The Extent of Conformational Rigidity Determines Hydration in Nonaromatic Hexacyclic Systems

ARTICLE *in* THE JOURNAL OF PHYSICAL CHEMISTRY B · FEBRUARY 2011

Impact Factor: 3.3 · DOI: 10.1021/jp110248j · Source: PubMed

THE JOURNAL OF PHYSICAL CHEMISTRY B

ARTICLE

pubs.acs.org/JPCB

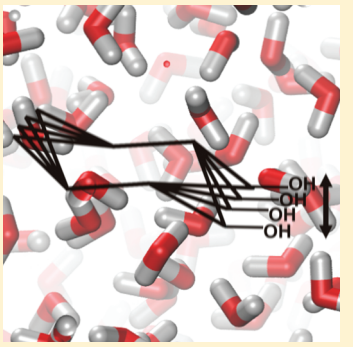
The Extent of Conformational Rigidity Determines Hydration in Nonaromatic Hexacyclic Systems

Annalisa Boscaino and Kevin J. Naidoo*

Scientific Computing Research Unit and Department of Chemistry, University of Cape Town, Rondebosch 7701, South Africa

Supporting Information

ABSTRACT: We conducted an ultrasonic study of the hydration number for hexacyclic systems. We find from these experiments that cyclohexane-based molecules such as cyclohexanol and myo-inositol show a very small increase in hydration number despite the large difference in the number of hydroxyl groups present in each of the molecules. There is however a dramatic increase in hydration number when shifting from molecules with a cyclohexane frame to molecules with a cyclopyranose frame particularly glucose. An analysis of classical and quantum molecular dynamics simulation trajectories reveal that the hydration number is strongly linked to the conformational flexibility within the molecule. Cyclopyranose is a more rigid ring system compared with cyclohexane and so its ring fluctuates in a smaller range and frequency. The effect of the ring rigidity is that the hydroxyls tethered to the cyclopyranose ring undergo less positional diffusion compared with those attached to the cyclohexane ring. This allows for long time intermolecular hydrogen bonds between the hydroxyls bonded to cyclopyranose rings and the surrounding waters, which leads to an increase in the hydration numbers of carbohydrates compared with those of hydroxylated cyclohexanes.



INTRODUCTION

The relative solubilities of simple atomic and molecular systems make it tempting to ascribe the underlying reasons for observed differences to primarily electrostatic interactions between solutes and solvents. Solubility equilibrium is reached when the solvation and preciptiation processes proceed at a constant rate. At the molecular length scale, this is reduced to an equilibrium that is achieved between the interactions of solvent molecules with that of the interactions between solvent and solute molecules (solvation). Analysis of the energetics of solvation effects in classical condensed phase simulations are therefore often described in terms of strong electrostatic (Coulomb) and weaker dispersion (van der Waal terms) interactions.[1] Although not simply electrostatic in nature, the hydrogen bond has been singled out as a key reason for carbohydrate hydration[2−4] and its subsequent behavior in water.[5,6] An early cautionary note made from a lengthy analysis of the character of organic compounds warned that only the presence of an hydroxyl group is insufficient to improve the favorability of solute water interactions.[7]

The most direct measure of the solute water interaction is through the measure of molecular hydration numbers. Numerous experimental techniques such as NMR[8] neutron scattering,[9] X-ray, and ultrasound interferometery[10,11] have been used to measure hydration numbers. These different experimental techniques lead to very different hydration number magnitudes while preserving the trend between a molecular series.[11] This is primarily because each method infers hydration by measuring physical properties unique to the experiment and on time scales restricted to the detection limits of the method. The definitions of hydration number range from the average number of waters that are rotationally slowed down by the solute[12] to the number of water molecules that are translationally bound to the solute.[13] Computer simulations are well suited to extract molecular events on the 10^{-10} to 10^{-6} s time scales. It is therefore very informative to combine experimental hydration trends with those calculated from computer simulations.[14]

Studies of linear[15,16] and aliphatic alcohols[17] concluded that the extent of hydration is dependent on the number of hydroxyl groups present in a similar set of molecules. For example the dynamic hydration numbers (n_{dhn}) measured by the ^{17}O spin relaxation times, ^{1}T, for ethanol and ethylene glycol are 10.5 and 5.7 respectively.[15] Ishihara et al. argued that the hydrophobic hydration for linear alcohols is disrupted by hydrophilic groups and so an increase in the number of hydroxyls for diols compared with their corresponding *n*-alcohols leads to a relative decrease in n_{dhn}. It is not only hydrophyllic groups that alter the diffusional behavior of the surrounding water molecules, hydrophobic groups have been shown by ^{2}H NMR relaxation[18] and femtosecond mid-infrared spectroscopy[19] to slow down the orientational mobility of water molecules in the surrounding hydration shell. Therefore, when attempting to associate the solute's molecular features responsible for the degree to which it is hydrated, both the hydrophyllic and hydrophobic nature of the solute must be taken into account.

In an experiment measuring the partial molar volumes of polyols and comparing them with those measured from their corresponding sugars, it was seen that the linear alcohols were more hydrated than the cyclic carbohydrates (sugars).[20] Carbohydrates

Received: October 26, 2010
Revised: January 12, 2011
Published: February 28, 2011

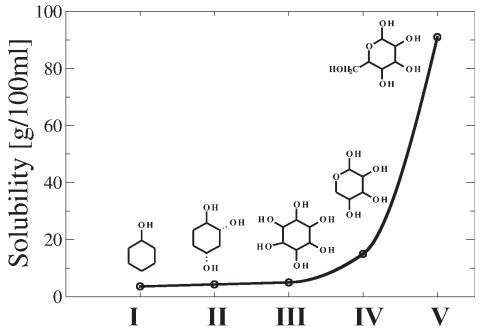


Figure 1. Comparison of the solubility of I cyclohexanol, II 1,2-cyclohexanediol and 1,4-cyclohexanediol, III myo-inositol, IV β-D-xylopyranose, and V β-D-glucopyranose.

are biologically important examples of hydroxylated cyclic systems. While the conformational flexibility of the linear alchohols is greater in range compared with the cyclic sugars, the rate of conformational motion of the linear alcohols is far slower than that of the cyclic sugars. In particular, the conformational motion in the case of linear alcohols is on the same time scale as that of the overall diffusion of the solute molecule while for cyclic carbohydrates the conformational motion is much faster compared the overall diffusion. These differences in hydration, between molecules of similar hydroxyl number and distribution, must be due to the variation in the rate of conformational motion exhibited by linear compared with cyclic systems. A classic study, by Shiio, of carbohydrate water interactions using ultrasonic interferometery concluded that the dominant effect in distinguishing the hydration behaviors of carbohydrates from each other were the number of hydroxyls and their relative positions to each other.[21] In a later study using a similar experimental technique, Galema and Høiland, proposed that carbohydrates with smaller hydration numbers have a stereochemistry more suited to fit into the three-dimensional hydrogen bonded structure of water.[10]

While there has been a great focus on the immediate effect of polarity on hydration properties, the conformational flexibility that underlies the internal molecular entropy has not been systematically investigated for carbohydrates. To uncover the nature of carbohydrate pyranose hydration, we consider hexacyclic systems where the number of hydroxyls is progressively varied on a homogeneous ring skeleton (cyclohexane) followed by the introduction of an oxygen atom into the ring to make a pyranose frame. The cyclohexane based molecules that were studied are cyclohexanol, 1,2-cyclohexanediol, 1,4-cyclohexanediol, and myo-inositol (in order of increasing number of hydroxyl groups) while β-D-xylopyranose and β-D-glucopyranose are the pyranose based molecules that were studied.

