Calculation of the Ramachandran Potential of Mean Force for a Disaccharide in Aqueous Solution

Kevin J. Naidoo*,† and J. W. Brady*,‡

Contribution from the Department of Food Science, Stocking Hall, Cornell University, Ithaca, New York 14853, and Department of Chemistry, University of Cape Town, Rondebosch 7700, South Africa

Received June 22, 1998. Revised Manuscript Received December 18, 1998

Abstract: Molecular dynamics simulations employing adaptive umbrella sampling have been used to calculate the Ramachandran conformational potential of mean force in aqueous (TIP3P) solution for the $\alpha(1\rightarrow4)$-linked dimer of D-xylopyranose (4-*O*-α-D-xylopyranosyl-α-D-xylopyranose), a pentose analogue of maltose and a useful general model for the effects of solvent structuring upon biopolymer hydration. The vacuum adiabatic energy map for this molecule closely resembles that for maltose, but the solution pmf is quite different, with one of the principal vacuum minima almost completely disappearing in solution and with the global minimum-energy conformation being a new minimum which does not occur at all on the vacuum surface. This conformation is apparently stabilized by a water molecule which hydrogen bonds to a hydroxyl group on each ring, bridging between the two rings. The new conformation also places the two hydrophobic methylene groups almost in van der Waals contact, reducing their exposed surface area. Unfortunately, the results reaffirm the dependence of hydration effects upon the specific details of each molecule's chemical structure, making the application of simple general models for hydration more difficult.

I. Introduction

Water is the most abundant molecule on the surface of the earth and is ubiquitous in biological systems. Although structurally simple, collectively it has quite unusual bulk properties.[1,2] Many of these unusual properties are due to the unique ability of water molecules to make four simultaneous hydrogen bonds, two as a donor involving their protons and two as an acceptor through their lone pairs. This parity between the number of interactions which can be formed as a donor and as an acceptor, along with their roughly tetrahedral arrangement, allows water to form space-filling hydrogen bond networks, in a way that other hydrogen-bonding molecules, such as HF or NH_3, cannot. For this reason, water is extensively structured in the liquid phase. As an integral part of living organisms, water is responsible, directly or indirectly, for much of the characteristic structure and chemistry of biomolecules. It is required for most globular proteins to fold into their native conformations and induces the formation of lipid bilayers, micelles, and vesicles. In many cases, it is impossible to understand the conformations of biopolymers without accounting for the effects of hydration.

Just as water is the most abundant inorganic molecule, the carbohydrates are the most abundant biological molecules. They play important structural and energy storage roles in both plants and animals and form the basis for several major industries. In addition, saccharides found in covalently bound complexes with other molecules such as lipids and proteins have been identified as essential in a number of molecular recognition processes. Recognition of carbohydrate structures is involved in cell binding and adhesion, fertilization, immune response to antigens, and host infection by pathogens. In all of these processes not only the primary structures but also the conformations of carbohydrates can be of considerable biological or medical significance.

In most biological systems carbohydrates are found in aqueous environments, and the results of many experiments indicate that the conformations of at least some disaccharides are sensitive to hydration.[3,4] The presence of the solvent water molecules can have a number of effects on the solute, which may be more important in some cases than in others. In all cases, water serves as a viscous medium, damping structural motions, and as a thermal bath, maintaining the overall internal temperature and occasionally providing random energy "kicks" which can promote transitions over internal energy barriers or serve as a sink for energy given off by exothermic processes. Furthermore, the high dielectric constant of liquid water strongly attenuates electrostatic interactions between solute charges and dipoles. The complex field of the solvent also interacts with the electronic structure of the solute, perturbing its electronic distribution and possibly favoring some structures and conformations over others.

In addition, the presence of carbohydrate solutes can also significantly affect the aqueous solvent. Certain specialized biological functions of carbohydrates directly involve these interactions with water, such as their antidesiccant and cryoprotective activity[5] and their ability to modify solution viscosity or prevent the drainage of water from gels. The polyfunctional, hydrogen-bonding carbohydrates would be expected to interact

---

(1) Eisenberg, D.; Kauzmann, W. *The Structure and Properties of Water*; Oxford University Press: Oxford, U.K., 1969; p 296.
(2) Stillinger, F. H. *Science* **1980**, *209*, 451–457.
(3) Rees, D. A.; Thom, D. *J. Chem. Soc., Perkin Trans. 2* **1977**, 191–201.
(4) Engelsen, S. B.; Herve du Penhoat, C.; Peréz, S. *J. Phys. Chem.* **1995**, *99*, 13334–13351.
(5) Ding, S.-P.; Fan, J.; Green, J. L.; Lu, Q.; Sanchez, E.; Angell, C. A. *J. Therm. Anal.* **1996**, *47*, 1391–1405.

strongly with water, and as is the case for the proteins and lipids, there may be significant interplay between the conformational structure of sugar solutes and the collective structure of the solvent. In an aqueous environment, a solute's functional groups must interact with the inherent structural requirements of the solvent, and its presence can impose an alternate structuring pattern on the adjacent water molecules. Such solvent structuring is often invoked to explain the properties of aqueous solutions,[2,6] although in practice, it has been very difficult to actually determine the nature of this structuring experimentally. The organization of water molecules around a particular solute will in general involve both positional and orientational correlations with the specific chemical architecture of the solute, and thus will vary from one molecule to the next. This solvent structuring can be quite complex, due to the complicated mix of chemical functionalities found in typical biopolymers, which frequently have polar or hydrogen-bonding functional groups in close proximity to nonpolar groups,[7−12] each with its own specific hydration requirements. When the relative positions of molecular functional groups can change through conformational transitions, the mutual compatibility of their hydration requirements can affect the conformational energy. The interaction of the hydration spheres of functional groups is the underlying phenomenon modeled in the Pratt−Chandler theory,[13] affecting the conformational equilibrium of molecules such as butane.[14] Similar types of hydration interactions could affect the solution conformations of disaccharides and polysaccharides and may well contribute to determining the oligosaccharide conformations of glycoconjugates involved in molecular recognition in disease mechanisms.[15,16]

The carbohydrates are useful general models for such hydration interactions in all types of biopolymers for several reasons, including the less complex sequence characteristics of the saccharides, with a predominance of only a few types of functional groups, and the relative rigidity of the monomer units. This rigidity means that the majority of the conformational flexibility is due to the glycosidic torsional degrees of freedom and hydroxyl rotations. However, potential configurational differences at the various asymmetric carbon atoms allow ample opportunity to test the response of hydration to small-scale changes in structure, since the rigidity of the rings at room temperature maintains the hydroxyl groups at fixed relative orientations.

The conformations of carbohydrates in solution have proven to be surprisingly resistant to experimental characterization. Crystal structures determined by X-ray or neutron diffraction experiments are available for many of the monosaccharides and disaccharides of biological interest,[17] but even for the simple sugars, the crystal conformation may not be the conformation in aqueous solution. In principle, NMR could provide much of the desired information about oligosaccharide conformations, but these studies suffer from several serious difficulties. For example, there are no vicinal proton coupling constants across glycosidic linkages, the heteronuclear $^3J_{C,H}$ values are small and difficult to measure, and the available Karplus-type equations for heteronuclear couplings may not always be generally reliable. Furthermore, there usually are only a few useful NOEs between the rings, so that there is insufficient information to uniquely specify linkage conformations. In addition, in many cases, a dynamic equilibrium might exist between more than one conformation, with the coupling or NOE data representing a virtual, average structure.[18,19] Thus, while the nearly complete conformation of the N-linked oligosaccharide from the human CD2 glycoprotein has recently been reported on the basis of NOE measurements,[20] for most disaccharides and oligosaccharides in solution, NMR data have only been used to determine whether a proposed conformer is possible. Few other structural probes are available. An innovative approach involving the correlation of optical rotation measurements with linkage conformation has been very useful in analyzing disaccharide solution structures,[21−23] but again, the chiroptical data generally cannot be used to make an unambiguous structure determination without additional input, particularly in cases involving conformational averaging.

