The Anomeric Equilibrium in D-Xylose: Free Energy and the Role of Solvent Structuring

R. K. Schmidt,[†] M. Karplus,*[,‡] and J. W. Brady*[,†]

Contribution from the Department of Food Science, Stocking Hall, Cornell University, Ithaca, New York 14853, and Department of Chemistry, Malinckrodt Hall, Harvard University, 12 Oxford Street, Cambridge, Massachusetts 02138

Received April 3, 1995[⊗]

Abstract: Standard molecular dynamics (MD) and free energy simulations have been used to analyze the anomeric equilibrium for D-xylose in aqueous solution. This molecule was selected as a simple model for the anomeric effect in sugars since there are no complications arising from the rotameric distribution of a primary alcohol. As in previous studies of glucose, the free energy calculations found a small free energy difference (0.15 kcal/mol favoring the α form) between the anomers; the experimental value is also small (−0.38 kcal/mol) but favors the β form. Thermodynamic integration analysis of the components contributing to this result showed that, as in the case of glucose, the free energy difference results from the balance between an internal term favoring the α anomer and a solvation term favoring the β anomer. The validity of the partitioning is confirmed by direct evaluation of the energy difference and its components from long standard MD simulations of each anomer. Hydrogen bonding analysis of the MD simulations provides a mechanistic explanation of the solvation preference for the β anomer. It results from improved hydrogen bonding of the anomeric hydroxyl group in the β anomer as the result of an increased accessible surface area. Three-dimensional averaging demonstrates the anisotropy in the solvent structure surrounding the two solutes and its dependence upon configuration.

I. Introduction

In aqueous solution the reactive aldehyde groups of reducing aldose sugars cyclize to produce five- and six-membered rings which are more stable than their open chain forms.[1,2] Closing the chain into a ring converts the C1 carbon atom into an asymmetric center. This results in two possible stereochemical species, called anomers. The reversible cyclization reaction is acid and base catalyzed and leads to an equilibrium mixture of the various possible forms, collectively called tautomers, with the proportions of the different species at equilibrium being determined by their relative free energies. The equilibrium concentrations have been measured for all of the commonly occurring sugars,[1] and in general, the so-called "pyranoid" six-membered rings predominate. For D-glucose in water, the equilibrium concentrations are approximately 64% for the β-pyranoid ring and 36% for the α-pyranoid ring, with negligible amounts of the open chain and "furanoid" (five-membered) ring forms. Other sugars have different ratios; mannose, for example, has almost exactly the opposite distribution, with 67% of the α-pyranoid form and 33% of the β-pyranoid tautomer. Table 1 lists the measured tautomeric concentrations for selected sugars. The atypical tautomeric distribution for D-idose apparently results from the all-axial hydroxyl distribution for the α form, which gives rise to alternate ring forms as well as to furanoid tautomers.

It has been noted that, for most C1-substituted sugars, the α anomer (axial in $^4C_1$ pyranoid rings) is more stable than the β.[3,4] This general phenomenon has been attributed to a so-called "anomeric effect" which stabilizes the α form. Several explanations for the anomeric effect have been advanced. The most widely accepted model is based on a back-donation of electronic density from the lone pairs of the ring oxygen atom into an antibonding σ* C1−O1 orbital, an explanation supported by crystallographic evidence of configuration-dependent variation in C1−O1 bond lengths.[5] Since the observed predominance of the β anomer for D-glucose is at variance with the anomeric effect, it has been suggested that the aqueous environment favors the β anomer over the α. The anomeric ratio would then result from a balance of internal energy terms favoring the α anomer and solvation terms favoring the β anomer. The precise balance of the two terms is expected to vary for the different sugars. This view is supported by the experimental observation that the anomeric ratio is solvent dependent; for glucose it can be inverted in DMSO solution at low temperature.[6]

Although the discussion of anomeric effects goes back many years, there has been interest recently in applying quantitative theoretical methods to the anomeric ratio for sugars,[7] as a way of better understanding the nature of this equilibrium and as a means of validating the computational models. Both isolated molecule and solvent effect calculations have been made. A number of semiempirical and *ab initio* quantum mechanical calculations of sugars[8] and analogous molecules[9−11] have

Table 1. Experimental Percentages of Each Tautomer in Aqueous Solution for Selected Sugars at 20 °C[1]

	α-pyranoid	β-pyranoid	α-furanoid	β-furanoid
D-glucose	36	64		
D-xylose	35	65		
D-mannose	67	33		
D-lyxose	75	25		
D-galactose	32	64	1	3
D-idose	31	37	16	16

[†] Cornell University.
[‡] Harvard University.
[⊗] Abstract published in *Advance ACS Abstracts,* December 15, 1995.