The first clue that the hydration of hexa cyclic systems cannot simply be described in terms of the number of hydroxyl groups can be seen from their solubilities (Figure 1). There is a minor difference in solubility between cyclohexanol, 1,2-cyclohexanediol, 1,4-cyclohexanediol, and 1,2,3,5/4,6-cyclohexanehexol (myo-inositol) where the only change is the increase in the number of hydroxyl groups. The big change is observed with the introduction of an ether linkage into the ring. The cycohexane based molecules (cyclohexanol through to myo-inositol) are far less soluble than the pyranose based rings (xylopyranose and glucopyranose).

In this paper, we compare the hydration properties of these six flexible, nonaromatic, cyclic molecules that are based on the hexane and pyranose rings. We explain the relationship between confomational properties and hydration for each molecule. We do this by measuring hydration numbers from ultrasound interferometery experiments and compare these with values calculated from MD simulations. To make sense of the hydration trend in relation to the molecular thermodynamics and relative conformational freedom we probe (1) the free energy, (2) the rotational distribution of the hydroxyls, (3) the conformational probability distribution of ring pucker, and (4) the librational rate of ring pucker for each molecule. These measures and computations show that molecular flexibility is a fundamental property from which the relative hydration properties of hexacyclic systems can be extrapolated.

MEASURING AND ANALYZING HYDRATION

Ultrasound interferometer experiments are used to measure dynamic hydration number. This is the number of solvent molecules interacting with the solute for times long enough so that they diffuse together through the solvent. The hydration number (n_h) is calculated from

$$n_h = \frac{n_w}{n_s\left(\frac{1-k_s}{k_{s,0}}\right)} \tag{1}$$

which is a relation between the number of moles of water (n_w), number of moles of the solute (n_s), and the isentropic coefficients of compressibility of water ($k_{s,0}$) and the solution (k_s). This relationship was first derived by Pasynski for electrolytes.[22] By passing repeated pulses of ultrasonic radiation through the solution followed by measuring the frequency of this radiation, the coefficients can be computed from the Laplace equation

$$k = \frac{1}{\rho u^2} \tag{2}$$

where ρ is the density of the aqueous solution and u is the ultrasonic speed. The assumption is that the compressibility of the hydration layer about the solute is zero. We present here hydration numbers that are for the solutes at infinite dilutions.

Water molecules located in popular configurations about the solute are radially and spatially identified from radial distribution (rdf) and spatial distribution functions (sdf), respectively. However, Impey et al. recognized that hydration is dynamic and so sought to calculate the persistent coordination of water molecules about ions.[11] To calculate a hydration number from a molecular dynamics (MD) computer simulation that is comparable to that measured from the ultrasound experiment, we expand on the expressions developed for ions.

First, we calculate the number of waters in the first hydration shell (n_{solute}) that can be found from the rdf

$$g_{AB}(r) = \frac{\rho_B(r|r_A = 0)}{\rho_B} \tag{3}$$

where the first index A refers to a solute atom at the origin, and the second index B refers to a solvent atom B, r is the distance between the atoms, and ρ_B is the bulk density of B in the solvent. Two types of rdfs, $g_{CO}(r)$ and $g_{OO}(r)$, were calculated for the hexacyclic molecules in water. Coordination numbers were obtained by integrating $g_{AB}(r)$ out to the first minimum. The $g(r)$ radial distributions show for most of the rdfs in the hexacyclic

molecular series that the first maximum appears at around 3.5 Å for C−O_w and at 2.8 Å for O−O_w interactions.[23]

Next we calculate the residence time, following Impey's work, using a Kronecker delta function $P_j(t,t_n; t^*)$. A value 1 is assigned if the solvent molecule j lies within the coordination shell of an atom on the solute for a time interval Δt without leaving the coordination shell for any period longer than t^*. A value of 0 is assigned if this condition is not met. Impey fixed the value of t^* for ions at 2 ps. Here, a comparison between different t^* was made and eventually the results with $t^* = 2$ ps were seen to better reproduce experimental results.

The probability $p(t)$ of a solvent molecule surviving in the coordination shell for time t is then

$$p(t) = \frac{\langle P_j(t, t_i; t^*)\rangle_{i,j}}{\langle P_j(0, t_i; t^*)\rangle_{i,j}} \quad (4)$$

This probability decays exponentially with time and can be fitted by $p(t) \sim \exp(-t/\tau)$, where τ is the residence time that is inserted into

$$n_h = n_{solute} \exp^{(-\tau_{bulk}/\tau_{solute})} \quad (5)$$

where n_{solute} is the number water molecules coordinated to the solute atom A in its first hydration shell calculated from the rdf relation (eq 3). The hydration number (n_h) from the MD simulation is then calculated for each of the hexacyclic molecules.

EXPERIMENTAL METHODS

All chemicals used were purchased from Sigma-Aldrich Inc. (St. Louis, MO). All the substances with purities higher than 99% were used without further purification. The water was doubly distilled. Solutions were prepared by weighting before measurements with a Afcoset-ER-120A with a precision of ±0.01 mg. The compositions of solutions are given in molarity (mol/L) throughout this paper. The ultrasonic speeds of all the solutions were measured using an Ultrasonic Interferometer Model M-81G set to an operating frequency of 3 MHz. Density measurements were carried out at 27 °C through an Anton Paar digital Model DMA 35 with a precision of 0.01 g cm^{-3}. The temperature of the samples during the measurements was maintained to an accuracy of ±0.1 °C in an electronically controlled thermostatic Huber water bath with compatible control CC1.