Because of the difficulties in experimental studies of carbohydrates, there have been many efforts to apply theoretical modeling calculations to these systems. The large size and structural complexity of carbohydrates makes quantum mechanical calculations troublesome,[24,25] so the majority of disaccharide modeling studies have been molecular mechanics calculations of various types.[26−28] For carbohydrates, as in analogous studies of proteins,[29,30] conformations are usually interpreted in terms of two-dimensional Ramachandran contour maps of the conformational energy for dimer repeat units as a function of the glycosidic linkage torsional angles ϕ and ψ, computed using empirical energy functions.[31] While the conformational energy depends on many other internal coordinates besides these two torsion angles, relaxed or adiabatic energy maps which plot the lowest possible energy for a particular ϕ and ψ are still useful.[32] Such disaccharide maps are widely employed to help interpret fiber diffraction data[33] and solution chain properties of polysaccharides such as characteristic ratios[34,35] and to analyze possible conformations for oligosaccharides.

(6) Israelachvili, J.; Wennerstrom, H. *Nature* **1996**, *379*, 219−225.

(7) Mezei, M.; Beveridge, D. L. *J. Comput. Chem.* **1984**, *5*, 523−527.

(8) Schmidt, R. K.; Karplus, M.; Brady, J. W. *J. Am. Chem. Soc.* **1996**, *118*, 541−546.

(9) Lounnas, V.; Pettitt, B. M. *Proteins* **1994**, *18*, 133−147.

(10) Lounnas, V.; Pettitt, B. M. *Proteins* **1994**, *18*, 148−160.

(11) Resat, H.; Mezei, M. *Biophys. J.* **1996**, *71*, 1179−1190.

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

(13) Pratt, L. R.; Chandler, D. *J. Chem. Phys.* **1977**, *67*, 3683−3703.

(14) Jorgensen, W. L. *J. Chem. Phys.* **1982**, *77*, 5757−5765.

(15) Naidoo, K. J.; Denysyk, D.; Brady, J. W. *Protein Eng.* **1997**, *10*, 1249−1261.

(16) Wyatt, R.; Kwong, P. D.; Desjardins, E.; Sweet, R. W.; Robinson, J.; Hendrickson, W. A.; Sodroski, J. G. *Nature* **1998**, *393*, 705−711.

(17) Jeffrey, G. A.; Sundaralingam, M. *Adv. Carbohydr. Chem. Biochem.* **1985**, *43*, 203−421.

(18) Shashkov, A. S.; Lipkind, G. M.; Kochetkov, N. K. *Carbohydr. Res.* **1986**, *147*, 175−182.

(19) Cumming, D. A.; Carver, J. P. *Biochemistry* **1987**, *26*, 6664−6676.

(20) Wyss, D. F.; Choi, J. S.; Li, J.; Knoppers, M. H.; Willis, K. J.; Arulanandam, A. R. N.; Smolyar, A.; Reinherz, E. L.; Wagner, G. *Science* **1995**, *269*, 1273−1278.

(21) Stevens, E. S.; Sathyanarayana, B. K. *J. Am. Chem. Soc.* **1989**, *111*, 4149−4154.

(22) Duda, C. A.; Stevens, E. S. *J. Am. Chem. Soc.* **1990**, *112*, 7406.

(23) Duda, C. A.; Stevens, E. S. *Carbohydr. Res.* **1990**, *206*, 347−351.

(24) Barrows, S. E.; Dulles, F. J.; Cramer, C. J.; French, A. D.; Truhlar, D. G. *Carbohydr. Res.* **1995**, *276*, 219−251.

(25) Van Alsenoy, C.; French, A. D.; Cao, M.; Newton, S. Q.; Schäfer, L. *J. Am. Chem. Soc.* **1994**, *116*, 9590−9595.

(26) Burkert, U.; Allinger, N. L. *Molecular Mechanics*; American Chemical Society: Washington, DC, 1982; Vol. 177, p 339.

(27) McCammon, J. A.; Harvey, S. C. *Dynamics of Proteins and Nucleic Acids*; Cambridge University Press: Cambridge, U.K., 1987; p 234.

(28) Brooks, C. L.; Karplus, M.; Pettitt, B. M. *Proteins: A Theoretical Perspective of Dynamics, Structure, and Thermodynamics*; Wiley-Interscience: New York, 1988; Vol. LXXI, p 259.

(29) Ramachandran, G. N.; Sasisekharan, V. *Adv. Protein Chem.* **1968**, *23*, 283−438.

(30) Zimmerman, S. S.; Pottle, M. S.; Némethy, G.; Scheraga, H. A. *Macromolecules* **1977**, *10*, 1−9.

(31) Brant, D. A. *Q. Rev. Biophys.* **1976**, *9*, 527−596.

(32) Brady, J. W. *Adv. Biophys. Bioeng.* **1990**, *1*, 155−202.

(33) *Fiber Diffraction Methods*; French, A. D.; Gardner, K. H., Ed.; American Chemical Society: Washington, DC, 1980; Vol. 141.

Although most conformational energy studies and early MD simulations omitted solvent, the possible influence of water upon disaccharide conformation may make it necessary to include the effects of hydration in theoretical simulations of these molecules.[27,28] A variety of models have been developed to represent the effects of hydration upon solutes,[36] with the direct incorporation of a large number of explicitly included water molecules being the most realistic, but also one of the most inconvenient, because of the enormous increase in computational expense. Various approximations represent the average effects of hydration, without the explicit inclusion of water molecules, through the use of a mean solvent field.[37,38] Many of these continuum solvent models contain terms related to the conformational dependence of the solvent-accessible surface area.[39−42] However, as already noted, liquid water is a highly structured fluid due to its extensive hydrogen bonding, and often the geometric requirements of the hydrogen bonds between water and solutes result in the solute imposing a very specific collective structuring upon the solvent.[7−12,43,44] For some molecules, these solute−water hydrogen bonds also favor a particular conformation, determining the conformational structure in solution,[8] as has been observed in MD simulations of both α,α-trehalose[44] and neocarrabiose.[45] Because such effects cannot generally be quantitatively reproduced using continuum models, it may often be necessary in modeling the behavior of disaccharides to use simulations with explicit water molecules.

