Modeling the α(1→6) Branch Point of Amylopectin in Solution

Robert B. Best, Graham E. Jackson, and Kevin J. Naidoo*

Department of Chemistry, University of Cape Town, Rondebosch 7701, South Africa

Received: October 10, 2001; In Final Form: February 10, 2002

We have investigated the structure and dynamics of the branching motif of amylopectin, the major component of starch. The trisaccharide panose and the tetrasaccharide 6^2 α-D-glucosylmaltotriose have been chosen as minimal model compounds for the α(1→6) branch point and molecular dynamics (MD) simulations of these have been run with explicit water. Calculation of NMR longitudinal relaxation times for panose shows good agreement with experimental values, thereby validating our simulation dynamics. Compared with the disaccharides maltose and isomaltose the addition of extra residues appears to provide more interresidue interactions, causing the branch models to explore a wider range of glycosidic conformational space, particularly for the α(1→6) linkage; nonetheless, there remain two distinct regions separated by a high-energy barrier. The correlation time for dynamics about the α(1→6) linkage torsion angles depends strongly on which of these regions is being sampled. We calculate the diffusion constants for water as a function of distance from the solute: the calculations show that these regions of structured water close to the solute diffuse more slowly than bulk water. We then show that the solvated saccharides are surrounded by a layer of structured, spatially localized water by comparing the diffusion of the solvent with its configurational probability distribution. Finally, the implications of the model compound properties for the full structure of amylopectin were investigated by calculating geometric properties of the branch point.

Introduction

Carbohydrates are responsible for a wide variety of biological functions such as cellular recognition and signaling processes. Some of the simplest carbohydrate polymers are found in starch, an important form of food storage in plants. Although the morphology and size of starch granules varies with source, all starches are composed of the polymers amylose and amylopectin in a ratio of approximately 30:70 by mass.[1] Amylose is a linear chain of α(1→4) linked glucose monomers, consisting of between 600 and 6000 residues. In amylopectin, these α(1→4) linked chains are joined by α(1→6) branch points with an average spacing of 20–30 residues and an overall degree of polymerization of 10^5. Diffraction methods such as SAXS and X-ray microfocus diffraction have led to a model for the starch granule in which there are alternate crystalline and amorphous layers; the crystalline layers are attributed to the packing of amylopectin single and double helices, while the α(1→6) branching regions are responsible for the amorphous regions.[2] Amylose is believed to be located mainly in the amorphous lamellae.[3]

The application of starch as a biodegradable material has attracted much interest, and it is already commercially available for simple packaging. More advanced materials applications require a better understanding of the molecular origin of its properties in order to design the modifications necessary for its practical use. The influence of water is particularly important, as polysaccharides are strongly solvated: the presence of water causes significant shifts in the glass transition temperature observed for both amylose and amylopectin.[4] It is also well-known that amylopectin is generally more soluble than amylose,[5]

which may be due to the greater flexibility of the branched polysaccharide.

Great advances in carbohydrate structure determination have been made in recent years thanks to experimental methods such as X-ray diffraction[6–9] and nuclear magnetic resonance (NMR)[10–12] and optical spectroscopies.[13–16] These methods each have their limitations: in the case of crystallography, it is the limited size of carbohydrates that can be crystallized, and for NMR the lack of sufficient structural nOe and dihedral coupling constraints, although new methods based on residual dipolar couplings hold much promise.[17–19] Furthermore, these methods provide limited direct structural information about amorphous structures such as that of starch. Computational approaches employing empirical force fields and molecular dynamics simulations have thus become a complementary tool for the investigation of carbohydrate structure and dynamics.[20–23] They are able to provide the molecular detail that can then be checked against experimental data.[15,24]

The conformational flexibility of glucopyranosyl carbohydrates is determined mainly by the linkages between residues, as the sugar rings themselves are virtually always found in the $^4C_1$ chair conformation.[25,26] Since starch is built entirely from α(1→4)- and α(1→6)-linked glucose residues, the simplest models are disaccharides containing these linkages alone. We have previously described the solution structure and dynamics of the α(1→4) linked disaccharide maltose and the α(1→6) linked isomaltose[22] and found the α(1→6) linkage to be the more flexible in water; in addition the water-structuring properties of amylose have been studied by molecular dynamics of a maltohexaose fragment.[27] In this paper, we extend the study to the branch points of amylopectin by investigating the solvent dynamics of tri- and tetrasaccharide models, panose and 6^2 α-D-glucosylmaltotriose (6GMT). A vacuum-based conformational analysis has previously been performed on these models of the

* To whom correspondence should be addressed. E-mail: knaidoo@science.uct.ac.za Fax: +27-21-689-7499.
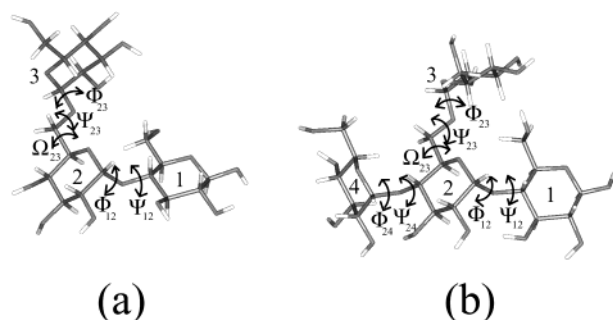
Figure 1. Schematic diagrams of the three residue model (panose) (a) and the four residue model (b) used to represent an amylopectin branch point. Residue numbering and torsion angle names are illustrated.

amylopectin branch point.[28,29] Here we show that the addition of residues on either side of the branch point modifies the conformational bias of each of the linkages over those of the disaccharides alone. As a check on the validity of the simulation dynamics, NMR T_1 relaxation times, which are sensitive to both the extent and time scale of molecular motion, have been directly calculated from the trajectories for comparison with experimental data. The effect of the branch model on the solvent structure was investigated using spatial distribution functions; the solvent dynamics were analyzed using a spatially resolved map of the solvent diffusion coefficients about the solute. We show that there is a layer of structured water immediately surrounding the carbohydrate, which is spatially localized (i.e., has a low diffusion coefficient). Finally, the branch geometry was analyzed by two methods in order to make predictions for the complete amylopectin structure.