While the rdfs are able to give information in one-dimension, that is, the number of neighbors about each site, the sdfs provide a three-dimensional location of water probability density about a solute. When generating a probability density map of the water structuring about a solute, all the rotational and translational diffusion throughout the simulation have to be removed. To achieve this, the instantaneous positions and orientations of the solute in each coordinate set, which make up the trajectory, are translated and rotated to obtain a best least-squares overlap with a reference frame. The coordinate transformations performed during this reorientation procedure are applied to all the atoms in the system. If two or more trajectories are to be compared, then the original reference frame is used that the resulting density maps will have the same relative orientation. Care must be taken when choosing the atoms for the least-squares fit; by using atoms that move around too much, an unsatisfactory overlay of solute molecules and a consequent possible blurring of the water densities will be obtained. Therefore, choosing only heavy atoms is the best option. Here, we calculated the distribution of water around the solute, using only the oxygen atoms of the water molecules and these are shown in blue. We have described the methods used to compute sdfs elsewhere.[6]

COMPUTATIONAL METHODS

We simulated infinite dilute conditions using a periodic cubic box of 24.8626 Å in length, containing 488 water molecules and a single solute molecule. Recently published CHARMM force field parameters were used for myo-inositol,[24] while the CSFF force field was used for all the other systems.[25] A study by Corzana et al. investigated the comparative solution behavior of several carbohhdrate force fields primarily in the context of glycosidic linkages.[9] In that study, the CSFF parameters was shown to produce reasonable hydration numbers compared with differential scanning calorimetry experiments for the methyl-α-D-maltoside molecule. The TIP3P model as implemented in CHARMM[26] was used to model the water molecules since this water model was used in the parametrization of both force fields. All simulations were performed as isothermal−isobaric ensembles (NPT) using the CHARMM program,[26,27] with chemical bonds to hydrogen atoms kept fixed using SHAKE.[28] The systems were then heated to 300 K over 2 ns and equilibrated for 2 ns. Each solute molecule was then analyzed over separate 5 ns trajectories.

The Self-Consistent Charge Density Functional Theory Tight Binding (SCC-DFTB). SCC-DFTB has recently become a popular method for QM/MM molecular simulations[29] including carbohydrates.[30] We ran 5 ns of QM/MM SCC-DFTB dynamics for cyclohexanol, 1,2-cyclohexanediol, 1,4-cyclohexanediol, myo-inositol, β-D-xylose and β-D-glucose using CHARMM[26] v33b2. These were run with the mio-0-1[31] parameters and an improvement to hydrogen-bonding interaction was included with the HBON keyword. QM/MM systems were run under the same conditions (24.8626 Å box length, 488 water molecules with one independent solute, heated to 300 K over 2 ns and equilibrated for 2 ns). However, in these simulations Ewald summations were used for the electrostatic interactions[32] along with the $TIP4P_{EW}$ water model.[33] We use the SCC-DFTB QM/MM method to generate trajectories, as this is the most accurate and computationally cost efficient simulation procedure available for the analyses of carbohydrate conformational motion specifically their ring puckering.[34,35]

Free Energy Perturbation. We used the Free Energy Perturbation (FEP) method to calculate the relative free energy difference between cyclohexane and each of the six cyclic systems. A nonphysical path was chosen ensuring that the difference between successive perturbed states remained less than 3 kcal/mol (values varied in the range of 1−2.5 kcal/mol).

Since the free energy depends only on the thermodynamic state, the results do not depend on the chosen path and the number of λ points however, the error can be minimized by increasing the overlap between the cyclohexane structure (A) and the structure B, that is, by choosing many λ points. The free energy difference (ΔG) between states A and B is given by[36]

$$\Delta G = G(\lambda_B) - G(\lambda_A) = -k_B T \ln \frac{Z(\lambda_B)}{Z(\lambda_A)} \quad (6)$$

with partition function (Z) for NVT ensemble that is used in our simulations. This expression can be rewritten in the form of

$$\Delta G = -k_B T \ln\langle e^{-[H(\lambda) - H(\lambda_A)]/k_B T}\rangle_{\lambda_A} \quad (7)$$

Here ⟨⟩ represents the ensemble average taken over the configurations representative of state A (cyclohexane) whereas H, k_B, and T represent the Hamiltonian of the system, the Boltzmann constant, and the temperature, respectively. The free energy difference between two states can be given as the sum of the free energy differences between the intermediate steps

$$\Delta G = -k_B T \ln \sum_{i=0}^{n-1} \langle e^{-[H(\lambda_{i+1}) - H(\lambda_i)]/k_B T} \rangle_{\lambda_i} \quad (8)$$

where, $H(\lambda_{i+1})$ and $H(\lambda_i)$ are the Hamiltonians of the system at $\lambda = i + 1$ and $\lambda = i$, respectively. The relative change in the binding free energy between two related (similar) compounds can be computed by using the FEP method in the thermodynamic cycle approach. All the systems for which FEPs have been performed start from the same structure, cyclohexane, thus referred to as A(λ=0), while cyclohexanol, 1,2-cyclohexanediol, 1,4-cyclohexanediol, myo-inositol, xylopyranose, and glucopyranose are referred as B(λ=1). The simulations where performed at λ = 0, 0.1, 0.2, 0.3, 0.4,...1, and the consistency of the results were confirmed by repeating the simulations in both directions $\lambda_i \rightarrow \lambda_{i+1}$ and $\lambda_{i+1} \rightarrow \lambda_i$ except at the two end points, so performing a so-called double-wide sampling.[37]

Ring Pucker. The Cremer−Pople method, used to define puckering of a six-membered ring, uses coordinates that are defined in terms of the distances z_j that are functions of 3 parameters (q_2, ϕ_2, and q_3).[38] Converting these to a spherical coordinate frame (Q, θ, ϕ)

$$\begin{aligned} q_2 \cos\phi_2 &= Q \sin\theta \cos\phi \\ q_2 \sin\phi_2 &= Q \sin\theta \sin\phi \\ q_3 &= Q \cos\theta \end{aligned} \quad (9)$$

is practically more useful since the canonical ring puckers can be located at specific points in a spherical volume. Here ϕ represents the azimuthal angle and lies in the range [0, 360°], θ is the puckering amplitude of range [0, 180°], and Q is the radial coordinate ($Q^2 \equiv \sum_{j=1}^{N} z_j^2 = \sum_m q_m^2$; $Q > 0$) representing the total puckering amplitude. Combinations of θ and ϕ describe canonical conformations that a hexacyclic molecule can adopt, namely the two chairs, $^4C_1$ and $^1C_4$, the six boats (B) and six twists/skew shapes (S), twelve half-chairs (H), and twelve envelopes (E).

Hydroxyl Rotational and Ring Pucker Librational rates. The time-dependent correlations in spontaneous fluctuations of various conformational properties provide a measure of the flexibility of the molecular system. The conformations of importance here are the hydroxyl rotations and the ring pucker functions. The correlations for the conformational property (A) where $\delta A(t)$ is the instantaneous or spontaneous fluctuation at time zero for the time correlation function (tcf) such that

$$c(t) = \langle \delta A(t) \delta A(0) \rangle = \langle A(t) A(0) \rangle - \langle A \rangle^2 \quad (10)$$

which can be normalized

$$C(t) = \frac{\langle \delta A(t) \delta A(0) \rangle}{\langle \delta A(0)^2 \rangle} \quad (11)$$

The correlation between two conformations, A and B, is found from the cross-correlation function

$$C(t) = \frac{\langle \delta A(t) \delta B(0) \rangle}{\langle \delta A(0) \delta B(0) \rangle} \quad (12)$$

Table 1. Number of Coordination Waters (n_{solute}) around Each Solute, Obtained by Applying the RDF to Each Heavy Atom (C and O) of the Molecules

compounds	n_{solute} [g(r)$_{min}$]	ratios
cyclohexanol	33	1
1,2-diol	38	1.2
1,4-diol	37	1.1
myo-inositol	67	2
xylopyranose	70	2.1
glucopyranose	89	2.7

Assuming that the decay of the tcf is exponential, it will fit a general form $C \exp(-t/\tau)$, where the time τ is the correlation time. Therefore the decay of the tcf gives an indication of the rate at which the fluctuations in conformational motion becomes uncorrelated, that is, the rate at which the conformational property of the molecule relaxes to an equilibrium.

