) at the end of a 10 ns MC/SD simulation. Arrows point to the positions of ϕ and ψ of selected X-ray structures with the xylobiose core, retrieved from the Cambridge Structural Database. The following X-ray structures are included (ϕ and ψ are given in parentheses): *O*-(4-methyl-α-D-glucopyranosyluronic acid)-(1−2)-*O*-β-D-xylopyranosyl-(1−4)-D-xylopyranose trihydrate (30.0°, 34.9°);[70u] β-D-(1−4)-xylobiose hexaacetate (22.5°, −30.4°).[70v]

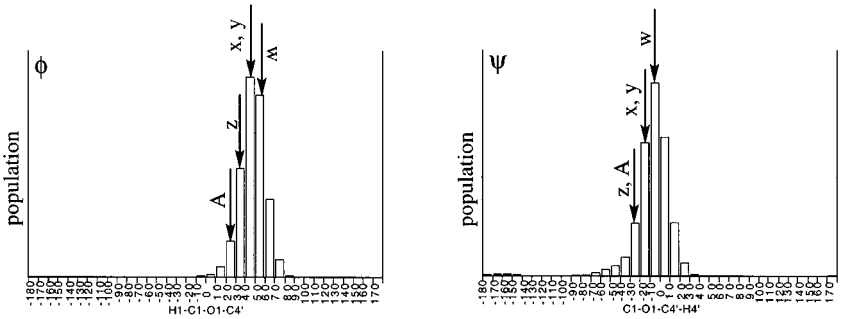


Figure 7. Distribution of the glycosidic (ϕ = H1−C1−O1−C4′) and aglyconic (ψ = C1−O1−C4′−H4′) torsions of galactosyl-β-(1−4)-glucose (**21**) at the end of a 10 ns MC/SD simulation. Arrows point to the positions of ϕ and ψ of selected X-ray structures with the galactosyl-β-(1−4)-glucose core, retrieved from the Cambridge Structural Database. The following X-ray structures are included (ϕ and ψ are given in parentheses): β-lactose (52.0°, −7.6°);[70w] lactose calcium bromide complex heptahydrate (47.8°, −14.8°);[70x] lactose calcium chloride heptahydrate (44.2°, −18.7°);[70y] α-lactose (33.2°, −27.5°);[70z] α-lactose (26.4°, −27.5°).[70aa]

for many naturally occurring glucosidic oligomers such as amylose and the cylodextrins. Experimentally, the solution conformations of cellobiose, maltose, and their methyl glycosides have been studied by NMR techniques[38,43,50,54−56] as well as by optical rotation methods.[57,58] While these studies generally agree that the most populated solution structures and crystalline structures are closely related, the conformational picture of disaccharides in solution is more complicated and is much less clear than one would like. In particular, these (and

