Published in final edited form as:
Biopolymers. 2012 May ; 97(5): 276–288. doi:10.1002/bip.22017.

Sucrose in Aqueous Solution Revisited: 1. Molecular Dynamics Simulations and Direct and Indirect Dipolar Coupling Analysis

Junchao Xia and **David A. Case**
BioMaPS Institute for Quantitative Biology and Department of Chemistry and Chemical Biology, Rutgers University, 610 Taylor Rd., Piscataway, NJ 08854

Abstract

Although the crystal structure of the disaccharide sucrose was solved more than 30 years ago, its conformational distribution in aqueous solution is still a matter of debate. We report here a variety of molecular dynamics simulations (mostly of 100 ns) using the GLYCAM06 force field and various water models, paying particular attention to comparisons to NMR measurements of residual dipolar couplings and electron-mediated spin-spin couplings. We focus on the glycosidic linkage conformation, the puckering phase angle of the fructose ring, and intramolecular hydrogen bonds between the two sugars. Our results show that sucrose is indeed a dynamic molecule, but the that crystal conformation is qualitatively the dominant one in dilute solution. A second conformational basin, populated in many force fields, is probably over-stabilized in the calculations.

1. Introduction

Sucrose, or table sugar, is a disaccharide with the sequence as α-D-Glu-(1→2)-β-D-Fru (see Figure 1). Although its crystal structure was reported in 1973,[1] the conformation in aqueous solution, particularly with respect to the glycosidic linkage, is still being investigated. For example, an early nuclear magnetic resonance (NMR) study of sucrose in D_2O and dimethyl sulfoxide-d_6 solvents by Bock and Lemieux [2] supported a fairly rigid conformation in solution, close to crystal structure, based on the measurements of coupling constants, chemical shifts, T_1 relaxation times and nuclear Overhauser effects (NOE). This conformational rigidity was reinforced by several other NMR results, such as ^{13}C relaxation as a function of concentration, temperature, and magnetic field strength,[3,4] relaxation rates in a viscous cryosolvent (D_2O–$DMSO$),[5,6] and residual dipolar couplings (RDC) in dilute liquid crystal media.[7]

Other NMR studies[8–13] have questioned this single-conformation conclusion. Steady state nuclear Overhauser effects and J coupling constant studies by Perez and coworkers[8] suggested that conformational averaging has to be included to give a good agreement between theoretical and experimental data. Poppe and Halbeek[9] found that the interglycosidic proton-proton distances derived from NOE measurements have both temperature and magnetic field dependence which is not consistent with a rigid glycosidic linkage. Residual dipolar couplings[13] were best fit by a model where the sugar rings have different alignment tensors, which can not be explained by a single structure. The flexibility of sucrose in aqueous solution has also been supported by optical rotation experiments[14] where data were better understood by the equilibrium mixture of two conformers including the crystalline structure. Inter-residue J coupling analysis[11] also showed that the second dihedral of glycosidic linkage (ψ) might sample two energy minima states.

The ambiguities of NMR experiments mostly stem from the fact that interpretation of NMR observables requires structural models, especially for flexible molecules. In principle,

computer simulations allow a detailed description of conformational transitions,[8,15–22] but results can depend critically on the force field used, and published results for sucrose vary a lot. Tran and Brady's early molecular dynamics studies[15,16] found the global minimum energy conformation to be dynamically stable with no transitions away from this minimum during short (20 ps) trajectories. Later simulations (of 1 ns duration) were consistent with this model, and suggested that no direct intramolecular hydrogen bond persists in aqueous solution.[18,19] Instead the sucrose conformation is stabilized by the dynamic presence of two bridging water molecules between residues. However, other molecular modeling results combined with NMR quantities are not consistent with this single minimum.[8]

The accuracy of computer simulations, however, depends on the accuracy of force fields, the reliability of solvation models, and the difficulty of sampling all relevant conformations in a finite simulation.[23–29] Several carbohydrate force fields[30–38] are available for molecular mechanics simulations of sugars. We have chosen the GLYCAM06 model, designed for both simple monosaccharides and complex sugars and glycoproteins.[30] There also exists a large variety of water models for MD simulations. A recent review[39] compared 47 major water models and found that all these models have their advantages and deficiencies. In addition, early simulations[15–22] were limited to sub-nanosecond time scales. Recent advances in computing power, however, have made feasible direct simulations of biomolecules in explicit water to hundreds of nanoseconds or microseconds.

Using the AMBER 10 molecular simulation package[40] and the GLYCAM 06 carbohydrate force field,[30] we performed 100 ns MD simulations of sucrose in aqueous solution under various conditions. We evaluated these simulation trajectories via calculation of residual dipolar couplings and indirect spin-spin couplings from simulation snapshots. We mainly focused on the global conformation change from the glycosidic linkage although we also consider the analysis of local fluctuations such as intramolecular hydrogen bond and sugar puckering. In a second paper[41] we will report systems with varying concentrations of sucrose in water and water-DMSO mixtures, for comparison with NMR relaxation measurements.

2. Computational Methods

1. Molecular Dynamics Simulation

Explicit solvent MD simulations were carried out using AMBER 10.[40] We used GLYCAM06,[30] in combination with several popular explicit water models including SPC,[42] TIP3P,[43] TIP4P,[44] TIP4P/Ew,[45] and TIP5P.[46] Most simulations were carried out at 300K; for the TIP4P/Ew model, additional simulations were made at 273, 283 and 315K. In each case sucrose (in its crystal conformation[1]) was solvated in a truncated octahedron of water molecules, such that the closest distance between any atom of the sucrose and the edge of the periodic box was 15 angstroms. All production simulations (100 ns) were under periodic boundary conditions in the constant volume, total energy and number of particles (NVE) ensemble. Before the production runs, equilibration processes in NVT (0.5 ns) and NPT (0.5 ns) ensemble were performed to adjust the systems into the desired temperatures and volumes. The integration time step was 0.001 ps. In all simulations, SHAKE algorithm was used to constrain bonds involving hydrogen. The non-bonded cutoff for evaluating electrostatic and van der Waals forces was set to 0.8 nm with no scaling of 1–4 interactions. To deal with long range electrostatic interactions the PME algorithm was used with the default settings, including a real space grid of 0.1 nm. We collected 100,000 conformation snapshots along each of the 100 ns trajectories.

Restrained MD simulations[47] (10 ns) were performed to analyze intramolecular hydrogen bonds as the function of glycosidic φ–ψ dihedrals and the puckering phase angle of fructose.

We restrained the phase angle (via the five dihedrals from the fructose ring according to the Altona-Sundaralingam scheme [48] as

$$\theta_j=\theta_m\cos(P+4\pi(j-1)/5), \quad (1)$$

where θ_m is an amplitude (≈ 40° from experiment[49]). P is phase angle, and j = 0, 1, 2, 3 and 4. 10 ns restrained MD simulations were carried out by retaining the phase angle (P = 0) and glycosidic dihedrals of four different free energy minima. We also did simulations with the same glycosidic dihedral restraints but various phase angle with P = 0, 18, 36, 54, 72 and 90. For all restraints we chose a force constant of 100 $kcal/Rad^2$.

