Structure of Aqueous Glucose Solutions as Determined by Neutron Diffraction with Isotopic Substitution Experiments and Molecular Dynamics Calculations

P. E. Mason,[†] G. W. Neilson,[‡] J. E. Enderby,[‡] M.-L. Saboungi,[§] and J. W. Brady*,[†]

Department of Food Science, Stocking Hall, Cornell University, Ithaca, New York 14853, H. H.Wills Physics Laboratory, University of Bristol, BS8 1TL, United Kingdom, and Centre de Recherche sur la Matière Divisée, 1 bis rue de la Férollerie, 45071 Orléans, France

Received: September 23, 2004; In Final Form: February 25, 2005

Neutron diffraction with isotopic substitution (NDIS) experiments and molecular dynamics (MD) simulations have been used to examine the structuring of solvent around D-glucose in aqueous solution. As expected, no significant tendency for glucose molecules to aggregate was found in either the experiments or the simulation. To the extent that solute pairing does occur as the result of the high concentration, it was found to take place through hydroxyl−hydroxyl hydrogen bonds, in competition with water molecules for the same hydrogen-bonding sites. A detailed analysis of the hydrogen-bonding patterns occurring in the simulations found that the sugar hydroxyl groups are more efficient hydrogen bond donors than acceptors. From the comparison of the MD and NDIS data, it was found that while the modeling generally does a satisfactory job in reproducing the experimental data the force fields may produce sugar rings that are too rigid and thus may require future revisions.

Introduction

It is now well-established that water structuring around solutes can vary significantly depending on the chemical character of the solute molecules, and that the nature of this solvent structuring plays an important role in solution properties, and can have profound consequences in biological systems. Characterizing this structuring, however, has proven to be difficult. Neutron diffraction is a powerful experimental method for probing the structure of aqueous solutions, but this method is beset with problems when applied to complex molecular solutes such as those that occur in biological systems. These molecules are usually large relative to water molecules and have little or no symmetry. They often contain a number of functional groups, each with their own specific hydration requirements, juxtaposed in close proximity by the topology of the solute, thus interfering with one another's hydration and imposing specific and often very complex structure on the adjacent solvent. For polyatomic molecules, the observed neutron scattering contains contributions from all of these distinct atoms in their individual environments and can be difficult or impossible to interpret in terms of the detailed anisotropic structure of the solvent. However, as has recently been demonstrated with simpler systems such as guanidinium salts,[1] molecular dynamics (MD) simulations can be useful in helping to interpret the results of neutron diffraction experiments.

We have previously reported the use of both MD simulations[2] and neutron diffraction with isotopic substitution (NDIS) experiments[3] to study aqueous solutions of D-glucose. This molecule is a good candidate for such studies for several reasons. It is significant since it is a prototypical biological molecule and important as an energy currency in living systems, in part because of its solubility. It contains several types of common biological functional groups but exhibits less conformational flexibility than other prototypical biological molecules such as peptides, thus reducing the amount of ambiguity resulting from the smearing of solvent densities due to conformational averaging. In addition, while it is quite hydrophilic, it is a neutral molecule, so that the hydration of charged groups will not swamp that of uncharged moieties. Perhaps more importantly from a practical standpoint, it is soluble at the moderately high concentrations needed for neutron diffraction experiments (1−5 *m*).

We describe here the use of new MD simulations of D-glucose to interpret details about the hydration of this solute as probed using NDIS experiments[3]. Due to the flux available from current neutron sources, higher sugar concentrations are needed than have been previously simulated. Also, in the experimental situations, the sugars undergo mutarotation to give a 36:64 mixture of the α and β anomers, respectively, while earlier simulations modeled only a single sugar anomer at a time. Accordingly, a new series of simulations was performed in which the primary simulation box contained a number of independent glucose molecules of both anomers in this appropriate ratio at concentrations of 1, 3, and 5 *m*.

Computational Procedures

In the MD simulations, a neutral periodic cubic system was created at 1, 3, and 5 *m* containing, in each case, a number of independent glucose molecules surrounded by explicit water molecules. The simulations employed the CSFF glucose force field with water molecules being represented by the TIP3P model.[4,5] All simulations were performed using the CHARMM program,[6] with chemical bonds to hydrogen atoms kept fixed using SHAKE[7] and a time step of 1 fs. The starting coordinates for the 1 *m* solution were prepared by placing 12 β- and 6 α-D-glucopyranose molecules (randomly orientated) in a 34 Å cube. The ratio of 2β/1α closely resembles a glucose solution at the

* Author to whom correspondence should be addressed. E-mail: jwb7@cornell.edu.
[†] Cornell University.
[‡] University of Bristol.
[§] Centre de Recherche sur la Matière Divisée.

10.1021/jp040622x CCC: $30.25
Published on Web 06/21/2005

TABLE 1: Composition of the Function ${}^{n}G_{Y}^{Y}(r)$ Measured in This Study for 1.00, 3.00, and 5.00 *m* Glucose Solutions[a]

glucose concentration (*m*)	*A* (mb)	*B* (mb)	*C* (mb)	*A* + *B* + *C* (mb)
1.00	238	27	−17	247
3.00	197	55	−36	216
5.00	171	69	−45	195

a) $${}_{n}G_{H}^{Y}(r) = \frac{Ag_{H_{ex}O}(r) + Bg_{H_{ex}C}(r) + Cg_{H_{ex}H_{non}}(r)}{A + B + C}$$

anomeric equilibrium. These solute coordinates were superimposed on a box of 1296 water molecules, and those that overlapped any solute heavy atom were discarded to produce the desired concentration; by design, this procedure produced a 1 *m* solution (18 glucose molecules and 995 TIP3P water molecules, 1.005 *m*). Finally, the box length was rescaled to 32.1527 Å, which yielded the correct physical number density (0.1028 atoms $Å^{-3}$). An identical method was used for preparing the starting coordinates for the 3 and 5 *m* solutions (3 *m*, 24 β and 12 α glucose molecules, 667 TIP3P water molecules, 2.999 *m*, box length 29.9447 Å, number density 0.1067; 5 *m*, 40 β and 20 α glucose molecules, 666 TIP3P waters, 5.005 *m*, box length 31.5474 Å, number density 1.095).

All van der Waals and electrostatic interactions were smoothly truncated on a group basis using switching functions from 13 to 15 Å. Initial velocities were assigned from a Boltzmann distribution (300 K) followed by 7 ps of equilibration dynamics with velocities being reassigned every 0.1 ps. The simulation was then run for 2 ns with no further velocity reassignment. The first 0.5 ns of this was taken as equilibration, and the remaining 1.5 ns was used for analysis.