(1) Shallenberger, R. S. *Advanced Sugar Chemistry*; AVI Publishing Co.: Westport, CT, 1982.
(2) Stoddart, J. F. *Stereochemistry of Carbohydrates*; Wiley-Interscience: New York, 1971.
(3) Edward, J. T. *Chem. Ind. (London)* **1955**, 1102.
(4) Lemieux, R. U.; Chu, P. *Abstracts of Papers*, 133rd National Meeting of the American Chemical Society, San Francisco, CA, American Chemical Society: Washington, DC, 1958.
(5) Jeffrey, G. A. In *Anomeric Effect: Origin and Consequences*; Horton, D., Ed.; ACS Symposium Series 87; American Chemical Society: Washington, DC, 1979.
(6) Franks, F. *Pure Appl. Chem.* **1987**, *59*, 1189.
(7) Maddox, J. *Nature* **1993**, *364*, 669.

0002-7863/96/1518-0541$12.00/0

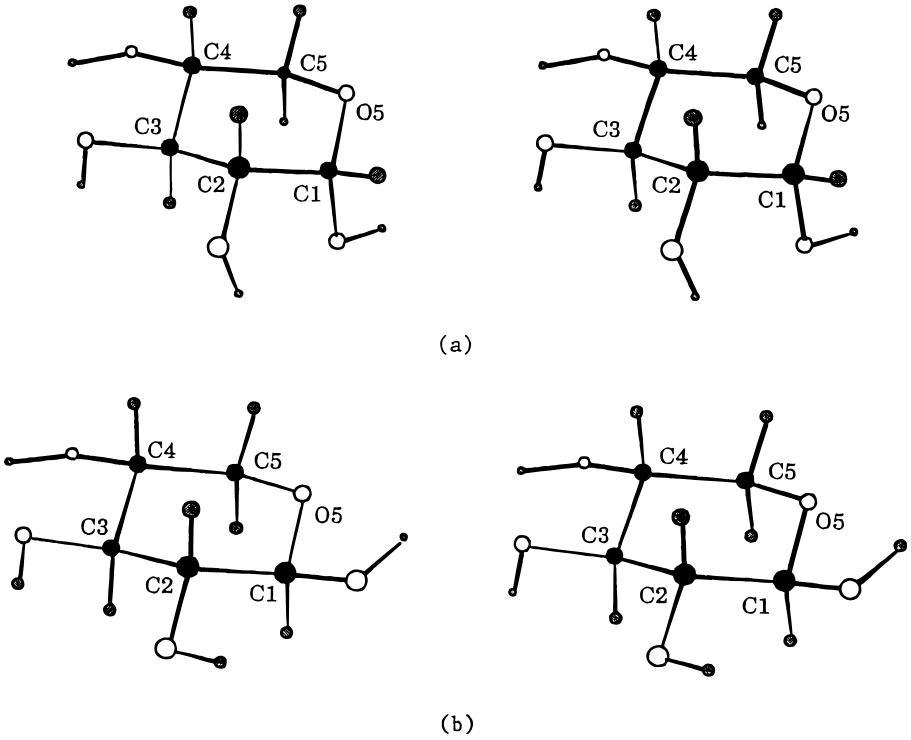

Figure 1. Stereodiagrams of (a) α-D-xylopyranose and (b) *β*-D-xylopyranose. Both molecules are shown in the counterclockwise geometry.

appeared in the literature. These calculations are hampered by the large number of heavy atoms in a sugar molecule, as well as by the many distinct conformational states (ring forms, hydroxyl orientations, and primary alcohol conformations). However, nearly every quantum mechanical calculation of sugars and their analogues has found that the isolated α anomer (or its analogues) is more stable and thus has confirmed the existence of an anomeric effect stabilizing α anomers in glucose-like molecules.

Molecular mechanics calculations have also been applied to the anomeric equilibrium problem. These simulations cannot represent effects arising from orbital interactions if they are not explicitly included in the parameterization of the potential energy functions, but they allow the direct modeling of solvation. In such calculations, and in reality, the simple picture of competing intramolecular and hydration terms is somewhat more complicated because averaging over many states is required. For example, apart from quantum mechanical effects of the solvent polarizing the solute, the presence of the solvent can result in conformational shifts and averaging of a different distribution of conformational states than that for the isolated molecule. This leads to a blurring of the distinction between intramolecular and hydration contributions. An important use of simulations is that they allow the direct evaluation of differences in hydration energies and solvent structures among the anomers. MD simulations have been made for both anomers of D-glucose,[12,13] and free energy simulations have been used to calculate the free energy difference between the two anomers in solution.[14,15] The results of these calculations have generally supported the model of competing solvation and anomeric terms of opposite signs. However, the very small magnitude of the experimental anomeric free energy difference (∼0.3 kcal/mol) makes these calculations difficult, as do uncertainties about the potential energy function and methodological protocols.