The conformation of maltose, the backbone disaccharide repeat unit of amylose, amylopectin, and glycogen, has often been the subject of study, since its importance as a general model for oligosaccharides and polysaccharides is comparable to that of the alanine "dipeptide" as a model for polypeptides and proteins. This disaccharide is the α(1→4)-linked dimer of D-glucopyranose. The crystal structure of maltose is known from neutron diffraction,[46] and the conformations of the maltose repeat units of polysaccharides have been studied by X-ray fiber diffraction methods.[47] Many Ramachandran energy maps have been prepared for maltose,[48−53] increasing in sophistication with
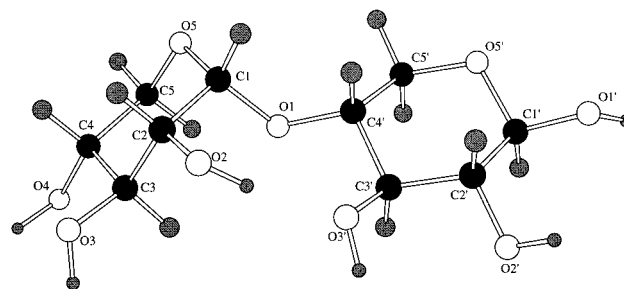

Figure 1. "Dixylose", the α(1→4)-linked dimer of D-xylopyranose, 4-*O*-α-D-xylopyranosyl-α-D-xylopyranose.

time, but all have generally produced qualitatively similar potential energy surfaces. These conformational energy maps usually have two main low-energy regions, one around (−50°, 180°), and another, larger, and lower-energy well running diagonally from around (+40°,+30°), in the vicinity of the crystal structure determined by diffraction,[46] to around (−60°, −40°). The principal reason for the overall similarity in energy maps calculated using different procedures and force fields is that many of the broad outlines of the energy surface are determined by van der Waals repulsions arising from steric clashes between atoms on different rings,[54] which are treated similarly by most theoretical force fields. Within these broad outlines, the finer details, such as whether the central valley is subdivided into several local minima, and their relative energies, are dependent upon the specific force field used. MD simulations of maltose, both in a vacuum and solution, have also been reported.[51,54−57]

An important limitation of Ramachandran energy maps of the type which have thus far been prepared for disaccharides is that they map only the conformational potential energy, while actual conformational equilibria are determined by the free energy, including solvation effects. Several Ramachandran energy maps for peptides have been calculated in recent years which estimate the conformational free energy in aqueous solution from various types of free energy simulations.[58−63] However, no conformational free energy maps for disaccharides in solution have yet been reported, although we recently described the calculation of such a free energy map for maltose in a vacuum.[54] Here we report the calculation of the Ramachandran potential of mean force (pmf) in aqueous solution for a pentose disaccharide analogue of maltose, the α(1→4)-linked dimer of D-xylopyranose (4-*O*-α-D-xylopyranosyl-α-D-xylopyranose, shown in Figure 1), using molecular dynamics (MD) simulations and adaptive umbrella sampling.[64,65] (Note that this molecule is not xylobiose, the β(1→4)-linked dimer which is

(34) Brant, D. A.; Christ, M. D. In *Computer Modeling of Carbohydrate Molecules*; French, A. D., Brady, J. W., Eds.; American Chemical Society: Washington DC, 1990; pp 42−68.

(35) Urbani, R.; Cesàro, A. *Polymer* **1991**, *32*, 3013−3020.

(36) Sharp, K. A. In *Computer Simulation of Biomolecular Systems, Vol. 2: Theoretical and Experimental Applications*; van Gunsteren, W. F., Weiner, P. K., Wilkinson, A. J., Eds.; ESCOM: Leiden, The Netherlands, 1993; pp 147−160.

(37) Hodes, Z. I.; Némethy, G.; Scheraga, H. A. *Biopolymers* **1979**, *18*, 1565−1610.

(38) Augspurger, J. D.; Scheraga, H. A. *J. Comput. Chem.* **1997**, *18*, 1072−1078.

(39) Jensen, J. H.; Gordon, M. S. *J. Am. Chem. Soc.* **1995**, *117*, 8159−8170.

(40) Cramer, C. J.; Truhlar, D. G. *J. Comput. Chem.* **1992**, *13*, 1089−1097.

(41) Still, W. C.; Tempczyk, A.; Hawley, R. C.; Hendrickson, T. *J. Am. Chem. Soc.* **1990**, *112*, 6127−6129.

(42) Cramer, C. J.; Truhlar, D. G. *J. Am. Chem. Soc.* **1991**, *113*, 8552−8554.

(43) Liu, Q.; Brady, J. W. *J. Phys. Chem.* **1997**, *B101*, 1317−1321.

(44) Liu, Q.; Schmidt, R. K.; Teo, B.; Karplus, P. A.; Brady, J. W. *J. Am. Chem. Soc.* **1997**, *119*, 7851−7862.

(45) Ueda, K.; Brady, J. W. *Biopolymers* **1996**, *38*, 461−469.

(46) Gress, M. E.; Jeffrey, G. A. *Acta Crystallogr.* **1977**, *B33*, 2490−2495.

(47) Blackwell, J.; Sarko, A.; Marchessault, R. H. *J. Mol. Biol.* **1969**, *42*, 379−383.

(48) Rees, D. A.; Smith, P. J. C. *J. Chem. Soc., Perkin Trans. 2* **1975**, 836−840.

(49) Brant, D. A. *Annu. Rev. Biophys. Bioeng.* **1972**, *1*, 369−408.

(50) Sathyanarayana, B. K.; Rao, V. S. R. *Biopolymers* **1972**, *11*, 1379−1394.

(51) Ha, S. N.; Madsen, L. J.; Brady, J. W. *Biopolymers* **1988**, *27*, 1927−1952.

(52) Tran, V.; Buleon, A.; Imberty, A.; Peréz, S. *Biopolymers* **1989**, *28*, 679−690.

(53) French, A. D. *Carbohydr. Res.* **1989**, *188*, 206−211.

(54) Schmidt, R. K.; Teo, B.; Brady, J. W. *J. Phys. Chem.* **1995**, *99*, 11339−11343.

(55) Brady, J. W.; Schmidt, R. K. *J. Phys. Chem.* **1993**, *97*, 958−966.

(56) Wang, C. X.; Chen, W. Z.; Tran, V.; Douillard, R. *Chem. Phys. Lett.* **1996**, *251*, 268−274.

(57) Ott, K.-H.; Meyer, B. *Carbohydr. Res.* **1996**, *281*, 11−34.

(58) Anderson, A.; Carson, M.; Hermans, J. *Ann. N. Y. Acad. Sci.* **1986**, *482*, 51−59.

(59) Anderson, A. G.; Hermans, J. *Proteins* **1988**, *3*, 262−265.

(60) Hermans, J. In *Computer Simulation of Biomolecular Systems: Theoretical and Experimental Applications*; van Gunsteren, W. F., Weiner, P. K., Eds.; ESCOM: Leiden, The Netherlands, 1989; pp 139−148.

(61) Elber, R. *J. Chem. Phys.* **1990**, *93*, 4312−4321.

(62) Marrone, T. J.; Gilson, M. K.; McCammon, J. A. *J. Phys. Chem.* **1996**, *100*, 1439−1441.

(63) Straatsma, T. P.; McCammon, J. A. *J. Chem. Phys.* **1994**, *101*, 5032−5039.

(64) Torrie, G. M.; Valleau, J. P. *J. Comput. Phys.* **1977**, *23*, 187−199.

the repeat unit of xylan; for convenience, this molecule, which does not appear to have a trivial name or to have been previously modeled, will be referred to as "dixylose" throughout the remainder of this paper.)

We have selected this molecule for study because it is a prototype for possible effects of hydration upon conformation for a wide variety of biopolymers, and particularly for saccharides. Its regular structure, with both monomer units the same, makes it more tractable than polypeptides for extracting specific types of interactions which contribute to conformational equilibria and, in particular, for identifying specific solvent structuring effects which could not be mimicked by continuum models. Dixylose will have many of the important features of maltose, but without the exocyclic primary alcohol groups, and it can be simulated much more easily, since it will not be necessary to average over the many combinations of rotamers for these slowly reorienting groups. The principal difference between the conformational energy map for dixylose and that for maltose is expected to be that any steric repulsions between the primary alcohol groups of the glucose rings in maltose will be absent, perhaps opening up additional regions of glycosidic angle space disallowed in maltose by the crowding of these groups. By comparing the solution pmf map with the conformational energy map for the isolated disaccharide, the effects of hydration can be determined. Because this molecule has apparently not previously been studied experimentally, it also offers a special opportunity to make specific predictions concerning its solution conformation before the experiments are performed, as a more demanding test of methods and models than calculating already-known conformations.