The methods used in the sample preparation, NDIS measurement, and data analysis of the 1, 3 and 5 *m* glucose solutions in this paper have been previously published[3]. Three functions were measured, $g_{H_{ex}H_{ex}}(r)$, where H_{ex} represents all exchangeable hydrogen atoms (as opposed to the nonexchangeable aliphatic protons H_{non}), ${}^{n}G_{H_{ex}}^{Y}(r)$, which contains correlation terms from the exchangeable hydrogen atoms to all other atoms (Y) in the system except for the exchangeable hydrogen atoms, and finally the functions ${}^{n}G_{Y}^{Y}(r)$, which contains terms from all the nonexchangeable hydrogen atoms to all the other non-exchangeable hydrogen atoms. The latter two functions both were normalized by division by the sum of the scattering prefactors and are more fully described in Tables 1 and 2.

Results and Discussion

Previous MD simulations have been used to characterize the structuring of water adjacent to glucose in solution, but these simulations have not been directly characteristic of the conditions of high concentrations and mixed anomeric composition that would exist in an experimental probe of this structuring using neutron diffraction. To replicate the experimental situation as closely as possible, the simulation box contained both anomers of glucose and experimental concentrations. At 3 *m* concentrations, there are only a little over 18 water molecules per sugar molecule, which is less than two complete hydration layers, although for a given pair of sugar molecules there are, on average, more than three complete solvation shells separating the solute molecules. For this reason, one might expect that any solvent structuring beyond the second hydration shell could exhibit the influence of other solute molecules. Previous simulations of single solutes in large water boxes have found evidence for weak correlations out beyond the third solvation layer,[8] even with a relatively unstructured water model like TIP3P, but the only significant structuring is in the first and second hydration layers, so calculations and experiments even at the high concentrations required by the diffraction sensitivity limits will be useful for the analysis of the effects of these solutes on water at more physiological concentrations.

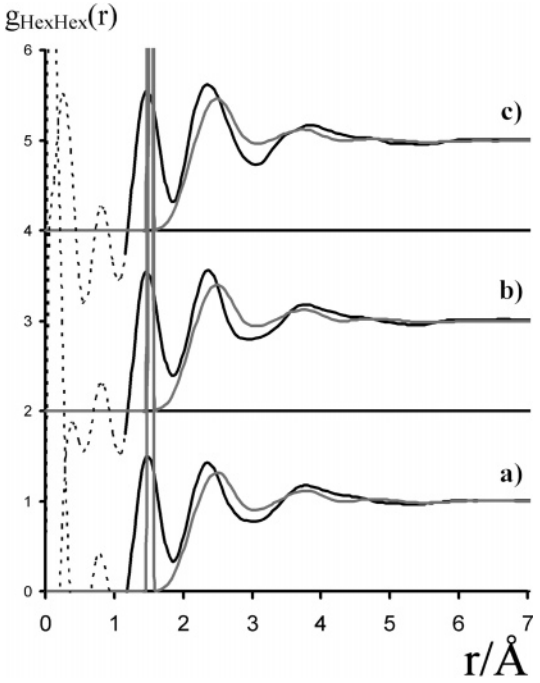

Figure 1. Radial distribution function $g_{H_{ex}H_{ex}}(r)$ for 1, 3, and 5 *m* solutions (a, b, and c, respectively). In each case, the experimental NDIS data is shown in black while the distributions calculated from MD simulations are shown in gray.

Figure 1 displays the proton−proton pairwise radial distribution function $g_{H_{ex}H_{ex}}(r)$ as obtained from the NDIS experiment and the MD simulations. The structure in the experimental radial distribution function at distances less than 1 Å arises from truncation errors and is shown as a dotted curve in this figure. Unfortunately, this function is informationally poor and difficult to interpret. The principal reason for this is that the water molecules, which contain almost all of the exchangeable hydrogen atoms, H_{ex}, in these solutions, while remaining

TABLE 2: Composition of the Function ${}^{n}G_{Y}^{Y}(r)$ Measured in This Study for 1.00, 3.00 and 5.00 *m* Glucose Solutions[a]

glucose concentration (*m*)	*D* (mb)	*E* (mb)	*F* (mb)	*G* (mb)	*H* (mb)	*I* (mb)	*D* + *E* + *F* + *G* + *H* + *I* (mb)
1.00	0.43	7.84	−0.57	−5.15	0.19	35.1	37.9
3.00	2.51	17.9	−3.30	−11.8	1.08	32.0	38.5
5.00	4.84	24.1	−6.36	−15.8	2.08	30.0	38.8

a) $${}_{n}G_{Y}^{Y}(r) = \frac{Dg_{CC}(r) + Eg_{CO}(r) + Fg_{CH_{non}}(r) + Gg_{OH_{non}}(r) + Hg_{H_{non}H_{non}}(r) + Ig_{OO}(r)}{D + E + F + G + H + I}$$

strongly hydrogen-bonded to each other, retain a great deal of rotational freedom, as has often been demonstrated from MD simulations. This rotational freedom is reflected as broad intermolecular features in $g_{H_{ex}H_{ex}}(r)$. The first peak at 1.5 Å, due to the geminal correlation of the water hydrogen atoms, is almost a singularity in the MD data but is a broad feature in the NDIS measurement. Part of this difference is due to the fact that the TIP3P water molecules employed in the simulation were kept completely rigid using the SHAKE constraint algorithm. Another reason for the difference in this peak between the diffraction and the simulation data is that the NDIS measurement is resolution-limited (because *Q*-space data is not available beyond 16.5 $Å^{-1}$). However, this proton−proton peak is also broadened in the experimental function by quantum mechanical uncertainty in the atomic positions in a manner that cannot be captured by a classical water model. However, since this peak probes only the internal structure of the water molecules, it is of little interest in the present context, where the goal is to characterize the solution structuring.

The peak in $g_{H_{ex}H_{ex}}(r)$ at 2.3 Å is due to the correlation of an H_{ex} on one water molecule with another H_{ex} on a different water molecule that is hydrogen-bonded to the first. It therefore reflects the hydrogen-bonding structure of the solution. In this range, the diffraction measurement is not resolution-limited. From Figure 1, it can be seen that there are differences in this peak between the diffraction and simulation data. These differences between modeling and experiment are similar to those observed for pure water. In general, MD simulations of water using standard models reproduce all of the correct structural features and at approximately the correct distances but differ in the details.