(43) (a) Lipkind, G. M.; Shashkov, A. S.; Kochetkov, N. K. *Carbohydr. Res.* **1985**, *141*, 191. (b) Lipkind, G. M.; Verovsky, V. E.; Kochetkov, N. K. *Carbohydr. Res.* **1984**, *133*, 1.
(44) (a) French, A. *Carbohydr. Res.* **1989**, *188*, 206. (b) French A. *Biopolymers* **1988**, 27, 1519.
(45) French, A. F.; Tran, V. H.; Perez, S. In *Computer Modeling of Carbohydrate Molecule;* French, A. D., Brady, J. W., Eds.; ACS Symposium Series 430; American Chemical Society: Washington, DC, 1990.
(46) Rees, D. A.; Smith, P. J. C. *J. Chem. Soc., Perkin Trans. 2* **1975**, 836.
(47) Melberg, S.; Rasmussen, K. *Carbohydr. Res.* **1979**, *69*, 27.
(48) Tvaroska, I. *Biopolymers* **1982**, *21*, 1887.
(49) Perez, S.; Taravel, F.; Vergelati, C. *Nouv. J. Chim.* **1985**, *9*, 561.
(50) Shashkov, A. S.; Lipkind, G. M.; Kochetkov, N. K. *Carbohydr. Res.* **1986**, *147*, 175.
(51) Dowd, M. K.; Zeng, J.; French, A. D.; Reilly, P. J. *Carbohydr. Res.* **1992**, *230*, 223.
(52) (a) Brady, J. W.; Schmidt, R. K. *J. Phys. Chem.* **1993**, *97*, 958. (b) Ha, S. N.; Madsen, L. J.; Brady, J. W. *Biopolymers* **1988**, *27*, 1927.
(53) Tran, V.; Buleon, A.; Imberty, A.; Perez, S. *Biopolymers* **1980**, *28*, 679.
(54) St.-Jacques, M.; Sundararajan, P. R.; Taylor, K. J.; Marchessault, R. H. *J. Am. Chem. Soc.* **1976**, *98*, 4386.
(55) Nardin, R.; Saint-Germain, J.; Vincendon, M.; Taravel, F.; Vignon, M.; *Nouv. J. Chim.* **1984**, *8*, 305.
(56) Casu, B.; Reggiani, M.; Gallo, G. G.; Vigevani, A. *Tetrahedron* **1966**, *22*, 3061.
(57) Rees, D. A.; *J. Chem. Soc. B* **1970**, 877.

other) disaccharides appear to be highly mobile in a conformational sense, and the most diagnostic observables (NOEs and $^3J_{C,H}$ coupling constants) are conformationally averaged. This averaging makes it difficult to define the populations of the various conformational states unambiguously. Thus quantitative comparisons between theory and experiment are problematic. These difficulties notwithstanding, a number of theoretical studies of cellobiose, maltose, and their glycosides have been carried out. In theoretical work on cellobiose, Rasmussen *et al.*[41] identified six minima on the potential energy surface, with predicted statistical weights of 60:34:3:2:1:<1. The first conformer was found to agree well with the X-ray structure of β-cellobiose. These minima were latter used by Tvaroska[42] as input for single-point PCILO calculations using a dielectric continuum reaction field to represent solvent. On the basis of the conformational molar fractions in water at rt, it was concluded that the structures of β-cellobiose are similar in the aqueous and crystalline states. Boltzmann-averaged $^3J_{C,H}$ coupling constants were found to be in reasonable agreement with experiment. Lipkind *et al.*[43] used the nonbonded potential energy function of Scott and Scheraga, together with torsional contributions, to generate the (ϕ, ψ) conformational map of methyl α-cellobioside. Electrostatic, hydrogen-bonding, and *exo*-anomeric effects were not utilized in the calculations which found four minima with predicted statistical weights of 40:44:6:10. Boltzmann-averaged observables (NOEs, J_ϕ, J_ψ) calculated from this conformational distribution were found to be in reasonable agreement with experiment. In contrast, the single conformation obtained from HSEA calculations could not account for the experimental findings. Relaxed and nonrelaxed conformational maps for cellobiose were also generated by French,[44] by Perez *et al.*,[45] and by Kouwizjer *et al.*,[8b] using energy minimization, and by Rees *et al.*,[46] using Metropolis Monte Carlo (MMC) simulations. These different force fields and computational procedures generally led to similar conformational maps, and experimentally observed X-ray and solution structures of related compounds were found to cluster around their low-energy regions.