II. Methods

As an initial step in the preparation of the conformational pmf for the dixylose molecule, an adiabatic Ramachandran energy map[32] was calculated for the disaccharide in a vacuum. This mapping, as well as the subsequent MD simulations, used the CHARMM molecular mechanics program[66] and a CHARMM-type force field for carbohydrates[67] which has been used to successfully reproduce known experimental properties of carbohydrates in a number of previous simulations.[4,8,12,44,45,51] The adiabatic map represents the molecular energy as a function of the glycosidic torsion angles ϕ and ψ (Figure 1), defined as $H1-C1-O1-C4'$ and $C1-O1-C4'-H4'$, respectively, where at every point on the map the lowest-energy combination of the other internal degrees of freedom was determined by constrained energy minimization. This relaxed energy was evaluated on a 20° grid over the entire range $[-180°,+180°]$ for both angles. To overcome the multiple-minimum problem,[68] four separate minimizations were carried out for each (ϕ,ψ) grid point, corresponding to all combinations of cases in which all hydroxyl groups of each ring pointed in either the clockwise or counterclockwise directions.[51] Since, in a vacuum, any other conformation for any of the hydroxyl groups would involve the disruption of an intramolecular hydrogen bond, all other arrangements were assumed to be of higher energy.[69] During the energy minimizations, the glycosidic torsion angles were constrained to remain at the grid point values through the application of stiff harmonic potentials centered on the desired angle. Because of the small size of this system, no long-range truncations of the nonbonded energy were employed in

this mapping.[66,70] The lowest energy found from the four minimizations for each (ϕ,ψ) point was used in the construction of the relaxed map.

In principle, the conformational free energy surface for a disaccharide in solution can be calculated from a very long MD or Monte Carlo (MC) simulation, since the potential of mean force as a function of the glycosidic angles, $W(\varphi,\psi)$, can be calculated from the probability distribution function $P(\varphi,\psi)$, the frequency with which each conformation occurs during the simulation, through the equation

$$W(\varphi,\psi) = -kT \ln P(\varphi,\psi) \qquad (1)$$

where $P(\varphi,\psi)$ can be computed from

$$P(\varphi,\psi) = \frac{\int \delta(\phi,\psi) e^{-\beta V(q)} \, dq}{\int e^{-\beta V(q)} \, dq} \qquad (2)$$

Unfortunately, this direct approach will in general not be successful, since by their very nature MD and MC simulations will tend to avoid high-energy regions, such as the barriers between low-energy conformations, and thus will give inadequate statistics for computing a meaningful $P(\varphi,\psi)$. This will clearly be true if a trajectory does not visit a particular region of conformation space, such as an isolated low-energy well. Similarly, if a trajectory is started in a high-energy conformation, which is separated from lower-energy regions by a barrier, it may spend some considerable time in that well before it succeeds in crossing the barrier. Unless it makes many crossings of this barrier, the time initially spent there will not be representative of the true relative probability of the initial conformation, and thus will distort the calculated $P(\varphi,\psi)$, giving an unrealistically low free energy for that conformer.

Carbohydrates are a particularly difficult case because their molecular energy can depend strongly on a number of degrees of freedom other than the glycosidic torsion angles, including the torsion angles specifying the directions of the many hydroxyl groups and the exocyclic hydroxymethyl groups.[69] While many of these groups rotate frequently in solution and amply sample their conformational range,[71,72] transitions in some of these angles can be very infrequent. Because both the structuring of the adjacent solvent and the solute energetics depend strongly on the proper averaging of these states, inadequate sampling due to short simulation times can give completely nonphysical results. This difficulty was clearly exhibited by a recent Car–Parrinello-type "first principles MD" [73] calculation of glucose in aqueous solution which lasted for only 3 ps.[74] This simulation produced results at variance with nearly all previous quantum mechanical and molecular mechanical studies, as well as with experiment, due to inadequate averaging over the hydroxyl rotational states. Molecular mechanics (MM) simulations can give similar results if averaged over such short periods, due to the unrealistic distribution of hydroxyl rotational states, but with averaging over periods 2 orders of magnitude longer, will converge to results which are stable, that is, which do not change with further increases in simulation averaging time.

Trajectory-averaged $P(\varphi,\psi)$'s have been reported from two standard MD simulations of disaccharides in solution, one for sucrose[4] and one for maltose.[57] In the sucrose example, the authors carefully avoided over-interpretation of this probability function, but in the maltose case, the authors claim that the $W(\varphi,\psi)$ computed from this probability function is the conformational free energy surface, even though themselves noting that this is not actually valid. Unfortunately, the complexity of carbohydrate energy surfaces and the infrequency of transitions make this approach incorrect even for simple disaccharides such as maltose or dixylose, and specialized techniques such as those

(65) Valleau, J. P.; Whittington, S. G. In *Statistical Mechanics. Part A: Equilibrium Techniques*; Berne, B. J., Ed.; Plenum Press: New York, 1977; pp 137–168.

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

(67) Ha, S. N.; Giammona, A.; Field, M.; Brady, J. W. *Carbohydr. Res.* **1988**, *180*, 207–221.

(68) Scheraga, H. A. *Biopolymers* **1983**, *22*, 1–14.

(69) Cramer, C. J.; Truhlar, D. G. *J. Am. Chem. Soc.* **1993**, *115*, 5745–5753.

(70) Tasaki, K.; McDonald, S.; Brady, J. W. *J. Comput. Chem.* **1993**, *14*, 278–284.

(71) Brady, J. W. *J. Am. Chem. Soc.* **1989**, *111*, 5155–5165.

(72) Schmidt, R. K. Ph.D. Thesis, Cornell University, 1995.

(73) Parrinello, M. In *Modern Techniques in Computational Chemistry: MOTECC-90*; Clementi, E., Ed.; ESCOM: Leiden, The Netherlands, 1990; pp 731–743.

(74) Molteni, C.; Parrinello, M. *J. Am. Chem. Soc.* **1998**, *120*, 2168–2171.

used in the peptide studies[58-63] are needed to calculate the relative free energies of different conformations.

To overcome the sampling problem in pmf calculations, a number of related free energy simulation techniques have been developed.[75-77] In the present simulations, the widely used method of umbrella sampling was employed to improve statistical convergence.[64,65] In umbrella sampling calculations the mechanical potential energy $V(\varphi,\psi)$ is augmented by the addition of an "umbrella" term $V_U(\varphi,\psi)$ selected so as to lower high-energy regions and thus increase the probability of sampling these regions. The unbiased probability $P(\varphi,\psi)$ needed for eq 1 can be obtained from the probability distribution $P^*(\varphi,\psi)$ observed from the MD trajectory calculated with umbrella potentials by employing a computational trick,[64,65] which simply multiplies the top and bottom of the standard probability expression (eq 2) by the same factors:

$$P(\varphi,\psi) = \frac{\delta(\phi,\psi)e^{-\beta(V+V_U)}(1/e^{-\beta V_U})\,dq}{\int e^{-\beta(V+V_U)}(1/e^{-\beta V_U})\,dq} \frac{\int e^{-\beta(V+V_U)}\,dq}{\int e^{-\beta(V+V_U)}\,dq} \quad (3)$$

$$= P^*(\varphi,\psi)e^{\beta V_U(\phi,\psi)}/\langle e^{\beta V_U}\rangle_U \quad (4)$$

The pmf without the biasing umbrella energy function, $W(\varphi,\psi)$, can then be calculated from

$$W(\varphi,\psi) = -kT \ln P^*(\varphi,\psi) - V_U(\varphi,\psi) + kT \ln \langle e^{\beta V_U(\phi,\psi)}\rangle_U + C \quad (5)$$

where C is an arbitrary constant and $\langle\ \rangle_U$ indicates averaging over the ensemble using the umbrella potential V_U.