Computational Methods

The model compounds used in this study are shown in Figure 1 together with the residue numbering and glycosidic torsion angle names. Atom names follow the convention for glucose, prefixed by the residue number. The torsion angles are defined by Φ_{xy} (xH1−xC1−xO1−yC4) and Ψ_{xy} (xC1−xO1−yC4−yH4) for the $\alpha(1\rightarrow4)$ linkages and by Φ_{xy} (xO5−xC1−xO1−yC6), Ψ_{xy} (xC1−xO1−yC6−yC5), and Ω (xO1−yC6−yC5−yC4) for the $\alpha(1\rightarrow6)$ linkages (where x and y are the residues on either side of the linkage), following previous convention.[22] The orientation of the hydroxymethyl groups is given by Θ (xC4−xC5−xC6−yO6), the gauche conformers being denoted by the standard notation gg, gt, and tg (corresponding to angles of Θ of −60°, 60°, and 180°, respectively). The trisaccharide panose is the minimal model for a branch point, while the tetrasaccharide 6^2 α-D-glucosylmaltotriose (6GMT) includes an additional glucose unit on the other side of the branch residue.

Simulation Conditions. All MD simulations were carried out using CHARMM[30] with a force-field specifically parametrized for carbohydrates.[31] The TIP3P water model[32,33] was used to represent the solvent explicitly. Initial conformations for each carbohydrate were chosen from the minima in the adiabatic maps of maltose and isomaltose.[22] Since the $\alpha(1\rightarrow6)$ linkage is known to have two major minima separated by a significant energy barrier in Ω, two initial conformations were chosen for each oligosaccharide: one from the previously described A_{iso} well ($\Phi = 80°$, $\Psi = 180°$, $\Omega = 180°$) and the other from the B_{iso} well ($\Phi = 80°$, $\Psi = 180°$, $\Omega = 60°$) of isomaltose: these have been denoted P−A and P−B for panose and G−A and G−B for 6^2 α-D-glucosylmaltotriose. In all cases, the $\alpha(1\rightarrow4)$ dihedrals were initially set to the values $\Phi = -50°$, $\Psi = -40°$, the lowest energy A_{malt} well of maltose. These

conformations were solvated in a truncated octahedral water box size of initial size 37.00 Å (distance between nearest faces) containing 846 water molecules. Waters overlapping with the solute were deleted, and the box resized to a density of 0.997 g/cm³, the experimental value for water at 298 K and 1 atm:[34] this is justified, since we are modeling a dilute oligosaccharide solution. After this adjustment, 810 and 816 waters were left in the P−A and P−B simulations and 805 and 810 in the G−A and G−B simulations, respectively.

All systems were heated from 100 to 300 K over 20 ps with constraints on the glycosidic dihedrals and equilibrated for a further 50 ps using velocity assignment to control temperature. Once the systems were equilibrated, 2 ns of dynamics was run in the microcanonical (N, V, E) ensemble using the leapfrog dynamics algorithm to integrate the equations of motion.[35] This ensemble does not require the rescaling of velocities to maintain temperature and, so, is more suited to the calculation of long time scale correlation functions necessary for comparison with experimental NMR data. The temperature was found to fluctuate stably about 300 K throughout the course of the simulations.

Theory for NMR Relaxation Data Analysis. NMR relaxation data for oligosaccharides are frequently interpreted using the model-free formalism of Lipari and Szabo.[36,37] However, this model is most suitable for the application to molecules where the rate of intramolecular motion is much more rapid than molecular tumbling, which is not the case for small oligosaccharides, complicating the analysis;[38] furthermore, the comparison between experiment and simulation is through indirectly fitted parameters. Here we adopt the approach we have used previously to calculate T_1 relaxation times directly from molecular dynamics simulation.[22]

Longitudinal T_1 relaxation times of natural abundance carbon-13 nuclei are determined from the dipolar interaction with directly attached protons, dependent on the motion of the internuclear r_{CH} vector. This can be expressed in terms of the spectral density of this motion by

$$T_1^{-1} = \frac{1}{4}D^2[J(\omega_H - \omega_C) + 3J(\omega_C) + 6J(\omega_H + \omega_C)] \quad (1)$$

where $D = \eta\gamma_C\gamma_H\langle r^3\rangle^{-1}$.

Having access to the full molecular dynamics trajectory allows evaluation of the appropriate correlation function (eq 2), whose Fourier transform (eq 3) is the desired spectral density.

$$C(t) = \frac{1}{5}\langle P_2(\mu_{LF}(0)\cdot\mu_{LF}(t))\rangle \quad (2)$$

$$J(\omega) = \int_0^\infty C(t) \cos \omega t \, dt \quad (3)$$

In practice the correlation function in (1) obtained from MD simulation is too short to accurately represent the contribution from molecular tumbling. Therefore $C(t)$ is factorized into a part dependent on the overall tumbling, represented by the exponential decay $C_O(t) = \exp(-t/\tau_M)$ and an intramolecular part $C_I(t)$ calculated from eq 2, but with the molecule reoriented to remove overall rotation. The value of τ_M, the correlation time for overall rotation, is determined from a vector considered to be unaffected by intramolecular motion. The resultant correlation function is given by eq 4; t_{max} is the length of the correlation function from molecular dynamics, which is used to evaluate the intramolecular correlation function.

α(1→6) Branch Point of Amylopectin

J. Phys. Chem. B, Vol. 106, No. 19, 2002 **5093**

$$C(t) = \begin{cases} C_I(t) \exp(-t/\tau_M) & \text{for } t \leq t_{max} \\ C_I(t_{max}) \exp(-t/\tau_M) & \text{for } t > t_{max} \end{cases} \quad (4)$$

Water Structure and Diffusion Methods. The three-dimensional water distribution, or spatial distribution function (SDF) about the oligosaccharides was calculated by a method similar to that used previously.[27,39] However, instead of using all frames of the trajectory, conformations of the solute were clustered using dihedral angles and the main clusters used for solvent density analysis. This was necessary to avoid meaningless averaging of the solvent density over different conformations. The volume around the solute was divided into a three-dimensional Cartesian grid with the width of each division about 0.5 Å. Each set of coordinates from the trajectory was translated to the solute center of mass and rotated by least-squares fitting to a reference structure.[40,41] The value of a normalized Gaussian function centered on each water oxygen was used for binning, to achieve a smoother density. Finally, the total oxygen density was normalized by dividing by the number of frames binned and the expected oxygen density per bin for a uniform distribution, such that a density of 1.0 corresponds to bulk density. Contours of 1.5 times larger than bulk water probability density were used to illustrate a solvent configurational probability about the solute that is 50% greater than that observed in the pure solvent.

