Integral Equation and ab Initio Study of the Effect of Solvation on Anomeric Equilibria in Aqueous Solution: Application to 4,6-Dimethyl-2-methoxytetrahydropyran

Sarah A. Maw, Richard A. Bryce, Richard J. Hall, Andrew J. Masters, and Ian H. Hillier*

Department of Chemistry, University of Manchester, Manchester, M13 9PL, U.K.

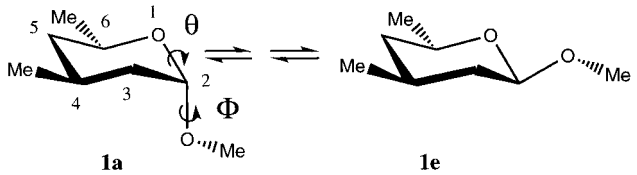
Received: December 17, 1997; In Final Form: March 12, 1998

A number of models are used to study the structure and energetics of the axial and equatorial anomers of 4,6-dimethyl-2-methoxytetrahydropyran in aqueous solution, to understand the observed preferential solvation of the equatorial form. Both quantum mechanical continuum and supermolecule treatments predict this effect qualitatively, but lead to its underestimation, suggesting the need for both long-range interactions and explicit solvent. A reference interaction site model, which includes both effects and uses a classical potential, is successful in quantitatively predicting such differential solvation, the result being less satisfactory when the solute is modeled quantum mechanically.

Introduction

Carbohydrate ligands occupy a central role in metabolism, as mediators of biological recognition. However, sugars present a unique challenge to the field of biomolecular structure due to their inherent flexibility, a multiplicity of polar groups, and subtle stereoelectronic behavior. In particular, the manifestation of the anomeric effect has invited intense theoretical and experimental investigation.[1−3] Reported originally in the work of Lemieux[4] and Edward[5] on sugars, the anomeric effect is general for any W−X−Y−Z system, where W and Y are groups of intermediate electronegativity, X has one or more lone pairs of electrons, and Z is an electron-withdrawing group. In the context of carbohydrates, this is generally believed to involve the hyperconjugation of oxygen lone pair density with an adjacent antiperiplanar C−O σ^* orbital, leading to an energetic stabilization and alteration in bond lengths of the O−C−O unit. It is useful to consider a simplified model system; therefore, without complicating hydroxyl and hydroxymethyl rotamers to consider, the anomeric equilibrium chosen for this study involved 4,6-dimethyl-2-methoxytetrahydropyran (DMT), depicted in Figure 1. The anomers are labeled variously as the axial, α or (sc,sc) anomer for **1a**, and equatorial, β or (ap,-sc) for **1e**. This latter descriptor refers to the antiperiplanar (ap) or synclinal (sc) orientation of the two dihedral angles, (θ,Φ), indicated in Figure 1. Donation of electron density can occur from the glycosidic aglycon oxygen into the pyranose ring σ^* orbital, termed the *exo*-anomeric effect,[6] resulting in a strong gauche, or synclinal, preference for the exocyclic dihedral, Φ. Conversely, the *endo*-anomeric effect refers to delocalization of a lone pair of electrons of the ring oxygen, which can only occur when the adjacent C−O bond is antiperiplanar to the lone pair of electrons; this orientation is achieved when θ is synclinal as in the axial anomer, **1a**. Controversy still exists as to the precise origin of this anomeric effect. Various explanations have been proposed and are outlined in a number of reviews.[6−9]

In aqueous solution, the position of anomeric equilibrium shifts to favor the equatorial anomer, **1e**. Experimentally, gas-phase studies show axial DMT to be lower in enthalpy by 1.2 kcal mol^{-1}. In aqueous solution the preference is reversed, with a ΔH_{eq-ax} of −0.3 kcal mol^{-1},[10] implying a significant differential solvation term of 1.5 kcal mol^{-1} favoring the equatorial form. This study observed negligible entropic components for nonpolar solvents, although due to hydrolysis, values for aqueous solvent were not obtained. If we take the enthalpies to approximate free energies, the corresponding distributions of the axial anomer are 88% in vacuo and 36% in water. The difficulty in experimentally quantifying this small free energy difference, however, is illustrated by measurements made for the closely related 2-methoxytetrahydropyran (MTP), found to be −0.7 kcal mol^{-1} from optical rotatory dispersion or 0.1 kcal mol^{-1} from NMR experiments.[11,12] The solvent dependence of the anomeric effect has also been studied experimentally,[12−14] with the conclusion that the axial preference is diminished by increasing solvent polarity. These results also indicate that hydroxylic solvents have an effect beyond a simple consideration of dielectric constant. Uncertainty still surrounds the origin of the preferential hydration: early NMR studies[15] on 2-methoxytetrahydropyran indicated entropy as the crucial factor. Subsequently, however, thermodynamic[10] and simulation studies[16,17] have attributed the solvation effect to enthalpic considerations, principally electrostatic in nature.



1a **1e**

Figure 1. Axial (**1a**) and equatorial (**1e**) anomers of 4,6-dimethyl-2-methoxytetrahydropyran (DMT).

To quantify this differential solvation, free energies may be calculated via atomistic simulations. The accumulation of the necessary ensemble averages can prove time-consuming and is coupled to the problem of insufficent sampling of higher energy configurations. Consequently, this approach has met with mixed success.[16−19] An alternative approach is to use continuum models, based on a reaction field approximation for the solvent. Here a solute molecule described quantum mechanically is embedded in a dielectric continuum. While also yielding solvation free energies, these models lack explicit solute−solvent interactions which would presumably be necessary for a complete description of sugar−water hydrogen bonding. An alternative and tractable route to free energies and equilibrium

S1089-5647(98)00547-1 CCC: $15.00
Published on Web 05/06/1998

constants is provided via the application of integral equation theories, which combine aspects of bulk and explicit solvation through structure present in the solute−solvent radial distribution function. One of these statistical models, the reference interaction site model, RISM,[20] and its modification for polar and charged solutes, the extended RISM, XRISM,[21] have been recently used to study a variety of solvation problems.[22−25] It is again possible to employ a quantum mechanical solute within a statistical solvent description of integral equation theories, which yield the RISM-SCF equations[26,27] for a Hartree−Fock description of the solute.

Classical simulation, continuum models, and integral equation methods have seen increasing application to solvation in recent years, principally to small solutes,[22,25] but also, in the case of the classical RISMs, to the proteins melittin[23] and calbindin.[24] However, it appears that no application of RISM to carbohydrates and their analogues has been reported. Consequently, we here describe a study of the use of XRISM and RISM-SCF in the treatment of the anomeric effect in aqueous DMT and compare these results with those from ab initio continuum models.