Selecting an appropriate umbrella potential function is not a simple task,[78,79] and in practice, it was necessary to progressively adapt the choice of this function as the calculation proceeded and more was learned about the conformational behavior in solution.[80-83] As a first approximation, the negative of the vacuum adiabatic energy map could be employed as the starting umbrella potential, since if the pmf in solution was exactly the same as this vacuum energy surface, this choice would have the effect, on average, of canceling the intramolecular conformational energy for all values of the torsion angles. Because of entropic effects, the conformational pmf will not be identical to the relaxed mechanical energy surface even in a vacuum, but this approach was found to work well for the calculation of the vacuum pmf for maltose.[54] If significant solvation effects are present, however, this approximation will work less satisfactorily. In the present case, preliminary simulations using this choice for the umbrella potential function found that it was a poor approximation, judging from the failure of the trajectories to widely explore the (ϕ,ψ) conformational space. Accordingly, alternate strategies for the selection of the trial umbrella functions, described shortly, were used in subsequent simulations.

In the various calculations, umbrella potential functions were applied to an otherwise conventional series of MD simulations of the disaccharide in aqueous solution. The disaccharide was modeled surrounded by water molecules in a cubic box 24.6481 Å in length subject to periodic boundary conditions.[55] The TIP3P water model[84] was used since the CHARMM force field was developed to be compatible with this model, particularly with respect to the exchange of solute−solute and water−water hydrogen bonds. Neither the CHARMM nor the TIP3P force field employs special hydrogen bonds functions, but rather uses appropriately parametrized van der Waals and electrostatic interactions to represent hydrogen bonds. The simulations were initiated by placing the disaccharide in the center of a previously equilibrated box of 512 water molecules and removing those solvent molecules which overlapped with the solute heavy atoms, which produced a primary system containing 492 water molecules and one solute molecule. The trajectories were integrated using a Verlet integration algorithm[85] with a step size of 1 fs. Chemical bond lengths involving hydrogen atoms, and the geometries of the water molecules, were kept rigid using the constraint algorithm SHAKE.[86] The trajectories were initiated by assigning starting atomic velocities from a Boltzmann distribution at 300 K, followed by an equilibration period of 20 ps (without the umbrella potential) during which atomic velocities were periodically scaled if needed to maintain the system temperature at 300 ± 3 K. During the subsequent dynamics used for data collection, the velocities were not scaled, the temperature was stable, and total energy was well conserved. The large number of nonbonded interactions in the solution simulations required that long-range interactions be smoothed to zero between 10 and 12 Å using switching functions applied on a neutral group basis.[66,70] In using the trajectories to compute the pmf, the allowed conformational region was partitioned into "boxes" $5° \times 5°$ in size, and the probability for each of these small regions was calculated from the fraction of time spent in it during the simulations.

Disaccharide conformational fluctuations are strongly damped in aqueous solution by collisions with the water molecules,[55] and large-scale conformational changes occur slowly. As a result, even with a perfect umbrella potential function the rate of convergence of the calculation will be slow if the Ramachandran pmf has more than one minimum-energy well, since many transitions covering the interesting regions of the map will be needed to ensure meaningful average probability densities. The total simulation time required might be reduced if a number of separate simulations, starting in different regions of the glycosidic angle space, could be used. In some of the first pmf calculations in water, for the approach of nonpolar spheres in solution, a series of simulations centered at different values of the single radial separation coordinate were performed, selected so that there was significant overlap in the separate simulations.[87,88] While the arbitrary constants in these simulations were in general independent, the overlap allowed the simulations to be appropriately scaled and combined. Unfortunately, in the present problem, the two-dimensional character makes such normalization more problematic.[89] In the present calculations, several independent simulations were conducted, starting in different regions of the vacuum adiabatic energy map. The conformational regions explored by these different trajectories overlapped significantly at the edges, allowing for consistent normalization of the densities calculated from each trajectory. This task was considerably aided by the fact that all of the overlaps were pairwise in character. That is, the conformational region explored by trajectory I intersected with that explored by simulation II, which also overlapped with the region sampled by trajectory III, but there were no cases where all three overlapped, or even where I overlapped with III. In principle, the normalization coefficient should be averaged over all parts of the overlap region, but in the present case, a single-point normalization was calculated for the overlap point (actually a grid "box" in torsion angle space $5° \times 5°$ in size) with the highest mutual occupancy. This

(75) Beveridge, D. L.; DiCapua, F. M. In *Computer Simulation of Biomolecular Systems: Theoretical and Experimental Applications*; van Gunsteren, W. F., Weiner, P. K., Eds.; ESCOM: Leiden, The Netherlands, 1989; pp 1−26.

(76) Reynolds, C. A.; King, P. M.; Richards, W. G. *Mol. Phys.* **1992**, *76*, 251−275.

(77) King, P. M. In *Computer Simulation of Biomolecular Systems, Vol. 2: Theoretical and Experimental Applications*; van Gunsteren, W. F., Weiner, P. K., Wilkinson, A. J., Eds.; ESCOM: Leiden, The Netherlands, 1993; pp 267−314.

(78) Beutler, T. C.; van Gunsteren, W. F. *J. Chem. Phys.* **1994**, *100*, 1492−1497.

(79) Harvey, S. C.; Prabhakaran, M. *J. Phys. Chem.* **1987**, *91*, 4799−4801.

(80) Mezei, M. *J. Comput. Phys.* **1987**, *68*, 237−248.

(81) Hooft, R. W. W.; van Eijck, B. P.; Kroon, J. *J. Chem. Phys.* **1992**, *97*, 1690−6694.

(82) van Eijck, B. P.; Hooft, R. W. W.; Kroon, J. *J. Phys. Chem.* **1993**, *97*, 12093−12099.

(83) Bartels, C.; Karplus, M. *J. Comput. Chem.* **1997**, *18*, 1450−1462.

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

(85) Verlet, L. *Phys. Rev.* **1967**, *159*, 98−103.

(86) van Gunsteren, W. F.; Berendsen, H. J. C. *Mol. Phys.* **1977**, *34*, 1311−1327.

(87) Pangali, C.; Rao, M.; Berne, B. J. *J. Chem. Phys.* **1979**, *71*, 2975−2981.

(88) Pangali, C.; Rao, M.; Berne, B. J. *J. Chem. Phys.* **1979**, *71*, 2982−2990.

(89) Mezei, M.; Mehrotra, P. K.; Beveridge, D. L. *J. Am. Chem. Soc.* **1985**, *107*, 2239−2245.
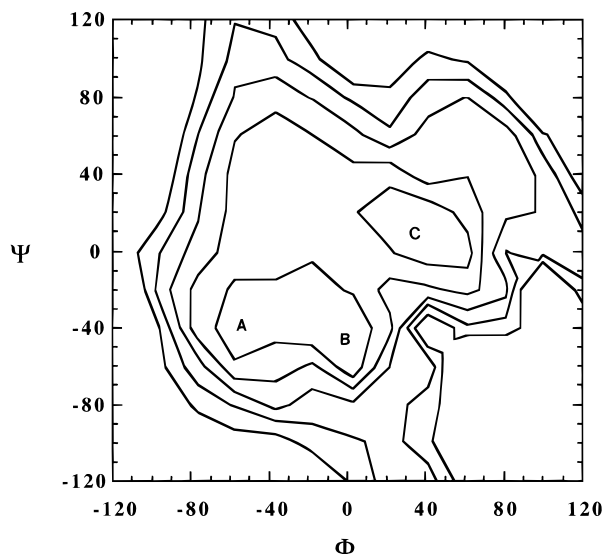
Figure 2. Vacuum adiabatic Ramachandran energy map for the dixylose molecule shown in Figure 1. Energy contours are indicated at 2, 4, 6, 8, and 10 kcal/mol above the global minimum in the "B" well.

approach was acceptable in the intermediate calculations since the resulting composite pmf was only used as an approximate trial umbrella potential energy term in the next cycle of simulations.