Using methods employed in their cellobiose studies, Rasmussen *et al.*[47] also studied the conformational behavior of β-maltose. These studies found four minima with predicted statistical weights of 59:19:15:7. The most populated conformer was found to be structurally similar to the X-ray structure of methyl β-maltoside. These minima were again used by Tvaroska[48] as input for single-point PCILO calculations in continuum water. Boltzmann-averaged $^3J_{C,H}$ coupling constants (also calculated by Perez *et al.*)[49] were found to be in good agreement with experiment. Lipkind *et al.*[50] extended their computational approach, originally reported for methyl α-cellobioside to methyl β-maltoside. These workers found four minima on their potential surface with predicted statistical weights of 40:45:12:3. Boltzmann-averaged NMR data (NOEs, J_ϕ, J_ψ) were in reasonable agreement with experiment, but no single conformer adequately accounted for the experimental data. Relaxed and nonrelaxed conformational maps for α- and/or β-maltose were generated by Merz *et al.*[7c] (modified AMBER), French[44] (MM2), French *et al.*[51] (MM3), Brady *et al.*[52] (CHARMM), Perez *et al.*[53] (MM2CARB), Kouwizjer *et al.*[8b] (CHEAT95), and Rees *et al.*[46] (MMC simulations). As in the cellobiose studies, these force fields produced similar conformational maps whose low-energy regions included X-ray structures of maltose derivatives. The calculations of Brady *et al.*[52] identified five minima for β-maltose. These minima served as starting points for several MD simulations *in vacuo* and in TIP3P water. Interconversions between the different conformational energy wells were not observed during the simulations. In vacuum, maltose was found to adopt a conformation having multiple intramolecular hydrogen bonds. In water, conformational fluctuations were found to be damped and most intramolecular hydrogen bonds were replaced by hydrogen bonds to the solvent.

Table 3. Experimental and Calculated $^3J_{C,H}$ Coupling Constants (Hz) across the Glycosidic (J_ϕ = H1−C1−O1−C4′) and Aglyconic ($J\psi$ = C1−O1−C′4−H4′) Linkages for Selected Disaccharides

	J_ϕ		$J\psi$	
	exptl	calcd	exptl	calcd
methyl β-cellobioside (**17**)	4.2[38]	2.7	4.3[38]	5.0
methyl β-maltoside (**18**)	3.5[50]	3.6	3.9[50]	4.1
methyl β-xylobioside (**20**)	4.7[68]	2.7	5.1[68]	4.4
β-lactose (**21**)	4.0[69]	2.7	5.1[69]	5.0

Table 4. Relative NOEs (%) (upon Irradiation of H1) and r^{-6} Averaged H1−Hx distances (Å^{-6}) for Disaccharide Derivatives

	relative NOE	$\langle r^{-6}\rangle$
α-Cellobiose 1-Phosphate[43a]		
H3 + H5	7.5	0.0095 (0.0054 + 0.0041)[a]
H4′	4.4	0.0057
H3′ + H6′	2.1	0.0031 (0.0002 + 0.0016 + 0.0013)[a]
H5′	1.0	0.0002
Methyl β-Maltoside (18)[50]		
H2	7.9	0.0058
H4′	5.9	0.0053
H3′	1.0	0.0005
Mannosyl-(α-1,3)-mannose-β-OMe (19)[62]		
H2′	1.8	0.0008
H5	1.0	0.0004
H4′	0.7	0.0003

[a] Values in parentheses are $\langle r^{-6}\rangle$ averages for individual proton−proton pairs.

The results of our 10 ns, 27 °C MC/SD simulations of aqueous methyl β-cellobioside (**17**) and methyl β-maltoside (**18**) with our new AMBER* force field are summarized in ϕ, ψ histogram plots (Figures 3 and 4) and in the first two entries of Tables 3 and 4. In contrast with energy minimization studies that yield discrete conformers, our rt simulations find **17** and **18** to be highly flexible and to exist in a continuum of conformational states that merge into a single, but very broad conformational well—at least so far as the ϕ and ψ interresidue angles are concerned. In both cases, the glycosidic torsion angles (ϕ) are all in the *exo*-anomeric conformational well, but the broad ϕ distributions indicate high conformational flexibility. Comparable flexibility is observed around the aglyconic torsion angles (ψ) which cluster around the *syn* conformation. There is also a small population (~1%) of the *anti* conformation for cellobioside **17**, in accord with the results of Lipkind *et al.*[43] The ϕ and ψ angles taken from X-ray structures of several di- and trisaccharides with cellobiose and maltose cores (Cambridge Structural Database, CSD[59]) are also indicated on the histogram plots. In all cases, their inter-

(58) Stevens, E. S.; Sathyanarayana, B. K. *J. Am. Chem. Soc.* **1989**, *111*, 4149.

residue torsion angles are found in populated regions of our ϕ, ψ histogram plots, though not always in the most heavily populated conformational regions.

