ELSEVIER

Carbohydrate Research 276 (1995) 219–251

CARBOHYDRATE RESEARCH

Relative stability of alternative chair forms and hydroxymethyl conformations of β-D-glucopyranose

Susan E. Barrows [a], Frederic J. Dulles [a], Christopher J. Cramer [a,*], Alfred D. French [b,*], Donald G. Truhlar [a,*]

[a] *Department of Chemistry and Supercomputer Institute, University of Minnesota, 207 Pleasant St. SE, Minneapolis, MN 55455-0431, USA*

[b] *United States Department of Agriculture, Agricultural Research Service, 1100 Robert E. Lee Blvd., P.O. Box 19687, New Orleans, LA 70179-0687, USA*

Received 26 January 1995; accepted 23 March 1995

Abstract

The relative energies of two hydroxymethyl conformers for each of the two chair forms ($^4C_1$ and $^1C_4$) of β-D-glucose were calculated at much more complete levels of quantum mechanical (QM) electronic structure theory than previously, and relative free energies in solution were calculated by adding vibrational, rotational, and solvent effects. The gas-phase results are based on very large basis sets (up to 624 contracted basis functions), and the coupled cluster method for electron correlation. Solvation Model 4 was used to calculate the effects of hydration or nonpolar solvation. Molecular mechanics (MM) and QM electronic structure theory have been applied to analyze the factors contributing to the relative energies of these conformers. Relative energies varied widely (up to 35 kcal/mol) depending on theoretical level, and several levels of theory predict the experimentally unobserved $^1C_4$ ring conformation to be the lower in energy. The highest level calculations predict the $^4C_1$ chair to be lower in free energy by about 8 kcal/mol, and we also find that the *gauche*$^+$ (*gt*) conformer of $^4C_1$ is lower than the *trans* (*tg*) conformer. Low-energy structures optimized by either quantum mechanical or molecular mechanical methods were commonly characterized by multiple intramolecular hydrogen bonds. Superior hydrogen bonding geometries are available in the $^1C_4$ chair, but are counteracted by increased steric repulsions between axial substituents; MM calculations also indicate increased torsional strain in the $^1C_4$ chair. Manifestations of greater steric strain in the calculated $^1C_4$ structures compared to the $^4C_1$ structures include longer ring bonds, a larger bond angle at the ring oxygen atom, and

* Corresponding authors.

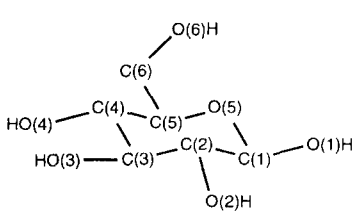
SSDI 0008-6215(95)00175-1

smaller puckering amplitudes. The MM and QM $^4C_1$ structures compare well with each other and with available X-ray diffraction data. The largest discrepancies between the two kinds of models occur for geometric parameters associated with the anomeric center — the QM structure agrees better with experiment. Greater differences between QM and MM structures are observed for $^1C_4$ structures, especially in the relative orientations of hydroxyl groups serving as hydrogen bond acceptors. In water, the $^4C_1$ chairs are better solvated than the $^1C_4$ chairs by about 5 to 9 kcal/mol because of both larger polarization free energies and improved hydrogen bonding interactions with the first solvation shell. In (a hypothetical) *n*-hexadecane solution, the $^4C_1$ chairs are better solvated by about 2 to 4 kcal/mol both because of larger polarization free energies and because the larger solvent accessible surface areas of the $^4C_1$ conformers allow increased favorable dispersion interactions. The differential polarization free energies are associated primarily with the hydroxyl groups; the greater steric congestion in the $^1C_4$ chairs reduces opportunities for favorable dielectric screening.

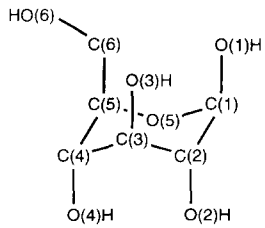
Keywords: Glucose; Conformational analysis; Chair; Ring puckering; Solvation; Quantum mechanics; Molecular orbital theory; MM3; Molecular modeling

1. Introduction

D-Glucose plays many important roles in biochemistry and biotechnology. Moreover, residues of D-glucopyranose are components of commercially important polysaccharides (e.g., starch and cellulose) and of bacterial and mammalian glycoproteins [1–3]. Because many of the properties of sugar-containing molecules depend on their conformation(s), conformational analysis of glucose has been extensively explored by a variety of computational methods [4–15]. Questions addressed in these studies include the ring conformation, the configuration at the anomeric center C-1 (see Chart 1), the orientations of the exocyclic substituents, and the effects of solvation. In general, however, the factors that determine the preferred structures of monosaccharides remain poorly understood.



$^4C_1$ chair



$^1C_4$ chair

Background.—For D-glucopyranose, one challenge has been to explain the 40:60 ratio [16] of the axial α anomer to the equatorial β anomer in aqueous solution. Considerable study has been devoted to simple model systems, like dimethoxymethane and 2-hydroxytetrahydropyran, where some conformers allow for $n_O \rightarrow \sigma^*_{CO}$ hypercon-

jugative delocalization (referred to as anomeric delocalization in systems where the donor oxygen atom resides in a ring), while others do not. In these model systems, both gas-phase theoretical studies and experimental studies in non-polar solvents predict conformers permitting this delocalization to be favored [17–34] (e.g., the axial anomer predominates in 2-hydroxytetrahydropyran). As a result, it has been suggested that the observed equilibrium anomeric population in glucose indicates more favorable solvation of the equatorial anomer by aqueous solvent [9,35]. In contrast, recent gas-phase calculations and aqueous solvation studies, the latter employing both continuum and explicit-solvent solvation models, all point to the differential anomeric solvation being less than 1 kcal [9,11–13,15] (all energies in the text are molar energies).

Another fundamental question is the issue of ring conformation, and this is the focus of the present investigation. It is well known that β-D-glucopyranose takes the all-equatorial chair conformation [36,37]. This shape is illustrated in Chart 1 in the conventional view and is designated $^4C_1$. The alternative chair, designated $^1C_4$, places all exocyclic substituents in axial positions. Accurately modeling the energetic consequences of so significant a difference in structure is a challenging test for any molecular modeling method.

All experimentally observed glucose moieties take the $^4C_1$ conformation. Thus, the free-energy difference between the $^4C_1$ and $^1C_4$ conformations must be fairly large, probably greater than 4 kcal, which would correspond to an equilibrium population of the $^1C_4$ conformation of less than 0.1% at 298 K. Using an empirical scheme, Angyal estimated an energy preference of 6.0 kcal for the all-equatorial conformation of β-D-glucopyranose in aqueous solution [16]. With a similar but more general approach based on the relative free energies of axially and equatorially monosubstituted cyclohexane rings [38], Franck proposed that axial vs. equatorial substitution of hydroxyl and hydroxymethyl groups should be destabilizing by 1.5 and 2.9 kcal, respectively [23]. Pearson and Rumquist [17], however, give a free-energy difference of 0.75 kcal (at 38°C) for 2-hydroxytetrahydropyran, with axial favored. This probably provides a more appropriate estimate for the anomeric hydroxyl, where stereoelectronic effects oppose steric effects. Summing these values for glucopyranose would provide a crude estimate that favors the $^4C_1$ conformer over the $^1C_4$ by about 7 kcal. Since these values are based on tetrahydropyran rings, they concern the unfavorable interactions of the axial substituents with *trans*-diaxial hydrogen atoms; interactions of OH or CH_2OH groups with each other should be markedly worse [39], so the 7 kcal value might be considered a lower bound to the combined steric and anomeric effects. On the other hand, the axial hydroxyl groups can take advantage of transannular hydrogen bonding, and there may also be non-additive effects, so the situation is not that clear.

Dowd, Reilly, and one of the authors recently studied the ring shapes of all of the aldopyranoses using molecular mechanics (MM) force field methods, in particular MM3(92) [40], which is based on the earlier MM3(90) [41–43] force field, but has improvements in the terms devoted to hydrogen bonding. This work [15] predicted an energy difference between the two chairs of β-D-glucose of 8.6 kcal. Predicted equilibria for ring forms of other sugars, for which two or more pyranose ring conformations are observed, were in good agreement with experimental data. Anomeric ratios were similarly well predicted, with less than 1 kcal of error by comparison to

experiment. Nevertheless, questions remain. First, molecular mechanics involves a parameterized force field, whose parameters are developed by optimizing performance for a training set of molecules. As such, the method works best when applied to molecules similar in structure and functionality to those found in the training set. However, while the MM3(92) training set included a number of alcohols and ethers, none approached the complexity of glucose. This limitation may underlie the variability of the MM3(92) force field's ability to correctly model the structure of some disaccharides [44]. For example, in the particular case of a non-reducing dimeric analog of sucrose, it has been suggested that the force field fails to adequately account for overlapping anomeric effects [44].