DISCUSSION AND RESULTS

Interatomic rdfs between non-hydrogen atoms (carbon and oxygen) in the solute molecules and the water oxygen atoms provide a radially averaged number of water molecules that are located about these solute atoms. The number of waters coordinated to the entire solute was calculated by summing the contributions from the carbon−water oxygen rdfs (Table 1 shows results calculated from the classical MD trajectory). This includes waters structured within a probability shell about the hydrophobic and hydrophyllic sites. These coordinated waters derived from the static structural distribution functions are large in number and represent the average hydration shell structure. They are, however, not dynamically associated with the solute or its directly calculable hydration properties. Another measure of the general static hydration structure about a solute is to use the sdfs calculated relative to an average ring conformation. This gives an anisotropic description of the most probable water positions about the solute. In Figure 2, the spatial distribution functions (sdfs) for the six molecules are shown calculated from the classical MD trajectory. In the simplest case where there is only one hydroxyl (cyclohexanol in Figure 2 a), the most probable water positions are located in the three (t, g^+, and g^-) sterically favored hydroxyl positions. When adding a second hydroxyl group on the opposite end of the cyclohexane ring, the same three high probility positions are seen about each of the two hydroxyls in the 1,4-cyclohexanediol sdf (Figure 2 b). However, when placing the two hydroxyls adjacent to each other as in 1,2-cyclohexanediol the high probability waters do not show up clearly in the sterically favored hydroxl positions, instead a new half moon-shaped probability density appears between the two hydroxyls (Figure 2c). Previously, we had shown that the intermolecular hydrogen bonds that form between functional groups on pyranose rings and the surrounding water are very close in energy to the intramolecular hydrogen bonds formed between these functional groups.[3] The intramolecular hydrogen bond strength between two hydroxyl groups is marginally stronger (0.5 kcal/mol) than the intermolecular hydroxyl-water hydrogen bond strength, leading to the two types of hydrogen bonds competing with each other.[4] When there are OH-groups on every ring atom, as is the case with myo-inositol, where all hydroxyls are in equatorial positions except the one on carbon 2 where the OH is axial, most of the high probability density

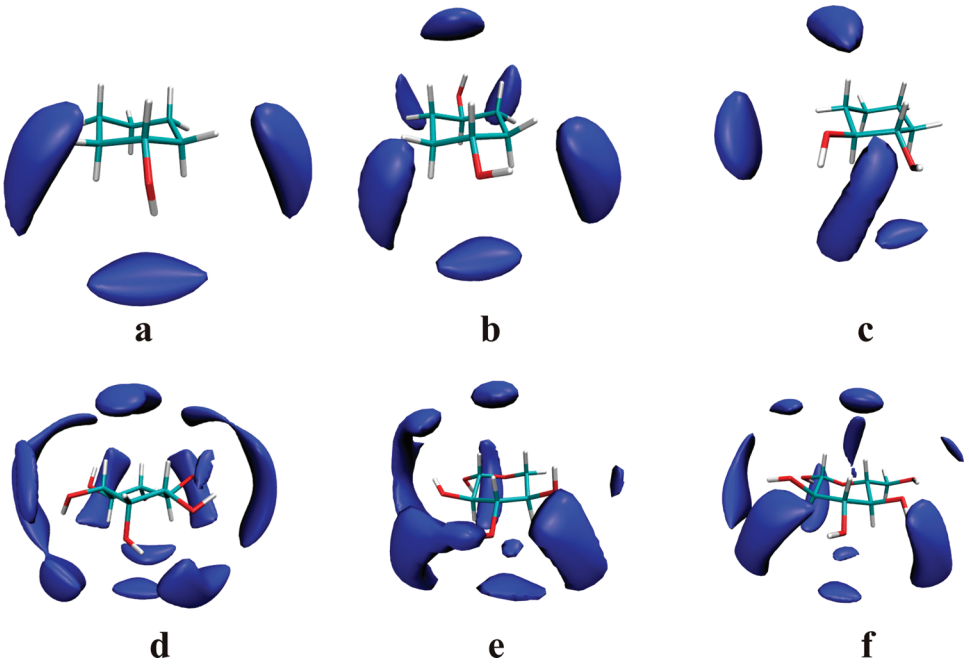


Figure 2. SDFs of the six aqueous systems contoured at 50% above bulk density: (a) cyclohexanol; (b) 1,4-cyclohexanediol; (c) 1,2-cyclohexanediol; (d) myo-inositol; (e) β-xylopyranose; (f) β-glucopyranose.

volumes are distributed in a half moon shape (Figure 2d). While the increases in the water probability density volumes for these four molecules correspond to the increase in their hydroxyls, there is no correlation between the increase in their water probability densities (Figure 2a−d) and their relative solubility (Figure 1).

Further evidence supporting this observation is bourne out when we analyze the probability densities of xylopyranose (Figure 2 e) and glucopyranose (Figure 2 f). Our results presented here provide the same observations of the anisotropic water structure about these carbohydrate molecules as have been previously reported by Brady et al.[39] Here our attention is drawn to the similarity in the overall water probability denisities of myo-inositol, xylopyranose, and glucopyranose while their solubilities are dramatically different. This leads us to conclude that insight into the hydration of these cyclic molecules must be found in the dynamic properties of the solute and solvent not in the static distributions such as rdfs and sdfs.

First, we measure from ultrasound experiments hydration numbers for the molecular series. For each aqueous system, seven different concentrations (except in case of cyclohexanol, due to its low solubility) in the range 0.1−0.7 M [mol/L] were prepared and 150 measurements were performed on each concentration. The hydration numbers at infinite dilution have been obtained by extrapolating the linear-fit of the $n_h(M)$ data down to infinite dilution. Plots of the experimental results for k_s and n_h versus M of the aqueous solutions of hydroxy-cyclic compounds at 300 K are reported in Figure 3a,b, respectively. Even though the ultrasound measurements took place over a very wide range of molarity, we did observe a monotonically decreasing linear change of compressibility with composition (Figure 3a), which is in agreement with previous reports.[10]

At face value, a straightforward relationship exists between the hydration properties of the entire series (Figure 3b), that is, an increase in number of OH groups delivers an increase in the n_h. However, the increase is not linearly proportional to the increase in the hydration property. When we normalize the data relative to n_h for a single hydroxyl that is, cyclohexanol (Table 2), we observe that the n_h for myo-inositol, which has 6 hydroxyls, is only twice that of cyclohexanol, which has one. This relatively small increase in n_h for a system having all its carbons hydroxylated illustrates that simply adding multiple hydrophyllic groups capable of forming hydrogen bonds with water does not result in a linear increase in hydration. The n_h for xylopyranose increases 3.4 times that of cyclohexanol, which is greater than the increase observed for myo-inositol even though it possesses only four hydroxyls. The greater change is therefore due to the introduction of an ether group into the ring. A further dramatic increase in hydration number continues for glucose where the structural change is the addition of a hydroxymethyl group.