To include more fully the effect of explicit hydrogen bonding on the structure and energetics of sugar solvation, the solvent may be represented quantum mechanically, by the placement of waters around the solute to form a supermolecule. Further insight into the important solute−solvent interaction can be gained from a consideration of the orientation of waters of crystallization in relation to hydrated solutes, using the Cambridge Structural Database as a source of information. We here also discuss the results of supermolecule calculations in light of such structural data.

Computational Details

Ab Initio Calculations. All electronic structure calculations were carried out with the Gaussian 94 suite of programs[28] with quite large basis sets (6-311++G**), both at the Hartree−Fock (HF) level and including electron correlation using density functional theory (DFT) methods. We here employ the widely used B3LYP[29] functional. Three models of solvation were considered: explicit water molecules described quantum mechanically; a continuum representation of the solvent; and RISMs using both a classical and quantum mechanical description of the solute.

In the first model, the position of a single water molecule was optimized in proximity to DMT with an initial orientation of the water H−O vector chosen along the solute CÔC bisectrix; this was performed for both the exocyclic and endocyclic oxgyens of DMT. Since only relative conformations were of interest, what small basis set superposition error[30] that might be present should largely cancel between anomers.

The second solvent model used a continuous charge distribution. Three distinct continuum models were employed: the self-consistent reaction field (SCRF) model,[31−33] the polarizable continuum model (PCM)[34] implemented within Gaussian 94,[28] and the semiempirical AM1-SM2 solvation model.[35] Within the SCRF model, the energy was evaluated to the $l = 7$ degree of multipole expansion with the aim of allowing sufficient flexibility to describe the fairly complex electron density of DMT. The multipolar expansion appeared to converge fairly satisfactorily in all cases, although in some instances convergence was slow. The PCM cavity, probably more realistic than the SCRF ellipsoid that we have used, was determined self-consistently from the $0.001e$ isodensity surface at a constant dielectric of 78.3. The cavitation free energy, the energy required to create the cavity within the solvent in which the solute resides, was obtained using the scaled particle theory of Pierotti,[36] truncated to second order, for a cavity of molecular shape.[37] The contribution from the dispersion interaction of the solute and solvent was found from the approach of Floris and Tomasi,[38] using the $d^{(6)}$ coefficients from the COSMIC force field.[39]

Integral Equation Theory. Our implementation of the third solvent model is based upon the XRISM integral equations[21] for a solute−solvent system described by site−site interactions, in which a solute molecule is at infinite dilution in solvent molecules. These take the form of site−site Ornstein−Zernike (OZ) equations supplemented by an HNC closure. The OZ equations take the form

$$\hat{\mathbf{h}}^{vv} = \hat{\omega}^{v}\hat{\mathbf{c}}^{vv}\hat{\omega}^{v} + \hat{\omega}^{v}\hat{\mathbf{c}}^{vv}\rho\hat{\mathbf{h}}^{vv} \quad (1)$$

$$\hat{\mathbf{h}}^{uv} = \hat{\omega}^{u}\hat{\mathbf{c}}^{uv}\hat{\omega}^{v} + \hat{\omega}^{u}\hat{\mathbf{c}}^{uv}\rho\hat{\mathbf{h}}^{vv} \quad (2)$$

where u and v label solute and solvent sites, respectively. $h_{\alpha\gamma}(r)$ is the total correlation function between sites α and γ and is a matrix of site−site correlation functions. $c_{\alpha\gamma}(r)$ is the direct correlation function, $\omega_{\alpha\gamma}(r)$ is the intramolecular correlation function, and ρ is the solvent number density. The caret implies a Fourier transform of the matrix. To solve this equation, a closure relation is provided by the hypernetted chain (HNC) approximation[40] (eq 3).

$$[\mathbf{h}]_{\alpha\gamma} = h_{\alpha\gamma}(r) = \exp(-\beta U_{\alpha\gamma}(r) + h_{\alpha\gamma}(r) - c_{\alpha\gamma}(r)) - 1 \quad (3)$$

Here, $U_{\alpha\gamma}$ is the site−site pair potential and β is $(k_B T)^{-1}$. From the site−site correlation function, we obtain a solute−solvent interaction energy, $\langle\epsilon_{uv}\rangle$, which corresponds to the average interaction energy per solute molecule in the presence of the unperturbed solvent at the limit of infinite dilution. We have implemented XRISM to allow for polar and charged solutes, using both a classical and quantum mechanical description of the solute.

Our implementation of the latter model follows that of Tenno et al.[26] in which the electronic structure of the solute and the statistical distribution of the solvent molecules are simultaneously optimized. In this RISM-SCF model, the solute is coupled to the solvent via interaction sites assigned to the solute and solvent. Optimization can then be carried out in a self-consistent manner within the framework of the microscopic mean field approximation. The solvated Fock operator, F_i^{solv}, is calculated from

$$F_i^{solv} = F_i - f_i \sum_{\lambda} V_\lambda b_\lambda \quad (4)$$

where b_λ is the population operator for solute site λ, f_i is the occupation number of orbital ϕ_i, and V_λ is the mean field,

$$V_{\lambda\in u} = \rho \sum_{\alpha\in v} q_\alpha \int_0^\infty 4\pi r^2 \, dr \frac{g_{\lambda\alpha}^{uv}(r)}{r} \quad (5)$$

Here $g(r)$ is the pair distribution function, related to $h(r)$ by the subtraction of unity, and q_α is the partial charge at site α. Within RISM theory, the free energy of solvation, $\Delta\mu^{HNC}$, is given as[41]

$$\Delta\mu^{HNC} = \frac{\rho}{\beta} \sum_{\alpha=1}^{n_u} \sum_{\gamma=1}^{n_v} \int_0^\infty 4\pi r^2 \, dr \left[\frac{1}{2} h_{\alpha\gamma}^{uv\,2}(r) - c_{\alpha\gamma}^{uv}(r) - \frac{1}{2} h_{\alpha\gamma}^{uv}(r)\, c_{\alpha\gamma}^{uv}(r) \right] \quad (6)$$