Another concern is how the experimental conditions are accounted for in the study of Dowd et al. [15]. They used a homogeneous dielectric constant of $\epsilon = 3.0$ to compute

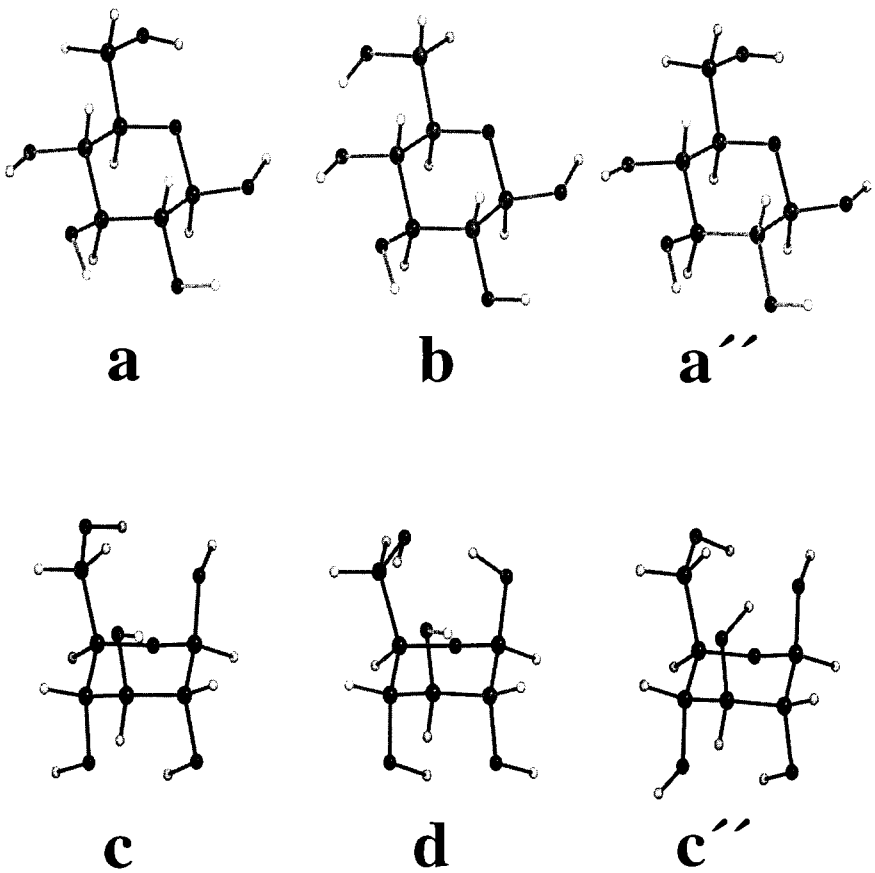


Fig. 1. $^4C_1$ G^+, $^4C_1$ T, $^4C_1$ G^+, $^1C_4$ G^+, $^1C_4$ G^-, and $^1C_4$ G^+ conformations of β-D-glucose optimized at several MM3 levels. See text (or Table 5) for force fields and dielectric constants used to optimize these structures and Chart 1 for atom numbering.

dipole–dipole and hydrogen-bonding interactions, independent of position in space. Such a treatment cannot take account of the fact that some parts of the low-dielectric solute screen other parts from the high-dielectric solvent. Furthermore, the extent of this screening depends on conformation. The importance of this effect, called differential solute descreening, in controlling relative free energies of alternative biochemical structures in aqueous solution has recently been emphasized [45]. (Some workers [46] propose using a larger ϵ for 1–4 interactions than for 1–3 interaction in six-membered rings. That approach is not pursued here.) Dowd et al. [15] also studied the effect of varying the homogeneous value of ϵ that they used, and their calculations indicated the α anomer to be increasingly preferred at lower dielectric constants. In contrast, the experimental result is that more of the β anomer of D-glucose is found in pyridine ($\epsilon = 12.5$) than in water ($\epsilon = 78.5$). One concludes that either the force field is inadequate to resolve this fine point, or a more sophisticated treatment of solvation is required.

Another unsettled issue, not addressed by Dowd et al., is the question of vibrational contributions to the molecular free energies.

Present approach.—When experimental data are unavailable and the reliability of MM calculations is questionable, information can be derived from quantum mechanical (QM) electronic structure calculations, and that is the approach used here. In particular we apply much higher levels of QM calculations than have been applied previously because high levels are required to obtain reliable results for small energy differences. However, there is a complication in that it is too expensive to apply high-level QM techniques to all the possible conformers. Glucose has six rotating groups, and were all torsional potentials to have three local minima, there would be $3^6 = 729$ different stereoisomers for each chair. Six such structures are shown in Fig. 1.

Thus we use a hybrid approach. We first use MM3(92) calculations to screen the 729 conformers for each chair and find leading candidates for the global minimum-energy conformation. Then we use high-level QM calculations to refine the geometries and energies of the leading candidates and to calculate rotational and vibrational contributions to the relative free energies.

We will use this approach to address the $^4C_1$ vs. $^1C_4$ conformational equilibrium for various ring and hydroxymethyl conformers of β-D-glucopyranose. Aqueous and nonpolar solvation effects on the conformational equilibria were calculated using the Solvation Model 4 [47–49], which accounts for both bulk dielectric and specific first-solvation-shell effects upon molecular solutes. Condensed-phase calculations are presented here for two solvents, water and *n*-hexadecane. The latter solvent is of particular interest because it has been found to mimic various organic biophases (e.g., cell membranes) and soil in its solvating properties [50–53].

Having obtained our best estimates, we will also explore the reliability of lower level approaches, which are of considerable interest for modeling other sugars and more complicated problems like glycoproteins and sugar interactions. At the QM level, we test both semiempirical molecular orbital theory and lower level ab initio approaches. At the MM level, we test the predictive capabilities not only of MM3(92) [40–43], but also of the more recent MM3(94) [54,55] method with various strategies for modeling the electrostatic effects.

Section 2 overviews the theoretical methods employed. Section 3 summarizes the highest level calculations and therein we provide our results for the structures and energetics of the lowest energy $^4C_1$ and $^1C_4$ conformers in the gas phase and in solution based on the best calculations. In Section 4 we discuss the energetic results with emphasis on electronic and steric effects in the gas phase and on identifying the dominant effects in solvation. In Section 5 we discuss structural aspects and provide an analysis of lower level approaches. Some additional computational details are given in Section 6. The paper ends with a summary of the main conclusions in Section 7.

2. Theoretical methods

Molecular mechanics.—MM3(92) is a general force field for organic conformational analysis [40–43]. The MM3(94) force field differs from the MM3(92) version in two respects that are important in the modeling of carbohydrates [54,55]. First, the parameters for the modified Buckingham potential used to describe hydrogen bonding are corrected for basis set superposition errors [56–59] present in the QM data used for the MM3(92) parameterization. This change reduces the strength of most hydrogen bonds by about 10%. Secondly, the original O–C–C–O torsional parameters [60] were changed both to increase the energy of eclipsed conformations and to slightly relocate the minima for *gauche* conformations.

To simulate gas-phase molecules, it has been recommended that one use an artificial dielectric constant of $\epsilon = 1.5$ for MM3(92) and MM3(94) calculations [15,40,44,61,62]. This accounts primarily for the dielectric response in the interior of the molecule, since polarization is not included explicitly (QM calculations include intramolecular polarization through a self-consistent field (SCF) optimization of the orbitals, and so they always have $\epsilon = 1$ interior to the solute). In the present work we also use higher values of the dielectric constant in some cases to include (as well as possible within the homogeneous dielectric approximation of MM methods) the effect of the higher dielectric constant in solution and crystalline phases.

Although MM3 is parameterized based on r_g values, i.e., the bond lengths observed from gas-phase electron diffraction studies, the MM3 program optionally uses vibrational frequencies to also calculate r_e, equilibrium bond lengths (for comparison to QM calculations), and r_α, low-temperature X-ray diffraction bond lengths [40]. These bond lengths are noted in some of the discussion below.