Diffusion constants around each residue of the carbohydrates were estimated from the mean squared displacement (MSD) correlation function (eq 5) using the appropriate Einstein relation

$$C_{msd}(t) = \langle |r_i(t) - r_i(0)|^2 \rangle \quad (5)$$

$$2tD = {}^1/_3 C_{msd}(t) \quad (6)$$

(eq 6).[35] The correlation function was averaged over all water molecules within a certain radius of a secondary alcohol and the value of D was estimated by a linear fit the last three-quarters of the data using eq 6.

This analysis of the diffusion constant was extended to a spatially resolved diffusion map, analogous to the extension of the radially averaged pair distribution function to the spatial distribution function. Each frame of the trajectory was reoriented with respect to a set of reference coordinates, following which each water molecule was assigned to a bin in a three-dimensional grid about the reference molecule. The correlation function (eq 5) was calculated for each water over a trajectory length of 1 ps from binned data, yielding a 3D set of correlation functions. After appropriate normalization, a diffusion constant was estimated for each bin by fitting the last three-quarters of the correlation function. The time over which the correlation function was collected was chosen as a tradeoff between reasonable convergence and the need to localize the calculated diffusion coefficients. A Gaussian smoothing function similar to that used for the SDF's was used to smooth the raw diffusion data.

Methods Used To Analyze Branch Dynamics. We are interested in the available orientations of the branching α(1→6) linkage relative to the main α(1→4) linked backbone. Two approaches have been adopted: first, a simple measurement of the angle between 1 C1 and 3 C4 from the center of geometry of ring 2 in each molecule. A more precise description of the motional extent was made by setting up a set of Cartesian axes, which are fixed relative to the branch residue and measuring the orientation of the α(1→6)-linked glucosyl residue within this coordinate frame. Since the distance between residues varies
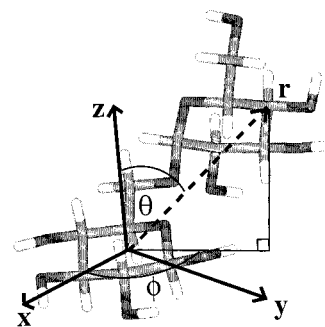
Figure 2. Definition of the spherical polar axis system used to describe the orientation of the α(1→6) linked branch residue relative to the central glucosyl ring.

relatively little, we have used a spherical polar coordinate system to describe the orientation of the branching residue, in which the coordinates θ and ϕ account for most of the motion. The motion of the C4 carbon of the α(1→6) linked residue relative to the central ring has been used to describe the branch geometry; the spherical polar coordinates are defined in Figure 2. The distance measured in this coordinate frame is quite constrained by the molecular geometry, so most of the variation is described by the spherical polar angles (θ, ϕ). The axes are defined about the branching ring (e.g., ring 2 in panose), using the convention established by Cremer and Pople in their description of ring puckering.[42] Denoting the coordinates of atom i by $\mathbf{R}_i$, the origin is placed at the center of geometry of the ring atoms (C1−C5 and O5) and the positive z-axis is uniquely defined by eq 7, with $\mathbf{R}'$ and $\mathbf{R}''$ as given by eq 8. The y-axis

$$z = \mathbf{R}' \times \mathbf{R}''/|\mathbf{R}' \times \mathbf{R}''| \quad (7)$$

$$\mathbf{R}' = \sum_{j=1}^{6} R_j \sin(2\pi(j-1)/6) \qquad \mathbf{R}'' = \sum_{j=1}^{6} R_j \cos(2\pi(j-1)/6) \quad (8)$$

is given by the projection of $\mathbf{R}_1$ onto the $x-y$ plane, from which the x-axis is defined as $x = y \times z$, to give a right-handed system. The motion of the α(1→6)-linked residue 3 is described by the motion of its C4 atom within this coordinate frame.

Results and Discussion

Conformational Analysis. Projections of the time evolution of the glycosidic dihedral angles of panose onto the adiabatic maps of maltose[27] and isomaltose[22] are shown in Figure 3, for simulations P−A and P−B. The α(1→4) linkage between residues 1 and 2 (Φ_{12}, Ψ_{12}) samples only the A_{malt} and B_{malt} wells in the maltose map (at approximately $\Phi_{12} = -50°$, $\Psi_{12} = -40°$, and $\Phi_{12} = -15°$, $\Psi_{12} = --20°$, which is reached in simulations of maltose itself, is only transiently entered in the P−B simulation. However, the overall sampling is very similar to that of maltose, given that the C_{malt} well is significantly higher in energy. In the P−A simulation there is a greater preference for the B_{malt} well, in which an intramolecular hydrogen bond between 1 O2 and 2 O3 is formed, compared with the P−B simulation.