2. Calculation of Residual Dipolar Couplings

Residual dipolar couplings (RDCs) arise in liquid crystalline media that partially align solute molecule so as to slightly alter the orientational averaging that eliminates direct dipolar couplings in isotropic solution. While it is always possible that the liquid crystal environment changes the conformational distribution, spectral similarities with the pure water case argue against this.[7] The simplest analysis of RDCs is based on the assumption that the liquid crystal molecules have only steric alignment effects and do not change the internal structure of the sugar. Due to the alignment effect, sugar molecules have certain order and after transforming to the space defined by the diagonalized alignment tensor the residual dipolar coupling can be written[50,51]

$$D_{ij}=-\left(\frac{\mu_0}{16\pi^2}\right)\frac{\gamma_i\gamma_j\hbar}{r_{ij}^3}\kappa[S_{zz}(3cos^2\theta_{ij}^z-1)+(S_{xx}-S_{yy})(cos^2\theta_{ij}^x-cos^2\theta_{ij}^y)], \quad (2)$$

where r_{ij} is the internuclear distance between spins, γ_i and γ_j are the gyromagnetic ratios of spins i and j. $S_{\alpha\alpha}$ are diagonal components of alignment tensor in the molecular frame and describe the transformation from the molecular frame of reference to the laboratory coordinate system of the liquid crystal medium. κ is a scaling factor related to internal motion, and θ_{ij}^{α} (α=x, y, and z) are the angles between the spin-spin vectors and the molecular coordinate frame defined by the diagonalized alignment tensor.

Experimental RDC data can be computed from a molecular model by treating the elements of the alignment tensor as parameters to be optimized; this leads to a set of linear equations that are usually solved by a singular value decomposition (SVD) method.[52] Such a fit, however, may be misleading if the assumed structure model is incorrect, and we show below some of the over-fitting problems that can arise. As an alternative, the alignment tensor can be estimated from the structure of a molecule and its postulated interactions with the alignment medium. Even though such models have limited accuracy, they do ensure that the alignment tensors are qualitatively consistent with the molecular model, which need not be the case if the tensor elements are determined only by a best-fit criteria to experiment.

We used two such alternative methods to estimate alignment tensors. The PALES program[53] calculates the alignment tensor by a Monte Carlo simulation of molecules randomly orientated around the area close to an infinite two dimensional plate. A second method estimates the alignment tensor from the radius of gyration tensor[54] or moment of inertia.[55,56] For the former, the diagonalized alignment tensor is estimated as:

$$(S_{xx},S_{yy},S_{zz})=(-\frac{1}{2}-\frac{1}{2}\delta,\delta-\frac{1}{2},1-\frac{1}{2}\delta), \quad (3)$$

where

$$\delta=\frac{\rho_{yy}-\rho_{xx}}{\rho_{zz}-\rho_{xx}}. \tag{4}$$

$\rho_{\alpha\alpha}$ ($\rho_{zz}>\rho_{yy}>\rho_{xx}$) represent characteristic lengths of the ellipsoid shape of a sugar molecule and can be obtained from the square roots of the eigenvalues of the radius of gyration tensor of the whole molecule defined as

$$R_{ij}^2=\frac{1}{N}\sum_{r=1}^{N}x_i^r x_j^r, \tag{5}$$

where x_i^r are the positions of the N atoms in the molecule. Note the molecular frame in Equation 2 corresponds to the eigenvectors of Equation 5.

One would not expect these methods to work so well on a single sugar residue, which is not very anisotropic in shape; but they should work better for a disaccharide, which has a more elongated shape and hence a larger and more obvious anisotropy. Estimation of alignment tensor from the gyration tensor has been applied to construct conformational ensembles of unfolded proteins by the Blackledge group[57] and to generate statistical coil model of unfolded ensemble of proteins by Freed and coworkers.[58] More recently, the Forman-Kay group[59] has used the local alignment from a fragment size of 15 residues to derive disordered state ensembles of proteins instead of global alignment from the whole molecule. The alignment tensor estimated from gyration tensor also has been proposed to refine ensembles of unstructured proteins combining with MD simulations.[60]

To evaluate the conformations sampled in MD trajectories we calculated an R factor[61] between the computed RDCs for each snapshot and those reported experimentally:

$$R=\sqrt{\frac{\sum_i^N(Q_i-Q_{0i})^2}{\sum_i^N {Q_{0i}}^2}}, \tag{6}$$

Here Q_i is our calculated RDC value for the i th spin nuclear pair and Q_{0i} is the related experimental value. The summation is over all available N experimental RDCs.

3. Calculation of indirect spin-spin coupling constants

Spin-spin couplings across the glycosidic torsion were estimated using a Karplus relation:

$$J(\theta)=A\cos^2\theta+B\cos\theta+C \tag{7}$$

where θ is a C–O–C–H or C–O–C–C torsion angle connecting the two sugar rings, and the constants (A=7.49, B=−0.96, C=0.15 for C–O–C–H, A=3.70, B=0.18, C=0.11 Hz for C–O–C–C) were taken from fits to DFT data by Cloran *et al.*.[62] As with the RDC calculations, calculations were made for each of the 100,000 snapshots spanning 100ns of simulation.

3. Results and Discussion

1. Free energy landscapes about the glycosidic linkage

We begin by showing the free energies as a function of the glycosidic torsion angles; these were computed from the observed probabilities from 100 ns MD simulations:

$$F(\varphi,\psi) = -k_b T \ln P(\varphi,\psi) \tag{8}$$

where k_b is Bolzmann's constant and T is temperature. For convenience we also shifted all surfaces to make the values of global minima equal to 0. Figure 2 shows the free energy landscape in φ–ψ glycosidic space for the conformations of sucrose at 300K in the TIP4PEW water model; similar figures for other water models are given in the Supplementary Material. [In the Appendix, we use metadynamics and umbrella sampling to explore convergence of these landscapes, showing that at these temperatures the raw results from 100 ns simulations are converged well enough for our purposes here. These similarities also support our expectation that the NVE results shown in Fig. 2 are close to those one would obtain for a canonical (NVT) simulation, as expected for a system with a large number of degrees of freedom.] For each water model, four important minima, M1 (105°, 300°), M2 (70°, 280°), M3 (85°, 195°), and M4 (90°, 50°), can be identified. The exact φ–ψ values of these minima are listed in Table 1 and change little with water model. As we will show in Section 3, M1 is closely related to S1 and S2 conformer, M3 to S3 and S4, and M4 to S5 in the early computational work from Brady's group.[15,16] These four minima have also been found by Freedberg and coworkers,[13] where the potential energy landscapes were constructed by a systematic grid search with CHARMM or AMBER-type carbohydrate force fields.[34,33,35,36] Earlier work by French *et al.*,[21,22] combining *ab initio* quantum mechanics and molecular mechanics, found similar minima. Hence, the general shape of the free energy landscapes should be reliable, but the relative populations of the four basins depends on the details of the force field, and needs confirmation from experimental data.

Table 1 also displays the well depths and barriers for the four minima. M1 is the global minimum for all five water models and its position is close to the crystal conformation (108°, 305°).[1] The relative order of these four minima is the same for four water models (SPC, TIP3P, TIP4P, and TIP4P/Ew), but is qualitatively different for TIP5P; we don't understand why TIP5P is an outlier here, and this probably deserves more investigation. We also integrated the probability densities within a circular region of radius of 20° around the four minima, and denoted the probabilities as P1, P2, P3 and P4 in Table 1. The probability of global minima M1 has the majority of around 60% and M2 takes more than 30% with small percentages of M3 and M4. This is similar for all models except TIP5P for which P1 only has 34% and there are significant populations of M3 and M4 (more than 15% for each). The free energy barriers (B12, B23, and B14) between different minima are also shown in Table 1. All the barriers are less than 4 kcal/mol which means that all these minima should interconvert on the simulation time scale of 100 nanoseconds. The orders of barriers are the same ($B12 < B14 < B23$) for all water models except TIP5P. We also notice that the barrier (B12) between M1 and M2 is rather small (≈ 0.5 kcal/mol) for the four water models except TIP5P. These small values of B12 (and low energies for M2) are consistent with other MM[16,17,20] and QM calculations.[21] The barriers of B23 and B14 for TIP5P model are much smaller than that of other four models and result in much larger populations of M3 and M4. The relative energy differences and barrier heights between minima are quite sensitive to the force field: surveys of potential energy surface by Freedberg and coworkers[13] showed that these quantities vary markedly in the CHARMM22, CSFF, Homans, and Accelrys force fields, although the φ–ψ values of the local minima are in good qualitative agreement. For this reason, careful comparison to experiment is an important adjunct to force field simulations.