Two recent studies of the anomeric effect in D-glucose have reported results which differ from this overall picture. van Eijck and co-workers have made free energy simulations employing the GROMOS force field and special simulation protocols; they found only small solvation effects on the glucose anomeric equilibrium.[16] Cramer and Truhlar have employed a combination of semiempirical quantum mechanical calculations with a solvation correction term to evaluate the glucose anomeric ratio,[17] and again found little solvation effect. Both of these studies appear to be at variance with the available experimental evidence, since *β* is the observed dominant form, in spite of the quantum mechanical predictions of a more stable α anomer for the isolated molecule. In contrast to these two calculations, a recent pair of experimental and theoretical studies of a prototypical pyranose analogue, 2-methoxytetrahydropyran, provide strong support for the existence of both an internal anomeric effect and a significant solvation contribution in aqueous solution.[11,18]

All of the previous studies of anomeric effects in sugars have involved hexoses in which the orientation of the primary alcohol may potentially affect the anomeric equilibrium. However, simpler anomeric sugars are available for study. D-Xylose (Figure 1) is the pentose analogue of D-glucose (lacking only the exocyclic primary alcohol group) and has the same anomeric ratio as D-glucose (see Table 1). Xylose has far fewer

(8) Polavarapu, P. L.; Ewig, C. S. *J. Comput. Chem.* **1992**, *13*, 1255−1261.
(9) Jeffrey, G. A.; Pople, J. A.; Binkley, J. S.; Vishveshwara, S. *J. Am. Chem. Soc.* **1978**, *100*, 373−379.
(10) Wiberg, K. B.; Murcko, M. A. *J. Am. Chem. Soc.* **1989**, *111*, 4821−4828.
(11) Wiberg, K. B.; Marquez, M. *J. Am. Chem. Soc.* **1994**, *116*, 2197−2198.
(12) Brady, J. W. *J. Am. Chem. Soc.* **1989**, *111*, 5155−5165.
(13) van Eijck, B. P.; Kroon-Batenburg, L. M. J.; Kroon, J. *J. Mol. Struct.* **1990**, *237*, 315−325.
(14) Ha, S.; Gao, J.; Tidor, B.; Brady, J. W.; Karplus, M. *J. Am. Chem. Soc.* **1991**, *113*, 1553−1557.
(15) Glennon, T. M.; Zheng, Y.-J.; Le Grand, S. M.; Shutzberg, R. A.; Merz, K. M. *J. Comput. Chem.* **1994**, *15*, 1019−1040.
(16) van Eijck, B. P.; Hooft, R. W. W.; Kroon, J. *J. Phys. Chem.* **1993**, *97*, 12093−12099.
(17) Cramer, C. J.; Truhlar, D. G. *J. Am. Chem. Soc.* **1993**, *115*, 5745−5753.
(18) Jorgensen, W. L.; Morales de Tirado, P. I.; Severance, D. L. *J. Am. Chem. Soc.* **1994**, *116*, 2199−2200.

Table 2. Free Energy, Enthalpy, and Entropy Changes for the Transformation of the α Anomer into the β Anomer at 25 °C as Determined from Calorimetry[19]

	α/β (%/%)	ΔG (kcal/mol)	ΔH (kcal/mol)	ΔS (cal/(mol K))	$-T\Delta S$ (kcal/mol)
D-glucose	37/63	−0.33	−0.27	+0.19	−0.06
D-xylose	35/65	−0.38	−0.54	−0.52	+0.16
D-mannose	67/33	+0.42	+0.45	+0.13	−0.04

conformational states than glucose and has two fewer heavy atoms, which also makes it more tractable for quantum mechanical calculations. Calorimetric experiments[19] indicate that the enthalpic term is dominant in both glucose and xylose (Table 2). However, the primary alcohol in glucose does appear to play a role in the anomeric equilibrium, since the similar ratios for glucose and xylose result from a combination of enthalpy and entropy terms of different magnitudes, with the small entropic contributions being of opposite signs (the difficulty of such measurements is illustrated by the disagreement concerning whether the anomeric effect in 2-methoxytetrahydropyran is entropy or enthalpy dominated[11]). We report here simulations of the anomeric equilibrium in D-xylose which support our earlier findings for D-glucose and provide a new qualitative picture of the origin of the solvation term.

II. Methods

Separate molecular dynamics simulations were performed for the two anomers of D-xylopyranose in aqueous solution. These simulations employed the potential energy function used in our previous glucose simulations[20] and the TIP3P water model,[21] with chemical bonds involving hydrogen atoms kept fixed using the SHAKE constraint algorithm.[22] All simulations were done using the CHARMM molecular mechanics program.[23] Periodic boundary conditions were applied to cubic boxes 18.4987 Å on a side containing one sugar molecule and 205 water molecules at 300 K. Group-based switching functions were used to truncate long-range interactions between 6.0 and 8.0 Å.[23] Both simulations were equilibrated for 20 ps, followed by an additional 200 ps for analysis. Free energy simulations[24,14] were also conducted to transform one of the two anomers into the other. The same energy function and water model were employed in a spherical stochastic boundary Langevin system[25] of radius 14.0 Å, using an atom-based shifting function for both electrostatic and van der Waals interactions with a cutoff of 13 Å, as in the previous glucose study. The water molecules between 12 and 14 Å were governed by the Langevin equation of motion, while all atoms less than 12 Å moved according to Newtonian dynamics. Both the standard MD simulations and the free energy study were repeated for the molecules in vacuum for comparison. All simulations used energy-minimized starting structures in the lowest energy chair form. These lowest energy conformers were identified by minimizing to convergence, using conjugate gradient minimization, stereotypical $^4C_1$ starting structures for each of the 81 possible combinations of hydroxyl conformations. No alternate ring conformations (boats, etc.) appear during the simulations, and these states were thus not considered in the present analysis.