The α(1→6) linkage of panose samples mutually exclusive regions of conformational space in both simulations. This is due to the fact that simulations P−A and P−B were started on either side of the high-energy barrier to transitions in Ω. Very similar regions of the isomaltose map are sampled by the trisaccharide (panose) compared with the isolated linkage in the disaccharide (isomaltose): the P−A simulation essentially samples only the A_{iso} and occasionally the E_{iso} well, while the
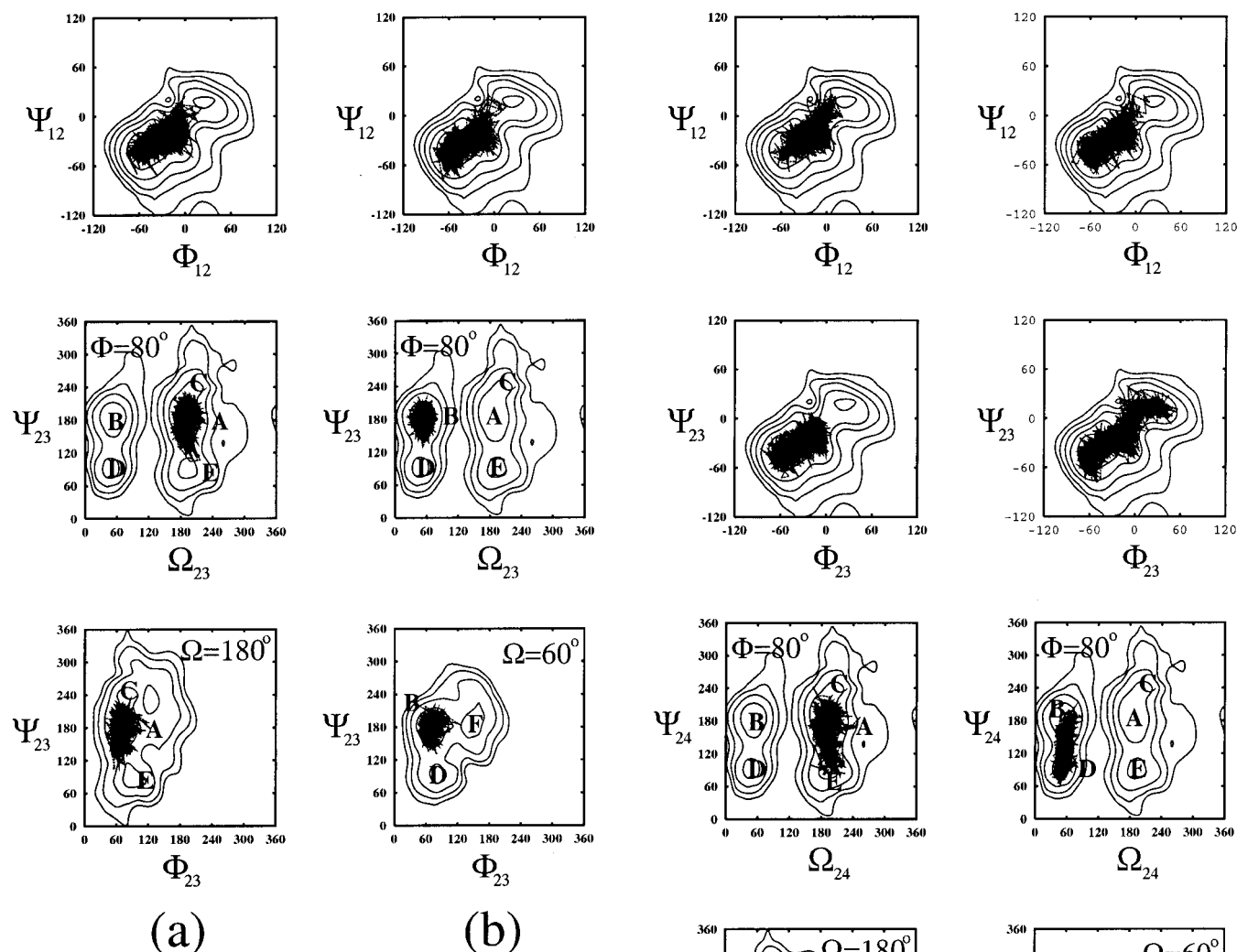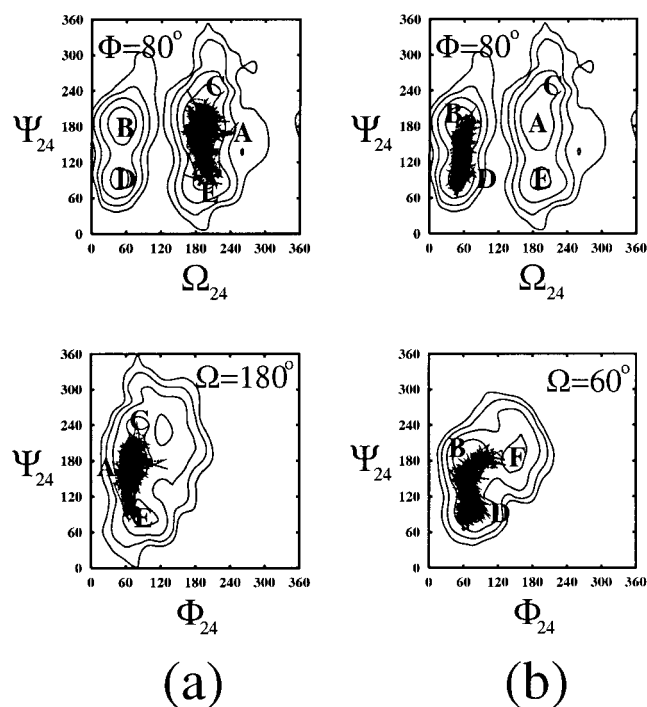
Figure 3. Projections of the dihedral angle dynamics trajectories for panose starting (a) from well A and (b) from well B in the adiabatic map of the $\alpha(1\rightarrow6)$ linkage.

P−B simulation favors the B_{iso} well. Therefore, from the conformational space analysis of panose, the $\alpha(1\rightarrow4)$ and $\alpha(1\rightarrow6)$ linkages appear to be essentially independent.

Figure 4 shows the analogous time series projections for 6^2 α-D-glucosylmaltotriose (6GMT). This molecule shows a marked deviation in the conformational preferences of its glycosidic linkages when compared with the component disaccharides maltose and isomaltose. The $\alpha(1\rightarrow4)$ linkage between residues 1 and 2 shows an even stronger preference for the B_{malt} energy well in simulation G−A than was the case for panose (simulation P−A); in simulation G−B it behaves similarly to panose's simulation P−B. The additional $\alpha(1\rightarrow4)$ linkage (between residues 2 and 4) samples all three maltose minima, with a marked bias toward the C_{malt} well in simulation G−B. The $\alpha(1\rightarrow6)$ linkage (2 to 3) shows significantly greater flexibility than it does in isomaltose or panose: frequent transitions are made in the Ψ_{24} to well E_{iso} for the G−A simulation and well D_{iso} for the G−B simulation. Addition of the extra residue in the tetrasaccharide provides alternative interactions on either side of the branch point increasing its flexibility.

A conformational clustering procedure (ART 2^{43}) was used to group the structures from the trajectories according to both $\alpha(1\rightarrow4)$ and $\alpha(1\rightarrow6)$ linkages. The principal clusters obtained by this method are listed in Tables 1 and 2 for panose and 6GMT, respectively. The results essentially confirm the above

Figure 4. Projections of the dihedral angle dynamics trajectories for the tetrasaccharide starting (a) from well A and (b) from well B in the adiabatic map of the $\alpha(1\rightarrow6)$ linkage.