Computer simulations 5 ns in length using a classical potential and parameters from CSFF[25] as well as hybrid quantum classical (QM/MM) simulations using the SCC-DFTB implementation in CHARMM[40] were performed for each molecule in the series. The n_h values were calculated from these simulations and then normalized relative to that of cyclohexanol (Table 2). The hydration number ratios from the classical MD agree with experimental results within 1-σ (cyclohexanol, 1,2-diol, myo-inositol) and 2-σ (1,4-diol, xylopyranose, glucopyranose). The QM/MM results compare with the experimental values within 2-σ (cyclohexanol, 1,2-diol, and 1,4-diol) and 3-σ (myo-inositol, xylopyranose, and glucopyranose). While magnitude of the experimental and simulation values differ, the trends are similar. They confirm that the n_h of a molecule has a minor dependence on the number of hydrophilic sites while there is a significant correlation with the introduction of a heteroatom to the ring. From this we conclude that a valid analysis of the underlying reasons for this trend can be deduced from the simulation trajectories.

The hydration free energies for each molecule in the series were calculated (using classical force fields) relative to cyclohexane

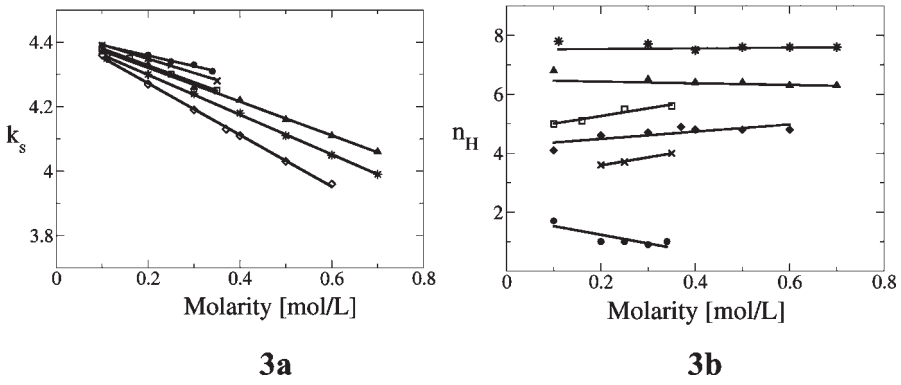


Figure 3. (a) Compressibility k_s [m^2/N] vs molarity [mol/L] for (•) cyclohexanol, (□) 1,2-cyclohexanediol, (×) 1,4-cyclohexanediol, (▲) xylopiranose, (∗) glucose, (◇) myo-inositol. b) Hydration numbers of (•) cyclohexanol, (□) 1,2-cyclohexanediol, (×) 1,4-cyclohexanediol, (▲) xylopiranose, (∗) glucose, (◇) myo-inositol.

Table 2. Hydration Numbers from Ultrasound Measurements and Computer Simulations[a]

	raw	ratios		
compounds	ultrasound n_h	ultrasound n_h	classical n_h	QM/MM n_h
cyclohexanol	1.9 ± 1.6	1	1	1
1,2-diol [b]	2.7 ± 1.3	1.4 ± 2.0	1.9	0.89
1,4-diol	3.8 ± 1.5	2 ± 2.1	1.7	0.98
myo-inositol	3.7 ± 0.6	1.9 ± 1.6	3.4	1.1
xylopyranose	6.4 ± 1.0	3.4 ± 1.8	4.8	1.4
glucopyranose [c]	7.2 ± 0.7	3.9 ± 1.7	6.3	2.6

[a] Errors on each measurement have been calculated from the propagation of errors (rms) checking for a χ^2 distribution. [b] 1,2-Cyclohexanediol may undergo polymerization in water affecting the ultrasound measurement. [c] The β-glucose hydration number agrees with studies in ref 10.

Table 3. Free Energies for Systems Are Perturbed from Cyclohexane (Cyclo) Calculated from Classical MD Simulations

	ΔΔG [kcal/mol]
$\Delta\Delta G_{cyclo\rightarrow cyclohexanol}$	−7.10
$\Delta\Delta G_{cyclo\rightarrow 1,2diol}$	−10.3
$\Delta\Delta G_{cyclo\rightarrow 1,4diol}$	−11.4
$\Delta\Delta G_{cyclo\rightarrow myo\text{-}inositol}$	−12.7
$\Delta\Delta G_{cyclo\rightarrow xylopyranose}$	−15.8
$\Delta\Delta G_{cyclo\rightarrow glucopyranose}$	−21.7

where they show a negative ΔG_h so indicating greater hydration (Table 3). Although cyclohexanol is mostly hydrophobic with only a single hydroxyl, it improves on the completely hydrophobic cyclohexane's free energy of hydration by 7.10 kcal/mol. By adding a second OH group to the cyclohexane frame, not adjacent to the first hydroxyl (1,4-cyclohexanediol), the ΔG_h improves by more than 4 kcal/mol compared with cyclohexanol. In the case of 1,2-cyclohexanediol where the hydroxyls are adjacent and capable of forming intramolecular hydrogen bonds, the increase in hydration free energy is less, that is, ∼3 kcal/mol. The least dramatic improvement in free energy is for the fully hydroxylated cyclohexane ring (myo-inositol) where ΔG_h = −12.70 kcal/mol increasing by ∼1 kcal/mol on the 1,4-cyclohexanediol value. The free energy pattern for the cases where hydroxyls are incrementally added to cyclohexane rings supports the hydration measurements and follows the solubility obervations seen in Figure 1.

Table 4. Linear Diffusion Coefficients for Hexacyclic Molecules Calculated from Classical MD Simulations

compounds	D/cm$^2\cdot$s^{-1}	normalized D
cyclohexanol	$1.51 \pm 0.34 \times 10^{-5}$	1
1,2-diol	$1.38 \pm 0.02 \times 10^{-5}$	0.913
1,4-diol	$1.47 \pm 0.02 \times 10^{-5}$	0.974
myo-inositol	$1.17 \pm 0.01 \times 10^{-5}$	0.774
xylopyranose	$1.13 \pm 0.003 \times 10^{-5}$	0.748
glucopyranose	$0.922 \pm 0.01 \times 10^{-5}$	0.610

The n_h measurements and calculations show similar although not as dramatic an increase, as observed for solubility, when progessing from cyclohexane based molecules to the pyranose based molecules. The free energy data supports the same pattern where ΔG_h = −15.80 kcal/mol for xylopyranose is more than 3 kcal/mol stronger than myo-inositol even though it has fewer hydroxyl groups. This increase in free energy of hydration must therefore be due to the introduction of an ether group into the cyclohexane ring especially since the number of hydrophyllic groups has decreased in the pyranose system. Finally, the most dramatic increase in the free energy of hydration, that is, 6 kcal/mol is due to the addition of a primary alcohol to the xylopyranose structure resulting in the glucosepyranose molecule. This simple addition of a single hydroxymethyl group gives a 3-fold increase in ΔG_h mimicking the dramatic increase in solubility seen in Figure 1.