It is interesting that the $g_{H_{ex}H_{ex}}(r)$ function, whether obtained from the diffraction experiments or the MD simulations, is very similar for all three concentrations studied. While the 1 *m* solution is predominantly water, the 5 *m* solution is 40 atom percent glucose, and the series of functions can be taken as a measure of the degree to which the increasing glucose concentration perturbs the water structure. The insensitivity of $g_{H_{ex}H_{ex}}(r)$ to concentration implies that the glucose solute is not strongly disrupting the normal structuring between water molecules and that those waters that interact directly with the solute do so through hydrogen bonds that closely resemble those to other water molecules. A more sensitive measure of the structuring of water can be obtained from the second peak of the oxygen−oxygen pair correlation function, $g_{OO}(r)$, which is shown as calculated from the MD simulations in Figure 2. One known shortcoming of the TIP3P water model is that this second peak in $g_{OO}(r)$ is relatively weak, indicating less tetrahedral structuring than in real water. However, given this limitation, it can be seen that this distribution is also only a weak function of concentration. While the first peak becomes stronger and broader with increasing concentration, the position and height of the second peak is relatively unchanged. It should be remembered that at 5 *m* there are only approximately 10 water molecules per glucose molecule. The relatively weak influence of the sugar solute on water structuring is interesting since sugars have traditionally been regarded as "structure makers" in a Hofmeister sense.[9] Some other solutes, such as LiCl, are completely disruptive to $g_{H_{ex}H_{ex}}(r)$ at atomic concentrations of 12 atom percent (10 *m*).[10] It has previously been observed that the array of sugar hydroxyl positions around glucose closely matches the water spacings in ice.[11] This geometric feature probably explains the weak concentration dependence of the structuring, particularly at the 5 *m* concentration, where nearly

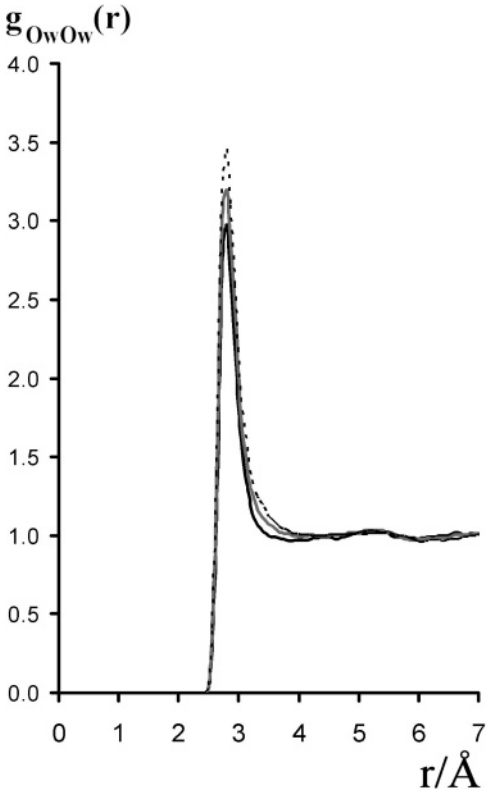

Figure 2. Water oxygen−oxygen radial distribution function $g_{OO}(r)$ calculated from the MD simulations of glucose solutions at 1 *m* (black), 3 *m* (gray), and 5 *m* (dotted).

all water molecules are directly hydrogen-bonded to glucose and second peak correlations are predominantly between water molecules hydrogen-bonded to sugar hydroxyl groups two carbon positions away on the ring (i.e., separated by an intervening sugar hydroxyl group and its hydrogen-bonded water molecule partner). It has also been found that the IR spectrum of water in the near infrared is totally insensitive to glucose, even at 5 *m* concentration, confirming that water can easily accommodate glucose as a solute.[12]

The function ${}^{n}G^{Y}_{H_{ex}}(r)$, shown in Figure 3, has contributions from three radial distribution functions, $g_{H_{ex}O}(r)$, $g_{H_{ex}C}(r)$, and $g_{H_{ex}H_{non}}(r)$, with the prefactors shown in Table 1. These may be further subdivided into functions due to exchangeable protons on the glucose molecules, ${}^{n}G^{Y}_{H_{ex}G}(r)$, and on the water molecules, ${}^{n}G^{Y}_{H_{ex}W}(r)$. The scattering prefactors for these substructure functions are directly related to the mole fraction of H_{ex} in each state. For the 5 *m* solution, 18% of the exchangeable protons H_{ex} are resident on the glucose solute molecules. As was the case for the $g_{H_{ex}H_{ex}}(r)$ function, it was found that the $H_{ex}O$ intramolecular correlation at 1.0 Å is large and resolution-limited in the diffraction experiment. Indeed, it is of such a large magnitude that if it is not removed from the NDIS data it has a propensity to induce errors in the Fourier transform in the region up to approximately 2.5 Å.[3] In this case, the comparison of the simulation and diffraction data is much more satisfactory. The first peak at 1.9 Å has two components, the hydrogen bond peak of $g_{H_{ex}O}(r)$ and the intramolecular correlation of $g_{H_{ex}GC}(r)$. As the concentration of glucose is increased, the $g_{H_{ex}GC}(r)$ component of the peak increases, and this feature gets sharper. The small decrease in correlation at 2.4 Å is quite interesting. This effect is mostly due to the negatively weighted (because of the negative scattering cross section of protium) intramolecular correlation of glucose H_{ex} with the nonexchangeable sugar protons H_{non}.

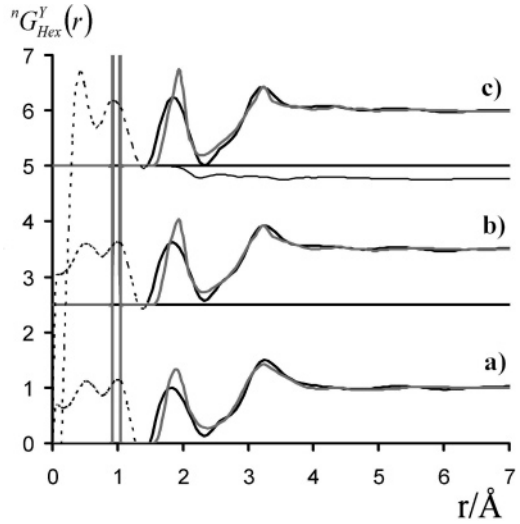

Figure 3. Total radial distribution function for all other atoms to the exchangeable hydrogen atoms, $^nG^Y_{H_{ex}}(r)$ for 1, 3, and 5 *m* solutions (a, b, and c, respectively). In each case, the NDIS data is shown in black while the distributions calculated from MD simulations are shown in gray. The negatively weighted $g_{H_{ex}H_{non}}(r)$ contribution to the $^nG^Y_{H_{ex}}(r)$ for the 5 *m* solutions is shown as a black line in the upper plot. The minor peak in this function at 2.3 Å implies that the $H_{non}-C-O-H_{ex}$ groups in glucose preferentially do not occupy the trans position (relative to the nonexchangeable glucose hydrogen).