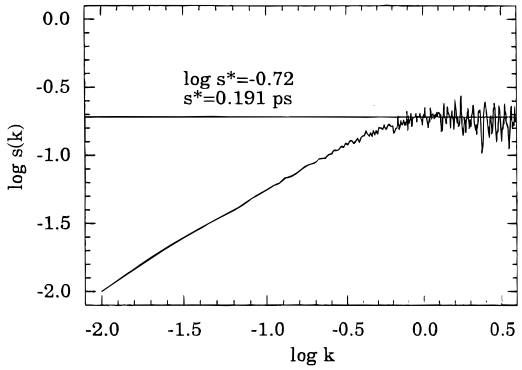

Figure 2. Statistical inefficiency error analysis of the history of the free energy calculated using the thermodynamic integration method. The statistical inefficiency s* is the limiting value of the ratio $s(k)$ of the observed variance in the average free energy as a function of k to the variance for uncorrelated Gaussian statistics, $s(k) = \{k\sigma^2[\langle A\rangle_k]\}/\{\sigma^2[A]\}$, where k is the number of successive dynamics steps in the average. Shown are the results for the $\lambda = 0.5$ simulation, in which configurations were saved every 0.01 ps (see ref 28).

The solution free energy simulations were performed using the general method described by Tidor and Karplus,[26] with one transformation beginning with the α anomer and one beginning with the β anomer. Comparison of the two simulations provides a measure of the convergence of the calculations. Both the exponential formula

$$\Delta G = -kT \ln \langle \exp[-\beta(V_1 - V_0)]\rangle_0 \quad (1)$$

and the thermodynamic integration equation

$$\Delta G = \int_0^1 \langle V_1 - V_0\rangle_\lambda \, d\lambda \quad (2)$$

were employed for a series of simulations using the mixed energy function

$$V_\lambda = (1 - \lambda)V_0 + \lambda V_1 \quad (3)$$

where V_0 is the initial state ($\lambda = 0$) and V_1 is the final state ($\lambda = 1$). A set of intermediate λ values of 0.1, 0.3, 0.5, 0.7, and 0.9 were introduced to improve the convergence of the simulation. It was found that the energy changes rapidly at the beginning and end of the integration, as has been found in other simulations.[14] Thus, additional simulations at $\lambda = 0.04$ and $\lambda = 0.96$ were run. If the angle and bond energy terms are scaled by numbers too close to the end points, the energy for these terms in one of the anomers is being decreased by more than 1 order of magnitude. Since these terms play a large part in keeping the stereochemistry of the molecule fixed, it is possible, once they are scaled to such a small value, for the stereochemistry at C1 to invert by "squeezing" the aliphatic proton through the plane of the C1 atom and two of the other substituent atoms, short-circuiting the free energy "mutation". For this reason, the bond and angle terms at the anomeric center were not scaled by the λ factor, and these terms were excluded from the final change in free energy. This exclusion is justified by the small magnitude of the terms; the contribution of bond terms to the total change in free energy is 0.012 kcal/mol (averaged over the forward and reverse simulations), while the angle terms contribute 0.14 kcal/mol. For each λ value 10 ps of equilibration and 50 ps of data collection dynamics were calculated.

An error analysis for this type of simulation is difficult because both statistical and systematic errors may be significant and only the former can be estimated.[27] Consequently, we believe it would be misleading to present the statistical errors in the tables. We have used the statistical efficiency method for exploring the statistical error.[28] An analysis was performed for the time history of the free energy during the 50 ps

(19) Kabayama, M. A.; Patterson, D.; Piche, L. *Can. J. Chem.* **1958**, *36*, 557−562. Takahashi, K.; Ono, S. *J. Biochem. (Tokyo)* **1973**, *73*, 763−770.
(20) Ha, S. N.; Giammona, A.; Field, M.; Brady, J. W. *Carbohydr. Res.* **1988**, *180*, 207−221.
(21) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926−935.
(22) van Gunsteren, W. F.; Berendsen, H. J. C. *Mol. Phys.* **1977**, *34*, 1311−1327.
(23) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187−217.
(24) Brooks, C. L.; Karplus, M.; Pettitt, B. M. *Proteins: A Theoretical Perspective of Dynamics, Structure, and Thermodynamics*; Advances in Chemical Physics, Vol. LXXI; Wiley-Interscience: New York, 1988.
(25) Brunger, A. T.; Brooks, C. L., III; Karplus, M. *Proc. Natl. Acad. Sci. U.S.A.* **1985**, *82*, 8458−8462.
(26) Tidor, B.; Karplus, M. *Biochemistry* **1991**, *30*, 3217−3228.
(27) Lau, F. T. K.; Karplus, M. *J. Mol. Biol.* **1944**, *236*, 1049−1066.
(28) Allen, M. P.; Tildesley, D. J. *Computer Simulation of Liquids*; Clarendon Press: Oxford, U.K., 1987; Section 6.4.1.