TABLE 1: Total Energies (au) and Relative Energies (kcal mol^{-1}) of Axial and Equatorial DMT in Vacuo. ZPVE+TC Are Zero-Point Vibrational Energy and Thermal Contributions

model	E_{ax}	E_{eq}	ΔE_{eq-ax}
AM1//AM1	−69.344 36	−69.340 16	2.64
HF/6-31G*//HF/6-31G*	−461.985 79	−461.983 50	1.44
ZPVE+TC+HF/6-31G*//HF/6-31G*	−461.768 10	−461.766 19	1.20
MP2/6-311++G**//HF/6-31G*	−462.107 55	−462.106 09	0.92
B3LYP/6-311++G**//HF/6-31G*	−465.066 78	−465.065 54	0.78
B3LYP/6-31G*//B3LYP/6-31G*	−464.934 60	−464.932 55	1.29
B3LYP/6-311++G**//B3LYP/6-31G*	−465.069 27	−465.068 01	0.79
B3LYP/6-311++G**//HF/6-31G*	−465.066 80	−465.065 66	0.72
B3LYP/6-311++G**//B3LYP/6-311++G**	−465.069 42	−465.068 03	0.87

In calculating the free energy, this equation uses the HNC closure (eq 3). However, with an alternative assumption of Gaussian density fluctuations (GF) in the behavior of the solvent bath,[42] the resulting solvation free energy, $\Delta\mu^{GF}$, becomes

$$\Delta\mu^{GF} = -\frac{\rho}{\beta}\sum_{\alpha=1}^{n_u}\sum_{\gamma=1}^{n_v}\int_0^\infty 4\pi r^2\, dr\left[c^{uv}_{\alpha\gamma}(r) + \frac{1}{2}h^{uv}_{\alpha\gamma}(r)\, c^{uv}_{\alpha\gamma}(r)\right] \quad (7)$$

In all RISM calculations, the solvent was modeled using a rigid TIP3P water model.[43] The solute geometries were taken to be the HF/6-31G*//HF/6-31G* stationary points. In the classical RISM theory, solute−solvent intermolecular interactions were described using an OPLS potential with a united-atom representation of the solute. To allow for an all-atom representation of the solute, partial charges for DMT were calculated using the ChelpG fitting procedure[44] to the HF/6-31G*//HF/6-31G* electrostatic potential and subsequently averaged to avoid any conformational dependence of the atomic charges. The XRISM equations were solved using the hybrid Newton−Raphson/Picard technique, discussed by Morriss et al.[45] The Coulomb interactions were treated using the methods introduced by Ng.[46] In all the calculations, a grid of 256 points was used, with a grid point separation of 0.0792 Å. Tests showed this to give accurate results. The density was taken to be 0.997 g cm^{-3} at a temperature of 298 K.

In the quantum mechanical RISM approach, we employ a 6-31G* basis for the solute and carry out the calculation at the HF level. It is necessary to derive partial atomic charges from the solute electron density via fitting the electrostatic potential at a number of grid points. A number of possible grids were investigated from the work of Ten-no et al.,[26] Lebedev,[47] Kollman,[48] Brennan,[44] and a grid based upon icosohedral geometry. The Kollman and Brennan grids correspond to the Merz−Kollman and ChelpG charge-fitting procedures, respectively. A comparison of charges and energies was made with a RISM-SCF study of solvent effects on acidity by Kawata et al.[49] The best agreement found was by use of the Kollman grid as implemented in Gaussian 94,[28] and it was this grid that was used for the study of the anomeric equilibrium. For both RISM and RISM-SCF procedures, an insight into the solvent structure is obtained via the pair distribution function, $g(r)$, as well as the average interaction energy between solute and solvent, $\langle\epsilon_{uv}\rangle$. Due to the ability of the solute to polarize in response to the solvent electric field, however, the solvation free energies require adjustment to account for the cost of polarizing the solute wave function. We may calculate this electron reorganization energy, denoted E^{reorg}, using the relation

$$E^{reorg} = E^{sol}_{intra} - E^{gas} \quad (8)$$

where E^{gas} is the electronic energy of the solute in the gas phase. E^{sol}_{intra} is the adjusted solute energy in solution, found from eq 9, corrected for solute−solvent interactions which are already accounted for in the calculation of $\Delta\mu$.

$$E^{sol}_{intra} = E^{sol} - \sum_{\lambda\in u} V_\lambda(Z_\lambda - \langle b_\lambda\rangle) \quad (9)$$

Here, E^{sol} is the total solute energy in solution. Subtraction of the reorganization energy from $\Delta\mu$ leads to a corrected free energy of solvation, $\Delta\mu_0$. The solvation free energies for both classical and quantum mechanical RISM can then be resolved into entropy and enthalpy components,[50] involving the first-order solvent density derivative of $h_{\alpha\gamma}$ using $\Delta\mu^{HNC}$.

Results

On the basis of previous studies,[16,51,52] the (sc,sc) and (ap,-sc) isomers were selected to be the initial geometries for optimization of the axial and equatorial anomers respectively in equilibrium **1**. The total and relative gas-phase energies of anomers **1a** and **1e** after geometry optimization are given in Table 1. Calorimetric measurements of Wiberg et al.[10] indicate an in vacuo preference for the axial isomer of 1.21 ± 0.07 kcal mol^{-1}. Although this agrees well with the value of ΔG_{eq-ax} (g) of 1.20 kcal mol^{-1} calculated at the HF/6-31G*//HF/6-31G* level, our higher level calculations employing a 6-311++G** basis with electron correlation at both MP2 and B3LYP predict energy differences close to 0.8 kcal mol^{-1}. This energy difference would be further reduced to 0.6 kcal mol^{-1} when zero-point and thermal corrections from the HF/6-31G*//HF/6-31G* calculation are considered. This is in accord with calculations performed on 2-methoxytetrahydropyran at the MP2/6-311++G**//MP2/6-31G* level,[51] which gave an energy difference of 0.65 kcal mol^{-1}.