Since not all of the relevant experiments are carried out at 300 K,[8,9,10,11,12,13] we also computed free energy surfaces at different temperatures (273, 283, 300, and 315K). These are discussed in the Table 1S and Figure 2S of Supplementary material. The qualitative features are the same as those shown in Figure 2.

2. RDC Analysis of Conformation Snapshots

The trajectories explore many possible configurations of sucrose, and we have computed the RDCs for each of these. Figure 3 displays the best 500 and worst 500 conformations of sucrose among the 100,000 snapshots extracted from the 100 ns simulation trajectory with TIP4P/Ew waters at T = 315 K. We obtained the alignment tensor of each snapshot first from the molecular shape via the approximation of Almond[54] or PALES,[53] and then calculated the RDCs of each conformation. [In a separate calculation (Supporting material) we divide the RDCs into different categories: (a) all 36 nuclear pairs, (b) 8 one-bond C-H pairs, (c) 10 one-bond C-C pairs, and (d) 18 two and three-bond pairs; results were generally consistent with those for the entire data set shown in Figure 3.] Conformations around the global free energy minimum (M1) have the smallest R factors, and those near M2 have the worst agreement with experiment. Results from the Almond or PALES algorithm are consistent with each other, although the computed R factors with PALES are slightly lower, as indicated in the Figure caption. In Figure 2S and Figure 3S, we also show the similar results using a different experimental RDC data set.[7]

In Figure 4, we display the experimental RDCs versus calculated RDC values using Almond's method for several structures, including the crystal structure[1] and the best MD conformation from the 100 ns trajectory with TIP4PEW waters at T = 315 K. The relatively large R factor of crystal structure (38.8%) suggests that some conformation change might happen. Allowing the (φ, ψ) values shift slightly from (108°, 305°) in crystal to (121.7°, 297.5°) reduces the R factor from 38.8% to 22.1%. Averaging over the 14 best MD snapshots slightly lowers the R-factor, as shown in the figure caption.

Averaging the RDCs over all 100,000 snapshots, however, yields a large R factor (49%; data is shown in Fig. 4) which implies that the simulation as a whole is not in good agreement with experiment, probably as a result of the sizable percentage of snapshots around M2, which have the worst R factors. Hence one way to improve the results would be to modify the force field and reduce the depth of the M2 conformational basin.

It is instructive to see what happens if the elements of the alignment tensor are treated as fitting parameters, rather than be estimated from the molecular structure. We rotated the crystal structure of sucrose in 1° increments about the φ and ψ dihedral angles, calculating the RDCs of each structure via best-fitting using a singular value decomposition method;[53] results are shown in the left side of Fig. 5. Note that in these calculations, the alignment tensors are fit to the data, and may not represent the actual shape of the molecule. There is little discrimination among conformers, with only very high energy conformers (near φ =300) excluded by the data. By contrast, the same calculation using the Almond model for the alignment tensor (Fig. 5, right) shows a sharp dependence on conformation, with only a few regions, some near the crystal structure, being favored.

We also tried to fit the alignment tensors for the two sugar rings individually as reported by Freedberg and coworkers,[13] and obtained similar R factors, 17.6% for glucose ring and 12.7% for fructose ring, using the crystal structure to define the individual ring conformations. However, the overal (axial) alignment was not the same for the two rings (0.00016 for glucose vs. 0.00026 for fructose), indicating the no common structure could accommodate both sets of data. In these circumstances, the use of separate alignment tensors, however, provides little insight into the nature of conformational ensemble

3. Indirect spin-spin couplings

Serianni and coworkers[11] analyzed the indirect three-bond spin-spin couplings for four inter residue ^{13}C-^{13}H and ^{13}C-^{13}C spin pairs (C2f-O1g-C1g-H1g, C2f-O1g-C1g-C2g, C3f-C2f-O1g-C1g, and C1f-C2f-O1g-C1g), finding conformer S4 in Brady's notation[15,16] to be in better agreement with the experimental J couplings than is the crystal structure (see Refs[15,16] and Table 4S). We ranked our 100,000 snapshots for agreement with this data, as we did for the RDC data. Figure 6 shows the best 500 conformers to be distributed in a region above M1 and a second region close to M3. In Table 4S and Figure 7 we list the J couplings and R factors of M1 to M4 from our MD simulation in TIP4P/Ew waters at T = 315 K, and S1 to S5 from computational results from Brady and coworkers.[15,16] We also find that M3 or S4 has a lower R factor (0.545 and 0.476 respectively) than the other three or four conformers, consistent with the conclusions of earlier work.[11] However we also find conformers in the M1 basin that have favorable R factors, as low as 0.287 as listed in Table 4S. The right side of Figure 6 displays the distribution of glycosidic dihedral angles from different conformers. The best MD conformer from J couplings is not very far away from the crystal structure and still belongs to the same M1 basin as the best MD conformer from the RDC analysis in Figure 3. The φ,ψ regions that best fit the RDCs are not identical to those for J-couplings, but have considerable overlap, Although the best fit region for J-couplings extends "upward" in ψ more toward the M4 region. In each case, the region that fits best to experiment is smaller and much closer to the crystal structure than is the totality of conformers visited by the MD simulation.

4. Ring Puckering of Fructose

In addition to the variation in φ–ψ glycosidic space, another likely mode of structural flexibility originates from the ring conformation of fructose. The ring pucker describes how C3 and C4 are arranged relative to the plane defined by C2, C5 and O5. Experimental studies such as RDCs,[49] J coupling,[11] and optional rotation,[63] have showed that the pucker is confined to the NE quadrant of the pseudo-rotational wheel, with the phase angle most likely in the range of 0° to 90°.

The phase angle was calculated by Altona and Sundaralingam algorithm.[48] Fig. 8(a) shows the probability density distribution of phase angle with different water models. The dependence of water models is small, and the phase angle dominantly fluctuates around −5° which is the value of crystal conformation.[1] We also noticed a slight peak presents at −150° for all models and the shift of peak for the TIP5P water model. Figure 8(b) shows the probability density distributions calculated from the best and worst 500 conformations as shown in Figure 3(a). The distributions are very similar to Figure 8(a) except that the distribution of the worst 500 shifts to more negative values. Hence, in all cases, our simulation results have large percentages of negative phase angles which are not preferred for experimental data: Optional rotation[63] likes the phase angle between 18° to 54°, J coupling[11] tends to choose values around 0° and recent RDCs predicted the range of 20° to 70°.[49] Hence adjustments of force field might be required to more favor positive phase angles.

5. Intramolecular Hydrogen Bonds

As depicted in Figure 1, there exist two intramolecular hydrogen bonds, O2g-H1Of and O5g-H6Of, in the crystal conformation.[1] Early NMR studies in aqueous solution by Bock and Lemieux[2] suggested that only one (O2g-H1Of) exists in solution. Molecular simulations by Engelsen and coworkers,[19] however, found no intramolecular hydrogen bond presence using a revised CHARMM-type force field,[34] which started from the crystal structure.