Table 3. Calculated Anomeric Free Energy Differences for D-Xylose[a]

transformation direction	exponential method		thermodynamic integration			
	ΔA	α/β(%/%)	ΔA total	α/β(%/%)	ΔA(solute−solute)	ΔA(solute−solvent)
$\alpha \rightarrow \beta$	0.3	62/38	0.45	68/32	0.90	−0.45
$\beta \rightarrow \alpha$	0.0	50/50	0.06	53/47	0.83	−0.77
mean	0.15	56/44	0.26	61/39	0.87	−0.61
$\alpha \rightarrow \beta$, in vacuum	0.05	52/48				

[a] The signs for all values are reported for the $\alpha \rightarrow \beta$ direction; values in kcal/mol.

Table 4. Energy Contributions (kcal/mol) Averaged over Conventional MD Simulations of D-Xylose for the Difference between α and β

	ΔE total	α/β (%/%)	ΔE(solute−solute)	ΔE(solute−solvent)	ΔE(solvent−solvent)
in solution	0.3	62/38	0.7	−1.2	0.7
in vacuum	0.24	60/40			

production periods. The statistical inefficiency s^* was determined to be 0.191 ps for the $\lambda = 0.5$ simulation as seen in Figure 2, indicating that an independent sample was drawn every 0.191 ps. Thus a total of 50.0 ps/0.191 ps = 262 independent sample points were obtained, so that the final result can be estimated to be accurate to within $1/(262)^{1/2}$, or 6%. Presumably, runs longer than 50 ps would be required for hexoses, where there are infrequent transitions between primary alcohol conformers.

III. Results and Discussion

The results of the free energy simulations of the "mutation" of D-xylopyranose were found to be qualitatively similar to our previous results for glucose (Table 3). As in that study, the simulations predict the wrong anomer (α) to be favored by a small amount, 0.15 kcal/mol, which is well within the uncertainty of the calculation. This very small total free energy difference again results from the approximate cancellation of an intramolecular term favoring the α form, which corresponds to an anomeric effect, and a solvent interaction term, which favors the β anomer. However, the magnitudes of both terms are significantly smaller than in the glucose study. Also, the present results exhibit less hysteresis between the forward and reverse calculations than in the glucose simulations. This probably is a consequence of longer simulation times and additional λ values at the end points of the calculation. The results of the thermodynamic integration and exponential formula are in satisfactory agreement.

That the solvation term favors the β anomer in the thermodynamic decomposition (Table 3) is in agreement with the average solute−solvent interaction energies from the separate standard MD simulations of the two anomers. Table 4 contains the results of direct averaging of the intramolecular energy, the solvent−solute interaction energy, and the solvent contribution to the difference between the α and β forms from these MD simulations. The solvent−solvent term converges slowly and therefore is less reliable than the first two. However, its magnitude is in agreement with theoretical results (refs 29 and 30 and G. Archontis and M. Karplus, unpublished calculations) showing that the solvent contribution to the enthalpy in polar molecules is approximately equal to minus one-half the solute−solvent term. Thus, as in the case of the free energy, the total solvent contribution to the enthalpy (solute−solvent plus solvent) leads to an overall stabilization of the β anomer. The calculated energy change has a magnitude similar to that of the experimental enthalpy but the opposite sign.[19] This suggests, in line with the free energy results, that the internal anomeric term from molecular mechanics is too large.

Table 5. Minimized Conformational Energy of D-Xylose as a Function of Anomeric Configuration and Conformation of the Hydroxyl Groups[a]

	α	β	$\beta - \alpha$	α/β (%/%)
counterclockwise	0.00	0.27	0.27	61/39
clockwise	2.57	5.71	3.14	99.5/0.5

[a] All values are in kcal/mol and are given relative to the lowest energy value.