TABLE 1: Clusters for Panose Simulations^a

| cluster	members	std dev	Φ_{12}	Ψ_{12}	Φ_{23}	Ψ_{23}	Ω_{23}
PA1	22140	12.0	−25.3	−24.2	71.9	−165.4	−168.5
PA2	17860	11.2	−33.4	−27.9	67.1	161.7	−170.9
PB1	24974	10.2	−52.2	−40.2	70.3	−177.9	55.4
PB2	15026	10.2	−20.3	−22.7	70.7	−176.4	54.9

^{a} PA1 and PA2 are from the panose A well and PB1 and PB2 are from the panose B well runs.

visual inspection. The two main clusters (PA1 and PA2) from the panose P−A simulation are very similar to each other, and equally there is a similarity between clusters (PB1 and PB2) from the P−B simulation where they differ only in the $\alpha(1\rightarrow4)$ linkage. However, the tetrasaccharide forms distinct clusters for

α(1→6) Branch Point of Amylopectin

J. Phys. Chem. B, Vol. 106, No. 19, 2002 **5095**

TABLE 2: Clusters for Tetrasaccharide Simulations[a]

cluster	members	std dev	Φ_{12}	Ψ_{12}	Φ_{24}	Ψ_{24}	Φ_{23}	Ψ_{23}	Ω_{23}
GA1	17837	12.3	−19.1	−20.4	−28.3	−28.0	70.9	−170.2	−169.6
GA2	18847	12.8	−31.8	−28.1	−33.9	−30.5	66.3	157.5	−172.2
GA3	3316	13.5	−17.3	−19.0	−21.5	−24.1	66.7	107.3	−161.8
GB1	15408	16.7	−43.4	−34.7	4.4	−0.40	70.4	148.6	56.2
GB2	24592	15.5	−35.5	−30.5	−42.1	−36.3	69.9	113.3	54.9

[a] GA1, GA2 and GA3 are from the A well run and GB1 and GB2 are from the B well run.
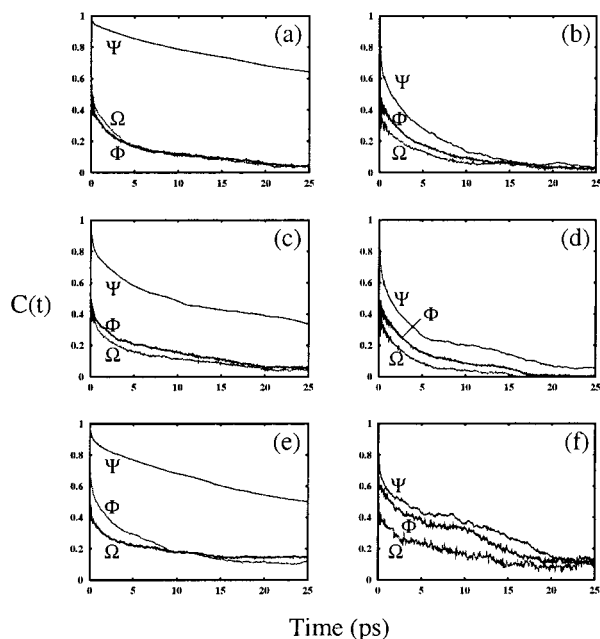
Figure 5. Dihedral angle correlation functions for the α(1→6) linkage for (a) isomaltose A well, (b) isomaltose B well, (c) panose A well, (d) panose B well, (e) tetrasaccharide A well, and (f) tetrasaccharide B well.

both G−A and G−B simulations. These are distinguished mainly by variations in the Ψ_{24} dihedral between the A_{iso} (GA1 and GA2) and E_{iso} (GA3) wells of isomaltose in simulation G−A and the B_{iso} and D_{iso} wells (GB1 and GB2) in simulation G−B.

Linkage Dynamics and Order Parameters. The assumed differences in flexibilities of the glycosidic linkages have been surmised by a number of experimentalists as the reason for solubility differences between amylose and amylopectin. To investigate this assumption, we compared the rate of motion about each glycosidic torsion angle in the model compounds panose and 6GMT. The rates can be assessed from time correlation functions for fluctuations in each dihedral. Comparison of correlation times for the α(1→4) and α(1→6) linkages confirmed our earlier finding that the motion about the α(1→6) linkage (correlation time 2−10 ps) occurs on a much shorter time scale than that about α(1→4) linkages (correlation time 10−20 ps). Here we limit our discussion to the α(1→6) dihedrals, which are responsible for most of the conformational variation. The α(1→6) dihedral correlation functions for isomaltose, panose, and 6GMT are plotted in Figure 5; the full trajectory has been used in the computation of each. It appears that the rate of dihedral motion is in fact relatively unaffected by the locally bonded molecular environment, as the correlation times for the three saccharides are very similar to one another. However, there is a notable difference between the dynamics of the A_{iso} and B_{iso} well conformers in the di-, tri-, and tetrasacchrides. In the A_{iso} well, (Figure 5a,c,e) the decay of

TABLE 3: T_1 Relaxation Times Calculated from Simulation and Experiment for Panose[a]

|	sim CH T_1	expt CH T_1	sim CH$_2$ T_1	expt CH$_2$ T_1
reducing ring	0.56	0.52(5)	0.27	0.21(7)
central ring	0.45	0.38(3)	0.23	0.21(1)
nonreducing ring	0.51	0.47(4)	0.26	0.21(1)

[a] Simulated results are the average value calculated for each of the A and B well runs.

the Ψ_{23} dihedral is much slower than is the case for the B_{iso} well (Figure 5b,d,f). The reason for this variation is due to the increase in volume of conformational space accessible to the Ψ_{23} dihedral in the A_{iso} well compared with that in the B_{iso} well; consequently, the fluctuations are slower in the former case.

A simple comparison of dynamics with experiment may be made by computing carbon-13 T_1 relaxation rates, which are sensitive to both intramolecular motion and overall tumbling. We calculated T_1 relaxation times for each of the carbons in panose, since experimental values for this compound are available.[44] The rotational diffusion time τ_M was determined by fitting an exponential function to the first part of the P_2 autocorrelation function for the vector between C1 and C4 of the central glucose residue, yielding a value of 130 ps. Longitudinal relaxation times were computed using this value together with the correlation functions for intramolecular motion in eq 4 and are listed in Table 3. Since each ring in a carbohydrate is relatively rigid, both experimental and simulated times have been averaged by residue, using separate averages for the methine (CH) and methylene (CH$_2$) groups. The terminal residues, which are more mobile, generally have slightly larger T_1 values. The match between experimental and simulated T_1 relaxation times is generally very good; the greatest discrepancy is for the methylene carbons, this could be due to the poor modeling of rotational frequencies by carbohydrate force fields. We have addressed this issue elsewhere.[45,46]

Saccharide-Induced Solvent Structure. Water has a strong influence on both structure and dynamics of carbohydrates in solution, particularly by making hydrogen bonds to the hydroxyl groups of the solute. However, when addressing the issue of solubility, not only must the effect of the solvent on the dynamics and conformation of the saccharide be taken into account but also the effect of the solvated saccharide on the solvent must be considered. In a saccharide solution the carbohydrate perturbs the equilibrium bulk water structure. Carbohydrates are strongly solvated due to their high density of polar groups; however, the structure of the surrounding water is nonetheless disrupted, most markedly near hydrophobic regions of the solute.[27] The distribution of water configurational probability about a saccharide is one measure of water structure, indicating how the solute causes deviations from bulk density. The spatial distribution function (SDF) for water oxygens, a three-dimensional analogue of the pair distribution function, calculated by orienting the sugar about the heavy atoms of the central ring is shown in Figure 6, contoured at 50% above bulk water density.