Figure 9 shows the probability density distributions for the two possible hydrogen bonds shown in Figure 1 The peak around 1.9 Å corresponds to the intramolecular hydrogen bond. The distribution with TIP5P is distinguished from that of the other four water models with the significantly lower peak of the hydrogen bond. There are only slight differences for the rest four models, and TIP3P has the highest peak of hydrogen bond. Hydrogen-bond occupations were calculated by integrating the first peaks in Figure 9 and are listed in Table 5S. For the TIP4P/Ew model, the O2g hydrogen bond has an occupation of about 12%, and the O5g bond an occupation of about 5%. These two values are increased to 18% and 14% as we restrain the conformation to M1 basin region and pucker phase angle to 0 as shown below (see Table 5S also).

From the schematic picture in Figure 1, we can see that the distance between O2g and H1Of not only depends on the glycosidic dihedrals, φ and ψ but also depends on the dihedrals from the 1-hydroxymethyl group of fructose. To see the effects of glycosidic dihedrals and puckering phase angle on this distance, we performed the restrained MD simulations shown in Figure 10. The peak corresponding the hydrogen bond only appears when the glycosidic dihedrals are restrained to the minimum M1 which is close to the crystal conformation. This means that the hydrogen bond will be broken when the sucrose molecule transfers to the other three minima. The presence of peaks at larger distances shows that even when the molecule fluctuates around the minimum M1, the rotation of 1-hydroxymethyl group could also break the hydrogen bond between O2g and H1Of, with the occupation of about 18%.

We also obtained the distributions with the φ-ψ dihedrals restrained to the minimum M1 but the puckering phase angle fixed to different values, P = 0, 18, 36, 54, 72, and 90. Figure 10 shows that the percentage of hydrogen-bond (35.53%, see Table 5S for other values) is greatest when P = 54. Based on Figure 10(c), similar arguments can also be reached for O5g-H6Of hydrogen bond since the distance also depends on other dihedral angles such from fructose ring and 6-hydroxymethyl group. Moreover, the O5g-H6Of is much more sensitive to the conformation of fructose ring as shown in Figure 10(d), namely the O5g-H6Of hydrogen bond breaks as the pucker angle of fructose moves to positive values, even when the glycosidic angles are restrained to crystal conformation.

4. Summary

These simulations show that the glycosidic conformation of crystal state also dominates in solution, although other free energy minima are visited. The quality of the simulation was evaluated by the calculation of residual dipolar couplings and indirect spin-spin couplings. Among the conformations sampled here, those in the M1 free energy basin (near the crystal structure, or with slightly larger values of ψ) give the best fits to both the RDC and J-coupling data. A straightforward average of RDCs and J-couplings over the entire trajectory yields poor results, mainly due to the high population near the M2 basin. Hence quantitative agreement with experiment requires a more accurate carbohydrate force field which would probably reduce the population of M2. We also note that many of the structures in the “M1 basin” fit the RDC data much better than the crystal structure itself (39% vs. 22% for the best MD conformation). This is due to the slight shift of (φ,ψ) values, from (108°, 305°) in crystal to (121.7°, 297.5°) in solution. This implies that water solvation modifies the crystal by weakening the intramolecular hydrogen bonds O2g-H1Of and O5g-H6Of.

The water solvation effects on the local structure flexibility of the fructose monosaccharide residue have also been investigated by the analysis of puckering phase angle. We found that the distribution of phase angles fluctuates around the −5° of crystal conformation. The best and worst 500 conformations from the RDC rank have similar distribution except that the distribution of worst 500 shifts to more negative values. This implies that there is only very

weak correlation between the glycosidic (φ,ψ) conformation from RDCs and the local structure flexibility of the fructose from the puckering phase angles. However the large percentage of negative puckering phase angles present in our simulation results seems to disagree with experiment, suggesting that changes to the force field torsion potentials to favor positive pucker angles should be explored.

The presence of solvent water appears to have a strong influence on the intramolecular hydrogen bond patterns O2g-H1Of and O5-H6Of present in crystal structure. They are dynamic processes of breaking and reforming in solution instead of a static picture in the crystal structure. Our simulations also show that these two intramolecular hydrogen bonds can only be maintained when the system has conformations around the M1 basin (which contains the crystal structure). Furthermore, the hydrogen bonds depend not only on the conformation of glycosidic linkage but also on the rotamer population of hydroxymethyl group at C1 or C6 of fructose. Hence one strategy to stabilize the M1 basin relative to M2 would be to increase the intramolecular H-bond strength, relative to H-bonds with water. Explorations along these lines are planned.

The results presented here suggest that the question, “is sucrose rigid or flexible?” needs to be answered with care. There is almost certainly averaging within the energy basin near the crystal stucture, and probably additional flexibility about ψ to regions “above” this M1 basin, as suggested in Figs. 3 and 6. On the other hand, these regions comprise only a small fraction of the φ–ψ map, and do not include many conformations favored by the GLYCAM force field. Overall, the NMR and compuational data seem most consistent with a limited conformational ensemble centered fairly near to the crystal conformation.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This research was funded by NIH grant GM45811. We thank Lachele Foley, Robert J. Woods, and Darón Freedberg for helpful discussions.

References

1. Brown GM, Levy HA. Acta Crystallogr Sect B. 1973; 29:790–797.
2. Bock K, Lemieux RU. Carbohydr Res. 1982; 100:63–74.
3. McCain DC, Markley JL. Carbohydr Res. 1986; 152:73–80. [PubMed: 3768917]
4. McCain DC, Markley JL. J Am Chem Soc. 1986; 108:4259–4264.
5. Kovacs H, Bagley S, Kowalewski J. J Magn Reson. 1989; 85:530–541.
6. Effemey M, Lang J, Kowalewski J. Magn Reson Chem. 2000; 38:1012–1018.
7. Neubauer H, Meiler J, Peti W, Griesinger C. Helv Chim Acta. 2001; 84:243–258.
8. Dupenhoat CH, Imberty A, Roques N, Michon V, Mentech J, Descotes G, Perez S. J Am Chem Soc. 1991; 113:3720–3727.
9. Poppe L, Vanhalbeek H. J Am Chem Soc. 1992; 114:1092–1094.
10. Adams B, Lerner L. J Am Chem Soc. 1992; 114:4827–4829.
11. Duker JM, Serianni AS. Carbohydr Res. 1993; 249:281–303. [PubMed: 8275501]
12. Batta G, Kover KE. Carbohydr Res. 1999; 320:267–272.
13. Venable RM, Delaglio F, Norris SE, Freedberg DI. Carbohydr Res. 2005; 340:863–874. [PubMed: 15780252]
14. Stevens ES, Duda CA. J Am Chem Soc. 1991; 113:8622–8627.
15. Tran VH, Brady JW. Biopolymers. 1990; 29:961–976. [PubMed: 2369624]