Following the preliminary simulation in which the umbrella potential was the negative of the vacuum adiabatic surface, a series of MD simulations was next performed in which the umbrella potentials used were iteratively adapted,[80−83] based upon the results of successive previous simulations. To explore the importance of the intrinsic torsional terms in the energy function

$$V_\phi = k_\phi \cos(n\phi)$$

$$V_\psi = k_\psi \cos(n\psi) \tag{6}$$

an umbrella potential which inverted, or "flattened", these two terms, $-V_\phi$ and $-V_\psi$, was employed. Five 100 ps MD simulations using these umbrella potentials were calculated, each starting from a different position on the vacuum Ramachandran map: (50°,0°), (0°,−20°), (−7°,−4°), (−20°,20°), and (−60°,60°). Each of these simulations was equilibrated using constraining potentials in ϕ and ψ which kept the molecule at this conformation; these constraints were then relaxed for data collection at the end of the equilibration. The combination of these five trajectories covered a substantial portion of the (ϕ,ψ) space, with −65° ≤ ϕ ≤ 65° and −65°≤ ψ ≤ 90°. The (ϕ,ψ) probability densities were then calculated from the combination of these trajectories using the one-point normalization procedures described above, and a preliminary approximation to the pmf, $W(\phi,\psi)$, was calculated from eq 2. This preliminary pmf contained three minima: two of approximately equal energy at (−50°,40°) and (50°,40°) and a third 0.5 kcal/mol higher in energy at (−60°,−30°).

This estimate of the pmf was used as the umbrella potential in a third round of simulations, with an additional modification. Because this intermediate estimate of the pmf de-emphasizes the low-energy wells, the molecule is now able to wander away from these wells and into both the saddle regions between these wells, where more sampling is highly desirable, and also into regions around the boundaries of the central region being studied, which are of less interest. The regions of the vacuum map which have ϕ angles in the ranges [−180°,−120°] and [120°,180°] have very high energies due to van der Waals overlaps[54] (see Figure 2) and thus are very unlikely to have low energies on the solution pmf. For this reason an additional constraining potential function was applied to keep the ϕ angle in the range [−120°,120°], thus improving the sampling frequency in the more probable low free energy regions. This constraining potential consisted of an additional energy term which increased linearly from 0 to 4.5 kcal/mol between

Table 1. Locations and Relative Energies of the Principal Minima on the Adiabatic Vacuum and Solution pmf Conformational Energy Maps for the Dixylose Molecule

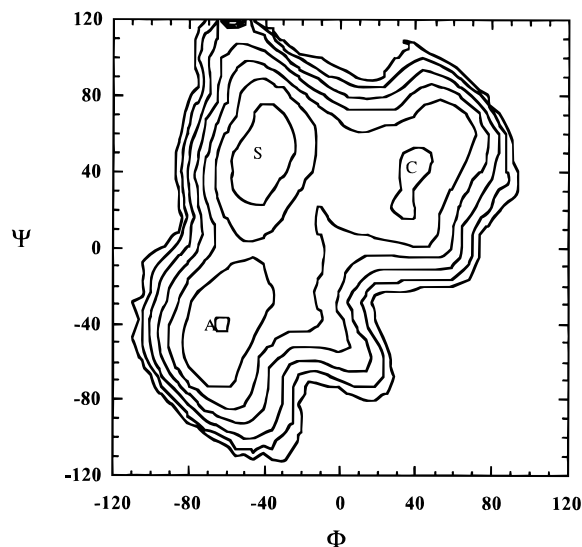| local minimum	ϕ (deg)	ψ (deg)	energy (kcal/mol)
	Vacuum Adiabatic Map		
A	−60	−40	0.89
B	0	−40	0.00
C	40	20	0.10
F	−40	180	6.08
	Solution PMF Map		
A	−62	−41	0.55
C	40	40	1.37
S	−37	52	0.00

Figure 3. Final calculated Ramachandran free energy map for the dixylose molecule in aqueous solution. Energy contours are indicated at 1, 2, 3, 4, 5, 6, and 7 kcal/mol above the global minimum in the "S" well.

100° and 120°. This was necessary to keep these areas from being the only ones sampled when the pmf was "inverted" as an umbrella function. Similarly, this same restraining function was applied to ψ, even though there is a local minimum in the (−40°,180°) region of the vacuum adiabatic map (see below). This was done since the available experimental evidence for the related maltose molecule indicates that the central region is the one most likely to be biologically important and since the slow rate of sampling of the conformational space precluded averaging over the larger range of ψ in the present calculations. The polar (−40°,180°) well is so much higher in energy on the vacuum surface (>6 kcal/mol) that it is unlikely to be important in solution, even if favored by hydration. In this next cycle of simulations, the inverses of the intrinsic torsional energy terms for ϕ and ψ were no longer used. Using the new combined umbrella potential function, three 500 ps simulations were conducted, one starting from each of the three local minima identified in the preliminary pmf described above. These minima were located at (50°,40°), (−50°,40°), and (−60°,−30°). Although some transitions between these wells were observed in the simulations, these were infrequent, so once again, the data from these three independent simulations were combined using a one-point normalization procedure as described above. From this joint probability distribution function, a new pmf was calculated which again possessed three minima, with the global minimum being at approximately (−50°,40°).

Next, another cycle of refinement of this pmf was calculated by computing an additional 4 ns trajectory using the negative of this second intermediate pmf as the umbrella potential, starting from the conformation at (−7°,−4°). This conformation was arbitrarily selected since it was approximately equally distant from each of the three wells on the intermediate pmf determined after the last cycle. In this simulation a number of transitions were observed between each of the wells and
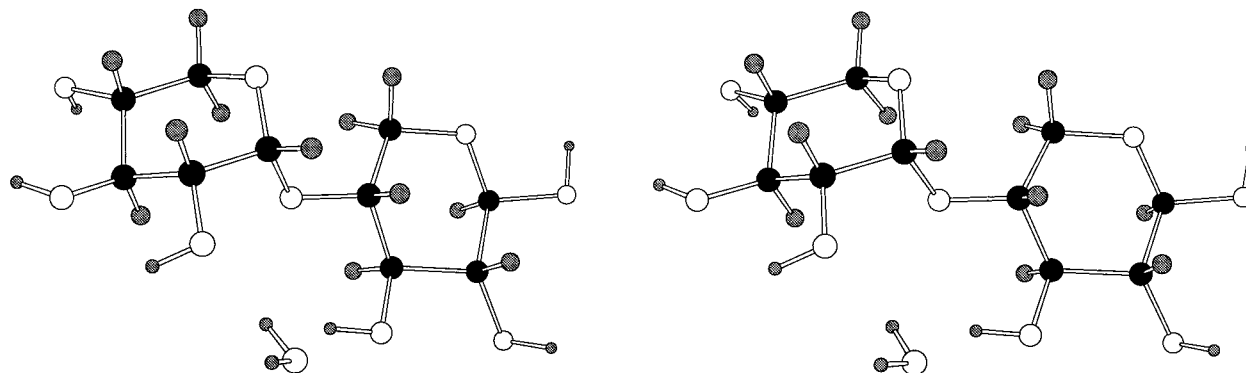
Figure 4. Representative "snapshot" configuration of the dixylose molecule in the "S" conformation, including a water molecule which is hydrogen-bonded to hydroxyl groups on both rings of the disaccharide.

the higher energy barrier regions between the wells were well explored. The probability distribution function $P^*(\phi,\psi)$ for this simulation was used to compute yet another estimate of the pmf. A final cycle of refinement of the pmf was then computed from an additional 2.7 ns trajectory which used the 4 ns pmf as its umbrella function. This final simulation also exhibited a number of spontaneous transitions between wells, indicating that the previous trial pmf used as an umbrella potential in this simulation is close to the actual pmf. Accordingly, the probability distribution function from this simulation was used to calculate the final pmf. A total of 8.7 ns of solution dynamics simulations were calculated in these various cycles of approximation. Following the calculation of the pmf, four standard 100 ps MD simulations were computed without umbrella functions, starting in each of the three minima of the vacuum surface, "A", "B", and "C", and also from the new "S" minimum of the pmf, to test whether the pmf accurately reflects the conformational free energy experienced by the molecule in solution.