Table 2 gives the effect of reaction field solvation on the two anomers. When the gas-phase anomers are solvated, the AM1-SM2 model of Cramer and Truhlar[35] shows the correct trend, but predicts only a small relative stabilization of the equatorial anomer of 0.3 kcal mol^{-1} from a calculation at the AM1 geometry. This is in accord with the results of AM1-SM2 calculations on MTP[16] and D-glucose.[53] A recent study at the SM5.4/PM3 level[54] found the equatorial form of DMT to be preferentially solvated by 0.6 kcal mol^{-1}. The SCRF model also correctly predicts the preferential solvation of the equatorial form, with a reduction in ΔE_{eq-ax} from 1.4 to 1.1 kcal mol^{-1} at the HF/6-31G* level. This is not unexpected, given that axial DMT has a HF/6-31G* dipole moment of 0.3 D and equatorial DMT a value of 1.8 D. At the same level of theory, the PCM is more successful, further reducing this difference to 0.7 kcal mol^{-1}. For a variety of calculations, application of the PCM produces a relative stabilization of **1e** greater than for the SCRF model by about 0.4 kcal mol^{-1}. A similar trend was found in a PCM study at the HF/4-31G level of 2-methoxytetrahydropyran,[55] solvation energetically favoring the equatorial anomer by 1.4 kcal mol^{-1}. However, as with

TABLE 2: Total Energies (au) and Relative Energies (kcal mol^{-1}) of Axial and Equatorial DMT in a Reaction Field. Note That the Energy Difference for the AM1-SM2 Model Contains Cavitation and Dispersion Terms and Therefore Does Not Correspond to the Total Electronic Energies

model	E_{ax}	E_{eq}	ΔE_{eq-ax}
AM1-SM2//AM1	−69.342 92	−69.339 22	2.33
SCRF/HF/6-31G*//SCRF/HF/6-31G*	−461.989 12	−461.987 32	1.13
SCRF/B3LYP/6-31G*//SCRF/B3LYP/6-31G*	−464.937 54	−464.935 75	1.12
SCRF/B3LYP/6-311++G**//SCRF/B3LYP/6-311++G**	−465.073 58	−465.072 64	0.59
PCM/HF/6-31G*//HF/6-31G*	−461.989 59	−461.988 47	0.70
PCM/HF/6-31G*//SCRF/HF/6-31G*	−461.989 60	−461.988 53	0.67
PCM/HF/6-31G*//SCRF/B3LYP/6-31G*	−461.986 13	−461.985 41	0.45
PCM/B3LYP/6-31G*//SCRF/B3LYP/6-31G*	−464.937 73	−464.936 63	0.69
PCM/B3LYP/6-311++G**//SCRF/B3LYP/6-31G*	−465.072 55	−465.072 26	0.18
PCM/B3LYP/6-311++G**//SCRF/HF/6-31G*	−465.070 10	−465.069 86	0.15
PCM/B3LYP/6-311++G**//SCRF/B3LYP/6-311++G**	−465.072 68	−465.072 39	0.18

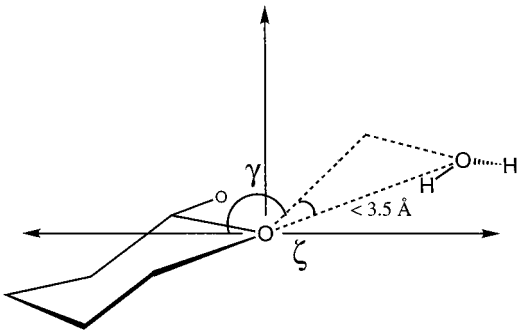


Figure 2. Definition of endocyclic substructure search query submitted to the Cambridge Structural Database. A hydrogen bond was defined as an interoxygen distance of 3.5 Å or less. Also defined are angles ζ and γ, used in the examination of the orientation of crystallographic hydrogen-bonding waters around the acetal oxygens of DMT.

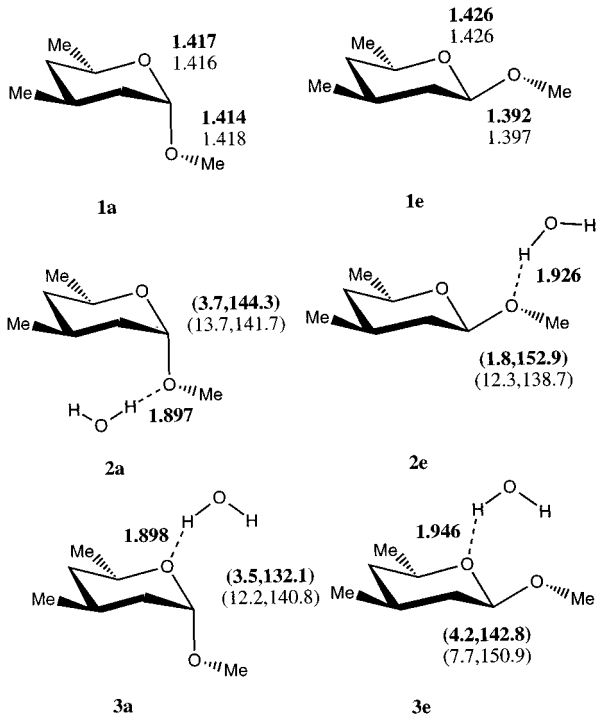


Figure 3. Comparison of calculated (B3LYP/6-311++G**) (bold) and experimental (crystal structure average) DMT structures: bond lengths of axial (**1a**) and equatorial (**1e**) anomers; solute−solvent bond lengths and angles (ζ,γ) for water bound to exocyclic (**2a**, **2e**) and endocyclic (**3a**, **3e**) oxygens. Lengths in angstroms and angles in degrees.

the gas-phase calculations, optimization using levels of theory higher than HF/6-31G* made little difference to the relative energy of the anomers and their geometry. A range of calculations at the PCM/B3LYP/6-311++G** level, with structures optimized using the SCRF method, predict the axial form to be preferred by 0.2 kcal mol^{-1}. Application of zero-point and thermal corrections at the HF level would make the anomers essentially isoenergetic. The dispersion energy correction was found to favor the axial anomer by 0.5 kcal mol^{-1}, and the cavitation term favored the equatorial anomer by 0.3 kcal mol^{-1}. Thus, including these effects leads to a final ΔG_{eq-ax} of about 0.2 kcal mol^{-1}. Therefore, although the trend exhibited by the reaction field models in preferentially stabilizing **1e** is in accord with experiment, the underestimation of the stabilization suggests the need to include explicit hydrogen bonding to obtain a quantitative model.