In both the free energy analysis and in the direct averaging from standard MD simulations, the magnitudes of the xylose internal energy differences in solution are larger than the energy difference calculated by static minimization of the isolated monomers (Table 5). This is due to the presence of the solvent, which can alter the equilibrium among the conformations and, thus, the average internal energy of the solute. *In vacuo*, the isolated molecules can only make strained hydrogen bonds between adjacent hydroxyl groups around the edge of the ring. As has been found in several other studies, only two arrangements of these hydrogen bonds are relatively low in energy: that in which all of the hydrogen bonds point in the same direction as the increasing carbon atom numbers, labeled clockwise; and that in which all of the hydrogen bonds point in the direction opposite to the ring atom numbering, labeled counterclockwise[31] (or reverse clockwise). As in glucose, the α anomer has the lowest minimized energy in both cases, and the counterclockwise arrangement (see Figure 1) is much lower in energy for both anomers and also gives the smaller difference between anomers. In solution, however, alternate hydrogen bond partners are available for the hydroxyl groups and the strained internal hydrogen bonds are frequently exchanged for more favorable ones with water at the expense of some increase in the internal energy of the molecule. The solution simulations were found to sample all possible hydroxyl conformations through frequent transitions. This averaging leads to the larger internal energy difference between the anomers. Ensembles of vacuum trajectories resulted in an average ΔE of 0.24 kcal/mol favoring the α anomer (Table 4), in almost exact agreement with the static minimization results. It was not possible to calculate the anomeric energy difference for vacuum ensembles of trajectories with clockwise hydrogen bonding since all trajectories started in this geometry spontaneously converted to the lower energy counterclockwise structures during the simulations. Free energy simulations were also made in vacuum. The results of these simulations are included in Table

(29) Yu, H.-A.; Karplus, M.; Pettitt, B. M. *J. Am. Chem. Soc.* **1991**, *113*, 2425−2434.
(30) Roux, B.; Yu, H.-A.; Karplus, M. J. *J. Phys. Chem.* **1990**, *94*, 4683−4688.
(31) Ha., S. N.; Madsen, L. J.; Brady, J. W. *Biopolymers* **1988**, *27*, 1927−1952.

Table 6. Number of Solute−Solvent Hydrogen Bonds

hydroxyl group	α	β	difference	difference using an angular cutoff of 100°
01	2.48	2.72	0.24	0.61
02	2.58	2.58	0.00	0.01
03	2.60	2.61	0.01	0.08
04	2.69	2.63	−0.06	0.08
05	0.88	0.95	0.07	0.06
totals	11.23	11.49	0.26	0.84

3. In vacuum the free energy difference favored the α anomer, but the magnitude of the difference (0.05 kcal/mol) is very small.

From the 200 ps standard simulations, the mean number of solvent−solute hydrogen bonds was calculated for each hydroxyl group using a geometrical hydrogen bond definition, with an O−O distance cutoff of 3.4 Å and an O−H−O angle cutoff of 120°. Table 6 list the number of observed hydrogen bonds for each anomer. As can be seen, the hydrogen bonding of each hydroxyl group is nearly identical for the two anomers except at the anomeric hydroxyl group. Each hydroxyl makes about 2.6 hydrogen bonds to the solvent, roughly one as a donor and two as an acceptor, except for the anomeric hydroxyl in the α anomer. This hydroxyl group makes an average of 0.2 hydrogen bonds less than in the β anomer. This difference at the anomeric hydroxyl increases to 0.6 hydrogen bonds if the angular cutoff in the hydrogen bond definition is decreased to 100°, while it remains unchanged at the other hydroxyls; changing the distance cutoff in the definition had little effect upon the numbers in Table 6. This difference in hydrogen bonds could account for the difference in solvent−solute interaction energy between the two forms. The angular dependence of the difference indicates that the additional hydrogen bonds in the β anomer are somewhat distorted.

An explanation for the observed greater degree of hydrogen bonding to solvent for the anomeric hydroxyl group of the β anomer can be based upon simple geometric considerations. Using a conventional Richards-type probe with a radius of 1.6 Å to estimate the accessible surface area of the two anomers,[32] no significant difference in the surface areas was found. However, there is a significant difference in the hydrophilic surface area (about 4% of the total), primarily due to the anomeric hydroxyl group. This result parallels that found for glucose, where there is also no difference in total surface area between the two anomers, but where there is a difference of approximately 6% in anomeric hydroxyl surface area favoring the β form. The larger anomeric hydroxyl surface area is likely to be the origin of the increased hydrogen bonding for β. This increase in hydroxyl accessibility results from its removal in the β anomer from contact with the C3 and C5 aliphatic protons on the "bottom" of the ring (see Figure 1). It is accompanied by a small increase in the exposed hydrophobic surface area of these protons in the β anomer.