A comparison between the experimental and calculated time-averaged $^3J_{C,H}$ coupling constants of **17** and **18** is provided in Table 3. Since the Karplus-type equation used to compute coupling constants from torsion angles is accurate to only ±1 Hz,[36] the calculations of J_ψ fall within the expected error limits of the experimental measurements. Similar agreement is found for J_ϕ of methyl β-maltoside; however, J_ϕ for methyl β-cellobioside is smaller than that from experiment by 1.5 Hz (*i.e.*, 0.5 Hz outside the expected error limits). This overall level of agreement between experimental and calculated interresidue coupling constants is comparable with results from the previous calculations described above. Table 4 shows a comparison of observed NOEs and calculated proton−proton r^{-6} distance averages ($\langle r^{-6}\rangle$) involving the anomeric hydrogen (H1) of cellobiose and maltose derivatives. In the case of cellobiose, the experimental measurements were made on α-1-phosphate while our simulations were on the β-1-methyl glycoside (**17**). Nevertheless, we find a good correlation between the relative strength of the NOEs and the $\langle r^{-6}\rangle$ average for the relevant hydrogens with both **17** and **18**. In each case, the hydrogen pairs having the strongest NOEs are found by the simulation to have the largest $\langle r^{-6}\rangle$.

Mannosyl-(α-1,3)-mannose-β-OMe (**19**) was taken as a model for the mannosyl-(α-1,3)-mannose linkage. This disaccharide and its α- and β-OMe analogs have been studied extensively both experimentally and theoretically due to the ubiquity of this system in glycoproteins. Homans *et al.*[60] used MNDO calculations to generate the conformational map for the mannosyl-(α-1,3)-mannose linkage. A single minimum was located at $\phi = -50°$; $\psi = 10°$, in good agreement with the experimentally deduced solution conformation of this moiety. Additional conformational maps for mannosyl-(α-1,3)-mannose were generated by Imberty *et al.*[61] (MM2CARB), Homans[5] (reparameterized AMBER), and Carver *et al.*[62] (MM2-85). Boltzmann-averaged NOEs were calculated either directly from the resulting maps (MM2CARB,[61] MM2-85[62]) or from subsequent short MD simulations (AMBER,[5] MM2-85[62]), leading to a reasonable agreement with experiment. Peter *et al.*[63] used MMC simulations and the HSEA force field to calculate NOEs for mannosyl-(α-1,3)-mannose-α-OMe. The results were compared with those obtained from Boltzmann-averaged grid search/energy minimization methods. As expected, the conformationally averaged NOEs from either MMC simulations or grid search/energy minimizations gave a better fit to experiment than did corresponding data derived from any single minimum energy conformation. Finally, Bernardi *et al.*[64] used Homans' carbohydrate parameters, the GB/SA continuum model for water, and the MC/SD simulation technique in 2 ns simulations of mannosyl-(α-1,3)-mannose-α-OMe and mannosyl-(α-1,3)-mannose-β-OMe. Key proton−proton distances were monitored and used to calculate relative NOEs which showed moderate agreement with experiment. The experimental values for the interresidue glycosidic and aglyconic torsion angles of the mannosyl-(α-1,3)-mannose linkage in water were estimated from NOE experiments on **19**, its α-OMe analog, and appropriate polysaccharides to be $-45 \pm 15°$ and $-15 \pm 15°$, respectively.[65−67]