16. Tran VH, Brady JW. Biopolymers. 1990; 29:977–997. [PubMed: 2369625]
17. Engelsen SB, Dupenhoat CH, Perez S. J Phys Chem. 1995; 99:13334–13351.
18. Engelsen SB, Perez S. Carbohydr Res. 1996; 292:21–38.
19. Engelsen SB, Monteiro C, de Penhoat CH, Perez S. Biophys Chem. 2001; 93:103–127. [PubMed: 11804720]
20. Casset F, Imberty A, duPenhoat CH, Koca J, Perez S. J Mol Struc-Theochem. 1997; 395:211–224.
21. French AD, Kelterer AM, Cramer CJ, Johnson GP, Dowd MK. Carbohydr Res. 2000; 326:305–322. [PubMed: 10890277]
22. French AD, Kelterer AM, Johnson GP, Dowd MK, Cramer CJ. J Comput Chem. 2001; 22:65–78.
23. French, AD.; Brady, JW., editors. ACS Symposium Series. Vol. 430. American Chemical Society; Washington, D C: 1990. Computer Modelling of Carbohydrate Molecules.
24. Vliegenthart, JFG.; Woods, RJ., editors. ACS Symposium Series. Vol. 930. American Chemical Society; Washington, DC: 2006. NMR Spectroscopy and Computer Modeling of Carbohydrates: Recent Advances.
25. Imberty A, Perez S. Chem Rev. 2000; 100:4567–4588. [PubMed: 11749358]
26. Dwek RA. Chem Rev. 1996; 96:683–720. [PubMed: 11848770]
27. Xia JC, Margulis CJ. J Biomol NMR. 2008; 42:241–256. [PubMed: 18953494]
28. Xia JC, Daly RP, Chuang FC, Parker L, Jensen JH, Margulis CJ. J Chem Theory Comput. 2007; 3:1620–1628.
29. Xia JC, Daly RP, Chuang FC, Parker L, Jensen JH, Margulis CJ. J Chem Theory Comput. 2007; 3:1629–1643.
30. Kirschner KN, Yongye AB, Tschampel SM, Gonzalez-Outeirino J, Daniels CR, Foley BL, Woods RJ. J Comput Chem. 2008; 29:622–655. [PubMed: 17849372]
31. Woods RJ, Dwek RA, Edge CJ, Fraserreid B. J of Phys Chem. 1995; 99:3832–3846.
32. Guvench O, Greene SN, Kamath G, Brady JW, Venable RM, Pastor RW, Mackerell AD. J Comput Chem. 2008; 29:2543–2564. [PubMed: 18470966]
33. Kuttel M, Brady JW, Naidoo KJ. J Comput Chem. 2002; 23:1236–1243. [PubMed: 12210149]
34. Ha SN, Giammona A, Field M, Brady JW. Carbohydr Res. 1988; 180:207–221. [PubMed: 3203342]
35. Homans SW. Biochemistry. 1990; 29:9110–9118. [PubMed: 2271581]
36. Momany FA, Rone R. J Comput Chem. 1992; 13:888–900.
37. Kony D, Damm W, Stoll S, van Gunsteren WF. J Comput Chem. 2002; 23:1416–1429. [PubMed: 12370944]
38. Damm W, Frontera A, TiradoRives J, Jorgensen WL. J Comput Chem. 1997; 18:1955–1970.
39. Guillot B. J Mol Liq. 2002; 101:219–260.
40. Case DA, Cheatham TE, Darden T, Gohlke H, Luo R, Merz KM, Onufriev A, Simmerling C, Wang B, Woods RJ. J Comput Chem. 2005; 26:1668–1688. [PubMed: 16200636]
41. Xia, Junchao; Case, David A. Sucrose in aqueous solution revisited: 2. Adaptively biased molecular dynamics simulations and computational analysis of NMR relaxation. Biopolymers. 2011 submitted.
42. Berendsen HJC, Grigera JR, Straatsma TP. J Phys Chem. 1987; 91:6269–6271.
43. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML. J Chem Phys. 1983; 79:926–935.
44. Jorgensen WL, Madura JD. Mol Phys. 1985; 56:1381–1392.
45. Horn HW, Swope WC, Pitera JW, Madura JD, Dick TJ, Hura GL, Head-Gordon T. J Chem Phys. 2004; 120:9665–9678. [PubMed: 15267980]
46. Mahoney MW, Jorgensen WL. J Chem Phys. 2000; 112:8910–8922.
47. Case DA, Darden TA, Cheatham TE III, Simmerling CL, Wang J, Duke RE, Luo R, Merz KM, Pearlman DA, Crowley M, Walker RC, Zhang W, Wang B, Hayik S, Roitberg A, Seabra G, Wong KF, Paesani F, Wu X, Brozell S, Tsui V, Gohlke H, Yang L, Tan C, Mongan J, Hornak V, Cui G, Beroza P, Mathews DH, Schafmeister C, Ross WS, Kollman PA. Amber. 2006:9.
48. Altona C, Sundaralingam M. J Am Chem Soc. 1972; 94:8205–8212. [PubMed: 5079964]

49. Freedberg DI. J Am Chem Soc. 2002; 124:2358–2362. [PubMed: 11878992]
50. Prestegard JH, Bougault CM, Kishore AI. Chem Rev. 2004; 104:3519–3540. [PubMed: 15303825]
51. Prestegard JH, Al-Hashimi HM, Tolman JR. Quart Rev Biophys. 2000; 33:371–424.
52. Losonczi JA, Andrec M, Fischer MWF, Prestegard JH. J Magn Reson. 1999; 138:334–342. [PubMed: 10341140]
53. Zweckstetter M, Bax A. J Am Chem Soc. 2000; 122:3791–3792.
54. Almond A, Axelsen JB. J Am Chem Soc. 2002; 124:9986–9987. [PubMed: 12188652]
55. Landersjo C, Jansson JLM, Maliniak A, Widmalm G. J Phys Chem B. 2005; 109:17320–17326. [PubMed: 16853211]
56. Azurmendi HF, Bush CA. J Am Chem Soc. 2002; 124:2426–2427. [PubMed: 11890777]
57. Bernado P, Blanchard L, Timmins P, Marion D, Ruigrok RWH, Blackledge M. Proc Natl Acad Sci U S A. 2005; 102:17002–17007. [PubMed: 16284250]
58. Jha AK, Colubri A, Freed KF, Sosnick TR. Proc Natl Acad Sci U S A. 2005; 102:13099–13104. [PubMed: 16131545]
59. Marsh JA, Baker JMR, Tollinger M, Forman-Kay JD. J Am Chem Soc. 2008; 130:7804–7805. [PubMed: 18512919]
60. Esteban-Martin S, Fenwick RB, Salvatella X. J Am Chem Soc. 2010; 132:4626–4632. [PubMed: 20222664]
61. Cornilescu G, Marquardt JL, Ottiger M, Bax A. J Am Chem Soc. 1998; 120:6836–6837.
62. Cloran F, Carmichael I, Serianni AS. J Am Chem Soc. 1999; 121:9843–9851.
63. Duda CA, Stevens ES. Carbohydr Res. 1990; 206:347–351.
64. Tvaroska I, Hricovini M, Petrakova E. Carbohydr Res. 1989; 189:359–362.

Appendix

The assumption of the calculation method of free energy in Equation 8 is that the MD trajectories are long enough to visit the relevant phase space sufficiently which is a good approximation for systems with low free energy barriers. For exploring the transitions with high engery barriers, more advanced molecular dynamics techniques are required to solve the sampling probleim in the short time simulations. The adaptively biased molecular dynamics (ABMD) (Babin, V.; Roland, C.; Sagui, C. J. Chem. Phys. 2008, 128, 134101) is a recent addition to the advanced molecular dynamics techniques for accelerating simulations and is capable of overcoming these high free energy barriers and estimating the potential of mean force in the relevant space. Figure 11 and Table 2 display the free energy results of T = 283 K constructed from the 100 ns normal MD simulation and the more advanced 25 ns ABMD simulation, The ABMD method is able to explore more high energy regions but the low energy parts are similar to that of normal MD simulations. This verifies the convergences of low energy parts of normal MD simulations which are more important for the studies of pure water. Namely, although the systems in normal MD simulations will visit more high energy regions as the simulation times are increased further, we do not expect that the simulations will change the free energy landscapes of low energy parts significantly. (See our second paper [41] for more detailed ABMD results about sucorse in high viscous solvent, water-DMSO mixture.)