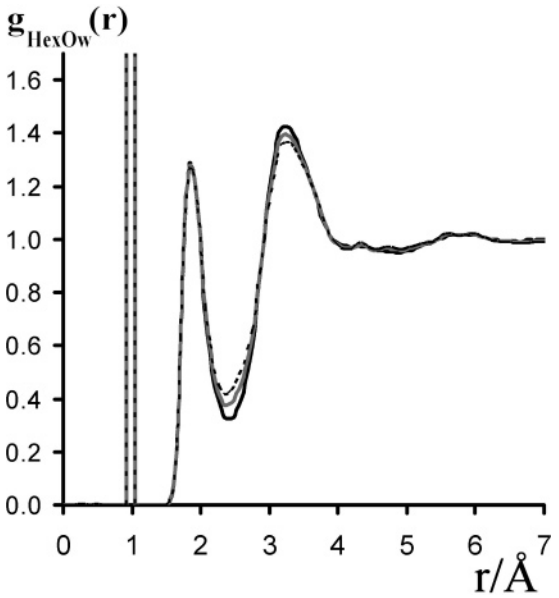

Figure 4. . Radial distribution function $g_{H_{ex}OW}(r)$ for the 1, 3, and 5 *m* solutions (black, gray, and dotted, respectively) as calculated from the MD simulations. This is one of the radial distribution functions representative of water structure. Given the fact that the 5 *m* solution is approximately 40 atom percent glucose, the glucose has a relatively minor effect upon water structure.

The fact that this intramolecular $H_{ex}-H_{non}$ contribution can be observed suggests that a suitably designed NDIS experiment to exploit this effect might be used to estimate hydroxyl rotameric populations. The H_{ex} of each hydroxyl group in the sugar solutes might be expected to exist in one of three staggered conformations. For two of these staggered positions, the $H_{exG}-H_{non}$ distance is approximately 2.3 Å, while in the trans position this distance is approximately 2.7 Å. From the MD simulation, it can be seen that these correlations are the only significant contributions to the structure in $g_{H_{ex}H_{non}}(r)$ in this distance range (Figure 3). The NDIS measurement seems to imply, consistent with the simulation results, that on average the conformations with the $H_{exG}-H_{non}$ distance of 2.3 Å are preferentially occupied.

The MD simulations also successfully predict the general sharpening of the peak at 3.2 Å with glucose concentration, again due to positively weighted correlations of $H_{exG}-C$ and negatively weighted $H_{exG}-H_{non}$ correlations. As was the case for the $g_{H_{ex}H_{ex}}(r)$ function, the diffraction data here can again be used to argue that the sugar solute has little effect on the long-range, bulk structure of the water. The simulation results can also be examined for information pertaining to this question. It was found that if only the $H_{exW}O_W$ correlation is extracted from the simulations, as is shown in Figure 4, the glucose is again confirmed to have only a very small effect on the bulk water structure.

The function $^nG^Y_Y(r)$, displayed in Figure 5, whose subcomponent composition is shown in Table 2, is the most complex of the structural functions measured in the NDIS experiments. This function probes not only the solution structure but also contains many significant intramolecular correlations. Glucose has no symmetry, and so its 24 different atoms may be expected to yield 300 different correlations (although only 190 of these are due to YY-style correlations). However, only about 8 peaks are observed in both the diffraction and simulation results. This is due in part to the relatively regular structure of the glucose molecules, with the carbon and oxygen atoms fitting approximately onto a tetrahedral lattice with a bond length of 1.4 Å (Figure 5), while the scattering prefactors (weightings) of the CO_G, (oxygen on glucose) CC, and O_GO_G correlations in the measurement are all similar in magnitude (relative weightings of 1.0, 0.57, and 0.44, respectively). In the sugar framework, the nonexchangeable protons H_{non} are aligned down the vertexes of the tetrahedral lattice but have a bond length of only 1 Å (Figure 5) and have a negative weighting in the experimental measurement (relative weightings, on the same scale as before, for the CH_{non} and OH_{non} correlations are -0.75 and -0.65, respectively). Since the glucose molecule is only about 5 lattice units long, it is expected that only about four positively and four negatively weighted correlations would be observed.

It can be seen that in some cases the prediction for molecular structure from the simulations produced features that are sharper than those observed in the diffraction measurement, although these features should not be resolution-limited in the NDIS experiment. To examine this matter explicitly, both sets of data were back-transformed in the range 1.8−20 $Å^{-1}$ (Figure 6). If these data were resolution-limited, then both sets of data would look identical up to the upper *Q*-range of the experimental measurement (16 $Å^{-1}$), yet the MD result would contain ripples that went out beyond the *Q*-range of the experiment. As can be seen from Figure 6, the *Q*-space data are qualitatively similar for both the simulation and diffraction studies but exhibit a discrepancy between the MD and NDIS measurements before the experimental high-*Q* limit that varies with the concentration of glucose. This is a clear demonstration that the NDIS measurement is not resolution-limited in this case and implies that the glucose potential used in the simulations may be somewhat too rigid.

This result is not entirely unexpected. The force field used in the modeling is a revised version of a CHARMM-type parameter set developed specifically for carbohydrates from vibrational, thermodynamic, ab initio, and structural data.[13,14] This force field is known to be more rigid than our previous carbohydrate force field.[15] Because the parameters for the valence terms in this force field were exhaustively matched to the vibrational data for both glucose and a number of saccharide small-molecule analogues,[13] it is unlikely that those terms