Quantum mechanics.—Semiempirical molecular orbital theory was employed to optimize structures with the Austin Model 1 [63] (AM1) and Parameterized Model 3 [64] (PM3) Hamiltonians, each of which uses 60 basis functions. Ab initio calculations were carried out at the restricted Hartree–Fock (HF) level [65] using several basis sets. The references for the basis functions and the number of contracted gaussian basis functions for glucose are as follows: minimal STO-3G [66], 72 functions; valence-double-ζ 3-21G [67], 120 functions; heavy-atom-polarized-valence-double-ζ 6-31G* [68–70], 204 functions; correlation-consistent-polarized-valence-double-ζ cc-pVDZ [71], 228 functions; correlation-consistent-polarized-valence-triple-ζ cc-pVTZ [71], 528 functions; and a

hybrid basis set that combines the s and p functions of the correlation-consistent-polarized-valence-quadruple-ζ [71] basis with the d and f polarization functions of the cc-pVTZ basis, 624 functions. We refer to this final basis as cc-p^{+}VQZ. HF calculations neglect the effect of electron correlation. Electron correlation effects were accounted for using many-body perturbation theory truncated at second order (MP2) [65] and using the more complete coupled cluster formalism [72] including single and double substitutions through infinite order (CCSD) [73].

Solvation effects.—Solvation free energies were calculated by Solvation Model 4 (SM4) [47–49]. This model is similar in spirit to the earlier Solvation Model 1 (SM1) [74] with one crucial distinction; it employs the Charge Model 1 (CM1) class IV charge model [75] to obtain atomic partial charges from the SCF wave functions in both the gas phase and solution. A critical element of the SM4 models, in light of uncertainties emphasized above for homogeneous-dielectric-constant models, is that they include a solute-shape-dependent (and hence conformation-dependent) treatment of the boundary between the $\epsilon = 1$ interior of the solute and the high-dielectric-constant external medium. They thereby avoid the arbitrariness of using effective homogeneous ϵ values to mimic solvation effects. The SM4 model was applied in the present paper in conjunction with the AM1 model for solute reorganization energies, the AM1–CM1 model for solute charge distributions, and MP2/cc-pVDZ geometries. Thus, the solvation free energies ΔG_S^0 include the calculated effects of electronic relaxation (i.e., redistribution of electronic density) but do not include any effects of geometry relaxation; the latter would be dominated by the inaccurate gas-phase potential energy surface at the semiempirical level. This approach of using high-level ab initio QM geometries for the calculation of aqueous solvation free energies gives results in good agreement with experiment for 1,2-ethanediol, a two-carbon fragment of glucose [76].

The free energy of solvation is given by [74,77,78]

$$\Delta G_S^0 = \Delta E_{\text{internal}} + G_P + G_{CDS}^0 \quad (1)$$

G_P accounts for the favorable effects of mutual electrostatic polarization of the solute and the solvent minus the associated cost of distorting the solvent structure and orientation of solvent molecules to achieve this. $\Delta E_{\text{internal}}$ accounts for the associated cost of internal distortion of the solute. G_{CDS}^0, which depends on solvent-accessible surface area (SASA) [79,80], includes the effects of cavitation, dispersion, and other interactions specific to the first solvation shell, such as the non-electrostatic components of solute–solvent hydrogen bonding and solute-induced changes in solvent–solvent hydrogen bonding, including the entropic effects of perturbing the structure of the first shell of solvent. G_{CDS}^0 also makes up empirically for some of the deficiencies of modeling the continuous electronic density using point charges.

SM4 requires different parameters for each solvent, and in the present study we considered *n*-hexadecane and water. Parameters for the former were taken from a general parameterization for that solvent that is published elsewhere [48]. Parameters for water were taken in part from another previous study; one parameter was determined especially for the present paper. This parameterization is discussed in more detail in Section 6. For brevity, the two different solvent models will be referred to as SM4–

hexadecane and SM4–water in the rest of this paper, or just as SM4 when no confusion is possible.

In SM4–hexadecane, G^0_{CDS} depends on two solvent-accessible surface areas. One, SASA(CD), is based on an effective solvent radius of 2.0 Å and accounts for cavitation and dispersion. The other, SASA(CS), is based on an effective solvent radius of 4.9 Å and accounts for cavitation and solvent structural effects. For SM4–water there is a single SASA, called SASA(CDS), calculated with an effective solvent radius of 1.4 Å, that accounts for all these effects.

3. Calculations and results

Nomenclature.—We use the following notation:

T O-6 *trans* to the ring oxygen

G^- C-6–O-6 bond *gauche* to both the C-5–O-5 and C-5–C-4 bonds

G^+ C-6–O-6 bond *gauche* to both the C-5–O-5 and C-5–H bonds

These are sometimes called *tg*, *gg*, and *gt*, respectively.

For identification purposes we will use boldface designations, e.g., **a, b, c, d**, etc., to identify a structure adopting a given chair and hydroxymethyl group conformation. In order to simplify the discussion, we will add one or more 'primes' to the designator when comparing structures having the same chair/hydroxymethyl conformation but optimized at different levels of theory, e.g., **a**, **a**′, and **a**″. In particular, a designator without primes connotes a structure optimized with the MM3(92) force field, a designator with one prime connotes a structure optimized at the MP2/cc-pVDZ level, and a designator with two primes connotes a structure optimized with the MM3(94) force field. We emphasize that the various hydroxyl group torsions may differ significantly when comparing structures optimized at different levels of theory, e.g., **c** and **c**″. We will use the designator enclosed in braces, e.g., {**a**}, when the subject is a group of structures all having the same chair/hydroxymethyl conformation but having potentially different hydroxyl group torsion angles. Figs 1 and 2 provide specific stereostructures for several conformers optimized at MM and QM levels of theory, respectively, and geometric details from levels of theory for which stereostructures are not provided are presented in tabular form where appropriate.

Structures and energies.—All 729 possible stereoisomers for each chair were energy minimized using the MM3(92) force field with dielectric constants of 1.5 and 3.0 (the latter being a repetition of the work of Dowd et al. [15]).

The minimum-energy $^4C_1$ conformer, **b**, as obtained by the MM3(92)/$\epsilon = 1.5$ calculations, has the T conformation, while the $^1C_4$ minimum, **d**, has the G^- conformation. The calculated energy difference favors the $^4C_1$ chair by 6.2 kcal. When a dielectric constant of 3.0 was employed, new MM3(92) global minima, **a** and **c**, each having the G^+ conformation, were located for each chair conformer. The stereostructures of these four minima are found in Fig. 1.

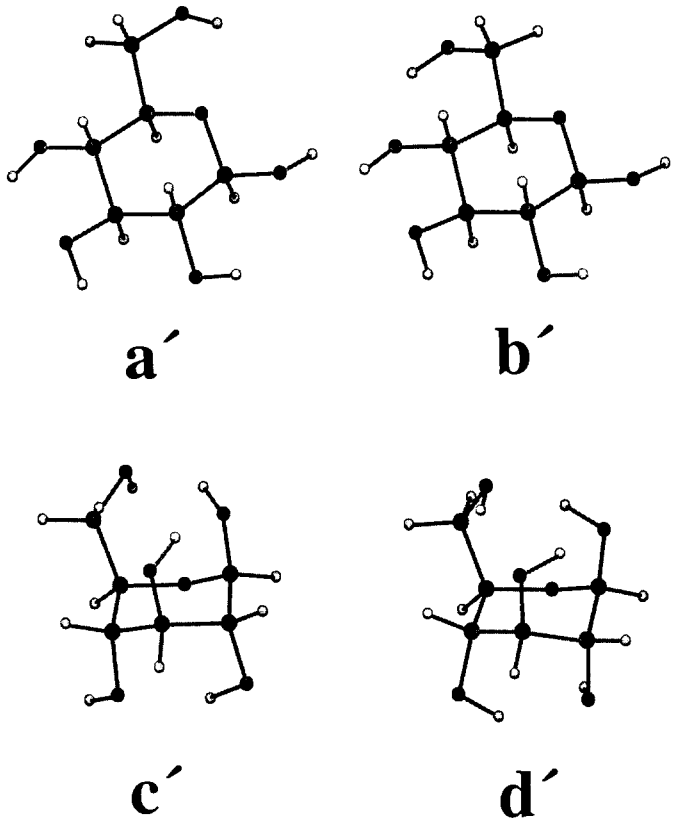


Fig. 2. $^4C_1$ G^+, $^4C_1$ T, $^1C_4$ G^+, and $^1C_4$ G^- conformations of β-D-glucose optimized at the MP2/cc-pVDZ level.