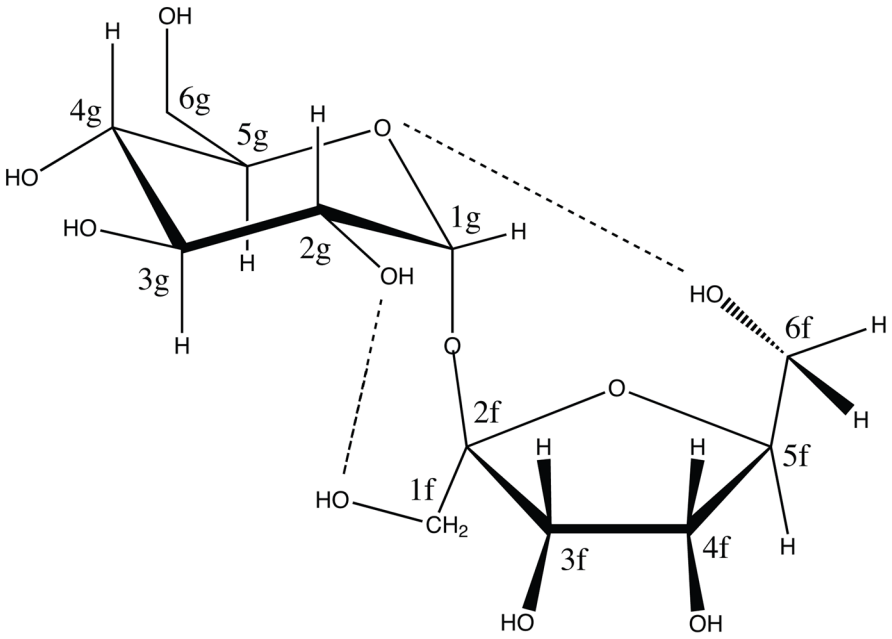

Figure 1.

Schematic representation of sucrose disaccharide, α-D-Glu-(1→2)-β-D-Fru. The glycosidic angles φ–ψ are defined as $\varphi = O5_g - C1_g - O2_f - C2_f$ and $\psi = C1_g - O2_f - C2_f - O5_f$, where g and f refer to glucopyranosyl and fructofuranosyl rings respectively. The dashed lines shows the two hydrogen bonds present in the crystal structure, O2g-H1Of and O5g-H6Of with the distances of 1.85 and 1.89 angstroms respectively.[1]

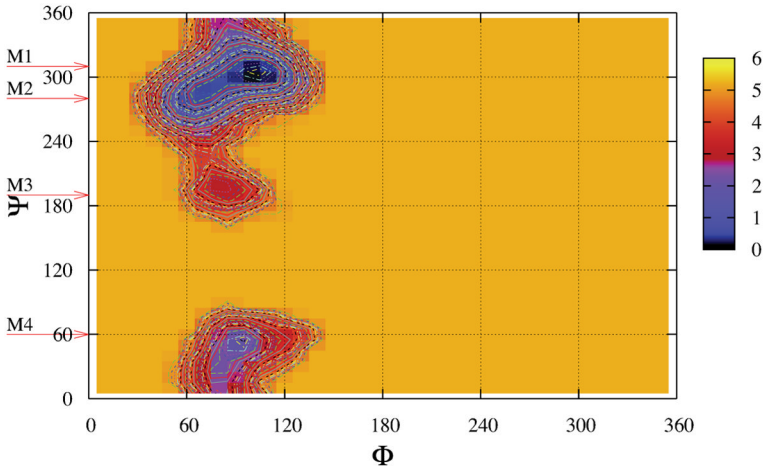

Figure 2.

Free energy landscapes in φ–ψ glycosidic space for sucrose in the TIP4P/Ew model. Contour lines are shown from 0.0 to 5.0 kcal/mol in steps of 0.1. The local minima are identified as M1, M2, M3 and M4; see also Table 1.

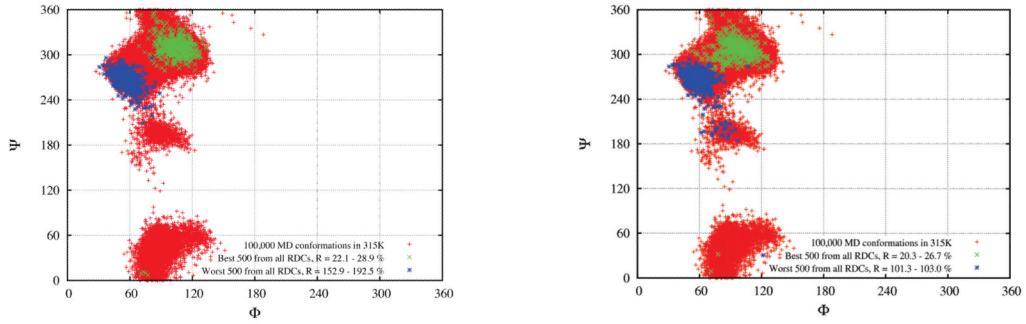

Figure 3.

The best 500 (green) and worst 500 (blue) conformations of sucrose among the 100,000 snapshots (red) extracted from the 100 ns simulation trajectory with TIP4PEW waters at T = 315 K. (a) The RDC ranks were produced by comparing the R factor of calculated RDCs by Almond method [54] with respect to the experimental values. [13] The R factors = 22.1%~28.9% for best 500, and 152.9%~192.5% for the worst 500. (b) The same, but using the PALES [53] method. All 36 proton pairs in Ref.[13] are included. The R factors = 20.3%~26.7% for best 500, and 101.3%~103.0% for the worst 500.

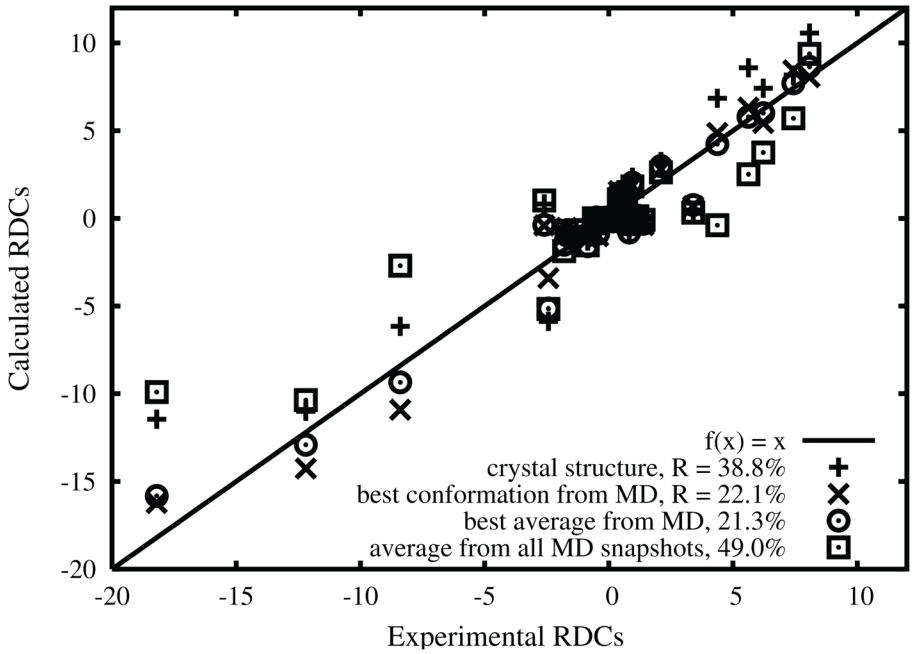

Figure 4.

Experimental RDCs .vs. calculated RDCs by Almond's method for the crystal structure, the best MD conformation 100 ns trajectory with TIP4P/Ew waters at T = 315 K, and the best MD average over the top 14 in the RDC rank.