The overall pattern of high water oxygen density is similar to previously studied cases.[27,39,47] The highest density is generally located in the plane of each glucosyl ring, between the secondary alcohols, where bridging hydrogen bonds can be formed with the solute. There is generally above average density close to the α(1→4) linked residues, as they are less mobile, but there is also higher density near the α(1→6) linkage in the simulation started from the B_{iso} well. This may result from the smaller accessible range of conformations it can adopt, compared
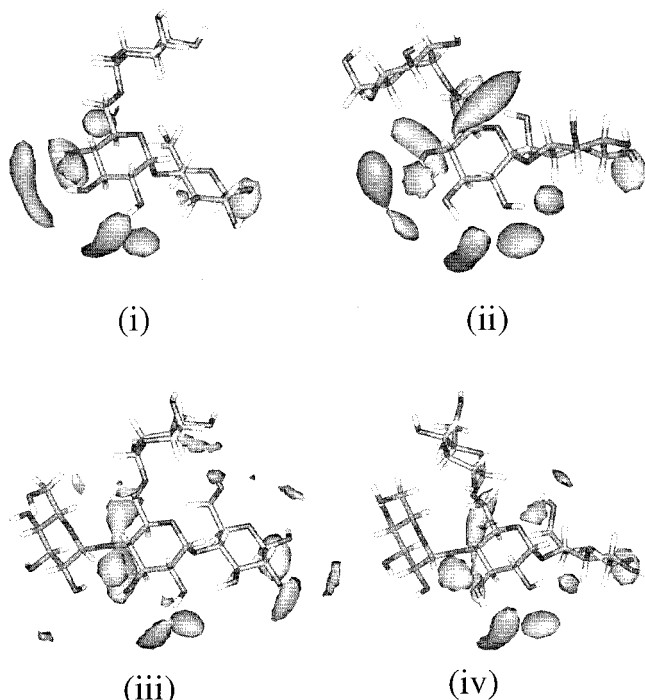
Figure 6. Spatial distribution functions for water about the carbohydrates, contoured at 50% above bulk water density: (i) panose A well; (ii) panose B well; (iii) tetrasaccharide A well; (iv) tetrasaccharide B well.


Figure 7. Water diffusion coefficients illustrated for the panose B well simulation: (a) radially averaged diffusion coeffients, calculated as a function of distance for the secondary alcohols of each residue; (b) spatially resolved map of diffusion coefficient, contoured at $0.50 \, \text{Å}^2 \cdot \text{ps}^{-1}$ overlayed on the SDF from Figure 6(ii) above.

with the A well. Interestingly, the water is not as densely localized for the tetrasaccharide 6GMT as it is for the trisaccharide panose, confirming the above finding that the additional residue increases flexibility about the branch point. Additionally, there are no regions of very high density (corresponding to "bound" water molecules) for our branch models, while densities of greater than 2.7 times the bulk density of water have been observed in the center of a maltohexaose fragment.[27] The branching residue therefore appears to have an overall disruptive effect on the structure of the solvent, favoring a higher solubility.

The exchange of water molecules between those located in the bulk and the first solvation shell of the saccharide may bring the conclusions of water solution structure around saccharide molecules into question. When the equilibrium water configurations are perturbed by the introduction of the saccharide, fluxes of various types arise in the solution. So, if the configurational probability densities as given by the SDF's (Figure 6) reflect an actual concentration gradient, then there should be a corresponding diffusion gradient. To address this issue, the diffusion coefficient, D, of the water molecules has been calculated as a function of distance from the solute secondary alcohols, averaged by residue (Figure 7a). The rate of solvent diffusion, while not significantly dependent on residue, shows a marked dependence on the distance from the solute. The limit of the diffusion constant at large distance is approximately 4.5 $\times 10^{-5} \, \text{cm}^2 \cdot \text{s}^{-1}$, which is close to the value expected for bulk TIP3P water.[48,49] A spatially resolved version of this radially averaged picture may also be calculated, in a similar way to generalizing a pair distribution function to a SDF. When overlaid on the SDF, this map (Figure 7b) shows that the high-density regions all lie within the region of solvent with low diffusion coefficient. Notably, the region with low diffusion coefficient extends over regions with no apparent high water densities. This is because the SDF's are averaged over space while the diffusion coefficients distributions are averaged over time. The averaging around a flexible oligosaccharide has a lower water density
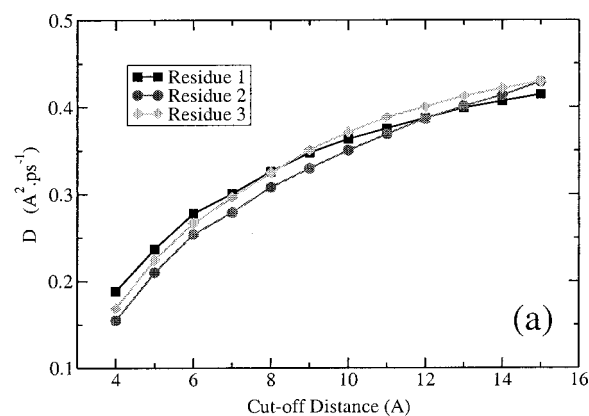
distribution since the water density count diminishes as the saccharide is displaced by the solvent while sweeping through conformational space. Our saccharide solution simulations therefore imply that a water concentration gradient exists and is located about the solvated sugar.