Table 1
Hydrogen-bond distances (Å) and valence angles (degrees) for MP2/cc-pVDZ optimized structures [a]

	∠O–H···O	rH···O		∠O–H···O	rH···O
$^4C_1$ G^+ **a′**			$^1C_4$ G^+ **c′**		
O-2–H···O-1	114	2.45	O-1–H···O-6	143	1.90
O-3–H···O-2	102	2.43	O-2–H···O-4	142	1.93
O-4–H···O-3	110	2.36	O-2–H···O-5	93	2.72
O-6–H···O-5	111	2.26	O-3–H···O-1	146	1.87
			O-6–H···O-5	92	2.48
$^4C_1$ T **b′**			$^1C_4$ G^- **d′**		
O-2–H···O-1	104	2.45	O-1–H···O-6	157	1.72
O-3–H···O-2	106	2.45	O-2–H···O-5	115	2.23
O-4–H···O-3	109	2.35	O-3–H···O-1	146	1.94
O-6–H···O-4	139	1.97	O-4–H···O-2	141	1.95
			O-6–H···O-3	142	1.93

[a] See Chart 1 for atom numbering.

Table 2
Calculated relative energies (kcal/mol) of D-glucose conformers in the gas phase [a]

Level of theory	$^4C_1\ G^+$ [b] **a′**	$^4C_1\ T$ **b′**	$^1C_4\ G^+$ **c′**	$^1C_4\ G^-$ **d′**
MP2/cc-pVTZ//MP2/cc-pVDZ	0.00	0.07	4.13	3.65
HF/cc-pVTZ//MP2/cc-pVDZ	0.00	0.40	12.21	15.73
HF/cc-p^{T}VQZ//MP2/cc-pVDZ	0.00	0.48	13.07	16.78
MP2/6-31G*	0.00	−0.45	0.66	−0.60
CCSD/6-31G*//MP2/6-31G*	0.00	−0.34	2.07	1.68
Composite E [c]	0.00	0.27	6.41	6.99
ZPVE [d]	0.00	0.18	1.05	1.09
ΔH_{298}(rot-vib) [d]	0.00	0.08	0.64	0.69
$-T\Delta S_{298}$(rot-vib) [d]	0.00	0.23	1.19	1.12
ΔG_{298}(rot-vib) [d]	0.00	0.31	1.83	1.81
Composite G_{298} [e]	0.00	0.58	8.24	8.80

[a] See text after Eq. (2) for $w/x//y/z$ notation.
[b] Absolute energies (E_h) for this column are −686.050 46; −683.606 17; −683.650 29; −685.176 03; −685.247 89; −686.166 44; 0.214 59; 0.224 75; −0.029 63; 0.195 12; −685.971 32.
[c] From Eq. (2).
[d] Zero-point vibrational energy (ZPVE) and thermal rotational–vibrational enthalpy and entropy terms calculated from unscaled HF/6-31G* frequencies.
[e] Composite $E + \Delta G_{298}$(rot-vib).

These four minima were taken as starting structures for QM optimizations. The MP2/cc-pVDZ optimized structures, **a′**, **b′**, **c′**, and **d′**, are presented in Fig. 2, and selected optimized hydrogen-bond distances and hydrogen-bond angles are in Table 1.

Table 3
Predicted solvation free energies and solvent-accessible surface areas (SASA) of β-D-glucose conformers in water and *n*-hexadecane [a]

	$^4C_1\ G^+$ **a′**	$^4C_1\ T$ **b′**	$^1C_4\ G^+$ **c′**	$^1C_4\ G^-$ **d′**
SM4 water				
SASA	380.4	375.9	356.2	352.0
G_P	−16.1	−15.5	−12.9	−8.5
$\Delta E_{internal} + G_P$	−14.7	−14.1	−11.0	−8.0
G^0_{CDS}	−5.1	−4.8	−3.6	−1.3
ΔG^0_S	−19.8	−18.9	−14.6	−9.3
relative ΔG^0_S	0.0	0.9	5.2	9.5
SM4 *n*-hexadecane				
SASA(CD)	425.5	423.7	400.3	394.0
SASA(CS)	958.9	957.6	919.2	909.2
G_P	−3.9	−3.7	−3.3	−2.2
$\Delta E_{internal} + G_P$	−3.7	−3.5	−3.0	−2.2
G^0_{CDS}	−6.4	−6.3	−5.1	−4.3
ΔG^0_S	−10.1	−9.8	−8.1	−6.4
relative ΔG^0_S	0.0	0.3	2.0	3.7

[a] Free energies in kcal/mol. Solvent-accessible surface areas are in Å^2.

Table 4
Predicted relative free energies of β-D-glucose conformers in the gas phase and in water and *n*-hexadecane solution [a,b]

	$^4C_1$ G^+ **a′**	$^4C_1$ T **b′**	$^1C_4$ G^+ **c′**	$^1C_4$ G^- **d′**
Gas phase	0.0	0.6	8.2	8.9
Aqueous solution	0.0	1.5	13.7	18.4
n-Hexadecane solution	0.0	0.9	10.2	12.6

[a] Free energies in kcal/mol.
[b] Calculated from composite G_{298} of Table 2 and relative ΔG_S^0 values of Table 3.

The highest level at which we optimized geometries is MP2/cc-pVDZ. Our best composite (C) electronic energies are obtained as:

$$\begin{aligned} E[\mathrm{C}] = {} & E[\mathrm{MP2/cc\text{-}pVTZ//MP2/cc\text{-}pVDZ}] \\ & + (E[\mathrm{CCSD/6\text{-}31G^*//MP2/6\text{-}31G^*}] - E[\mathrm{MP2/6\text{-}31G^*}]) \\ & + (E[\mathrm{HF/cc\text{-}p^TVQZ//MP2/cc\text{-}pVDZ}] \\ & - E[\mathrm{HF/cc\text{-}pVTZ//MP2/cc\text{-}pVDZ}]) \end{aligned} \quad (2)$$

The notation $w/x//y/z$ implies that the energy is evaluated with level of theory w using basis set x at the geometry calculated with level of theory y and basis set z. The required calculations are summarized in Table 2. Composite relative electronic energies are 0.0, 0.3, 6.4, and 7.0 kcal for **a′**, **b′**, **c′**, and **d′**, respectively.

Rovibrational contributions to the free energies at 298K calculated at the HF/6-31G* level further favor the $^4C_1$ conformers **a′** and **b′** by about 1.8 kcal. Adding the rovibrational contributions to the composite electronic energies yields the best calculated relative free energies in the gas phase, which are 0.0, 0.6, 8.2, and 8.8 kcal for **a′**, **b′**, **c′**, and **d′**, respectively.

Solvation free energies without geometry relaxation were calculated for each of the four conformers using the SM4 solvation models, and these results are given in Table 3. For the conformers in the same order as above, the relative solvation free energies are 0.0, 0.9, 5.2, and 9.5 kcal in water and 0.0, 0.3, 2.0, and 3.7 kcal, respectively, in *n*-hexadecane.

Table 4 presents the best calculated relative free energies of the four glucose conformers in the gas phase, aqueous solution, and *n*-hexadecane solution. The predicted relative energies of the $^1C_4$ chairs in either solvent (and in the gas phase) are sufficiently high to preclude experimental observation with currently available techniques.

4. Discussion of energetics

Table 4 shows that 8 kcal of the predicted free-energy difference between the best $^4C_1$ chairs and the $^1C_4$ chairs come from gas-phase contributions and 2–5.5 kcal come from differential solvation effects. It is noteworthy that the stronger favorable interactions of hydroxyl groups with aqueous solvent cause a much larger solution free-energy difference between the two chairs than is found in the non-polar *n*-hexadecane solvent.

We now examine the energetic differences between the two chair forms based on our best QM level of theory in the gas phase and solution. The lower-level QM calculations and MM3(94) calculations are used to provide some insight into the electronic origin of the effects.