The comparative diffusion rates (Table 4) supports the n_h and ΔG_h trends such that cyclohexanol diffuses the fastest and glucose diffuses the slowest in water. Since these systems are of similar shape and size, the slower diffusion of glucose implies that it drags a larger sheath of waters along with it compared with molecules that have greater numbers of hydroxyl groups (i.e., myo-inositol). Waters are able to form long-term hydrogen bonds with hydroxyls that diffuse slower than the solvent diffusion rate. The diffusional rate of the solute hydroxyls depends on the solute's translational and rotational diffusion as well as the conformational motion leading to further hydroxyl positional diffusion. Solute molecules that are several times larger than water diffuse slower than the water diffusion rate. Therefore this

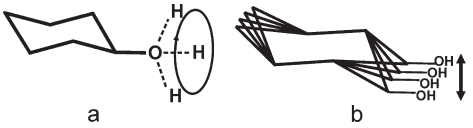


Figure 4. The (a) rotational and (b) the librational diffusion of an hydroxyl group tethered to a hexacyclic ring.

Table 5. Auto and Cross Correlation Functions for Hydroxyl and Hydroxymethyl Rotations Calculated from Classical Force Field Simulations

compounds	auto correlation	τ [ps]	cross correlation	τ [ps]
cyclohexanol	ϕ_{11} = H1−C1−O1−HO1	1.34		
1,2-diol	ϕ_{11} = H1−C1−O1−HO1	4.73	ϕ_{12}	3.71
	ϕ_{22} = H2−C2−O2−HO2	3.11		
1,4-diol	ϕ_{11} = H1−C1−O1−HO1	1.95	ϕ_{12}	0.98
	ϕ_{22} = H4−C4−O4−HO4	1.56		
myo-inositol	ϕ_{11} = H1−C1−O1−HO1	6.72	ϕ_{12}	10.8
	ϕ_{22} = H2−C2−O2−HO2	9.75	ϕ_{23}	5.70
	ϕ_{33} = H3−C3−O3−HO3	8.28	ϕ_{34}	5.70
	ϕ_{44} = H4−C4−O4−HO4	10.5	ϕ_{45}	4.50
	ϕ_{55} = H5−C5−O5−HO5	5.93	ϕ_{56}	7.60
	ϕ_{66} = H1−C1−O1−HO1	5.94	ϕ_{61}	10.9
xylopyranose	ϕ_{11} = H1−C1−O1−HO1	4.70	ϕ_{12}	8.90
	ϕ_{22} = H2−C2−O2−HO2	5.64	ϕ_{23}	8.88
	ϕ_{33} = H3−C3−O3−HO3	6.40	ϕ_{34}	8.89
	ϕ_{44} = H4−C4−O4−HO4	5.14	ϕ_{41}	14.9
glucopyranose	ϕ_{11} = H1−C1−O1−HO1	4.45		
	ϕ_{22} = H2−C2−O2−HO2	4.60	ϕ_{12}	6.76
	ϕ_{33} = H3−C3−O3−HO3	6.47	ϕ_{23}	12.7
	ϕ_{44} = H4−C4−O4−HO4	12.2	ϕ_{34}	13.4
	ϕ_{55} = C5−C6−O6−HO6	7.80	ϕ_{45}	36.7
	ϕ_{66} = H61−C6−O6−HO6	8.10	ϕ_{56}	10.9
	ϕ_{77} = O5−C5−C6−O6	82.5	ϕ_{68}	69.4
	ϕ_{88} = C4−C5−C6−O6	75.8		

will not be the cause of prematurely interrupted intermolecular hydrogen bonding between the solute and water. All evidence therefore points to a hydration model that does not depend on the number of hydrophyllic groups but on the conformational properties of the system. The two sources of conformational motion that affect the rotational and positional diffusion of the hydroxyls are the rotation of the hydroxyl and the puckering of the hexacyclic ring (Figure 4). *The greater the rotational and positional diffusion of the hydroxyls, the shorter the lifetime of the intermolecular hydrogen bonds that they make with the surrounding water molecules.*[2] *This is the major factor that determines hydration number.*

To address the relative rotational behavior of secondary and primary alcohols on cyclohexanol, 1,2-cyclohexanediol, 1,4-cyclohexanediol, myo-inositol, β-D-xylopyranose, and β-D-glucopyranose we calculated tcfs for the hydroxyl groups' rotations from the force field simulations as these have been accurately parametrized againsts experimental data for this property. From the tcfs, we then calculated relaxation times (τ), which are listed in Table 5. We labeled hydroxyl groups in sequence around the ring so hydroxyl i is adjacent to $i+1$ and for nonpyranose rings $i = 6$ is adjacent to $i = 1$. It can be seen that hydroxyl groups that

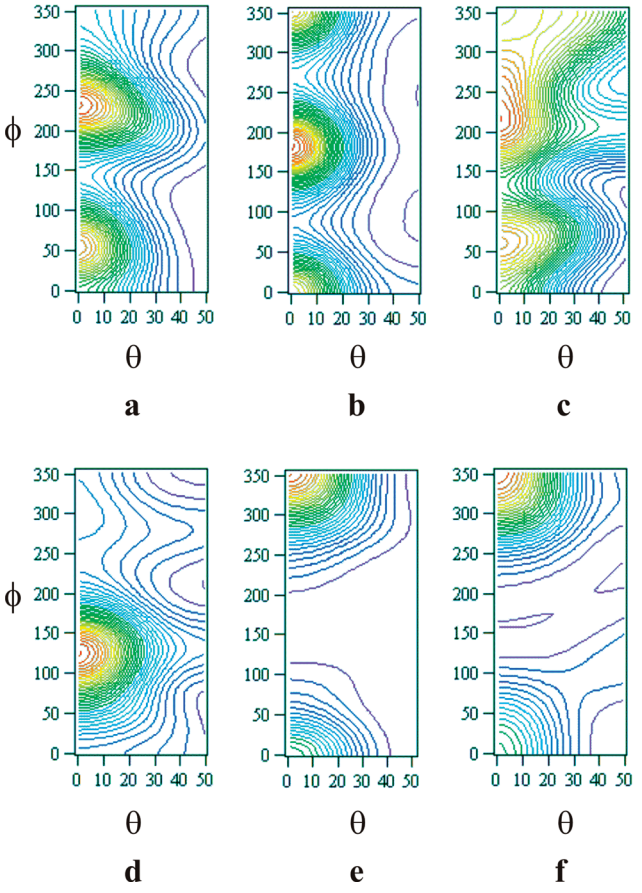


Figure 5. Pucker parameter probability distributions for (a) cyclohexanol; (b) 1,4-cyclohexanediol; (c) 1,2-cyclohexanediol; (d) myo-inositol; (e) β-xylopyranose; (f) β-glucopyranose.

are adjacent to each other rotate more slowly because of the bridging water hydrogen bonds[4] that form the half moon shaped high probability water densities. This accounts for 1,2-cyclohexanediol diffusing slower than 1,4-cyclohexanediol. However, the doubling of hydroxyl groups only marginally slows down the diols compared with cyclohexanol, particularly for myo-inositol where the cyclohexane ring is fully hydroxylated the diffusion rate is only 23% slower than that of cyclohexanol. The even slower diffusion of xylopyranose indicates that a slow down in conformational motion of the hydroxyls must be the cause of an increase in the lifetime of the hydration about the ring.