Implications for an Amylopectin Branch Point. Panose and 6GMT are minimalist models for an $\alpha(1{\rightarrow}6)$ branch point in amylopectin, representing its immediate environment. Conclusions drawn from these models can used to comment on the effect of the immediate molecular environment on the dynamics of the $\alpha(1{\rightarrow}6)$. The above study of the dihedral angle time correlation distribution functions shows that the Ψ_{23} dihedral is the torsion most responsible for the difference in the rate of $\alpha(1{\rightarrow}6)$ motion. To gain insight into the volume of conformational space that the branch glucose explores, we analyze the branch geometry in terms of polar coordinates (θ, ϕ), defined in Figure 2, for panose and 6GMT. Simulations from both the tri- and tetrasaccharide in Figures 8a–d reveal distinct patterns for the simulations initialized in the A_{iso} (Figure 8a,c) probability density is shifted, particularly in ϕ, and is much more localized in the case of the B_{iso} well, in agreement with the above finding that the $\alpha(1{\rightarrow}6)$ linkage has greater flexibility in the A_{iso} well. By contrast with the difference between A_{iso} and B_{iso} wells, the distribution shows relatively little dependence on the model used, although the tetrasaccharide has a somewhat broader distribution due to its greater flexibility.

While the description of the $\alpha(1{\rightarrow}6)$ linkage is useful, it supplies no information on the relative motions of the main $\alpha(1{\rightarrow}4)$ linked chain and the branch residue. We traced the angle between C1 of residue 1 and C4 of residue 3 in each molecule as a measure of the relative orientation of the backbone and branch residue. The distribution of this angle for the simulations P−A, G−A (solid lines) and P−B, G−B (dotted lines) are illustrated in Figure 9. It is clear that the A and B well

α(1→6) Branch Point of Amylopectin

J. Phys. Chem. B, Vol. 106, No. 19, 2002 **5097**
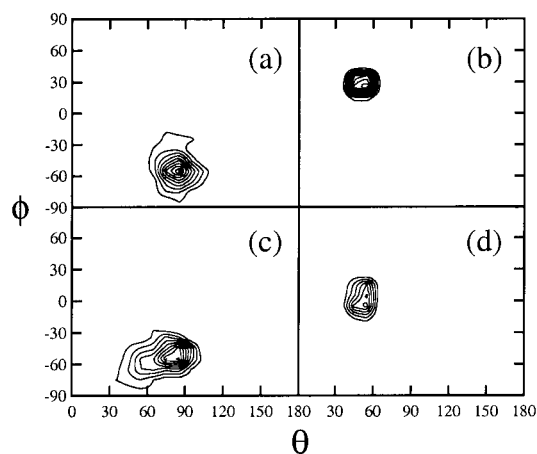
Figure 8. Distribution of the spherical polar angles (θ, ϕ) for each simulation: (a) panose A well; (b) panose B well; (c) tetrasaccharide A well; (d) tetrasaccharide B well. The probability density $p(\theta, \phi)$ is contoured from 0.01 in steps of 0.005 for the A wells and 0.04 in steps of 0.02 for the B well simulations.
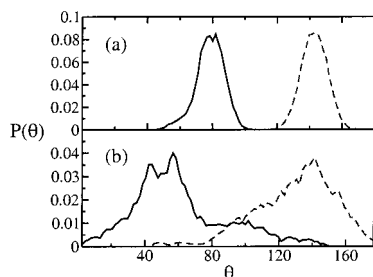
Figure 9. Probability distributions for the angle at the center of the middle ring between C1 of the α(1→6) linked residue and C4 of the α(1→4) linked residue for (i) panose A well simulation, (ii) panose B well simulation, (iii) tetrasaccharide A well simulation, and (iv) tetrasaccharide B well simulation.

simulations have distinct distributions with maximum positions dependent on the model chosen.

This analysis indicates that the tetrasaccharide has considerably more flexibility in each simulation, spanning a much wider range of angles. This is due to the alternate interactions that the branching α(1→6) residue can make with each of the adjacent α(1→4)-linked residues. In the case of panose there is only one such set of interactions. The considerable overlap of the distributions in Figure 9b suggests a subset of angles which could accommodate either the A_{iso} or B_{iso} well conformations of the α(1→6) linkage.

Amylopectin forms double helices in the crystalline lamellae of the starch granule, which are believed to be composed of parallel α(1→4) linked strands. Such an arrangement would be easily accommodated by the B_{iso} well, in which the branch "folds back" parallel to the main chain. The A_{iso} well branch arrangement is more perpendicular but could also be accommodated by the starch structure at the cost of a large loop at the neck of the double helix; this is consistent with studies of these model compounds in a vacuum.[28]

Conclusion

The solution simulations of amylopectin branch point models presented here provide insight into their favored conformations. Just as for isomaltose, the α(1→6) glycosidic conformational space of the two model compounds is divided into two separate regions divided by a high-energy barrier, one comprising the A_{iso} and E_{iso} wells and the other the B_{iso} and D_{iso} wells. For

both regions, the addition of extra residues promotes greater flexibility through the interactions with the carbohydrate backbone. This is manifested (i) in the conformational sampling, where both of the branch models made multiple interwell transitions, (ii) by the water structure, which is less ordered for the tetrasaccharide, and (iii) by the increased variation in branch to backbone angle from trisaccharide to tetrasaccharide (Figure 8). Further, the effect of the branch residue appears to be an overall disruption of the water structure relative to α(1→4)-linked amylose, since there are no regions of exceptionally high water density. This suggests a reason for the higher solubility of amylopectin.

Another significant observation was that the two regions of conformational space have very different properties. The A_{iso} well has a larger volume of linkage geometries available to it, as seen in Figure 8, and related to this is the longer correlation of the Ψ_{23} torsion angle when in the A_{iso} well. The accessible branch geometries are furthermore quite different for the A_{iso} and B_{iso} wells. The B_{iso} well would naturally accommodate the high-density packing in the starch granule, as it places the two strands in a parallel arrangement suitable for the formation of amylopectin double helices and must be considered most likely in the absence of new experimental data. Unfortunately, a comparison of the relative free energies the A_{iso} and B_{iso} well regions was not possible as no transitions were made between them on the time scale of the simulation. We are presently investigating this issue through the use of free energy perturbation methods.

Finally, we have shown that there exists a shell of spatially localized water centred about the dissolved saccharides. We arrived at this conclusion by matching spatial distribution functions for each carbohydrate solution with a novel analysis of the water diffusion coefficients that have been radially averaged (Figure 7a) and anisotropically resolved (Figure 7b). The high-density configurational probability volumes observed in close proximity to the hydroxyls in the SDF plots therefore represents a pockets of greater than bulk solvent concentration.

References and Notes

(1) Zobel, H. F. *Starch/Stärke* **1988**, *40*, 44−50.

(2) Waigh, T. A.; Hopkinson. I.; Donald. A. M.; Butler, M. F.; Heidelbach, F.; Riekel, C. *Macromolecules* **1997**, *30*, 3813−3820.