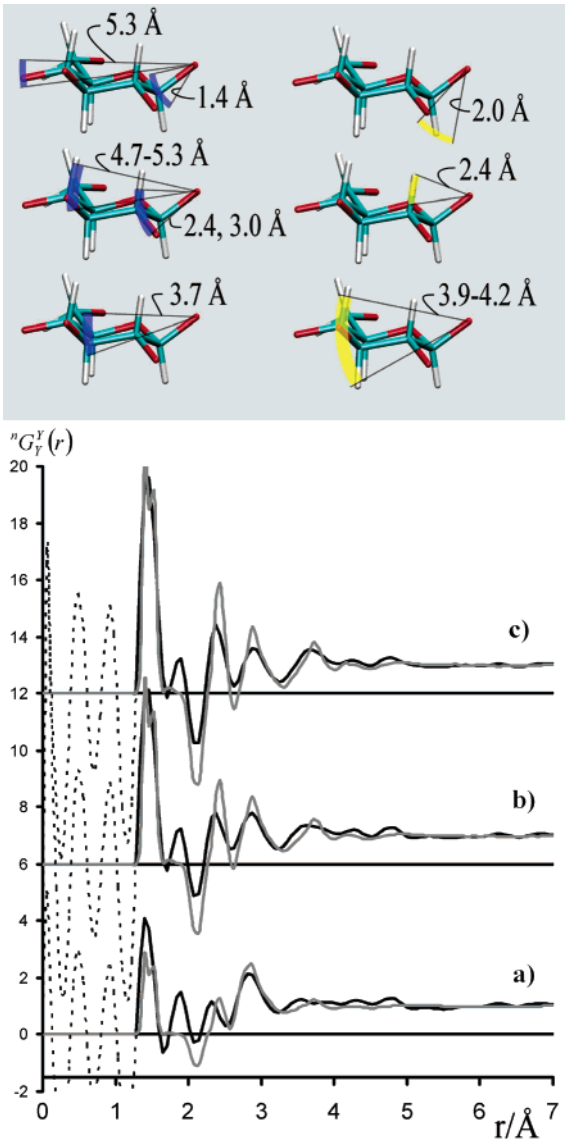

Figure 5. ${}^nG_Y^Y(r)$ for 1, 3, and 5 *m* solutions (a, b, and c, respectively). In each case, the NDIS data is shown in black while the distributions calculated from MD simulations are shown in gray. The figures at the top show the origins of the glucose molecule correlations, with positively weighted correlations being indicated in blue and negatively weighted correlations being shown in yellow. As the structure factors contain no correlations to H_{ex}, these have been removed from the glucose molecules for clarity.

contain significant errors. It is more likely that any excessive rigidity arises from the torsional degrees of freedom. Indeed, the original version of this force field has already been revised to make glucose more flexible around the exocyclic C5−C6 dihedral angle.[14] It is possible that the force field remains too rigid in various torsional terms, including the hydroxyl rotations and ring torsions (e.g., O−C−C−O), as well as for the exocyclic primary alcohol rotations.

It is unfortunate that the data obtained from diffraction experiments with isotopic substitutions on a large, asymmetric solute like glucose are too complex to be interpreted in terms of the detailed first and second shell water structuring around the solute, since the scattering is averaged over all of the distinct sites in the molecule, each of which will be somewhat different in its hydration characteristics due to the local topological structure. However, the results of MD simulations can be used to map the density of solvent water around complex solutes such as glucose, as we have demonstrated in previous studies

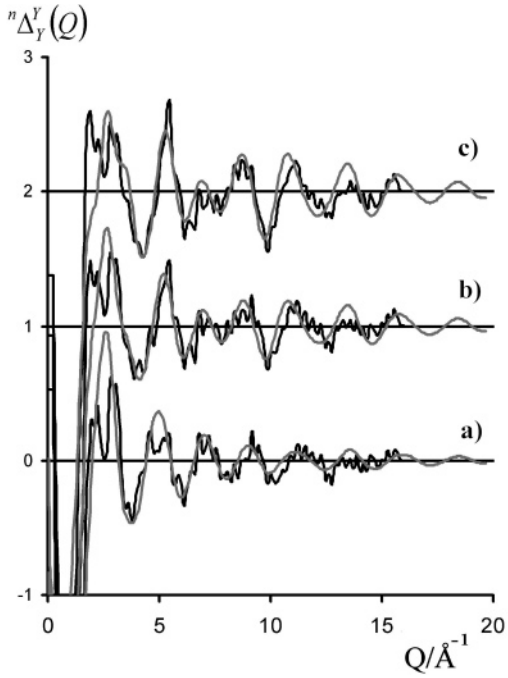

Figure 6. . Back-transformation of the NDIS and MD data in the range 1.8−20 Å^{-1} (regeneration of the experimental data recorded from the diffractometer). If the NDIS data were resolution-limited, then the two data sets would look identical up to the *Q*-limit of the machine (approximately 16 Å^{-1} in this case). The fact that there is a significant discrepancy between the back-transformed data at *Q*-ranges of less than the experimental range of the machine shows that the experimental data is not resolution-limited.

of this and other sugar molecules.[2,16,17] Such a mapping was carried out for both types of anomers in the present simulations. In general, the density contours were the same as in our previous studies, indicating that the much higher concentrations simulated here do not significantly perturb the first shell solvent structuring for the individual molecules, even at the highest concentration. The density distribution patterns are statistically converged, since not only do they reproduce the distributions found earlier for shorter simulations of smaller systems but they are also the same for the different independent sugar molecules and for shorter portions of the total simulation time. Figure 7 displays the average density of O_W atoms around the β sugar anomers as calculated from the 1, 3, and 5 *m* solutions. Even in the 5 *m* solution, where there are only about 10 water molecules for each sugar molecule, so that many waters are shared between two sugars, the strict structural requirements of the water−sugar hydrogen bonds and their packing dictate similar hydration patterns. As in previous studies, water molecules tend to straddle between adjacent hydroxyl groups. In addition, the hydrophobic "top" and "bottom" surfaces, dominated the non-hydrogen-bonding aliphatic protons, structured water molecules more weakly and with a different orientational pattern. The total average numbers of hydrogen bonds (see definition below) for the β anomer were found to be 10.80, 10.69, and 10.68 for the 1,3, and 5 *m* solutions, respectively, and the total hydration numbers for these sugars, being those water molecules on average in van der Waals contact, were 20.67, 18.86, and 17.01 for the three concentrations. For the α anomer, these numbers were 10.47, 10.42, and 10.38 for the hydrogen bond partners and 19.41, 18.05, and 16.45 for the total hydration numbers.

The simulations did not find any significant preferential tendency for the sugar molecules to interact with one another, even at the highest concentrations studied, as would be expected from the fact that even a 5 *m* concentration is below the