The hydroxyl rotation of xylopyranose is similar to the cyclohexane-based molecules (Table 5) The major change from cyclohexane to pyranose ring systems is due to the slow down in the positional diffusion of the hydroxyl groups. Since this is not sourced from the rotational frequency of the hydroxyls the change in their positional motion must be due to the ring to which they are tethered (Figure 4b). An analysis of the extent of the ring puckering for cyclohexane and pyranose ring systems was done using a Cremer−Pople spherical polar coordinate definition[38] for hexacyclic ring pucker as described above. Here, we only used the SCC-DFTB trajectories as these have been shown to be the most reliable for this property.[35] The most stable pyranose conformations are $^4C_1$ and $^1C_4$ chairs.[41] Our puckering results are in agreement with this so that all the rings studied here adopt the expected $^4C_1$ chair conformation (i.e.,

Table 6. Relaxation Time of θ and ϕ Angles for Puckering Analysis Calculated from the SCC-DFTB Trajectories

compounds	τ_θ [ps]	τ_ϕ [ps]
cyclohexanol	$0.072 \pm 3.2 \times 10^{-4}$	2.6 ± 0.037
1,2-diol	2.4 ± 0.13	3.9 ± 0.18
1,4-diol	3.8 ± 0.13	3.9 ± 0.17
myo-inositol	4.8 ± 0.21	4.5 ± 0.23
xylopyranose	4.6 ± 0.15	6.6 ± 0.86
glucopyranose	6.7 ± 0.49	6.8 ± 2.7

$\theta \sim 0°$ and $\phi \in (0, 360)$. The probability distribution of cyclohexane ring pucker has a wide range for ϕ (Figure 5a−d) while in the case of the pyranose rings the distribution is more confined (Figure 5e,f). We calculated the frequency of ring conformational motion using the pucker tcfs and then derived the corresponding relaxation times for both the zenithal (τ_θ) and azimuthal angles (τ_ϕ). The difference in ring pucker flexibility between the cyclohexanes and pyranoses is very clear (Table 6). The pyranose rings pucker more slowly compared with compounds that have a cyclohexane ring base. The slow down from myo-inositol ($\tau_\phi = 4.5$) to xylopyranose ($\tau_\phi = 6.6$) is significant and accounts for water associating longer with the slower moving hydroxyls on the monosaccharide leading to the observed increase in hydration from $n_h = 3.7$ to $n_h = 6.4$. The glucose ring pucker is about the same as that of xylopyranose ($\tau_\phi = 6.8$). However, it is the only molecule with a slowly rotating hydroxymethyl group ($\tau \sim 80$). Previously, we had shown that the hydroxymethyl group participates in numerous hydrogen bonds with water, particularly when in the gt conformation water molecules are trapped between the primary alcohol and the ring oxygen, leading to water forming bridging hydrogen bonds.[42] The combination of these relatively rigid conformational features found in glucose is the reason for its high hydration ($n_h = 7.2$) and correlates strongly with its high solubility in water.

CONCLUSIONS

We have shown that the number of hydroxyls on a flexible nonaromatic ring is not linearly proportional to its relative hydration. The H_2O molecules shared between adjacent OH groups tend to form complex bridging hydrogen bonds that are observed in the rdf and sdf probability densities. However, the residence time of the water molecules in the volumes close to the solute is more fundamental to the definition of hydration than is the probability density of finding the water in a hydrogen-bonding location. This is because the energetic difference between intermolecular water hydroxyl hydrogen bonds and intramolecular hydroxyl−hydroxyl hydrogen bonds is less than 1 kT. Therefore the competition between these two hydrogen bond types leads to frequent exchanges between them when the hydroxyl group has fast rotational and positional diffusion.

The most significant contribution to sustained intermolecular (water−hydroxyl) hydrogen bonds and therefore increased n_h, is the positional diffusion of the hydroxyls. This is closely related to the frequency with which the ring, to which it is tethered, puckers. Here frequency of pyranose ring pucker is significantly slower than that of cyclohexane ring pucker. The hydroxyls therefore librate slower on the former ring and so increase the possibility of hydrogen bonding with surrounding water molecules for longer times. This rationale for hydration is bourne out when we measure the water residence times about hydroxyls bonded to a cyclohexane frame compared with water residence times about hydroxyls bonded to pyranose rings. Conformational rigidity is therefore the reason for improved hydration in these cyclic systems bearing hydrophyllic groups. Moreover, the presence of a conformationally rigid ring and a rigid hydroxymethyl group that slowly rotates accounts for the increase in hydration number of glucose compared with that of xylopyranose. These measurements and calculations show that conformational rigidity is responsible for an enhanced interaction between water and monosaccahride molecules that corresponds well with the observed solubility trends for hexacyclic molecules.

ASSOCIATED CONTENT

Supporting Information. Details of the FEP intermediate hybrid simulations and tables of residence times are provided. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

*E-mail: Kevin.Naidoo@uct.ac.za. Fax: +27-21-686-4333.

ACKNOWLEDGMENT

This work is based upon research supported by the South African Research Chairs Initiative (SARChI) of the Department of Science and Technology and National Research Foundation to K.J.N. A.B. thanks the Scientific Computing Research Unit and the University of Cape Town for postgraduate study support. We thank Bhajan Lal for advice in our selection of the ultrasonic experiment.