The ϕ and ψ population histograms from our 10 ns, 27 °C MC/SD simulations of **19** are given in Figure 5. As with **17** and **18**, the conformational behavior of **19** can be described as being defined by a single but broad energy well with the interresidue ϕ and ψ distributions centered around the *exo*-anomeric and *syn* geometries. Both the solution conformation ($\phi = -45 \pm 15°$, $\psi = -15 \pm 15°$)[65] and the crystal conformation ($\phi = -57.6°$, $\psi = -19.4°$) of the mannosyl-(α-1,3)-mannose linkage are found in highly populated regions on the corresponding simulation histograms. Disaccharide **19** has also been studied by NMR, and several conformationally diagnostic NOE signals have been reported.[62] Again, we find a good correlation between the relative NOE intensities and the $\langle r^{-6}\rangle$ averages from our simulations (see Table 4, third entry).

The solution conformation of methyl β-xylobiose (**20**) has been studied by NMR. J_ϕ and J_ψ values were measured, and their solvent and temperature dependencies were interpreted in term of high conformational flexibility around the interresidue glycosidic and aglyconic linkages.[68] PCILO calculations were used to generate a two-dimensional conformational map for **20**. Nine minima were located and used as starting points for energy minimization using PCILO, MM2, and MM2CARB. Solvent effects were introduced by means of the continuum reaction field. All methods identified seven minima with similar geometries though their relative energies varied. The PCILO results were used to calculate Boltzmann-averaged coupling constants which were found to be in reasonable agreement with experiment. The results of our 10 ns, 27 °C MC/SD simulation for **20** are summarized in Figure 6 and Table 3. A broad angle distribution is again found for the glycosidic ϕ torsion and an even broader one for the aglyconic ψ torsion. Only two structures with the appropriate xylobiose core were found in the CSD, both suffering from steric hindrance which may have distorted their conformations relative to the parent. Nevertheless, the ϕ and ψ angles of both structures correspond to significantly populated regions of our ϕ, ψ histograms. Our calculated coupling constants are compared with the experimentally measured ones in Table 3, entry 3. As in the case of cellobioside **17**, J_ψ is well-reproduced by the simulation while J_ϕ is again too small.

We also carried out related free energy simulations of β-lactose (**21**).[69] Its ϕ, ψ histograms are shown in Figure 7 and are similar to those of the other disaccharides we studied. X-ray crystal structures of related oligosaccha-

(59) Allen, F. H.; Bellard, S.; Brice, M. D.; Cartwright, B. A.; Doubleday, A.; Higgs, H.; Hummelink, T.; Hummelink-Peters, B. G.; Kennard, O.; Motherwell, W. D. S.; Rodgers, J. R.; Watson, D. G. *Acta Crystallogr.* **1979**, *B35*, 2331.
(60) Homans, S. W.; Pastore, A.; Dwek, R. A.; Rademacher, T. W. *Biochemistry* **1987**, *26*, 6649.
(61) Imberty, A.; Tran, V.; Perez, S. *J. Comput. Chem.* **1989**, *11*, 205.
(62) Carver, J. P.; Mandel, D.; Michnick, S. W.; Imberty, A.; Brady, J. W. In *Computer Modeling of Carbohydrate Molecules;* French, A. D., Brady, J. W., Eds.; ACS Symposium Series 430, American Chemical Society: Washington, DC, 1990.
(63) Peter, T.; Meyer, B.; Stuike-Prill, R.; Somorjai, R.; Brisson, J-R. *Carbohydr. Res.* **1993**, *238*, 49.
(64) Bernardi, A.; Raimondi, L.; Zanferrari, D. THEOCHEM, in press.
(65) Brisson, J. R.; Carver, J. P. *Biochemistry* **1983**, *22*, 1362.
(66) Brisson, J. R.; Carver, J. P. *Biochemistry* **1983**, *22*, 3671.
(67) Homans, S. W.; Dwek, R. A.; Rademacher, T. W. *Biochemistry* **1987**, *26*, 6553.
(68) Hricovíni, M.; Tvaroska, I.; Hirch, J. *Carbohydr. Res.* **1990**, *198*, 193.
(69) Poppe, L.; Van Halbeek, H. *J. Magn. Reson.* **1991**, *93*, 214.