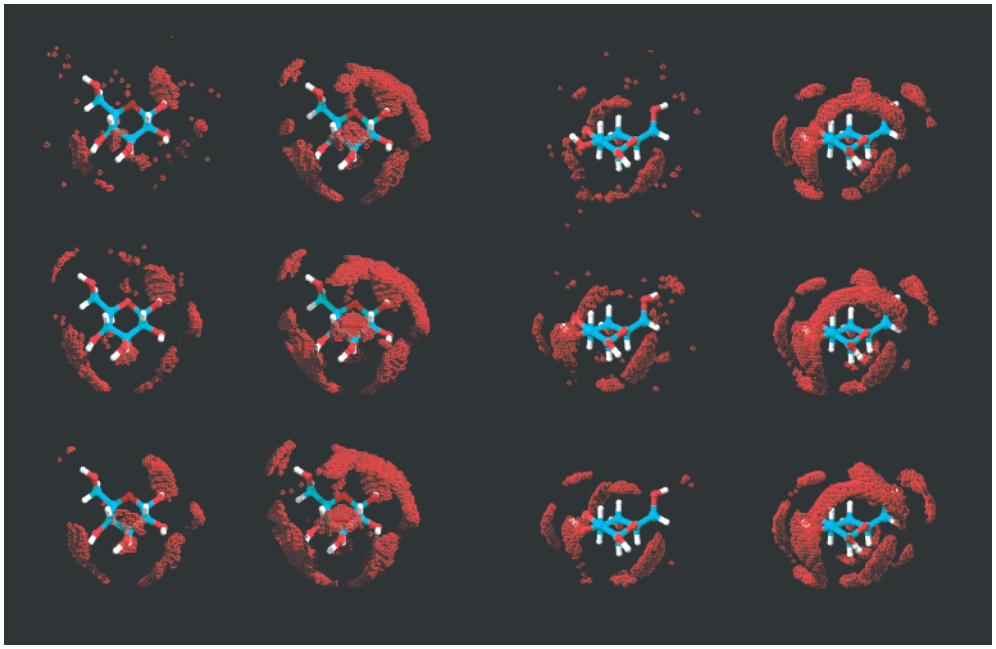

Figure 7. . O_w and O_G densities around β-D-glucopyranose as calculated from the MD trajectories (α-D-glucopyranose exhibits the same general trends). The first two columns and last two columns are different views of the same data. The first column is 4.5 times the average number density of O_G while the second column is the 3.5 times the average number density of O_W. The first, second, and third rows are the 1, 3, and 5 *m* solutions, respectively. These figures illustrate the preference of glucose to interact with O_W rather than O_G.

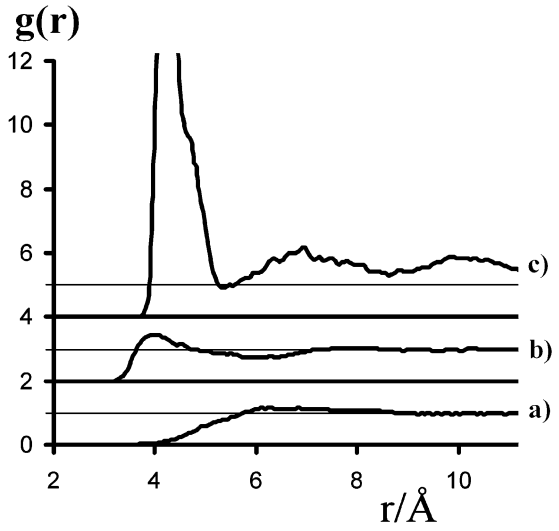

Figure 8. . (a) Function $g_{gluc-C}(r)$ in 5 *m* glucose solution (exhibiting nonassociation), where "gluc" here represents the average coordinate of the heavy ring atoms C1−C5 and O5. (b) $g_{CC}(r)$ in 3.0 *m* guanidinium chloride solution (weak pairing). (c) $g_{SC}(r)$ in 1.5 *m* guanidinium sulfate solution (exhibiting strong ion pairing).

solubility limit for glucose at this temperature. However, at these concentrations there was inevitably some direct sugar−sugar interaction, and when this occurred, it was found to be dominated by hydroxyl−hydroxyl hydrogen bonding, similar to the sugar−water interactions. Figure 7 also displays the density of O_G atoms, from neighboring solute molecules, contoured in the same fashion as for the water oxygen atoms. As can be seen, the pattern is very similar to that observed for the O_W atoms, but at a significantly lower density, reflecting the tendency for the sugar molecules to preferentially interact with water.

Figure 8 presents the radial distribution function for sugar carbon atoms around the geometric center of the molecule. This function displays little useful structure due to the extensive orientational averaging involved in calculating such a function for an extended asymmetric solute like this sugar. Previously, we have shown that species such as guanidinium sulfate that strongly associate in solution display characteristic structural signatures observed in the NDIS data and predicted by MD simulations.[18] On the basis of this experience, it might be expected that if glucose was strongly associating in solution, then this would be manifested structurally as a concentration-dependent signature in the NDIS data (principally observed as a peak in ${}^{n}G_{Y}^{Y}(r)$ at $Q < 5$ Å). The lack of such an observation implies that the NDIS experimental data supports the MD observation that glucose molecules do not significantly associate in solution even at these relatively high concentrations.

Another advantage of the extraordinary detail available in MD simulations is that the trajectories can be examined for information about whether the hydrogen bonding to water is preferentially as hydrogen bond donor or acceptor, since both water and the sugar hydroxyl groups can serve in either capacity. In principle, sugar hydroxyl groups can make two hydrogen bonds as an acceptor through the oxygen atom and one as a hydrogen bond donor, although the actual number of hydrogen bonds made to water is always less than this number,[19] probably due to the interference of the hydration requirements of adjacent groups and to intramolecular hydrogen bonding. For the purpose of analyzing the hydrogen bonding in our systems, an appropriate donor−acceptor pair are defined as hydrogen-bonded if their heavy atom distance is less than 3.5 Å and the O−H···O bond angle is greater than 120°. Table 3 lists the average number of hydrogen bonds of each type to and from water as found in the simulations, and Table 4 gives the number of hydrogen bonds to and from the sugar molecules. As can be seen from Table 3, in mixed water−sugar hydrogen bonds, the water molecule is more likely to be the hydrogen bond donor. The glucose hydroxyl groups are essentially equivalent to water as hydrogen bond donors but are poorer hydrogen bond acceptors. These differences probably arise because the sugar hydroxyl hydrogen atom charge of 0.43 is closer to the modified TIP3P water model of 0.417[5] than the oxygen charge of −0.66 is to the water value of −0.834.

TABLE 3: Hydrogen Bonds to and from Water in Various Glucose Systems[a]

	1 m		3 m		5 m	
	water as H bond acceptor	water as H bond donor	water as H bond acceptor	water as H bond donor	water as H bond acceptor	water as H bond donor
from glucose	0.0786	0.1253	0.2333	0.3253	0.3292	0.5071
from water	1.902	1.9028	1.708	1.7271	1.5912	1.5672
total	1.9806	2.0281	1.9413	2.0524	1.9204	2.0743
fraction to glucose	0.040	0.062	0.120	0.158	0.171	0.244
fraction to water	0.960	0.938	0.880	0.842	0.829	0.756

[a] For the definition of what is counted as a hydrogen bond, see text.