REFERENCES

(1) Jensen, K. P.; Jorgensen, W. L. *J. Chem. Theory Comput.* **2006**, *2*, 1499.
(2) Astley, T.; Birch, G. G.; Drew, M. G. B.; Rodger, P. M. *J. Phys. Chem. A* **1999**, *103*, 5080.
(3) Chen, Y.-J.; Naidoo, K. J. *J. Phys. Chem. B* **2003**, *107*, 9558.
(4) Naidoo, K. J.; Chen, Y.-J. *Mol. Phys.* **2003**, *101*, 2687.
(5) (a) Franks, F. *Pure Appl. Chem.* **1987**, *59*, 1189. (b) Brady, J. W.; Schmidt, R. K. *J. Phys. Chem.* **1993**, *97*, 958.
(6) Naidoo, K. J.; Kuttel, M. M. *J. Comput. Chem.* **2001**, *22*, 445.
(7) Hine, J.; Mookerjee, P. K. *J. Org. Chem.* **1975**, *40*, 292.
(8) (a) Corzana, F.; Motawia, M. F.; Du Penhoat, C., H.; Perez, S.; Tschampel, S. M.; Woods, R. J.; Engelsen, S. *J. Comput. Chem.* **2004**, *25*, 573. (b) Engelsen, S. B.; Monteirob, C.; Herve de Penhoatb, C.; Perez, S. *Biophys. Chem.* **2001**, *93*, 103.
(9) Mason, P. E.; Neilson, G. W.; Enderby, J. E.; Cuello, G.; Brady, J. W. *J. Chem. Phys.* **2006**, *125*, No. 224505.
(10) Galema, S. A.; Hoiland, H. *J. Phys. Chem.* **1991**, *95*, 5321.
(11) Impey, R. W.; Madden, P. A.; McDonald, I. R. *J. Phys. Chem.* **1983**, *87*, 5071.
(12) (a) Buchner, R.; Capewell, S. G.; Hefter, G.; May, P. M. *J. Phys. Chem. B* **1999**, *103*, 1185. (b) Kaatze, U.; Hushcha, T. O.; Eggers, F. *J. Solution Chem.* **2000**, *29*, 299.
(13) Bockris, J. O.; Reddy, A. K. N. *Modern Electrochemistry 1: Ionics*; Plenum: New York, 1998; Vol. 1.
(14) (a) Hajduk, P. J.; Horita, D. A.; Lerner, L. E. *J. Am. Chem. Soc.* **1993**, *115*, 9196. (b) DiBari, M.; Cavatorta, F.; Deriu, A.; Albanese, G. *Biophys. J.* **2001**, *81*, 1190. (c) Brady, J. W. *J. Am. Chem. Soc.* **1989**, *111*, 5155.

(15) Ishihara, Y.; Okouchi, S.; Uedaira, H. *J. Chem. Soc., Faraday Trans.* **1997**, *93*, 3337.
(16) Barnett, C. B.; Naidoo, K. J. Calculating Ring Pucker Free Energy Surfaces from Reaction Coordinate Forces. In *Theory and Applications of Computational Chemistry*; Wei, D.-Q., Ed.; American Institute of Physics: Shanghai, China, 2008; 2009, Vol. 1102; pp 214.
(17) Graziano, G. *Phys. Chem. Chem. Phys.* **1999**, *1*, 3567.
(18) Qvist, J.; Halle, B. *J. Am. Chem. Soc.* **2008**, *130*, 10345.
(19) Rezus, Y. L. A.; Bakker, H. J. *Phys. Rev. Lett.* **2007**, 99.
(20) Chavez, A. L.; Gordon, G. B. *Chem. Senses* **1997**, *22*, 149.
(21) Shiio, H. *J. Am. Chem. Soc.* **1958**, *80*, 70.
(22) Burakowsky, A.; Glinski, J. *Chem. Phys.* **2007**, *332*, 336.
(23) Vishnyakov, A.; Lyubartsev, A. P.; Laaksonen, A. *J. Phys. Chem.* **2001**, *105*, 1702.
(24) Guvench, O.; Hatcher, E.; Venable, R. M.; Pastor, R. W.; MacKerell, A. D. *J. Chem. Theory Comp.* **2009**, *5*, 2353.
(25) Kuttel, M. M.; Brady, J. W.; Naidoo, K. J. *J. Comput. Chem.* **2002**, *23*, 1236.
(26) Brooks, B. R.; Brooks, C. L.; Mackerell, A. D.; Nilsson, L.; Petrella, R. J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S.; Caflisch, A.; Caves, L.; Cui, Q.; Dinner, A. R.; Feig, M.; Fischer, S.; Gao, J.; Hodoscek, M.; Im, W.; Kuczera, K.; Lazaridis, T.; Ma, J.; Ovchinnikov, V.; Paci, E.; Pastor, R. W.; Post, C. B.; Pu, J. Z.; Schaefer, M.; Tidor, B.; Venable, R. M.; Woodcock, H. L.; Wu, X.; Yang, W.; York, D. M.; Karplus, M. *J. Comput. Chem.* **2009**, *30*, 1545.
(27) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4* (2), 187.
(28) van Gunsteren, W. F.; Berendsen, H. J. C. *Mol. Phys.* **1977**, *34*, 1311.
(29) (a) Zhou, H.; Tajkhorshid, E.; Frauenheim, T.; Suhai, S.; Elstner, M. *Chem. Phys.* **2002**, *277*, 91. (b) Xu, D.; Guo, H.; Cui, Q. *J. Am. Chem. Soc.* **2007**, *129*, 10814. (c) Woodcock, H. L.; Hodoscek, M.; Brooks, B. R. *J. Phys. Chem. B* **2007**, 5720.
(30) Barnett, C. B.; Wilkinson, K. A.; Naidoo, K. J. *J. Am. Chem. Soc.* **2010**, *132*, 12800.
(31) Elstner, M.; Porezag, D.; Jungnickel, G.; Elsner, J.; Haugk, M.; Frauenheim, T.; Suhai, S.; Seifert, G. *Phys. Rev. B* **1998**, *58*, 7260.
(32) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. *J. Chem. Phys.* **1995**, *103*, 8577.
(33) Horn, H. W.; Swope, W. C.; Pitera, J. W.; Madura, J. D.; Dick, T. J.; Hura, G. L.; Head-Gordon, T. *J. Chem. Phys.* **2004**, *120*, 9665.
(34) (a) Barnett, C. B.; Naidoo, K. J. Calculating Ring Pucker Free Energy Surfaces From Reaction Coordinate Forces.*Theory and Applications of Computational Chemistry*; Shanghai, 2009. (b) Barnett, C. B.; Naidoo, K. J. *Mol. Phys.* **2009**, *107*, 1243.
(35) Barnett, C. B.; Naidoo, K. J. *J. Phys. Chem. B* **2010**, *114*, 17142.
(36) Zwanzig, R. W. *J. Chem. Phys.* **1954**, *22*, 1420.
(37) Jorgensen, W. L.; Ravimohan, C. *J. Chem. Phys.* **1985**, *83*, 3050.
(38) Cremer, D.; Pople, J. A. *J. Am. Chem. Soc.* **1975**, *96*, 1354.
(39) (a) Schmidt, R. K.; Karplus, M.; Brady, J. W. *J. Am. Chem. Soc.* **1996**, *118*, 541. (b) Liu, Q.; Brady, J. W. *J. Am. Chem. Soc.* **1996**, *118*, 12276. (c) Lui, Q.; Brady, J. W. *J. Phys. Chem. B* **1997**, *101*, 1317.
(40) Cui, Q.; Elstner, M.; Kaxiras, E.; Frauenheim, T.; Karplus, M. *J. Phys. Chem. B* **2001**, *105*, 569.
(41) Angyal, S. J. *Angew. Chem., Int. Ed.* **1969**, *8*, 157.
(42) Barnett, C. B.; Naidoo, K. J. *J. Phys. Chem. B* **2008**, *112*, 15450.