We now turn to a discussion of the effect of explicit water on the two anomers **1a** and **1e**. First, to obtain information regarding water orientation with respect to the sugar moiety, a survey of crystallographic data was performed. X-ray diffraction (XRD) data have served as an important experimental source of structural information about solutes in the presence of water and other solvents. Indeed, after the first observation of C−O bond length changes in crystal structures of acetals[56] due to the anomeric effect, several fruitful surveys of crystal structures and comparisons with ab initio geometries have been made in studies of the anomeric effect.[57−63] Two substructure search queries to explore the orientation of hydrogen-bonding waters around the solute were defined: the first substructure search query is illustrated in Figure 2 and found waters with an oxygen within a distance of 3.5 Å from the solute endocyclic oxygen. The second substructure differed from the first in searching for exocyclic oxygen hydrogen bonds to the solvent. Both queries were defined and submitted using the search and retrieval package QUEST3D, employing the April 1996 release of the Cambridge Structural Database.[64] The queries were found to match 99 and 35 entries, respectively, after having filtered out data with an *R*-factor of less than 0.10, nonorganic compounds, polymers, or structures with disorder, as well as compounds lacking a synclinal or antiperiplanar C−O−C−O dihedral. A coordinate system was then defined using angles ζ and γ, shown in Figure 2, for the endocyclic query. A further search was performed to compare bond lengths, using a query similar to Figure 2 but without the requirement of an explicit water. The search yielded 1552 matches from the database, and the results are compared in Figure 3. The comparison of bond lengths is made using the gas-phase ab initio structures on account of the close correspondence of the gas-phase and solution geometries. Indeed, at the HF/6-31G* level, the maximum difference between the SCRF and HF acetal bond lengths is only 0.004 Å. This correspondence is also reflected

TABLE 3: Total Energies (au) and Relative Energies (kcal mol^{-1}) of Axial and Equatorial DMT with Water Molecules Initially Orientated toward the Exocyclic Oxygen (2)

model	E_{ax}	E_{eq}	ΔE_{eq-ax}
AM1//AM1	−82.162 00	−82.159 29	1.70
HF/6-31G*//HF/6-31G*	−538.004 83	−538.002 87	1.23
HF/6-31G**//HF/6-31G*	−538.042 25	−538.040 35	1.19
HF/6-31+G**//HF/6-31G*	−538.056 06	−538.054 30	1.10
B3LYP/6-311++G**//HF/6-31G*	−541.533 27	−541.531 54	1.09
B3LYP/6-311++G**//B3LYP/6-311++G**	−541.536 84	−541.535 14	1.07
PCM/HF/6-31G*//HF/6-31G*	−538.015 36	−538.013 87	0.94
PCM/B3LYP/6-311++G**//HF/6-31G*	−541.541 12	−541.539 97	0.72

TABLE 4: Total Energies (au) and Relative Energies (kcal mol^{-1}) of Axial and Equatorial DMT with Water Molecules Initially Orientated toward the Endocyclic Oxygen (3)

model	E_{ax}	E_{eq}	ΔE_{eq-ax}
AM1//AM1	−82.161 59	−82.159 89	1.07
HF/6-31G*//HF/6-31G*	−538.005 38	−538.004 88	0.31
HF/6-31G**//HF/6-31G*	−538.042 86	−538.042 20	0.41
HF/6-31+G**//HF/6-31G*	−538.056 28	−538.055 61	0.42
B3LYP/6-311++G**//HF/6-31G*	−541.533 55	−541.532 94	0.38
B3LYP/6-311++G**//B3LYP/6-311++G**	−541.537 37	−541.536 43	0.59
PCM/HF/6-31G*//HF/6-31G*	−538.014 88	−538.014 90	−0.01
PCM/B3LYP/6-311++G**//HF/6-31G*	−541.540 17	−541.540 58	−0.26

in the closeness of the energies for the PCM/HF/6-31G*//SCRF/HF/6-31G* and PCM/HF/6-31G*//HF/6-31G* stationary points in Table 2.

Tables 3 and 4 summarize the results of supermolecular calculations. Optimized geometries at the HF/6-31G* level are illustrated in Figure 3, including a comparison of the water orientations with waters of hydration from the database search. The agreement between the B3LYP/6-311++G** and average crystallographic bond lengths about the anomeric center is excellent (**1a**, **1e**), illustrating their conformational dependence on the angles (θ,Φ). Both the supermolecule calculations and crystallographic data agree that the hydrogen-bonded water molecules lie along the CÔC bisector, where the oxygen nonbonding density is located. That there is some variation in the values of (ζ,γ) from theory and experiment is a reflection of the hydrogen-bonding potential energy surface, known to be quite flat in form.[65]

As far as the solvation energies are concerned, we see that inclusion of a single water molecule, hydrogen bonded to the appropriate oxygen atom, generally reduces the energy difference between the two anomers, as found experimentally for bulk water. Thus, using the semiempirical AM1 Hamiltonian, ΔE_{eq-ax} reduces from 2.6 kcal mol^{-1} in the gas phase to 1.7 kcal mol^{-1} for **2** and 1.1 kcal mol^{-1} for **3**, indicating immediately specific differential interactions with the water molecule over and above the nonspecific AM1-SM2 model.

We now discuss the relative energies of the **a** and **e** anomers from ab initio calculations when explicit hydrogen bonding to water occurs at the exo (**2a**, **2e**) and endo (**3a**, **3e**) oxygen atoms. These calculations (Tables 3, 4) reveal a greater preferential stabilization of the **e** anomer, when hydrogen bonding occurs at the endo oxygen. The stabilization is more pronounced for the HF calculations than at the B3LYP level due to the formation of cyclic structure, **3e**, with major and minor bifurcated O−H hydrogen bond lengths of 2.2 Å to the endocyclic oxygen and 2.6 Å to the exocyclic oxygen of DMT. Using B3LYP, however, this cyclic structure was not found, the water molecule instead forming a single hydrogen bond to the endocyclic oxygen of length 1.95 Å, as indicated in Figure 3. For water binding to the exo oxygen, the preferential stabilization is quite small (0.2 kcal mol^{-1} at the HF level), and indeed at the B3LYP level, explicit solvation favors **2a**. Further solvation of the supermolecules can be readily achieved via a continuum description of sugar−water complexes **2** and **3**. The presence of an explicit water results in an increased volume of the PCM cavity (Table 5), although for all cavities employed, there is little difference in volume for the two anomers. The PCM results are summarized in Tables 3 and 4. The continuum description results in further preferential differential solvation of the equatorial anomer, this effect being greater for hydrogen bonding at the endocyclic oxygen. Indeed, this model now predicts that the equatorial anomer is favored in solution, by less than 1 kcal mol^{-1}.

TABLE 5: Cavity Volumes (Å^3) Employed in Continuum Model Calculations

model	V_{ax}	V_{eq}
SCRF/HF/6-31G*//SCRF/HF/6-31G*	268.3	268.2
PCM/HF/6-31G*//SCRF/HF/6-31G*	239.9	240.3
PCM/HF/6-31G*//HF/6-31G*[a]	264.9	266.2
PCM/HF/6-31G*//HF/6-31G*[b]	264.4	264.9

[a] Water molecule initially orientated toward the exocyclic oxygen. [b] Water molecule initially orientated toward the endocyclic oxygen.