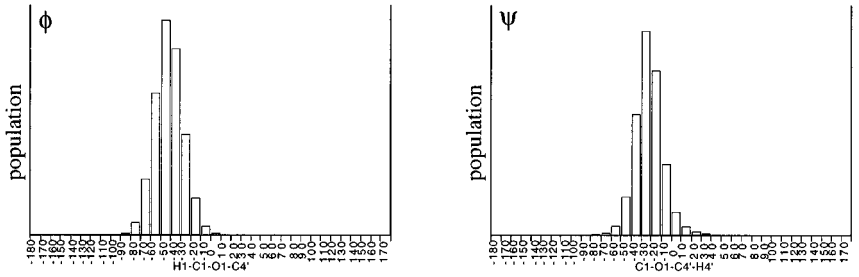


Figure 8. Distribution of the glycosidic (ϕ = H1−C1−O1−C4′) and aglyconic (ψ = C1−O1−C4′−H4′) torsions of methyl β-maltoside (**18**) as calculated with the mixed united-atom/all-atom reparameterized AMBER* force fields.

rides from the CSD are again found to occupy heavily populated regions of our ϕ, ψ histograms. Finally, we obtain roughly the same level of agreement between experimental and calculated interresidue NMR coupling constants as described above with the other disaccharides (see Table 3, entry 4).

While our all-atom carbohydrate force field performs well for both mono- and disaccharides, calculations with it are slower than those with our previous united-atom field by roughly a factor of 2. We therefore carried out one additional study to investigate a mixed representation in which only those atoms having close contacts with atoms in other residues were represented using the all-atom model. For oligosaccharides having a 1,4-linkage, the all-atom representation would thus be used for C1, C2, C3′, C4′, C5′. For systems having a 1,3-linkage, the all-atom model would be used for C1, C2, C2′, C3′, C4′. Other carbon atoms, *e.g.*, those remote from interresidue linkages, would be included in the computationally more efficient united-atom representation. To test such a mixed representation, we carried out a 10 ns, 27 °C MC/SD simulation of methyl β-maltoside (**18**) in GB/SA water. The ϕ, ψ histograms generated by this simulation are given in Figure 8. They compare well with the corresponding all-atom histograms for the same molecule given in Figure 4. The two representations of **18** also predict similar interresidue coupling constants ($J_{\phi(\text{all-atom})}$ = 3.6 Hz, $J_{\phi(\text{mixed})}$ = 3.1 Hz; $J_{\psi(\text{all-atom})}$ = 4.1 Hz, $J_{\psi(\text{mixed})}$ = 4.5 Hz). Thus, both force fields yield similar results with **18**, but the mixed representation required ~70% of the time needed for the all-atom calculation. For larger molecular systems, the speed advantage of the mixed representation is expected to be more substantial.

Conclusion

The work described above shows that *in vacuo*, *ab initio* quantum calculations can be used to produce a molecular mechanics force field that, when coupled with a solvation model for water and an efficient free energy simulation algorithm, gives reasonable conformational free energies and geometries for simple sugars and oligosaccharides. The most significant difficulty we encountered was validating the performance of the force field—especially with disaccharides. The problem is that there is virtually no purely experimental measurement of a conformational free energy difference with an oligosaccharide. Thus there is little reliable data against which to compare a calculation other than X-ray crystal structures and conformationally averaged spectroscopic data such as NMR coupling constants and NOE signals. While conformational free energy information is contained in available conformationally averaged spectroscopic data, many assumptions are necessary in order to extract it. Furthermore, most disaccharides in aqueous solution appear to exist in single, but highly flexible, conformational families without significant energetic barriers between various contributing forms. Taken together, these features make it difficult to obtain reliable conformational energy differences experimentally. And until such data is available, force field calculations on carbohydrates must be considered to have only qualitative accuracy.