Relative stability of gas-phase chair conformations.—To provide some insight into the relative importance of electrostatic and hydrogen-bonding interactions in the chair–chair energy difference, we carried out MM3(94) optimizations for **a**″ and **d**″ (not pictured) with dielectric constants of 1.0 and 4.0. The energy of the $^1C_4$ G^- chair relative to the $^4C_1$ G^+ chair increases from 8.9 kcal at $\epsilon = 1.0$ to 15.5 kcal at $\epsilon = 4.0$. Since increasing the dielectric constant decreases the magnitudes of the dipole–dipole and hydrogen-bonding terms in the force field, it is apparent that these interactions are stronger in the $^1C_4$ chair **d**″. The dominant factor in this difference appears to be shorter and more nearly linear hydrogen bonds in the $^1C_4$ chairs (Table 1) than in the $^4C_1$ chairs, a subject to which we return below.

Electron correlation also preferentially stabilizes the $^1C_4$ structures, as deduced from examination of the first six lines in Table 2. Thus, a best Hartree–Fock difference of 13 kcal is reduced to 6 kcal at the correlated level.

Another quantitatively important energetic effect is that the shorter hydrogen bonds in the $^1C_4$ chairs compared to the $^4C_1$ chairs (see Table 1) tend to tighten up the structures. As a result, the zero-point vibrational energies of the $^1C_4$ chairs are about 1 kcal larger than those of the $^4C_1$ chairs; this increase is spread quite evenly throughout all 66 normal modes. These tightened normal modes also decrease the vibrational entropy of the $^1C_4$ structures, and at 298 K that effect is worth about another 1 kcal; about 50% of this energy is contained in the five lowest frequency normal modes, another 25% in the next 17 modes, and the remainder in the higher frequency modes. The total rotational and vibrational corrections to the gas-phase free energies favor the $^4C_1$ chairs by slightly less than 1.5–1.8 kcal compared with the $^1C_4$ chairs.

Solvation.—Table 3 indicates that, in both water and *n*-hexadecane, the $^1C_4$ conformers are significantly further destabilized relative to the $^4C_1$. This result may be understood as a combination of differential electrostatic polarization effects and differential interactions of the two chairs with their respective first solvation shells.

We begin by considering the polarization effect, as measured by G_P. Electrostatic polarization favors the $^4C_1$ chairs by 0.5–1.5 kcal in *n*-hexadecane, and by 3–7 kcal in water. In general, systems which allow for intramolecular hydrogen bonding have less negative polarization free energies for conformers in which such hydrogen bonding is present than for conformers where the donor and acceptor groups interact exclusively with the surrounding solvent. This phenomenon has been noted in previous studies on the aqueous solvation of both glucose [12] and 1,2-ethanediol [76]. The less negative polarization free energy for the $^1C_4$ chairs is thus a consequence of their shorter and stronger intramolecular hydrogen bonds. The difference in G_P may be apportioned into individual atomic (or functional group) contributions [81], and, when this is carried out, the sums of the aqueous polarization free energies for the five hydroxyl groups are −17.1, −15.7, −13.3, and −7.2 kcal for the $^4C_1$ G^+ and T and $^1C_4$ G^+ and G^- conformers, respectively. These relative free energies of polarization track closely with those derived from the entire molecule.

The hydrogen-bonding effect on polarization is particularly evident on going from the G^+ to the G^- conformer of the $^1C_4$ chair (i.e., from **c′** to **d′**). Rotation of the C-6 hydroxyl group so that it lies over the ring allows for improved cyclic intramolecular hydrogen bonding between the hydroxyl groups on C-1, C-3, and C-6 in **d′**, but it makes the total aqueous polarization free energy less negative. In particular, the polarization free energy G_P associated with the hydroxyl groups on C-1, C-3, and C-6 changes from −6.4 to −3.3 kcal on going from the $^1C_4$ G^+ to the $^1C_4$ G^- conformer, which is about three-quarters of the total difference in G_P between the two conformers. We emphasize that this is not an example of conformational charge dependence. With no clear dependence on conformation, the hydroxyl oxygen CM1 partial atomic charges range from −0.53 to −0.57, and the hydroxyl proton CM1 charges range from 0.39 to 0.42 for all of the structures in Fig. 2. More strikingly, on going from **c′** to **d′** no charge changes by more than 0.01 in the hydroxyl groups on C-1, C-3 and C-6. This 3.1 kcal difference in polarization energies is entirely an example of the differential solute descreening discussed in Section 1, and it illustrates the potential energetic magnitude of this effect.

Another factor contributing to the greater stabilization of the $^4C_1$ chairs in solution is improved interaction with the first solvation shell, as measured by G^0_{CDS}. The SM4–hexadecane model recognizes two distinct ranges over which specific first-shell solvation effects operate, and this leads to two different solvent-accessible surface areas (SASA) to which individual components of these effects are assumed proportional. The larger area, SASA(CS), is a measure of the region over which solvent structure is disrupted. The smaller area, SASA(CD), corresponds to the range over which London dispersion forces operate. Moreover, the favorable CD components of G^0_{CDS} associated with the latter region usually outweigh the unfavorable CS components associated with the former, and that proves to be the case for glucose. Thus, a combination of the stronger contributions from CD terms and the greater CD SASA are sufficient to reverse the CS preference for the $^1C_4$ chairs and to contribute 1–2 kcal to relative preference for the $^4C_1$ chairs. The combination of roughly equal relative polarization and first-solvation-shell effects leads to the net differential free energies of *n*-hexadecane solvation presented in Table 3.

For glucose in water, the dominant aqueous contributions to the CDS term are the very favorable interactions associated with the hydroxyl groups. The increased exposure of these groups in the $^4C_1$ conformers contributes in part to the increased overall SASA in these conformers relative to the $^1C_4$, and as a result they have more negative G^0_{CDS} free energies by 1–4 kcal. The portions of G^0_{CDS} assignable to the five hydroxyl groups sum to −4.8, −5.0, −3.8, and −1.9 kcal for **a′**, **b′**, **c′**, and **d′**, respectively. Again, the effect of placing the C-6 hydroxyl group over the tetrahydropyran ring is large; that portion of G^0_{CDS} assignable to the C-6 hydroxyl group changes from −1.3 to −0.3 kcal on going from **c′** to **d′**.

The relative energies in solution are also in part a result of the stronger intramolecular hydrogen bonding, mentioned above, in the $^1C_4$ chair than in the $^4C_1$ chair. It has been experimentally established that monosaccharides manifest strong intramolecular hydrogen bonds in relatively nonpolar solvents like chloroform [82] and that these hydrogen bonds, though weakened, are maintained in polar aprotic solvents like dimethyl sulfox-

ide [83,84]. In aqueous solution, however, there appears to be no energetic preference for intramolecular ones over intermolecular hydrogen bonds to solvent molecules [85–88]. Given that situation, the relative energies calculated for the present conformers, which exhibit maximal intramolecular hydrogen bonding, might not be expected to be useful in predicting relative energies in aqueous solution. However, since the cost of breaking the intramolecular hydrogen bonds in the $^{1}C_{4}$ chair would seem to be larger than in the $^{4}C_{1}$, and since intermolecular hydrogen bonds would not be expected to be very different for the two chairs, it would appear that at worst the present results may be considered a lower bound on the chair–chair energy difference in solution, at least with regard to this issue.

The energy difference between {**a**} and {**b**} is particularly interesting because, in aqueous solution, NMR experiments identify the former as an equilibrium contributor, but not the latter [89]. This is consistent with the present predicted relative free energies in aqueous solution. However, while the differential solvation effect for the two conformers was noted earlier using the SM2 solvation model [12], this work provides the first instance where gas-phase calculations have found {**b**} to be higher in energy than {**a**}. Previous lower level ab initio QM studies [11] and some MM studies [6,9,13] have all predicted the opposite ordering, illustrating the importance of well converged QM calculations. As discussed below, the MM3(94) force field also predicts an ordering consistent with our best (composite) QM results.