TABLE 6: Thermodynamic Properties (kcal mol^{-1}) Obtained from RISM Using United-Atom and All-Atom Representations of DMT. Solvation Free Energy Differences, $\Delta\mu^{GF}$ and $\Delta\mu^{HNC}$ Are Obtained via the GF and HNC Approximations. The Solute−Solvent Interaction Energy Is Denoted $\langle\epsilon_{uv}\rangle$

	united-atom			all-atom		
property	ax	eq	difference	ax	eq	difference
$\Delta\mu^{GF}$	−4.86	−6.60	−1.74	−10.16	−12.28	−2.12
$\Delta\mu^{HNC}$	28.80	27.08	−1.72	51.13	48.65	−2.48
$\langle\epsilon_{uv}\rangle$	−28.84	−32.77	−3.93	−24.36	−29.38	−5.02
Δh	−9.01	−11.12	−2.11	3.42	0.63	−2.79
$T\Delta S$	−37.81	−38.20	−0.39	−47.71	−48.02	−0.31

The bridge between explicit and continuum methods is provided by interaction site models, which seek to describe the effect of solvation through a sum of pairwise additive spherically symmetric potentials. The energetics of the anomeric equilibrium using a united-atom and all-atom description of the anomers is given in Table 6, with a comparison of the solute oxygen−solvent hydrogen pair correlation functions in Figure 4. The use of the Gaussian fluctuations (GF) closure with XRISM (eq 7) appears to show more realistic absolute solvation thermodynamics than the free energy expression derived from

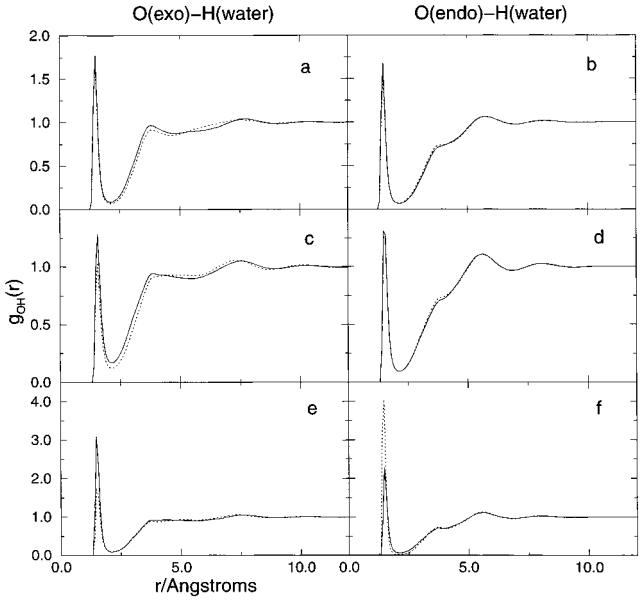


Figure 4. Comparisons of the pair correlation functions for united-atom (a, b), all-atom (c, d) and SCF-RISM (e, f) calculations of the axial and equatorial isomers of DMT in TIP3P water at 298 K. Axial is represented by a dotted line, equatorial by a solid line.

the HNC closure (eq 6), as observed previously.[22,41] However, the calculated free energies from GF appear too large, given an experimental free energy of hydration for dimethoxymethane of −2.9 kcal mol^{-1}.[66,67] In contrast, both closures yield a very reasonable solvation free energy difference, with $\Delta\Delta\mu$ of −1.7 kcal mol^{-1} in close agreement with the experimental value of −1.5 kcal mol^{-1}. For the all-atom DMT simulation, $\Delta\Delta\mu$ for both closures is about 0.5 kcal mol^{-1} greater than for the united-atom case. However, for both approaches, from the thermodynamic decomposition into enthalpic and entropic components, the enthapic term, $\Delta\Delta h$, is seen to dominate, in accord with previous studies.[10,16,17] A possible explanation of this energetic preference might be found from a consideration of the solvent structure in the solute−solvent pair correlation function. An early theory concerning the behavior of anomeric equilibria in aqueous solution proposed that the most solvated anomer was the one in which the hydrogen-bonding functionalities of the sugar were aligned in an optimal fashion relative to a tetrahedral lattice of water molecules.[68,69] While a gross simplification of liquid structure, the idea of compatability of the solute and its induced cooperative solvent network is useful and has been explored recently.[18,70] Indeed, from Figures 4a and 4c, the exocyclic pair distribution function indicates the equatorial anomer has a slightly more compact second solvation shell. More prominent, however, is the nearest neighbor peak for the hydrogen-bonding waters at the equatorial exocyclic oxygen in Figure 4c, found to be participating in 0.1 hydrogen bond more than the axial counterpart. It must be noted that the RISM total solvation free energy is a function of all possible solute−solvent pairwise interactions and not just of the atom pairs considered in the correlation functions of Figure 4. However, such preferential solvation is in good agreement with recent simulation studies on sugars. A molecular dynamics study of D-xylose by Schmidt et al.[18] also observed preferential hydrogen-bonding at the equatorial aglycon oxygen, where a differential of 0.2 hydrogen bonds was observed. At the endocyclic oxygen, however, there appears to be little discrimination. This feature was also borne out by the xylose study, but appears to be in contradiction with our supermolecule calculations, which indicate the greatest preferential solvation at the endocyclic oxygen in **3**.

TABLE 7: Thermodynamic Properties Obtained from a RISM-SCF Calculation on DMT. $\Delta\mu_0$ Refers to a Solvation Free Energy Difference Including the Solute Polarization Cost, $E_{\text{intra}}^{\text{sol}}$ Is the Corrected Solute Energy and E^{reorg} Is the Solute Reorganization Energy. See Table 5 for an Explanation of Other Notation. Thermodynamic Energies and Relative Energies Are in kcal mol^{-1}. Absolute Energies for $E_{\text{intra}}^{\text{sol}}$ and E^{reorg} Are in au

property	axial	equatorial	difference
$\Delta\mu_0^{\text{GF}}$	−16.97	−16.74	0.23
$\Delta\mu_0^{\text{HNC}}$	44.76	44.53	−0.23
$\Delta\mu^{\text{GF}}$	−25.74	−28.51	−2.77
$\Delta\mu^{\text{HNC}}$	36.00	32.77	−3.23
$\langle\epsilon_{\text{uv}}\rangle$	−54.99	−62.16	−7.17
Δh	−14.94	−19.14	−4.20
$T\Delta S$	−50.94	−51.90	−0.96
$E_{\text{intra}}^{\text{sol}}$	−461.971 78	−461.964 74	4.42
E^{reorg}	0.013 97	0.018 75	3.00

The results from the RISM-SCF calculation are given in Table 7. Although the uncorrected values of $\Delta\Delta\mu$ indicate a strong equatorial preference of about 3 kcal mol^{-1}, once the electron reorganization cost, E^{reorg}, has been included, this differential solvation reduces to essentially zero. For $\Delta\Delta\mu_0^{\text{GF}}$, the solvation free energy in fact favors the axial anomer by 0.2 kcal mol^{-1}. The enthalpy term, however, still dwarfs the contribution of entropy. This lack of discrimination by the solvent is concomitantly found in the pair distributions of Figures 4e and 4f. The equatorial exocyclic oxygen makes a greater fraction of 0.2 hydrogen bond, as in the classical RISM calculations, whereas axial endocyclic oxygen appears more favorably bonded, again by 0.2 hydrogen bond. Clearly, further work is required in the application of the RISM-SCF approach to the subtle problems presented by anomeric equilibria.