TABLE 4: Total Number of Hydrogen Bonds to Glucose, Broken Down as Donors and Acceptors

	α-D-glucopyranose			β-D-glucopyranose			
	acceptor	donor	total	acceptor	donor	total	Δ(β−α)
			1 *m*				
to water	6.00	4.27	10.27	6.46	4.34	10.80	0.53
to glucose	0.34	0.37	0.71	0.21	0.27	0.48	−0.23
sum	6.34	4.64	10.98	6.66	4.61	11.27	0.29
			3 *m*				
to water	5.43	3.86	9.29	5.74	3.90	9.64	0.35
to glucose	0.76	0.87	1.33	0.75	0.75	1.50	0.17
sum	6.19	4.73	10.91	6.49	4.65	11.15	0.24
			5 *m*				
to water	4.94	3.58	8.52	5.36	3.58	8.94	0.42
to glucose	1.11	1.20	2.31	1.06	1.13	2.19	−0.12
sum	6.06	4.78	10.83	6.42	4.71	11.13	0.30

It can be seen from Table 4 that the total number of hydrogen bonds made by sugar molecules is essentially the same at all three concentrations studied. As the concentration increases, fewer of these hydrogen bonds are to water and more are to other sugar molecules, but the hydrogen-bonding requirements of the sugar remain satisfied. As has been found before, there is a consistent tendency for the β anomer to make more hydrogen bonds to water than the α anomer. This equatorial anomer also makes more hydrogen bonds to other glucose molecules as well. Since hydrogen bonds of this type typically have an energy of approximately 5 kcal/mol, this difference of from 0.3 to 0.5 hydrogen bonds favors the β anomer by as much as 1.5−2.5 kcal/mol. This result is entirely consistent with previous studies in this laboratory[2,16] and others.[20,21] This hydration preference balances out the internal stereoelectronic anomeric effect favoring the α anomer and explains why the β anomer is the slightly preferred form in solution.

Table 5 breaks down the hydrogen-bonding behavior of each individual oxygen atom, including the ring oxygen atom, in both anomers of the glucose solutes. As would be expected, the main difference in the hydrogen-bonding behavior of these two anomers is in the number of hydrogen bonds to the anomeric hydroxyl and the adjacent ring oxygen atom. The β isomer forms approximately 30% more hydrogen bonds as an acceptor at these atoms than the α anomer. There is also a smaller but still significant difference between the two anomers at the O2 atom but with the α anomer forming approximately 7% more hydrogen bonds than the β anomer.

From the MD simulations, it was found that despite the increase in the barrier height for hydroxyl rotation over that in our previous force field all of the sugar hydroxyl groups rotated many times over the course of the simulation. The hydroxyl groups OH1 and OH6 for both anomers occupied for the most part one of the three staggered minima, while the O2, O3, and O4 hydroxyl groups mainly occupied only two of these positions in the ratio of approximately 2:1, with the third conformer essentially unoccupied.

TABLE 5: Number of Hydrogen Bonds to Each Individual Oxygen Atom on Glucose

	O1	O2	O3	O4	O5	O6
1 m						
α-D-Glucose as Hydrogen Bond Donor						
H bonds to glucose	0.081	0.058	0.073	0.080	0.000	0.093
H bonds to water	0.952	0.900	0.929	0.909	0.000	0.944
total H bonds	1.032	0.957	1.002	0.989	0.000	1.036
fraction of H bonds to glucose	0.078	0.060	0.073	0.081	0.000	0.089
fraction of H bonds to water	0.922	0.940	0.927	0.919	0.000	0.911
β-D-Glucose as Hydrogen Bond Donor						
H bonds to glucose	0.078	0.040	0.052	0.059	0.000	0.051
H bonds to water	0.975	0.927	0.963	0.932	0.000	0.992
total H bonds	1.053	0.967	1.015	0.991	0.000	1.043
fraction of H bonds to glucose	0.074	0.041	0.051	0.059	0.000	0.049
fraction of H bonds to water	0.926	0.959	0.949	0.941	0.000	0.951
α-D-Glucose as Hydrogen Bond Acceptor						
H bonds to glucose	0.049	0.065	0.102	0.047	0.028	0.080
H bonds to water	1.219	1.418	1.117	1.030	0.635	1.430
total H bonds	1.268	1.482	1.219	1.077	0.663	1.511
fraction of H bonds to glucose	0.039	0.044	0.084	0.043	0.042	0.053
fraction of H bonds to water	0.961	0.956	0.916	0.957	0.958	0.947
β-D-Glucose as Hydrogen Bond Acceptor						
H bonds to glucose	0.037	0.052	0.037	0.028	0.027	0.062
H bonds to water	1.636	1.332	1.222	1.017	0.869	1.418
total H bonds	1.672	1.383	1.259	1.045	0.896	1.480
fraction of H bonds to glucose	0.022	0.037	0.029	0.027	0.030	0.042
fraction of H bonds to water	0.978	0.963	0.971	0.973	0.970	0.958
5 m						
α-Glucose as Hydrogen Bond Donor						
H bonds to glucose	0.27	0.22	0.22	0.23	0.00	0.21
H bonds to water	0.74	0.72	0.75	0.74	0.00	0.82
total H bonds	1.01	0.94	0.97	0.97	0.00	1.03
fraction of H bonds to glucose	0.27	0.23	0.23	0.24	0.00	0.21
fraction of H bonds to water	0.73	0.77	0.77	0.76	0.00	0.79
β-Glucose as Hydrogen Bond Donor						
H bonds to glucose	0.30	0.24	0.23	0.16	0.00	0.24
H bonds to water	0.71	0.71	0.74	0.82	0.00	0.78
total H bonds	1.01	0.95	0.97	0.97	0.00	1.02
fraction of H bonds to glucose	0.29	0.25	0.23	0.16	0.00	0.24
fraction of H bonds to water	0.71	0.75	0.77	0.84	0.00	0.76
α-Glucose as Hydrogen Bond Acceptor						
H bonds to glucose	0.17	0.22	0.25	0.17	0.11	0.18
H bonds to water	1.00	1.14	0.88	0.82	0.43	1.15
total H bonds	1.16	1.37	1.12	0.99	0.54	1.34
fraction of H bonds to glucose	0.14	0.16	0.22	0.17	0.20	0.14
fraction of H bonds to water	0.86	0.84	0.78	0.83	0.80	0.86
β-Glucose as Hydrogen Bond Acceptor						
H bonds to glucose	0.22	0.25	0.21	0.16	0.09	0.25
H bonds to water	1.32	1.02	0.93	0.79	0.65	1.07
total H bonds	1.54	1.28	1.15	0.94	0.74	1.32
fraction of H bonds to glucose	0.14	0.20	0.19	0.17	0.13	0.19
fraction of H bonds to water	0.86	0.80	0.81	0.83	0.87	0.81