III. Results and Discussion

The calculated vacuum adiabatic energy map for dixylose is shown in Figure 2. As might be expected, this map closely resembles the various maps previously produced for maltose.[51] The global minimum is located in the well labeled "B", and the locations and relative energies of the various wells are given in Table 1. The barrier between the "A" and "B" wells is less than 0.4 kcal/mol above the "A" well and less than 0.9 kcal/mol above the "B" well. The map is only shown for ϕ and ψ angles in the range [$-120°,120°$] to facilitate comparison with the more restricted pmf map in solution, but like maltose, it also has a local minimum, called "F" to maintain consistency with the maltose nomenclature,[51] around ($-40°,180°$) with an energy 6.08 kcal/mol above the global minimum at "B". This energy is much higher than for the maltose map because in the maltose case the "F" conformation is stabilized by a hydrogen bond between the exocyclic hydroxyl group O′(6)H of the reducing residue and the O2 group of the nonreducing residue. Without these exocyclic hydroxymethyl groups, such a stabilizing hydrogen bond is not possible in dixylose. The other significant difference between this map and that for maltose is that the energy of the region around ($-40°,+40°$), while still about 3 kcal/mol above the minima at "A" and "B", has been lowered by about 3 kcal/mol relative to the maltose case (compare Figure 2 of ref 54), due to the absence of crowding of the exocyclic primary alcohol groups.

Figure 3 displays the final calculated conformational pmf for the dixylose molecule in aqueous solution. As can readily be seen, this Ramachandran map is significantly different from the vacuum adiabatic map of the mechanical potential energy (Figure 2). The most striking differences are that a new well has appeared in the upper left quadrant, which contains the global minimum "S" located at ($-37°,52°$), while the low-

energy region labeled "B" in the vacuum map of Figure 2 has disappeared. In addition, the shape and relative energy of the "C" well have been significantly affected; this well is now much narrower and also much higher in energy than on the vacuum adiabatic surface. The positions and relative energies of these minima are also listed in Table 1.

These changes in relative conformational energies are apparently due to differences in how the conformers structure adjacent water molecules. Examination of instantaneous molecular positions from the trajectories suggests reasons for the appearance of the "S" geometry as the global free energy minimum in solution. In this conformation, the disaccharide can make two hydrogen bonds with a single water molecule which bridges between the O2 and O3′ hydroxyl groups located on the different rings. Figure 4 shows a representative "snapshot" from the trajectory illustrating a water molecule making such simultaneous hydrogen bonds. As can be seen from the figure, this "S" conformation is also stabilized by hydrophobic interactions between the two aliphatic C(5)H$_2$ groups, since this conformation brings them almost into contact. The shared water molecule bridging between two rings of a disaccharide is almost exactly similar to the situation observed in previous simulations of neocarrabiose,[45] in which the presence of a doubly-hydrogen-bonded bridging water molecule induced a conformational shift away from the vacuum minimum energy geometry to that expected from diffraction experiments. It is not clear why the "B" geometry in the present dixylose case is destabilized by hydration, since it would appear that a water molecule could also bridge between the two rings in this geometry as well; perhaps additional geometric requirements make this unfavorable. However, in this "B" conformation, the two CH$_2$ groups are not in close proximity, and it is likely that they are not at an optimal distance for hydrophobic interaction. These two groups are approximately 5 Å apart in the "B" conformation; this distance does not correspond to a free energy minimum on the pmf calculated by Pangali et al. for the approach of two nonpolar spheres in aqueous solution.[87,88] Although the systems are not exactly the same, this suggests that the "B" conformation might be destabilized by an unfavorable hydrophobic contribution, regardless of whether a favorable water-bridging interaction is possible.

As already noted, the bridging water molecule seen here which is hydrogen bonded to two solute hydroxyl groups is another example of such a bridging situation for sugars in solution using the present energy function and the TIP3P or SPC/E water models.[8,12,43−45] Ben-Naim has proposed that such interactions should be common and important in biological

solutes and should have a large free energy of association,[90] and MM simulations appear to confirm this view.[91] However, Kollman has recently proposed that this hydrophilic association contributes only a very small amount to the total free energy.[92] Unfortunately, the present study provides no direct information on how the free energy could be broken down into component contributions and no estimates of the relative magnitudes of any possible hydrophobic and hydrophilic solvation effects. However, the frequent appearance of such water bridges between hydroxyl groups in our simulations suggests that these interactions have a significant energy.

Separate conventional MD simulations of the disaccharide without any umbrella potential were conducted starting from the "A", "B", "C", and "S" geometries. These trajectories are independent simulations which were not used in the calculation of the pmf. Figure 5 displays the superposition of several of these trajectories onto the calculated pmf surface. As can be seen from these plots, the pmf is a good description of the free energy which the molecule experiences in solution, with the trajectories confined to the low-energy regions of the surface. Without the umbrella potential, the molecule was found to be stable in the "S" geometry and not stable in the "B" geometry. The trajectory which began in the "A" well divided its time between the "A" and "S" wells, making several transitions over the low-energy barrier. Clearly, this trajectory and the one in the "S" well would make little sense projected onto the vacuum adiabatic energy map (Figure 2).

The implications of these results for the solution conformation of maltose are unclear. Since the reason for the disappearance of the "B" well is uncertain, it is possible that this effect might also occur in maltose, or even other oligosaccharides. Such a conformational shift toward the "A" well would be consistent with the experimental data for maltose,[18,21] which is thought to imply a shift away from the crystal structure in the "C" well to the "A" conformation. On the other hand, while the "S" conformation is quite important in the present simulations for the dixylose molecule, it is probable that this conformation is not an important one for maltose itself. For maltose in this conformation, steric hindrance between the bulky exocyclic primary alcohol groups raises the conformational energy by at least 1−3 kcal/mol on the vacuum map[51,54] and might also disrupt the hydrophobic association of the C(5)H moieties as well, thus disfavoring this conformation significantly. It is interesting to note, however, that the "S" conformation would be consistent with both the NOE[18] and chiroptical[21] data for maltose.