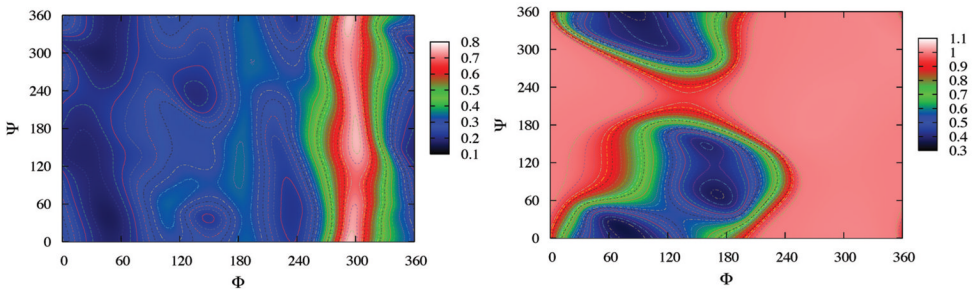

Figure 5.
R-factors for RDC's, as a function of the glycosidic dihedral angles, for the SVD (*left*) and Almond (*right*) models of the alignment tensor. See the text for details of the calculation.

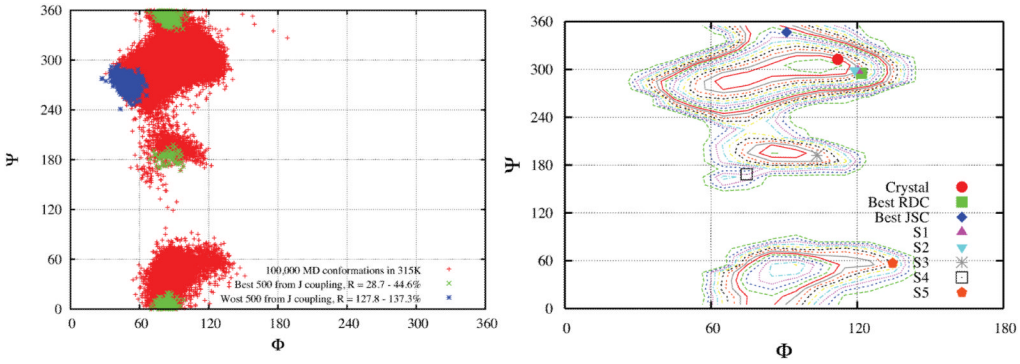

Figure 6.

Left: Same as Figure 3, using four inter-residue J coupling constants.[11] Right: Locations of glycosidic dihedral angles from different conformations on the free energy landscape constructed from the MD simulation with TIP4P/Ew at T = 315K.

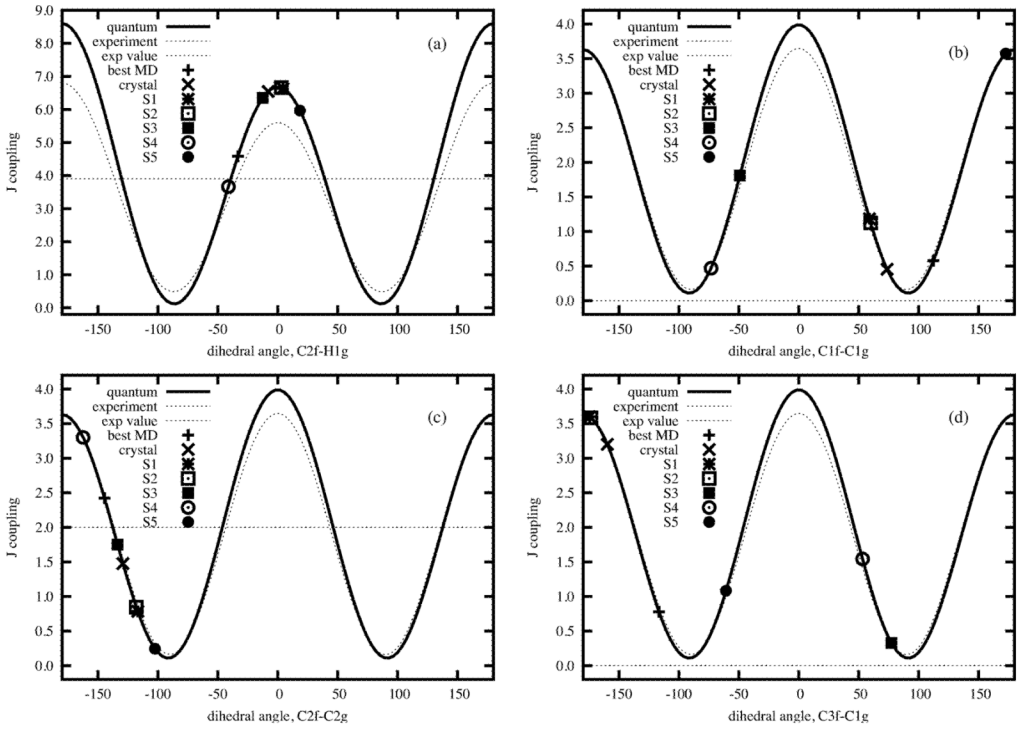

Figure 7.

Individual J-couplings vs torsion angle. “Quantum” represents that the Karplus relation curve is fitted from DFT calculations,[62] and “Experiment” for the curve fitted from experimental data.[64] Torsion angles are defined as (a) C2f-O1g-C1g-H1g, (b) C1f-C2f-O1g-C1g, (c) C2f-O1g-C1g-C2g, and (d) C3f-C2f-O1g-C1g.

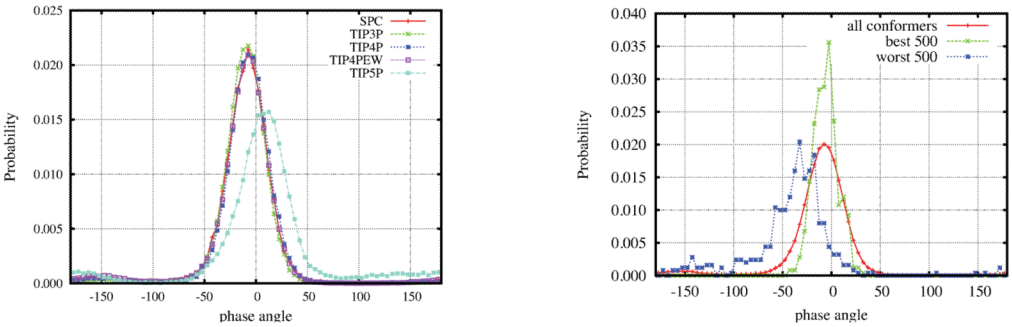

Figure 8.
Puckering phase angle probability density distributions of the fructose monosaccharide residue of sucrose. (left) Different water models at T = 300K, (right) All 100,000 conformers, best and worst 500 from RDC rank in Figure 3.

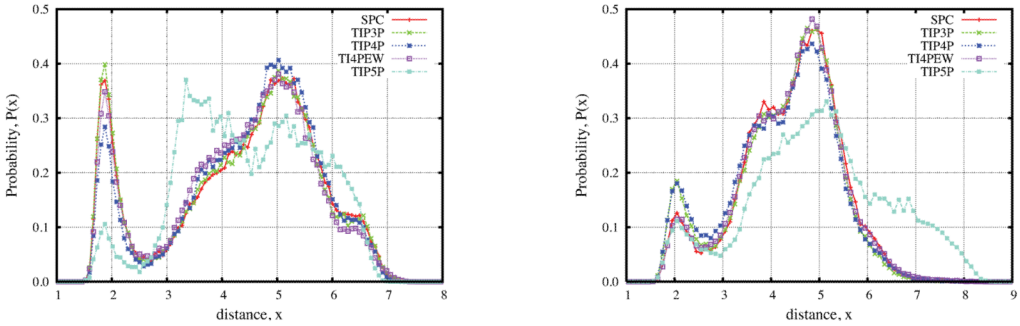

Figure 9.
Probability density distributions of the distances between two atoms from MD trajectories. Left: O2g and H1Of; right: O5g and H6Of.