Conclusions

The new MD simulations reported here were undertaken for use in making direct comparisons with experimental data obtained from NDIS experiments as a test of the reliability of

the various approximations in the simulations, including primarily the MM force fields. The calculated anisotropic solvent distributions were found to be substantially the same as those found previously from simulations of a single sugar molecule in a water box,[2,16,17] corresponding to the unrealizable state of infinite dilution at a small finite concentration. Similar solvent distributions have also been found by others using different force fields,[22] indicating that the distributions probably result from the geometric constraints of hydrogen-bonded packing. The present simulations at high concentrations demonstrate that this first hydration shell structuring, dominated by hydrogen bonding, is not significantly perturbed even at high concentrations (5 *m*). In agreement with our previous analysis of the scattering data, the long-range structure of the water was not found to be significantly affected by the presence of the sugar solute. This is not to say that the water is not "structured" relative to the molecular architecture of the solute but rather that the water spacing imposed by the sugar is not substantially different from that of pure bulk water. In addition, the interaction of first shell water molecules with the solute, primarily with sugar hydroxyl groups, is sufficiently like hydrogen bonds to water as to be indistinguishable in the radial pair correlation functions.

In general, the correspondence of the calculated solvent distribution functions with the experimental ones is satisfactory and gives confidence that the modeling can be used to study the details of solution structuring that cannot be probed experimentally. However, in the case of the function ${}^{n}G_{\mathrm{Y}}^{\mathrm{Y}}(r)$, there are observed discrepancies between the calculated and the experimental data that cannot be attributed solely to quantum mechanical uncertainty or resolution limitations in the diffraction data. The results imply that the simulated glucose molecule is more rigid than the actual glucose molecules in the experiments, possibly due to ring torsional terms that may be too high. Although several parameters may need revision, the term that is most likely to be too rigid is the OH−CT−CT−OH term in the ring structure. The O−C−C−O atomic sequence is subject to the stereoelectronic gauche effect, which has already necessitated revision of the exocyclic torsional term too make it less rigid by lowering the transition barrier. However, at the time that this correction was made, no comparable revision was made to the similar ring torsional term. This intramolecular energy term may also be affected by electronic polarization due to water molecules in the first hydration shell. In principle, this and other terms could be arbitrarily adjusted, and the simulations could be repeated until the agreement with the diffraction data was deemed satisfactory. This approach has been referred to as empirical potential structure refinement (EPSR) by Soper.[23] This method has long been used in the development of empirical force fields for water.[4,24,25] In the case of complex molecular species such as glucose, however, the information available from diffraction is sparse relative to the large number of parameters that must be fit. A better approach would seem to be to revise the terms in question using improved ab initio calculations and experimental data for sugars and small-molecule analogues (if available) and then to let the comparison with the diffraction data serve as an independent test of the quality of the parameter set. We hope to pursue this approach soon.

Acknowledgment. The authors gratefully acknowledge the assistance of G. Cuello and P. Palleau of the ILL. This project was supported by grant GM63018 from the National Institutes of Health.

References and Notes

(1) Mason, P. E.; Neilson, G. W.; Enderby, J. E.; Saboungi, M.-L.; Dempsey, C. E.; MacKerell, A. D.; Brady, J. W. *J. Am. Chem. Soc.* **2004**, *126*, 11462−11470.
(2) Ha, S.; Gao, J.; Tidor, B.; Brady, J. W.; Karplus, M. *J. Am. Chem. Soc.* **1991**, *113*, 1553−1557.
(3) Mason, P. E.; Neilson, G. W.; Barnes, A. C.; Enderby, J. E.; Brady, J. W.; Saboungi, M.-L. *J. Chem. Phys.* **2003**, *119*, 3347−3353.
(4) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926−935.
(5) Neria, E.; Fischer, S.; Karplus, M. *J. Chem. Phys.* **1996**, *105*, 1902−1919.
(6) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187−217.
(7) van Gunsteren, W. F.; Berendsen, H. J. C. *Mol. Phys.* **1977**, *34*, 1311−1327.
(8) Kanchanawong, P.; Brady, J. W. Unpublished work.
(9) Franks, F. In *Water: A Comprehensive Treatise*; Franks, F., Ed.; Plenum Press: New York, 1973; Vol. 2, pp 1−54.
(10) Tromp, R. H.; Neilson, G. W.; Soper, A. K. *J. Chem. Phys.* **1992**, *96*, 8460−8469.
(11) Kabayama, M. A.; Patterson, D. *Can. J. Chem.* **1958**, *36*, 563−573.
(12) Enderby, J. E. Unpublished work.
(13) Palma, R.; Zuccato, P.; Himmel, M. E.; Liang, G.; Brady, J. W. In *Glycosyl Hydrolases in Biomass Conversion*; Himmel, M. E., Baker, J. O., Saddler, J. N., Eds.; ACS Symposium Series 769; American Chemical Society: Washington, DC, 2000; pp 112−130.
(14) Kuttel, M.; Brady, J. W.; Naidoo, K. J. *J. Comput. Chem.* **2002**, *23*, 1236−1243.
(15) Ha, S. N.; Giammona, A.; Field, M.; Brady, J. W. *Carbohydr. Res.* **1988**, *180*, 207−221.
(16) Schmidt, R. K.; Karplus, M.; Brady, J. W. *J. Am. Chem. Soc.* **1996**, *118*, 541−546.
(17) Liu, Q.; Brady, J. W. *J. Am. Chem. Soc.* **1996**, *118*, 12276−12286.
(18) Mason, P. E.; Dempsey, C. E.; Neilson, G. W.; Brady, J. W. *J. Phys. Chem. B*, to be submitted for publication.
(19) van Eijck, B. P.; Kroon-Batenburg, L. M. J.; Kroon, J. *J. Mol. Struct.* **1990**, *237*, 315−325.
(20) Senderowitz, H.; Parish, C.; Still, W. C. *J. Am. Chem. Soc.* **1996**, *118*, 2078−2086.
(21) Zuccarello, F.; Buemi, G. *Carbohydr. Res.* **1995**, *273*, 129−145.
(22) Sidhu, K. S.; Goodfellow, J. M.; Turner, J. Z. *J. Chem. Phys.* **1999**, *110*, 7943−7950.
(23) Soper, A. K. *Mol. Phys.* **2001**, *99*, 1503−1516.
(24) Stillinger, F. H.; Rahman, A. *J. Chem. Phys.* **1974**, *60*, 1545−1557.
(25) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; Hermans, J. In *Intermolecular Forces*; Pullman, B., Ed.; Reidel Publishing: Dordrecht, The Netherlands, 1981; pp 331−342.