Current experimental limitations notwithstanding, what experimental data we have is well-reproduced by our free energy simulations with the new force field. For monosaccharides, anomeric free energies for nine solvated systems were calculated with an average unsigned error of 0.32 kcal/mol. For disaccharides, the distribution histograms for the glycosidic and aglyconic torsion angles showed the expected conformational preference and flexibility. Furthermore, all relevant X-ray structures fell within populated regions of these ϕ, ψ histograms, and experimental and calculated NMR coupling constants were in qualitative agreement. We also found a strong, semiquantitative correlation between experimental relative NOE signals and r^{-6} averaged proton−proton distances computed from our simulations.

The conformational picture of 1,3- and 1,4-linked disaccharides we have found suggests them to spend the vast majority of their time oscillating around a single, but broad, conformational energy well that corresponds to a *syn*, *exo*-anomeric geometry in aqueous solution at

(70) (a) Leung, F.; Chanzy, H. D.; Perez, S.; Marchessault, R. H. *Can. J. Chem.* **1976**, *54*, 1365. (b) Taga, T.; Sumiya, S.; Osaki, K.; Utamura, T.; Koizumi, K. *Acta Crystallogr.* **1981**, *B37*, 963. (c) Jacobson, R. A.; Wunderlich, J. A.; Lipscomb, W. N. *Acta Crystallogr.* **1961**, *14*, 598. (d) Brown, C. J. *J. Chem. Soc. A* **1966**, 927. (e) Chu, S. S. C.; Jeffrey, G. A. *Acta Crystallogr.* **1968**, *B24*, 830. (f) Nikitin, A. V.; Andrianov, V. I.; Myasnikova, R. M.; Firgang, S. I.; Usov, A. I.; Sopin, V. F.; Pertsin, A. I. *Kristallografiya* **1986**, *31*, 676. (g) Nikitin, A. V.; Shibanova, T. A.; Firgang, S. I.; Usov, A. I.; Sopin, V. F.; Myasnikova, R. M. *Kristallografiya* **1987**, *32*, 896. (h) Ham, J. T.; Williams, D. G.; *Acta Crystallogr.* **1970**, *B26*, 1373. (i) Pangborn, W.; Longs, D.; Perez, S. *Int. J. Biol. Macromol.* **1985**, *7*, 363. (j) Taga, T.; Inagaki, E.; Fugimori, Y.; Nakamura, S. *Carbohydr. Res.* **1993**, *240*, 39. (k) Taga, T.; Inagaki, E.; Fugimori, Y.; Nakamura, S. *Carbohydr. Res.* **1994**, *251*, 203. (l) Jeffrey, G. A.; Huang, D. *Carbohydr. Res.* **1991**, *222*, 47. (m) Quigley, G. J.; Sarko, A.; Marchessault, R. H. *J. Am. Chem. Soc.* **1970**, *92*, 5834. (n) Gress, M. E.; Jeffrey, G. A. *Acta Crystallogr.* **1977**, *B33*, 2490. (o) Takusagawa, F.; Jacobson, R. A. *Acta Crystallogr.* **1978**, *B34*, 213. (p) Chu, S. S. C.; Jeffrey, G. A. *Acta Crystallogr.* **1967**, *23*, 1038. (q) Tanaka, I.; Ashida, T.; Kakudo, M. *Acta Crystallogr.* **1976**, *B32*, 155. (r) Imberty, A.; Perez, S. *Carbohydr. Res.* **1988**, *181*, 41. (s) Brisse, F.; Marchessault, R. H.; Perez, S.; Zugenmaier, P. *J. Am. Chem. Soc.* **1982**, *104*, 7470. (t) Warin, V.; Baert, F.; Fouret, R.; Strecker, G.; Spik, G.; Fournet, B.; Montreuil, J. *Carbohydr. Res.* **1979**, *76*, 11. (u) Moran, R. A.; Richards, G. F. *Acta Crystallogr.* **1973**, *B29*, 2770. (v) Leung, F.; Marchessault, R. H. *Can. J. Chem.* **1973**, *51*, 1215. (w) Hirotsu, K.; Shimada, A. *Bull. Chem. Soc. Jpn.* **1974**, *47*, 1872. (x) Bugg, C. E. *J. Am. Chem. Soc.* **1973**, *95*, 908. (y) Cook, W. J.; Bugg, C. E. *Acta Crystallogr.* **1973**, *B29*, 907. (z) Noordik, J. H.; Beurskens, P. T.; Bennema, P.; Visser, R. A.; Gould, R. O. *Z. Kristallogr.* **1984**, *168*, 59. (aa) Fries, D. C.; Rao, S. T.; Sundaralingam, M. *Acta Crystallogr.* **1971**, *B27*, 994.