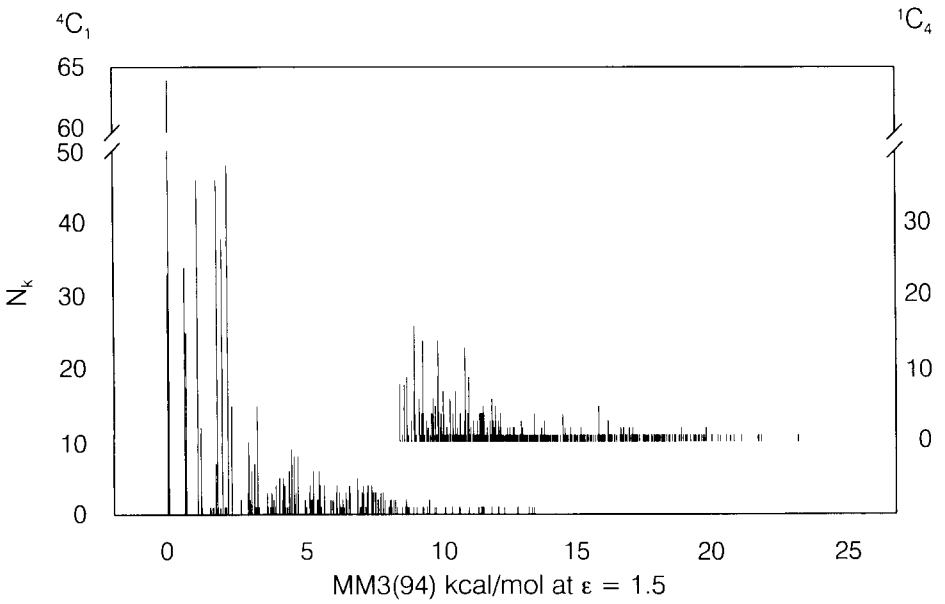


Fig. 3. Histogram illustrating the number of conformers (N_k) having similar relative energies for the 185 stationary $^{4}C_{1}$ conformers (left ordinate) and for the 480 stationary $^{1}C_{4}$ conformers (right ordinate, inset histogram) as optimized at the MM3(94)/ϵ = 1.5 level. The energies for both sets of conformers are relative to the $^{4}C_{1}$ global minimum.

Table 5
Relative energies (kcal/mol) of β-D-glucose conformers for different force-field calculations

Force field	Structure [a]					
	$^4C_1$			$^1C_4$		
	G^+ from MM3(92)/ϵ = 3.0 **a**	T from MM3(92)/ϵ = 1.5 **b**	G^+ from MM3(94)/ϵ = 1.5 **a″**	G^+ from MM3(92)/ϵ = 3.0 **c**	G^- from MM3(92)/ϵ = 1.5 **d**	G^+ from MM3(94)/ϵ = 1.5 **c″**
MM3(92)/ϵ = 1.5						
Unrelaxed geometries	1.29	0.00	1.31	9.97	6.21	8.91
Relaxed geometries	1.08	0.00	1.08	9.81	6.21	7.85
MM3(92)/ϵ = 3.0						
Unrelaxed geometries	0.00	1.11	0.74	8.58	9.76	9.23
Relaxed geometries	0.00	0.65	0.00	8.58	9.37	8.84
MM3(94)/ϵ = 1.5						
Unrelaxed geometries	0.70	1.89	0.00	11.13	17.54	8.63
Relaxed geometries	0.00	0.04	0.00	10.54	10.06	8.63

[a] This is the actual structure used for rows labeled 'unrelaxed geometry' and the starting structure from which optimization proceeded for rows labeled 'relaxed geometry'. For each row, results are relative to lowest absolute energy for that row.

A previous study on many different conformers of $^4C_1$ α-D-glucose allows us to estimate that the energetic effect of geometric relaxation in water will be about −0.2 kcal for every conformer [12]. We expect any correction in *n*-hexadecane to be smaller and similarly constant. Such corrections would, of course, have no effect on the relative free energies in solution.

5. Structural aspects and comparison of theoretical models

In this section we discuss structural and methodological issues in more detail.

MM3(94).—All 729 potential stereoisomers for each chair were also optimized using the MM3(94) force field with a dielectric constant of 1.5. Fig. 3 presents in histogram format the relative energies of all of the stereoisomers found to be stationary with the full matrix optimizer. The ordinate corresponds to the number of potential stereoisomers that converge to the same local minimum, i.e., N_k, where k denotes a unique local minimum. The 1458 starting structures in this case optimized to 185 unique $^4C_1$ chairs, ranging in relative energy from 0.0 to 13.6 kcal, and 480 unique $^1C_4$ chairs, ranging in relative energy from 8.6 to 24.1 kcal. The remaining 544 $^4C_1$ conformers and 249 $^1C_4$ conformers were not stationary on the hypersurface, i.e., for at least one rotatable bond there failed to be a torsional barrier separating the non-stationary conformer from another conformer. As noted in Section 6, MM3 optimizations may fail to fully converge in very flat regions of the potential energy hypersurface, so it is conceivable that some of the 'unique' chairs would converge to the same stereostructure with a more robust minimizer. However, the qualitative trend, i.e., the similar relative spread of $^4C_1$ vs. $^1C_4$ energies, would remain unchanged.

The lowest energy stereostructures for the $^4C_1$ chair, **a″**, and the $^1C_4$ chair, **c″**, are presented in Fig. 1. The Boltzmann-averaged energy difference between the two chair forms, based on the 185 unique $^4C_1$ chairs and the 480 unique $^1C_4$ chairs, is 9.0 kcal. This is somewhat greater than the energy separation of 8.6 kcal between **a″** and **c″**

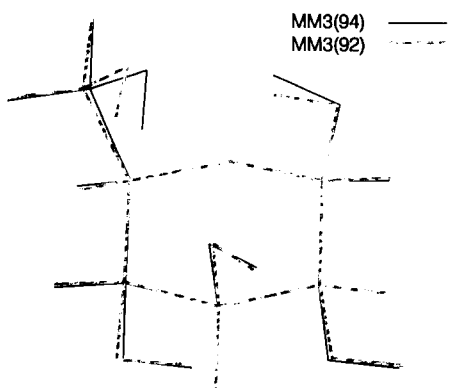


Fig. 4. Overlapped $^1C_4$ G^- structures **{d}** optimized at the MM3(92)/ϵ = 1.5 and MM3(94)/ϵ = 1.5 levels. The overlap minimizes the RMS deviation in ring atom positions.

because there are more $^4C_1$ chairs lying close in energy to **a**″ than there are $^1C_4$ chairs lying close in energy to **c**″.

Choice of homogeneous dielectric constant in MM calculations.—Table 5 presents the relative energies of **a**, **b**, **c**, **d**, **a**″, and **c**″ as calculated with three different force fields, MM3(92)/$\epsilon = 1.5$, MM3(92)/$\epsilon = 3.0$, and MM3(94)/$\epsilon = 1.5$. Energies were calculated for the frozen structures pictured in Fig. 1, and the six stereoisomers were also reoptimized using each of the other two force fields (**a** and **a**″ optimize to identical structures for all three choices of force field/dielectric constant). Table 5 illustrates the

Table 6
Comparison of selected geometrical data for QM, MM, and X-ray crystal structures of $^4C_1$ β-D-glucose [a]

	MP2/ cc-pVDZ **a**′	MM3(94) **a**″			X-ray [b]	Differences [c]		
		r_e	r_g	r_α		MP2 − MM3	MM3 − X-ray	MP2 − X-ray
Bond lengths								
C-1–C-2	1.523	1.513	1.523	1.517	1.525	0.010	−0.008	−0.002
C-2–C-3	1.517	1.510	1.520	1.515	1.520	0.007	−0.005	−0.003
C-3–C-4	1.520	1.514	1.524	1.518	1.511	0.006	0.007	0.009
C-4–C-5	1.526	1.519	1.529	1.524	1.529	0.007	−0.005	−0.003
C-5–C-6	1.520	1.	1.527	1.519	1.513	0.003	0.006	0.007
C-5–C-1	1.422	1.415	1.423	1.417	1.433	0.007	−0.016	−0.011
C-1–O-1	1.394	1.413	1.421	1.412	1.383	−0.019	0.029	0.011
C-2–O-2	1.420	1.426	1.434	1.425	1.429	−0.006	−0.004	−0.009
C-3–O-3	1.421	1.426	1.434	1.426	1.432	−0.005	−0.006	−0.011
C-4–O-4	1.418	1.426	1.434	1.425	1.419	−0.008	0.006	−0.001
C-5–O-5	1.437	1.433	1.442	1.436	1.437	0.004	−0.001	0.000
C-6–O-6	1.416	1.425	1.433	1.416	1.419	−0.009	−0.003	−0.003
Mean absolute error						0.008	0.008	0.006
Valence angles								
C-5–O-5–C-1	111.3		112.1		112.7	−0.8	−0.6	−1.4
O-5–C-1–O-1	109.3		105.6		107.0	3.7	−1.4	2.3
Dihedral angles								
O-5–C-1–C-2–C-3	58.4		60.6		53.7	−2.2	6.9	4.7
C-1–C-2–C-3–C-4	−54.0		−56.8		−50.8	2.8	−6.0	−3.2
C-2–C-3–C-4–C-5	53.7		56.3		53.4	−2.6	2.9	0.3
C-3–C-4–C-5–O-5	−57.4		−58.3		−59.8	0.9	1.5	2.4
C-4–C-5–O-5–C-1	63.5		63.6		66.3	−0.1	−2.7	−2.8
C-5–O-5–C-1–C-2	−63.9		−65.1		−62.8	1.2	−2.3	−1.1
Mean absolute error in tabulated angles						1.8	3.0	2.3
Cremer – Pople puckering parameters								
q	0.589		0.609		0.584	−0.020	0.025	0.005
θ [d]	3.0		2.0		7.0	1.0	−5.0	−4.0