Discussion

We have here described the results of a number of theoretical approaches to model the subtle, yet definite, effect that water has on the relative stabilities of the two anomers of DMT. The success of continuum models involving a quantum mechanical description of the solute in treating a range of solvation problems is well-known, and here the preferential solvation of the equatorial form, found experimentally, is predicted by two similar continuum approaches. However, the differential solvation is somewhat underestimated. Despite its obvious limitations, a supermolecule model of the solvation involving a single water molecule, hydrogen bonded to the appropriate oxygen atom of the solute, yields local structure in broad agreement with structural data. Energetically, such explicit solvation generally favors the equatorial form, with this effect being more pronounced for interactions involving the endocyclic oxygen atom. Further solvation of the supermolecules using a continuum model provides little further insight. Thus, the bulk solvation of the hydrogen-bonded structures involving the exocyclic oxygen atom, using the PCM (Table 3), results in less preferential solvation of the equatorial form than is given by a PCM treatment of the solute with no explicit hydrogen bonding (Table 2). In the case of the supermolecule involving explicit hydrogen bonding to the endocyclic oxygen, further solvation at the PCM level (Table 4) does result in the equatorial form being preferred.

Continuum and supermolecule calculations thus suggest the need for inclusion of both explicit solvent molecules and interaction beyond the first solvation shell. These are provided by the RISMs, which can also include the solute molecule modeled quantum mechanically. The classical RISM calcula-

tions are successful in quantitatively predicting the observed differential solvation, which is largely assigned to enthalpic effects. At a structural level, this preferential solvation is reflected to some extent by a greater solvent interaction with the exocyclic oxygen, for the first and possibly second solvation shells, the former in agreement with simulation studies,[18] but not with our supermolecule studies.

The RISM-SCF calculations are less satisfactory, with essentially little difference in the solvation energies of the two anomers when the internal polarization energy of solute is included. These values are quite large (9−12 kcal mol^{-1}) and suggest that this model overestimates the solute−solvent interactions. Similar effects have been found for hybrid schemes involving a quantum mechanical description of the solute and a molecular mechanical (MM) description of the solvent, which have led to a reparametrization of the MM force field.[71] A similar problem may have been encountered here, which requires further investigation.

Acknowledgment. We thank EPSRC and Glaxo Wellcome Plc. for support of this research.