Previous studies of similar systems have found that many of the water molecules hydrogen bonded to sugar molecules are simultaneously hydrogen bonded to two adjacent hydroxyl groups.[33] Such doubly hydrogen bonded water molecules account for approximately one-third of all of the solvent−solute hydrogen bonds. Because of the geometric requirements of hydrogen bonds and simple packing considerations, the distribution of water molecules around sugar solutes is expected to be highly anisotropic, as in crystallographic surveys of peptides,[34]

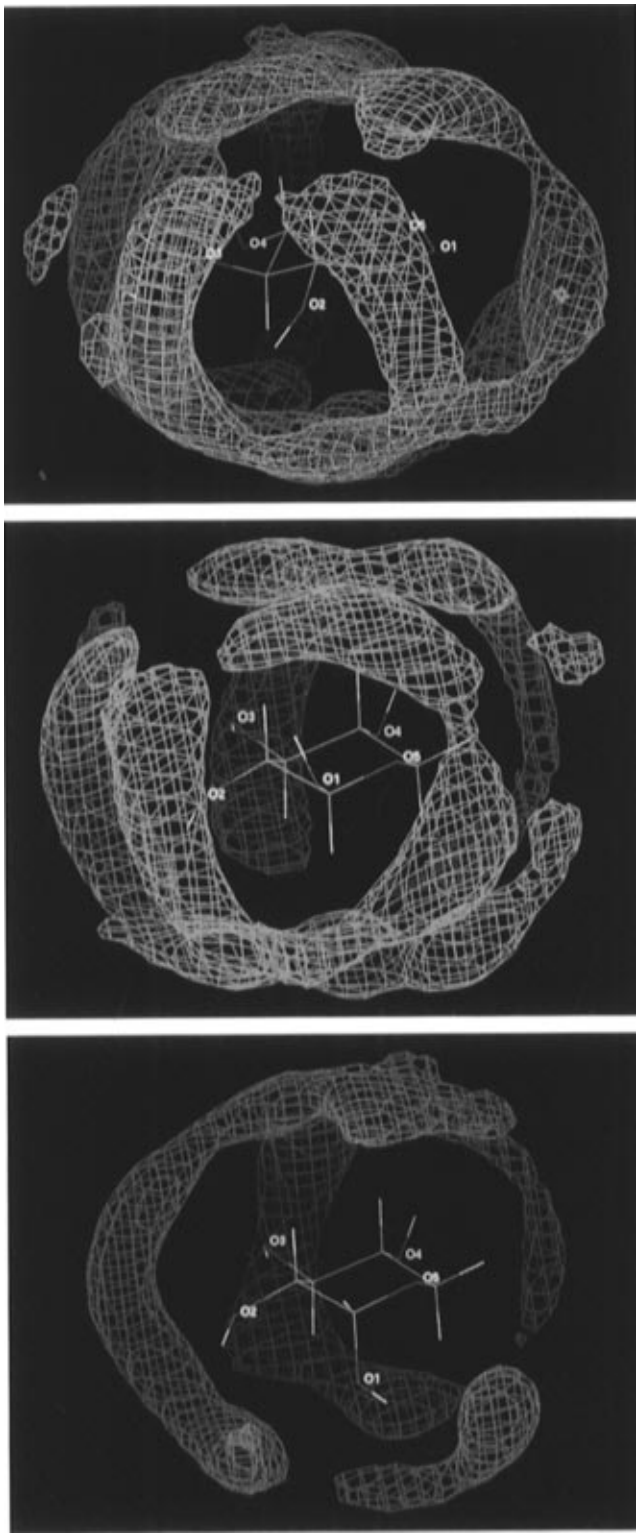

Figure 3. Contours of the probability density for the oxygen atoms of the water molecules around D-xylose as calculated from MD simulations and displayed using the program CHAIN.[36] The mesh surface represents the regions with 50% greater than bulk density. (top) Superposition of the density contours for both anomers. The molecular structure for only the β anomer is shown, positioned in the conventional orientation for pyranoses.[2] The blue contours are for the β anomer, and the red contours are for the α anomer. (bottom and center) Display of these distributions individually, rotated slightly from the conventional orientation to better illustrate the differences at the anomeric center. The reduced density around the anomeric hydroxyl for the α anomer (bottom) is clearly visible.

and could vary with anomeric configuration. To investigate this possibility, we computed spatially resolved solvent

(32) Lee, B.; Richards, F. M. *J. Mol. Biol.* **1971**, *55*, 379−400.
(33) Brady, J. W.; Schmidt, R. K. *J. Phys. Chem.* **1993**, *97*, 958−966.
(34) Thanki, N.; Thornton, J. M.; Goodfellow, J. M. *J. Mol. Biol.* **1988**, *202*, 637−657.

densities.[35] First, rotations and translations were removed from the dynamics trajectory by using the relatively rigid ring to establish a solute-fixed coordinate frame. Solvent densities relative to the sugar structure were averaged over the entire course of the simulation and were then contoured using the program CHAIN[36] in a manner analogous to the contouring of electron density in crystallographic diffraction studies. Figure 3 displays the results for the two anomers of xylose in different colors with the selected contour enclosing the areas which exceed the average density by 50%. The long, banana-shaped regions between adjacent hydroxyl groups have a roughly constant distance from the center of the pyranose ring and a constant azimuthal angle around the edge of the ring. These primarily result from water molecules which are simultaneously hydrogen bonded to two adjacent hydroxyl groups. The clouds also have the character of a bent dumbbell, with greater density at the top and bottom of the lobes. As can be seen from the overlap of these lobes for the two sugar anomers, both forms hydrate very similarly at most hydroxyl groups. The increased hydrogen bonding at the anomeric hydroxyl in the β anomer can be readily seen in the figure, with significantly less density around the α group at this contour level. The difference in these two densities is consistent not only with a greater amount of hydrogen bonding for the β anomer but also with the experimental observation of a lower entropy for the β form, since the distribution implies greater water structuring for this anomer.