[a] Bond lengths and puckering amplitude q are in Å and valence angles, dihedral angles, and puckering parameter θ are in degrees.
[b] From ref. [90], $R = 4.3\%$.
[c] For bond length differences, MM3 r_e is used to compare to MP2 values and MM3 r_α is used to compare to X-ray values.
[d] The small values for θ make comparison of ϕ values pointless.

sensitivity of the conformational analysis problem to the choice of MM3 force field and the dielectric constant. The predicted $^1C_4$ structures are much more sensitive to the choice of force field and dielectric constant than are the $^4C_1$ structures: the energy difference between frozen and relaxed structures (i.e., the difference as a result of switching force fields and/or dielectric constants) ranges from 0.2 to 7.5 kcal for $^1C_4$, compared with only 0.2–1.9 kcal for $^4C_1$. This sensitivity is caused by the steric crowding in the $^1C_4$ structures. Fig. 4 illustrates the overlap of {**d**} structures calculated with the MM3(92)/$\epsilon = 1.5$ and MM3(94)/$\epsilon = 1.5$ force fields, where the largest change of 7.5 kcal is observed on relaxation (i.e., rows 5 and 6 of column 6 in Table 5). As can be seen in the figure, little change in the ring geometry is observed and the minimized root-mean-square (RMS) deviation for the six ring atoms is only 0.02 Å. Instead, most of the deviation in the structures is in the location of the ring substituents: the total heavy-atom RMS deviation is 0.08 Å, and the all-atom RMS deviation is 0.13 Å. The large hydroxymethyl group is particularly sensitive to changing the force field—C-6 positions deviate by 0.12 Å, O(6) positions by 0.26 Å, and HO-6 positions by 0.37 Å. This change causes movement in the hydroxyl group on C-1 as well, which is hydrogen bonded to O-6; O-1 positions deviate by 0.04 Å, and HO-1 positions by 0.32 Å. The large energy change derives almost entirely from the non-1,4 van der Waals interaction term in the force field, i.e., from non-bonded interactions not associated with torsion angles. The small changes in structure required to induce so large an energetic effect illustrate how sensitive the van der Waals potentials are to structural changes in situations where steric crowding is severe.

In addition to being more sensitive to the version of the MM3 force field, the $^1C_4$ energies are also more sensitive to the choice of dielectric constant. For example, the predicted difference in relative energies between relaxed **a** and **b** changes by 1.7 kcal when the MM3(92) dielectric constant is changed from 1.5 to 3.0 (rows 2 and 4 of columns 2 and 3 in Table 5). On the other hand, the predicted difference in relative energies between **c** and **d** changes by 4.4 kcal for the same change in dielectric constant. This sensitivity can be attributed to the shorter range electrostatic interactions associated with the shorter hydrogen bonds found in the $^1C_4$ chairs compared to the $^4C_1$ chairs.

Bond lengths and valence angles.—The geometries optimized at different levels of theory differ primarily in hydroxyl group orientations. Bond lengths and bond angles show far less variation. For a $^4C_1$ G^- structure of β-D-glucose, experimental data are available in the form of an X-ray crystal structure [90]. Table 6 compares the X-ray structure to {**a**} geometries calculated at the MP2/cc-pVDZ and MM3(94)/$\epsilon = 1.5$ levels. The QM and X-ray geometries are in close agreement on bond lengths, with the largest deviation being 0.01 Å and the mean absolute deviation being about half that. The largest deviations are associated with bonds to the anomeric center, C-1. MM3(94) predicts the C-1–O-1 bond to be 0.03 Å longer than it is in the crystal structure and the C-1–O-5 bond to be 0.02 Å shorter. These discrepancies appear to reflect a weakness in accounting for the exo-anomeric effect (i.e., the lone pair delocalization from the anomeric hydroxyl oxygen atom into the σ^*_{CO} orbital of the tetrahydropyran [20,22,25,30,44,91,92]). At the QM level the bond length variation associated with the exo-anomeric effect is reproduced, although not to the extent present in the crystal structure.

That MM3(94) does not fully account for anomeric delocalization is also apparent in the O-5–C-1–O-1 bond angle. This angle is predicted to be smaller than is observed in the crystal structure. Anomeric delocalization is known to increase this bond angle relative to O–C–O linkages where stereochemical orientation of the oxygen substituents does not permit the delocalization to occur [18–22,25,27,29,30,91–95]. Endocyclic torsional angle variations between the theoretical and experimental structures range up to 7°; however, such variations in structure cost only tenths of a kcal energetically [96], and they might easily be accounted for by crystal packing effects.

Ring puckering.—The internal coordinates that vary most when pyranose rings adopt different conformations are the six ring torsion angles. However, simply listing these angles does not easily convey the ring shape, e.g., chair, boat, skew, half-chair, or envelope. Puckering parameters [97], on the other hand, readily connect to these more qualitative descriptors. Moreover, they provide a precise description of structures intermediate to the 38 characteristic conformers (2 chairs, 6 boats, 6 skews, 12 half-chairs, and 12 envelopes) for pyranose rings. We use here the Cremer–Pople system [97], typically used by carbohydrate crystallographers [98]. This system is described in the original literature [97,98] and in two recent papers by one of the present authors [15,61]. Briefly, the three puckering parameters for pyranose rings are q, the amplitude of puckering (the RMS deviation from the mean plane), θ, indicating the type of ring conformation (e.g., chair, skew, etc.), and ϕ, indicating the particular atoms that deviate the most from the mean plane. These three parameters constitute a spherical polar coordinate system, with q playing the role of the radius, θ serving as the colatitude, and ϕ serving as the azimuthal angle. By convention, 'perfect' $^4C_1$ and $^1C_4$ chairs have θ values of 0° and 180°, respectively. These positions on the Cremer–Pople surface may be thought of as corresponding to the north and south poles on a sphere. Near the poles, variations in ϕ do not significantly affect overall position.

The experimental ring puckering amplitude is in excellent agreement with the QM results, while MM3(94) predicts a chair which is less flattened by 0.02 Å. All structures are quite close to being 'perfect' chairs based on their Cremer–Pople θ values, although the predicted θ values are slightly too small at both the MM and QM levels. In addition to the θ value of 7.0° in Table 6, the average value of θ for nine other β-glucoside residues is also 7° [90,99–102].

The MM3(94) Cremer–Pople puckering amplitude, q, of **a** decreases from 0.624 to 0.614 as the assumed dielectric constant in the force field calculations is increased from 1.0 to 4.0. As mentioned above, increasing the dielectric constant decreases the strength of dipole–dipole interactions and hydrogen bonds, and this evidently loosens the 'noose' of hydroxyl groups around the ring. Since this noose squeezes the ring atoms out of the mean plane, the value of q decreases as ϵ increases. This dependence suggests that puckering amplitudes for glucose residues in crystal structures where the individual saccharide units do not exhibit extensive intramolecular hydrogen bonding (typically because they are indulging in intermolecular hydrogen bonding) might be expected to be smaller than those calculated (or observed) for the alternative possibility.