References and Notes

(1) Cramer, C. J. *J. Org. Chem.* **1992**, *57*, 7034.
(2) Marcos, E. S.; Pappalardo, R. R.; Chiara, J. L.; Domene, M. C.; Martinez, J. M.; Parrondo, R. M. *J. Mol. Struct. (THEOCHEM)* **1996**, *371*, 245.
(3) Cramer, C. J.; Truhlar, D. G.; French, A. D. *Carbohydr. Res.* **1997**, *298*, 1.
(4) Lemieux, R. U., Chū, N. J., Eds. *Abstracts of Papers*, 133rd National Meeting of the American Chemical Society, Washington, DC, 1958.
(5) Edward, J. T. *Chem. Ind. (London)* **1955**, 1102.
(6) Juaristi, E.; Cuevas, G. *Tetrahedron* **1992**, *48*, 5019.
(7) Kirby, A. J. *The Anomeric Effect and Related Stereoelectronic Effects at Oxygen;* Springer-Verlag: New York, 1983.
(8) Tvaroska, I.; Bleha, T. *Adv. Carbohydr. Chem. Biochem.* **1989**, *47*, 45.
(9) Juaristi, E.; Cuevas, G. *The Anomeric Effect;* CRC Press: Florida, 1994.
(10) Wiberg, K. B.; Marquez, M. *J. Am. Chem. Soc.* **1994**, *116*, 2197.
(11) Praly, J.-P.; Lemieux, R. U. *Can. J. Chem.* **1987**, *65*, 213.
(12) Lemieux, R. U.; Pavia, A. A.; Martin, J. C.; Watanabe, K. A. *Can. J. Chem.* **1969**, *47*, 4427.
(13) Eliel, E. L.; Giza, C. A. *J. Org. Chem.* **1968**, *33*, 3754.
(14) Anderson, C. B.; Sepp, D. T. *Tetrahedron* **1968**, *24*, 6873.
(15) Booth, H.; Khedair, K. A. *J. Chem. Soc., Chem. Commun.* **1985**, 467.
(16) Jorgensen, W. L.; de Tirado, P. I. M.; Severance, D. L. *J. Am. Chem. Soc.* **1994**, *116*, 2199.
(17) Ha, S.; Gao, J.; Tidor, B.; Brady, J. W.; Karplus, M. *J. Am. Chem. Soc.* **1991**, *113*, 1553.
(18) Schmidt, R. K.; Karplus, M.; Brady, J. W. *J. Am. Chem. Soc.* **1996**, *118*, 541.
(19) Koehler, J. E. H.; Saenger, W.; van Gunsteren, W. F. *Eur. Biophys. J.* **1987**, *15*, 197.
(20) Chandler, D.; Anderson, H. C. *J. Chem. Phys.* **1972**, *57*, 1930.
(21) Hirata, F.; Rossky, P. J. *Chem. Phys. Lett.* **1981**, *83*, 329.
(22) Yu, H.; Montgomery-Pettitt, B.; Karplus, M. *J. Am. Chem. Soc.* **1991**, *113*, 2425.
(23) Kitao, A.; Hirata, F.; Gō, N. *J. Phys. Chem.* **1993**, *97*, 10231.
(24) Svensson, B.; Woodward, C. E. *J. Phys. Chem.* **1995**, *99*, 1614.
(25) Lee, P. H.; Maggiora, G. M. *J. Phys. Chem.* **1993**, *97*, 10175.
(26) Ten-no, S.; Hirata, F.; Kato, S. *Chem. Phys. Lett.* **1993**, *214*, 391.
(27) Ten-no, S.; Hirata, F.; Kato, S. *J. Chem. Phys.* **1994**, *100*, 7443.
(28) Frisch, M. J.; Trucks, G. W.; Schlegal, H. B.; Gill, P. M. W.; Johnson, B. G.; Robb, M. A.; Cheeseman, J. R.; Keith, T. A.; Petersson, G. A.; Montgomery, J. A.; Raghavachari, K.; Al-Laham, M. A.; Zakrzewski, V. G.; Ortiz, J. V.; Foresman, J. B.; Cioslowski, J.; Stefanov, B. B.; Nanayakkara, A.; Challacombe, M.; Peng, C. Y.; Ayala, P. Y.; Chen, W.; Wong, M. W.; Andres, J. L.; Replogle, E. S.; Gomberts, R.; Martin, R. L.; Fox, D. J.; Binkley, J. S.; Defrees, D. J.; Baker, J.; Stewart, J. P.; Head-Gordon, M.; Gonzalez, C.; Pople, J. A. *Gaussian 94*; Gaussian, Inc.: Pittsburgh, PA, 1995.
(29) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648.
(30) Boys, S. F.; Bernardi, F. *Mol. Phys.* **1970**, *19*, 531.
(31) Rivail, J. L.; Rinaldi, D. *Chem. Phys.* **1976**, *18*, 233.
(32) Rinaldi, D.; Ruizlopez, M. F.; Rivail, J. F. *J. Chem. Phys.* **1983**, *78*, 834.
(33) Rinaldi, D.; Rivail, J. L.; Rguini, N. *J. Comput. Chem.* **1992**, *13*, 675.
(34) Miertus, S.; Scrocco, E.; Tomasi, J. *J. Chem. Phys.* **1981**, *55*, 117.
(35) Cramer, C. J.; Truhlar, D. G. *Science* **1992**, *256*, 213.
(36) Pierroti, R. A. *Chem. Rev.* **1976**, *76*, 717.
(37) Claverie, P.; Daudey, J. P.; Langlet, J.; Pullman, B.; Piazzola, D.; Huron, M. J. *J. Phys. Chem.* **1978**, *82*, 405.
(38) Floris, F.; Tomasi, J. *J. Comput. Chem.* **1989**, *10*, 616.
(39) Vinter, J. G.; Davies, A.; Saunders, M. R. *J. Comput. Aid. Mol. Des.* **1987**, *1*, 31.
(40) Hirata, F.; Pettitt, B. M.; Rossky, P. J. *J. Chem. Phys.* **1982**, *77*, 509.
(41) Lee, P. H.; Maggiora, G. M. *J. Phys. Chem.* **1993**, *97*, 10175.
(42) Chandler, D.; Singh, Y.; Richardson, D. *J. Chem. Phys.* **1984**, *81*, 1975.
(43) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926.
(44) Wiberg, K. B.; Breneman, C. M. *J. Comput. Chem.* **1990**, *11*, 361.
(45) Morriss, G. P.; MacGowan, D. *Mol. Phys.* **1986**, *58*, 745.
(46) Ng, K. C. *J. Chem. Phys.* **1974**, *61*, 2680.
(47) Lebedev, V. I. *Zh. Vychisl. Mater. Mat. Fiz.* **1977**, *18*, 132.
(48) Singh, U. C.; Kollman, P. A. *J. Comput. Chem.* **1984**, *5*, 129.
(49) Kawata, M.; Ten-no, S.; Kato, S.; Hirata, F. *J. Phys. Chem.* **1996**, *100*, 1111.
(50) Yu, H.; Roux, B.; Karplus, M. *J. Am. Chem. Soc.* **1990**, *92*, 5020.
(51) Tvaroska, I.; Carver, J. P. *J. Phys. Chem.* **1994**, *98*, 9477.
(52) Salzner, U.; Schleyer, P. v. R. *J. Org. Chem.* **1994**, *59*, 2138.
(53) Cramer, C. J.; Truhlar, D. G. *J. Am. Chem. Soc.* **1993**, *115*, 5745.
(54) Giesen, D. J.; Chambers, C. C.; Hawkins, G. D.; Cramer, C. J.; Truhlar, D. G. In *Computational Thermochemistry: Prediction and Estimation of Molecular Thermodynamics*; Irikura, K., Frurip, D. J., Eds.; ACS Symposium Series 285; American Chemical Society: Washington, DC, 1998.
(55) Montagnani, R.; Tomasi, J. *Int. J. Quantum Chem.* **1991**, *39*, 851.
(56) Berman, H. M.; Jeffrey, G. A. *Science* **1967**, *157*, 1576.
(57) Jeffrey, G. A.; Pople, J. A.; Binkley, J. S.; Vishveshwara, S. *J. Am. Chem. Soc.* **1978**, *100*, 373.
(58) Jeffrey, G. A.; Yates, J. H. *J. Am. Chem. Soc.* **1979**, *101*, 820.
(59) Jeffrey, G. A.; Yates, J. H. *Carbohydr. Res.* **1981**, *96*, 205.
(60) Longchambon, F. Ph.D. Thesis, University of Paris-Nord, Bobigny, France, 1984.
(61) Fuchs, B.; Schleifer, L.; Tartakovsky, E. *Nouv. J. Chim.* **1984**, *8*, 275.
(62) Tvaroska, I.; Kozár, T. *Chem. Zvesti* **1981**, *35*, 425.
(63) Perez, S.; Marchessault, R. H. *Carbohydr. Res.* **1978**, *65*, 114.
(64) Allen, F. H.; Kennard, O. *Chem. Des. Autom. News* **1993**, *8*, 31.
(65) Taylor, R.; Kennard, O. *Acc. Chem. Res.* **1984**, *17*, 320.
(66) Hine, J.; Mookerjee, P. K. *J. Org. Chem.* **1975**, *40*, 292.
(67) Cabani, S.; Gianni, P.; Mollica, V.; Lepori, L. *J. Solution Chem.* **1981**, *10*, 563.
(68) Franks, F. *Water;* Royal Society of Chemistry: Letchworth, 1983.
(69) Kabayama, M. A.; Patterson, D. *Can. J. Chem.* **1958**, *36*, 563.
(70) Liu, Q.; Brady, J. W. *J. Am. Chem. Soc.* **1996**, *118*, 12276.
(71) Freindorf, M.; Gao, J. *J. Comput. Chem.* **1996**, *17*, 386.