(3) Smith, A. M. *Biomacromolecules* **2001**, *2*, 335−341.

(4) DiBari, M.; Cavatorta, F.; Deriu, A.; Albanese, G. *Biophys. J.* **2001**, *81*, 1190−1194.

(5) Burchard, W. *Biomacromolecules* **2001**, *2*, 342−353.

(6) Imberty, A.; Perez, S. *Carbohydr Res.* **1988**, *181*, 41−55.

(7) Imberty, A.; Buleon, A.; Perez, S. *Starch/Stärke* **1991**, *43*, 375−385.

(8) Quigley, G. J.; Sarko, A.; Marchessault, R. H. *J. Am. Chem. Soc.* **1970**, *92*, 2.

(9) Liu, J. H. Y.; Brant, D. A.; Kitamura, S.; Kajiwara, K.; Mimura, R. *Macromolecules* **1999**, *32*, 8611−8620.

(10) Kjellberg, A.; Rundlof, T.; Kowalewski, J.; Widmalm, G. *J. Phys. Chem. B* **1998**, *102*, 1013−1020.

(11) Martin-Pastor, M.; Bush, C. A. *Biochemistry* **1999**, *38*.

(12) Duus, J. Ø.; Gotfredsen, C. H.; Bock, K. *Chem. Rev.* **2000**, *100*, 4589−4614.

(13) Rees, D. A.; Thom, D. *J. Chem. Soc., Perkin Trans. 2* **1976**, 191−201.

(14) Stevens, E. S.; Sathyanarayana, B. K.; Morris, E. R. *J. Phys. Chem.* **1989**, *93*, 3434−3436.

(15) Hardy, B. J.; Bystricky, S.; Kovac, P.; Widmalm, G. *Biopolymers* **1997**, *41*, 83−96.

(16) Stroyan, E. P.; Stevens, E. S. *Carbohydr Res.* **2000**, *327*, 447−453.

(17) Rundlof, T.; Landersjo, C.; Lycknert, K.; Maliniak, A.; Widmalm, G. *Magn. Reson. Chem.* **1998**, *36*, 773−776.

(18) Almond, A.; Bunkenborg, J.; Franch, T.; Gotfredsen, C. H.; Duus, J. O. *J. Am. Chem. Soc.* **2001**, *123*, 4792−4802.

(19) Martin-Pastor, M.; Bush, C. A. *Biochemistry* **2000**, *39*, 4674−4683.

(20) Rutherford, T. J.; Homans, S. W. *Biochemistry* **1994**, *33*, 9606−9614.

(21) Imberty, A.; Pérez, S. *Chem. Rev.* **2000**, *100*, 4567−4588.

(22) Best, R. B.; Jackson, G. E.; Naidoo, K. J. *J. Phys. Chem. B* **2001**, *105*, 4742−4751.

(23) Staaf, M.; Hoog, C.; Stevensson, B.; Maliniak, A.; Widmalm, G. *Biochemistry* **2001**, *40*, 3623−3628.

(24) Xu, Q. W.; Mohan, S.; Bush, C. A. *Biopolymers* **1996**, *38*, 339−353.

(25) *Carbohydrate Chemistry*; Kennedy, J. F., Ed.; Oxford University Press: Oxford, U.K., 1988.

(26) O'Donoghue, P.; Luthey-Schulten, Z. A. *J. Phys. Chem. B* **2000**, *104*, 10398−10405.

(27) Naidoo, K. J.; Kuttel, M. *Comput. Chem.* **2001**, *22*, 445−456.

(28) Imberty, A.; Perez, S. *Int. J. Biol. Macromol.* **1989**, *11*, 177−185.

(29) O'Sullivan, A. C.; Perez, S. *Biopolymers* **1999**, *50*, 381−390.

(30) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4* (2), 187−217.

(31) Palma, R.; Himmel, M. E.; Liang, G.; Brady, J. W. Molecular Mechanics Studies of Cellulases. In *ACS Symposium Series: Glycosyl Hydrolases in Biomass Conversion*; Himmel, M. E., Ed.; American Chemical Society: Washington, DC, 2000; No. 769, Chapter 7, pp 112−130.

(32) Jorgensen, W. L. *J. Phys. Chem.* **1982**, *77*, 4156−4163.

(33) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926−935.

(34) Franks, F. *The physics and physical chemistry of water*, 1st ed.; Plenum Press: New York, 1972; Vol. 1.

(35) Allen, M. P.; Tildesley, D. J. *Computer Simulation of Liquids*, 1st ed.; Oxford Science Publications: Oxford, U.K., 1989.

(36) Lipari, G.; Szabo, A. *J. Am. Chem. Soc.* **1982**, *104*, 4559−4570.

(37) Lipari, G.; Szabo, A. *J. Am. Chem. Soc.* **1982**, *104*, 4546−4559.

(38) Kowalewski, J.; Widmalm, G. *J. Phys. Chem.* **1994**, *98*, 28−34.

(39) Liu, Q.; Brady, J. W. *J. Am. Chem. Soc.* **1996**, *118*, 12276−12286.

(40) Kabsch, W. *Acta Crystallogr.* **1976**, *A32*, 922−923.

(41) Kabsch, W. *Acta Crystallogr.* **1978**, *A34*, 827−828.

(42) Cremer, D.; Pople, J. A. *J. Am. Chem. Soc.* **1975**, *96*, 1354−1358.

(43) Carpenter, G. A.; Grossberg, S. *Appl. Opt.* **1987**, *26*, 4919−4930.

(44) Jackson, G. E.; Best, R. B.; Naidoo, K. J. *Spec. Lett.*, submitted for publication.

(45) Kuttel, M. M.; Naidoo, K. J. Towards realistic primary alcohol rotation in carbohydrate models. 20th International Carbohydrate Symposium, 2000, Hamburg, Germany.

(46) Kuttel, M. M.; Brady, J. W.; Naidoo, K. J. *J. Comput. Chem.*, submitted for publication.

(47) Hoog, C.; Widmalm, G. *J. Phys. Chem. B* **2001**, *105*, 6375−6379.

(48) Spoel, D. v. d.; Maaren, P. J. v.; Berendsen, H. J. C. *J. Chem. Phys.* **1998**, *108*, 10220−10230.

(49) Levitt, M.; Hirschberg, M.; Sharon, R.; Laidig, K. E.; Daggett, V. *J. Phys. Chem. B* **1997**, *101*, 5051−5061.