IV. Conclusions

The present results demonstrate that hydration can significantly affect the conformational equilibrium for biopolymers in solution.[58−63] This result is consistent with previous pmf studies of peptides in solution. For the alanine "dipeptide", the C_7^{eq} conformation, which is almost never observed in protein crystal structures, is the lowest-energy conformer in vacuo due to a strained intramolecular hydrogen bond. In solution, however, competition by water molecules to hydrogen-bond with the carbonyl oxygen and amide proton result in significant stabilization of the α-helical and β-sheet conformations. Even for simple prototypical species such as the dixylose molecule studied here, the complex interplay of solvation contributions such as bridging by hydrogen-bonded water molecules and
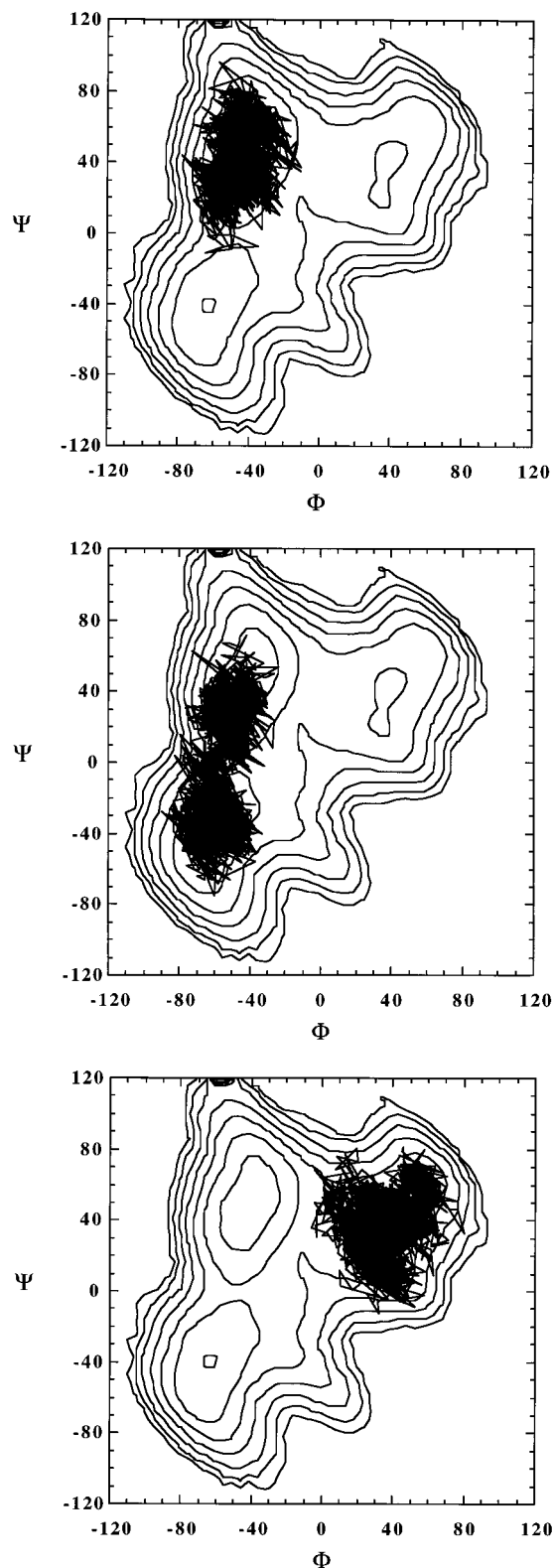
(90) Ben-Naim, A. *J. Chem. Phys.* **1989**, *90*, 7412−7425.
(91) Mezei, M.; Ben-Naim, A. *J. Chem. Phys.* **1990**, *92*, 1359−1361.
(92) Sun, Y.; Kollman, P. *J. Phys. Chem.* **1996**, *100*, 6760−6763.

Figure 5. Examples of standard 100 ps simulations, without the umbrella potential functions, superimposed on the solution pmf. The unconstrained trajectories closely follow the contours of the pmf, illustrating that it is indeed the free energy surface which the molecule experiences in solution: (a) trajectory for a simulation initiated in the "S" conformation; (b) trajectory for a simulation initiated in the "A" conformation (note that this trajectory made transitions between the "A" and "S" wells, following the low-energy col between the two regions); (c) trajectory for a simulation initiated in the "C" conformation.

hydrophobic interactions between the methylene groups makes it difficult to anticipate from inspection the conformational

effects of hydration. Unfortunately, in some cases, these effects may also be too complex to be adequately represented by continuum models. The compromises made by the solvent in simultaneously structuring relative to several nearby functional groups, on both the same and neighboring monomer residues, inevitably will favor some parts of conformation space over others and may lead to important conformational shifts.

The present results demonstrate that, through the application of successive iterations of adaptive sampling, it is possible to calculate conformational pmfs in solution for disaccharides, although the computer time required is very large. As was found by Straatsma and McCammon,[63] the intrinsic torsional potential associated with the two torsional angles ϕ and ψ contributes only a small amount to the total variation in the free energy with these two angles and, thus, makes a poor choice for the umbrella potential function, although its inclusion in the umbrella potential is simple and does smooth out the free energy surface somewhat. Similarly, the vacuum relaxed map, or even the vacuum pmf, would be poor choices for the umbrella function in any system in which there are large solvation effects, as in the present case. In general, the actual solution pmf may not be obvious, just as the stability of the "S" conformation was not expected a priori in the present case, making it harder to "guess" an appropriate initial umbrella function. For these reasons an adaptive strategy is necessary. Because of the slow rate of diffusive motions in aqueous solution, even with the umbrella potentials, the rate of convergence can be very slow and the very long simulation times required prevented us from considering the polar regions of the angles, even though there is a local minimum on the vacuum map around $(-40°, 180°)$. Future efforts will attempt to explore this area and connect it with the central low-energy region, perhaps using constrained dynamics approaches. Alternatively, a reaction path between these regions could be sought using perturbation approaches.[59] The present pmf map may also be useful as an initial trial umbrella potential for calculating the pmf for maltose in solution; such a calculation is presently underway. It would also be interesting to explore the dependence of the results on the choice of carbohydrate potential energy function and on the water model used, since the magnitude of the sugar−water hydrogen bond energy could obviously affect the stability of the "S" conformation.

The enormous effort required in the preparation of this limited pmf clearly demonstrates the desirability of using various implicit solvation models whenever possible.[69,93] Unfortunately, the surprising and unexpected conformational shift occasioned by hydration, which gives rise to the "S" conformation, also demonstrates that in some cases simple implicit solvation models cannot account for all of the effects of hydration. For this reason, MC and MD simulations which explicitly include solvent water molecules will continue to be needed until implicit models can more reliably include the specific structuring effects of hydrophobic and hydrophilic hydration. However, the present results also indicate that the inclusion of only a very small number of specific water molecules, perhaps only one or two in some cases, combined with the continuum models, could be quite successful.[94]

Finally, it should be pointed out that the results of this calculation stand as an a priori prediction of the conformation for this molecule in solution. As discussed in the Introduction, this molecule apparently has not been previously studied experimentally, although it is the simplest prototype for a number of important saccharide polymers. However, although the experiments might be moderately difficult, there is no major obstacle to carrying out several probes of dixylose solution conformation. The synthesis of this molecule from xylose would be relatively straightforward, although product isolation and preparing variants specifically deuterated at particular positions for NMR study might be more challenging.[95] Chiroptical measurements of the type pioneered by Stevens[21−23] should also be possible, facilitated by the absence of exocyclic primary alcohol groups. It is to be hoped that the unambiguous and surprising predictions of the present simulations will prompt experimental tests of these calculations soon.

(93) Senderowitz, H.; Parish, C.; Still, W. C. *J. Am. Chem. Soc.* **1996**, *118*, 2078−2086.
(94) Rick, S. W.; Berne, B. J. *J. Am. Chem. Soc.* **1994**, *116*, 3949−3954.
(95) Cumming, D. A.; Dime, D. S.; Grey, A. A.; Krepinsky, J. J.; Carver, J. P. *J. Biol. Chem.* **1986**, *261*, 3208−3213.