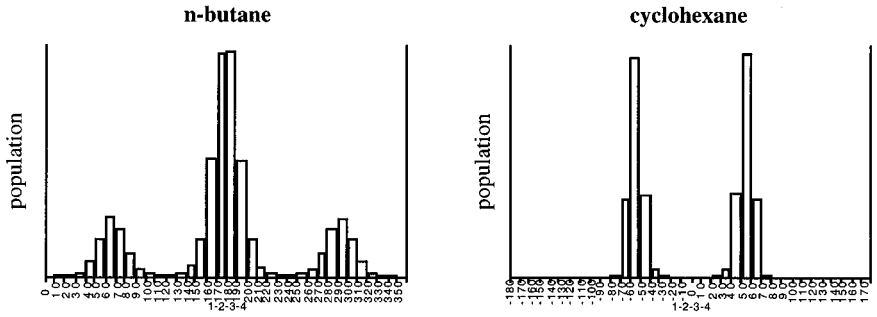


Figure 9. Distribution of C−C−C−C torsion angles for *n*-butane and cyclohexane at rt on the MM3 potential surface.

rt. Small populations of ψ *anti* conformers were observed in a few of our simulations, but they correspond to free energies greater than 2 kcal/mol above the more heavily populated *syn* conformers with our force field. The torsional ϕ, ψ oscillations in disaccharides are generally greater than the C−C−C−C torsional variations of simple saturated hydrocarbons such as *n*-butane and cyclohexane at rt (*cf.* Figure 9). However, unlike our disaccharides, these alkanes (and most other organic molecules) exist in multiple, stable conformational states (defined as *conformers*), and this property, the presence of multiple, energy barrier separated conformers, has been used traditionally to define molecular flexibility. Thus while it would be wrong to describe oligosaccharides as rigid, they do populate essentially a single energy well and must therefore be regarded as less flexible than typical organic molecules such as *n*-alkanes and peptides.

While quantitative experimental data on carbohydrate conformational energies are sorely needed, we believe that the parameters and methodologies described here have been validated as well as possible with available data and with a wide range of systems including five disaccharides. In every system we studied, we found good correlation between experiment and our free energy simulation derived results. Thus these methods and the force field described here should find useful applications in semiquantitative modeling of carbohydrate-based molecules.

Acknowledgment. This work was supported by grants from the National Science Foundation (CHE95 44243), the Kanagawa Academy of Science and Technology, and the National Institutes of Health (P41-RR06892).

Supporting Information Available: Atomic partial charges and new torsional parameters for **1**−**9** and the glycosidic and aglyconic linkages and a comparison between *ab initio* and AMBER* rotational profiles for model compounds 2-methoxyethanol, 2-propanol, **1**, **2**, and **12** (12 pages). This material is contained in libraries on microfiche, immediately follows this article in the microfilm version of the journal, and can be ordered from the ACS; see any current masthead page for ordering information.

JO9612483