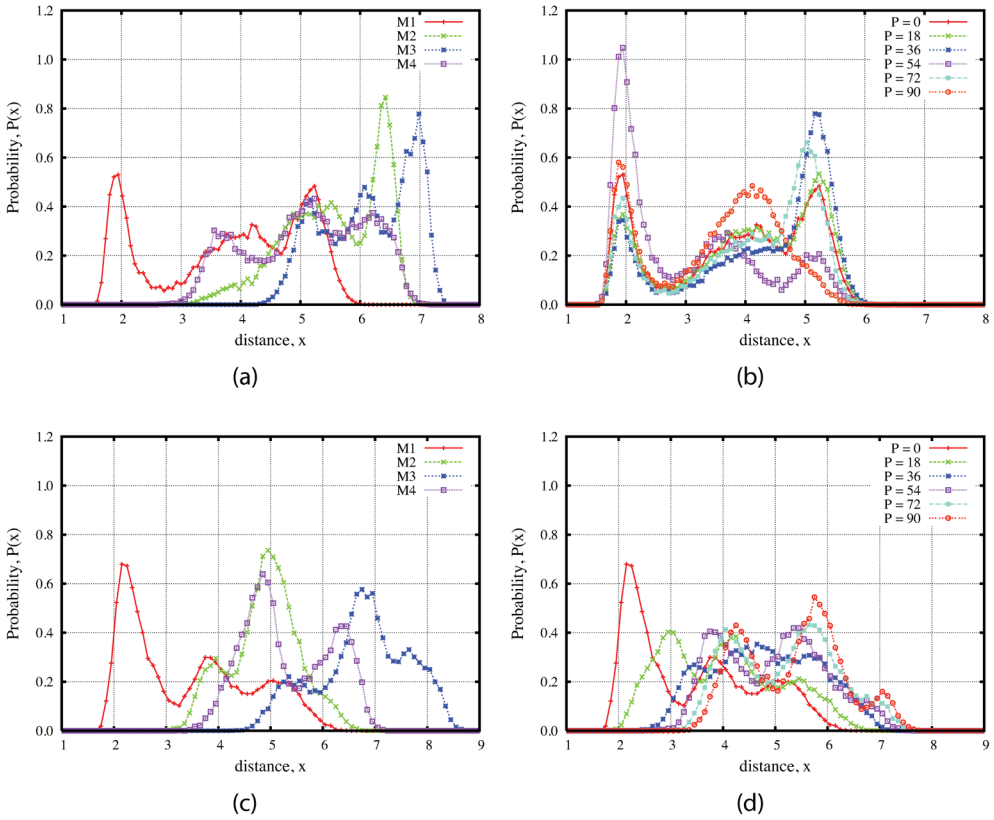

Figure 10.

Probability density distributions of the distance between two atoms from 10ns restrained MD simulations with the TIP4P/Ew waters at T = 300 K. (a) and (b) for O2g and H1Of, (c) and (d) for O5g and H6OfO2g. For (a) and (c) the dihedrals of φ–ψ glycosidic linkage are restrained to the four different free energy minima as shown in Figure 2 and five dihedrals from the fructose ring are restrained in order to make the puckering phase angle fixed to P = 0. For (b) and (d) the values of φ–ψ are restrained to the minimum M1 but the puckering phase angle are fixed to different values (P = 0, 18, 36, 54, 72, and 90).

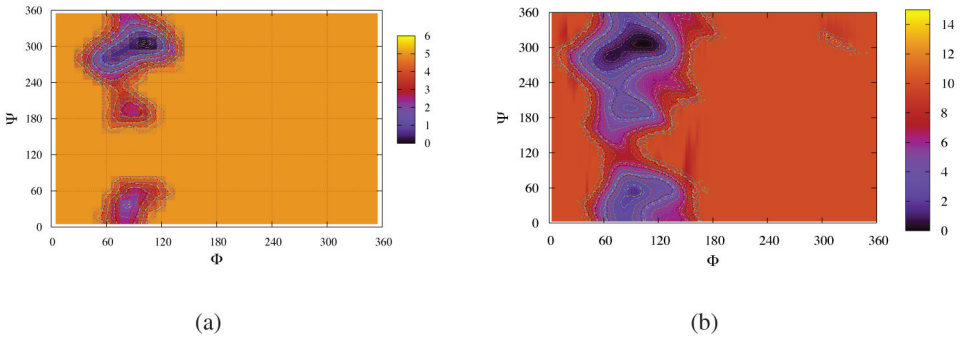

Figure 11.

Free energy landscapes in φ–ψ glycosidic space for sucrose in TIP4P/Ew waters at T =283 K. (a) from 100 ns normal MD simulation, the contour lines were generated from 0.0 to 5.0 KCal/mol by the difference of 0.1. (b) from 25 ns ABMD simulation, the contour lines were generated from 0.0 to 10.0 KCal/mol by the difference of 0.25. The ABMD method is able to explore more high energy regions but the low energy parts are similar to that of normal MD simulations. Note that we let all global minima have the values of zero and this results in that ABMD has different background color corresponding to the regions with higher free energy values.

Table 1

Free energy information of four minima of sucrose in different explicit water models at T = 300K.

	SPC	TIP3P	TIP4P	TIP4P/Ew	TIP5P
			positions (φ,ψ) (in degree)		
M1	(104.5, 305.1)	(105.1, 304.9)	(104.9,305.2)	(104.8,305,4)	(99.8, 316.5)
M2	(69.8, 285.2)	(69.9, 285.6)	(69.4, 285.5)	(69.9, 284.7)	(65,6, 285.5)
M3	(83.9, 195.0)	(86.8, 194.9)	(85.3, 195.4)	(83.3, 196.4)	(95.2, 194.6)
M4	(90.3, 50.5)	(89.9, 51.0)	(93.6,54.9)	(94.8. 54.9)	(86.7, 38.6)
			well depths (in kcal/mol)		
F1	0.0	0.0	0.0	0.0	0.0
F2	0.3	0.55	0.65	0.5	1.1
F3	3.4	3.1	3.1	2.9	0.1
F4	2.3	2.7	1.6	1.8	0.4
			probabilities within a radius of 20°		
P1	0.581	0.634	0.605	0.577	0.341
P2	0.379	0.314	0.327	0.355	0.053
P3	0.002	0.004	0.003	0.005	0.216
P4	0.015	0.008	0.039	0.024	0.185
			barriers between minima (in kcal/mol)		
B12	0.4	0.6	0.7	0.6	1.3
B23	3.9	3.1	3.7	3.4	2.3
B14	2.6	3.0	2.8	2.4	0.8

Table 2

Free energy information of four minima of sucrose in pure water at T = 283K from the 100 normal MD simulation and the 25 ns ABMD simulation resepectively.

	ABMD	Normal MD
positions (φ,ψ) (in degree)		
M1	(104.5, 306.1)	(104.1, 305.4)
M2	(69.7, 283.3)	(71.2, 285.8)
M3	(88.5, 195.9)	(85.8, 194.9)
M4	(92.9, 54.3)	(84,4, 35.5)
heights (in KCal/mol)		
F1	0.0	0.0
F2	0.3	0.5
F3	2.9	2.3
F4	1.4	1.8
probabilities within a radius of 20 degree		
P1	0.557	0.606
P2	0.241	0.365
P3	0.003	0.008
P4	0.018	0.022
barriers between minima (in KCal/mol)		
B12	0.6	0.6
B23	3.8	3.0
B14	2.5	2.4