IV. Conclusions

The results of the present simulations of D-xylose in solution are qualitatively consistent with our previous simulations of D-glucose[14] and support the model that the observed anomeric equilibrium results from a competition of internal and solvation terms of opposite signs. The 200 ps standard trajectories for the two anomers provide a structural explanation of the preference for the β form in solution. There is a difference in the solvent structure around the anomeric hydroxyl group with increased hydrogen bonding for the β anomer. The geometric analysis is in accord with the greater exposed hydrophilic surface area for the β anomer. The thermodynamic integration method has been used to determine the contribution of different components to the free energy. Of particular interest is the result that in solution there are significant intrasolute and solute−solvent terms of opposite sign, with the former stabilizing the α anomer and the latter the β anomer. Since the vacuum simulation difference between α and β is essentially zero, the "large" intrasolute term calculated in solution is due to a "solvent effect", as is the explicit solvent contribution. The existence of this solvation shift for the internal energy indicates that care must be used in interpreting sugar properties using vacuum calculations.

Contrary to the comments of van Eijck et al.[16] and others,[37,38] the separation of the contributions to the free energy into components used for the present analysis is a valid procedure,[39] although it can depend on the path used in the thermodynamic integration. It has been pointed out recently[40] that the alchemical path, in which one molecule is transmuted into another with a single λ for all interactions (eqs 2 and 3), yields a well-defined decomposition of the contributions to the free energy change. By expanding the exponential in the high-temperature approximation, it was shown that this integration path leads to a systematic distribution of the coupling terms between the free energy components. This is analogous to a Mulliken population analysis of the electronic charge distribution. In the present case, it permits the separation into internal terms, in the presence of the solvent, and solvation terms. Many of the arguments concerning the anomeric effect have involved attempts to use experiments to estimate effects of this type. The strong correlation of the component analysis with the internal energy averages from standard simulations and the correlation with the qualitative mechanistic explanation in terms of solvent structuring indicate that the chosen integration path is a useful one. Further, since the dominant contribution is enthalpic, the path dependence is expected to be small. The approach used for bond and angle terms is also consistent with many previous free energy calculations, which have neglected internal contributions altogether[41] or have used fixed internal geometries.[42]

Like the previous glucose studies,[14,15] the present simulations predict the wrong anomer to be preferred in solution. The carbohydrate force field used is known to have several limitations and is under revision. However, the central finding of a hydration term favoring the β anomer is unlikely to be strongly dependent upon the particular force field since it arises primarily from solvent structuring determined by geometric features of the solute. One question concerns the effects of excluding bond and angle contributions. For the present calculations, the effects, apart from improving statistical convergence, were probably small, since the carbohydrate force field was not developed to be able to produce the small structural changes with configuration observed for the anomeric bond length and angle. However, *ab initio* studies suggest that these effects could also contribute to the anomeric equilibrium, so they will need to be considered in the next generation of revised carbohydrate molecular mechanics parameters. The recent *ab initio*, experimental and Monte Carlo results for 2-methoxytetrahydropyran[11,18] may offer some guidance in this regard.

Acknowledgment. The authors thank B. Teo, K. Naidoo, P. A. Karplus, and H.-A. Yu for helpful discussions. This work was supported by NSF Grant CHE-9307690. R.K.S. would like to thank the NSF for a graduate student fellowship.

JA951066A

(35) Teo, B.; Schmidt, R. K.; Karplus, P. A.; Brady, J. W. Manuscript in preparation.
(36) Sack, J. S. *J. Mol. Graphics* **1988**, *6*, 224−225.
(37) van Gunsteren, W. F.; Beutler, T. C.; Fraternali, F.; King, P. M.; Mark, A. E.; Smith, P. E. In *Computer Simulation of Biomolecular Systems: Theoretical and Experimental Applications*; van Gunsteren, W. F., Weiner, P. K., Wilkinson, A. J., Eds.; Escom: Leiden, The Netherlands, 1993; Vol. 2.
(38) Smith, P. E.; van Gunsteren, W. F. *J. Phys. Chem.* **1994**, *98*, 13735−13740.
(39) Boresch, S.; Archontis, G.; Karplus, M. *Proteins* **1994**, *20*, 25−33.
(40) Boresch, S.; Karplus, M. *J. Mol. Biol.*, in press.
(41) Rao, B. G.; Singh, U. C. *J. Am. Chem. Soc.* **1991**, *113*, 4381−4389.
(42) Jorgensen, W. L.; Ravimohan, C. J. *Chem. Phys.* **1985**, *83*, 3050−3054.