Experimental data are not available for the $^1C_4$ conformers. Nevertheless, it is interesting to compare QM and MM3(94) results for equivalent structures. Specifically for such a comparison, MM3(94) optimizations were carried out starting from the QM

Table 7
Hydroxyl and hydroxymethyl group torsional angles (degrees) and hydrogen bonding in glucose conformers at selected levels of theory

Geometry	Torsion angle and hydrogen bond acceptor (HBA) atom(s) [a]										
	HO-1 [b]	HBA	HO-2 [b]	HBA	HO-3 [b]	HBA	HO-4 [b]	HBA	C-6–C-5 [c]	HO-6 [d]	HBA
$^4C_1$ G^+ {a}											
MP2/cc-pVDZ [e]	51.3		−61.1	O-1	65.8	O-2	−71.1	O-3	−59.1	−52.4	O-5
MM3(94)/ϵ = 1.5 [f,g]	75.2		−65.4	O-1	69.0	O-2	−67.8	O-3	−58.0	−54.1	O-5
$^4C_1$ T {b}											
MP2/cc-pVDZ [e]	55.2		−61.5	O-1	65.8	O-2	−69.3	O-3	48.4	50.6	O-4
MM3(94)/ϵ = 1.5 [f]	79.0		−66.0	O-1	69.5	O-2	−69.7	O-3	52.6	47.2	O-4
$^1C_4$ G^+ {c}											
MM3(92)/ϵ = 3.0 [g]	51.3		167.4	O-4,O-5	52.0	O-1	37.2		−52.8	−61.2	O-5
AM1 [h]	56.0		171.3	O-4,O-5	−178.7	O-1	59.9		−55.6	−49.9	O-5,O-1
PM3 [h]	39.5		162.6	O-4,O-5	177.4	O-1	69.7		−45.1	−59.5	O-5
HF/3-21G [h]	90.7	O-6	157.5	O-4	141.3	O-1	54.9		−92.2	47.3	
HF/6-31G* [f]	79.5	O-6	164.4	O-4,O-5	146.3	O-1	59.6		−82.8	48.5	
HF/cc-pVDZ [f]	78.2	O-6	164.7	O-4,O-5	147.1	O-1	60.3		−82.4	48.3	
MP2/3-21G [f]	94.2	O-6	157.9	O-4,O-5	143.2	O-1	64.4		−92.3	33.5	O-5
MP2/6-31G* [f]	87.4	O-6	163.8	O-4,O-5	147.9	O-1	65.9		−87.1	39.5	O-5
MP2/cc-pVDZ [e,f]	87.7	O-6	163.2	O-4,O-5	149.8	O-1	68.4		−88.2	41.9	O-5
MM3(94)/ϵ = 1.5 [f]	77.8	O-6	176.8	O-4,O-5	139.1	O-1	−39.8		−66.3	107.9	

$^1C_4$ G^- {d}											
MM3(92)/$\epsilon = 1.5$ [g]	149.0	O-6–O-3	−37.5		48.1		−164.6	O-2–O-5	177.8	−60.1	O-3
AM1 [h]	102.4	O-6	−62.2		153.9	O-1,O-6	−171.7	O-2	173.2	−72.7	O-3
PM3 [h]	94.7	O-6	−58.7		161.3	O-1,O-6	−173.8		175.7	−75.1	O-3
HF/3-21G [h]	124.0	O-6	−61.7		121.2	O-1	−162.4	O-2	−176.7	−58.9	O-3
HF/6-31G* [f]	122.8	O-6	−62.9		117.2	O-1	−166.7	O-2	−175.4	−60.8	O-3
HF/cc-pVDZ [f]	149.7	O-6,O-3	−64.5		68.2		−167.2	O-2,O-5	−177.8	−58.9	O-3
MP2/3-21G [f]	128.9	O-6	−137.6	O-5	127.6	O-1	−154.8	O-2	−179.7	−51.9	O-3
MP2/6-31G* [f]	128.2	O-6	−135.2	O-5	126.1	O-1	−150.4	O-2	−179.1	−54.6	O-3
MP2/cc-pVDZ [e,f]	133.0	O-6	−135.4	O-5	127.6	O-1	−153.2	O-2	−179.9	−53.7	O-3
MM3(94)/$\epsilon = 1.5$ [f]	110.6	O-6	−134.8	O-5	130.4	O-1	−138.2	O-2	176.6	−67.1	O-3
$^1C_4$ G^+ c″											
MM3(94)/$\epsilon = 1.5$ [g]	53.3		173.9	O-4,O-5	141.0	O-1	−34.7		−58.5	−44.1	O-5
Other $^1C_4$ structure											
PM3 [i]	−77.1		174.9	O-4,O-5	−160.8		179.3		−173.2	62.8	

[a] Torsion angles in degrees. Listed acceptor oxygen atom(s) are within 3 Å of the donor group proton and the O–H · · · O angle is at least 90°.
[b] Torsion angle defined as H–O(n)–C(n)–H.
[c] Torsion angle defined as O-6–C-6–C-5–H.
[d] Torsion angle defined as H–O-6–C-6–C-5.
[e] Stereostructure in Fig. 2.
[f] Starting geometry was optimized geometry of preceding line in table.
[g] Stereostructure in Fig. 1.
[h] Starting geometry was corresponding MM3(92) optimized geometry.
[i] Stereostructure in Fig. 5.

structures **c**′ and **d**′. The most significant deviations between the QM and resulting MM structures are for the torsion angles of the hydroxyl groups—in several cases these groups rotate up to 120° away from the starting MP2 positions during MM3(94) optimizations (a complete listing of hydroxyl group torsional angles for multiple levels of theory is provided in Table 7). The primary cause for this deviation in hydroxyl group orientations between the MM and QM levels of theory is probably related to the manner in which MM3 is parameterized for hydrogen bonding. Rather than explicitly recognizing the interaction between a proton and an oxygen lone pair, the hydrogen-bonding term involves only (*i*) the interatomic distance between the hydrogen and oxygen atoms and (*ii*) the O–H · · · O angle. As a result, the energy associated with the hydrogen bonding term is insensitive to the orientation of the acceptor hydroxyl group. For the $^1C_4$ structures in particular, steric interactions become quite severe for conformations where hydroxyl groups are oriented so as to direct one lone pair (assuming an sp^3 type hybridization) towards the donor group proton, and as a result these structures are not stable. Thus, for **c**′, MM3(94) optimization starting from the MP2/cc-pVDZ geometry rotates the C-4 and C-6 hydroxyl groups, which serve only as acceptors and not as donors, away from any steric interactions with the crowded ring.

The difference in predicted QM and MM ring puckering is quite a bit larger for **c**′ than for **a** (Cremer–Pople q values differ by 0.05 Å and θ values differ by 3°). We note here the physical effect that, because of the *trans*-diaxial interactions, the $^1C_4$ structures have reduced puckering amplitudes compared to the $^4C_1$ conformers at both levels of theory (from 0.02 Å at the MM level to 0.05 Å at the QM level). MM3(94) predicts considerably flatter rings for **c**′ and **d**′ than are found at the QM level. It should be noted, in regard to this latter point, that levels of correlation beyond MP2 destabilize the $^1C_4$ form relative to the $^4C_1$, and this effect was not possible to include in geometry optimization. Thus, some further flattening of the $^1C_4$ ab initio structures might be expected upon optimization with the higher levels of theory.

Crystal field effects.—A miniature crystal of β-D-glucose based on published [90] atomic coordinates, space group, and unit cell parameters, was constructed using the computer program CHEM-X as described previously [103]. The entire crystal was minimized with MM3(94)/$\epsilon = 3.0$. The central molecule in the optimized array still had the same discrepancies in bond lengths and valence angles observed in the optimized single-molecule structure; disappointingly, the geometry at the anomeric center became slightly worse. This confirms that these discrepancies arise from errors in the force field, not from crystal-packing effects. The puckering parameters, on the other hand, are brought into closer agreement with experiment: q is reduced by 0.01 Å and θ increases by 0.4°. In addition, the agreement between the model and the crystal structure is considerably improved with respect to the endocyclic torsion angles, where the mean absolute deviation drops from 3.7° to 2.2°. The central molecule in the optimized minicrystal, treated as an isolated molecule, has an MM3(94)/$\epsilon = 3.0$ energy that is 4.3 kcal higher than fully optimized **a**, to which it corresponds. This provides an estimate of the strain from crystal packing. Analysis of the force field terms shows that 2.1 kcal of this strain comes from torsion angle distortions, 1.1 kcal from dipole–dipole and hydrogen-bonding electrostatic interactions, 0.8 kcal from van der Waals interactions, and other terms in the force field all contribute less than 0.4 kcal.

Hydroxyl and hydroxymethyl orientations.—Although the hydroxyl and hydroxymethyl group orientations (as measured by their corresponding torsion angles) of the $^4C_1$ conformers remained reasonably constant at all theoretical levels, this did not prove to be the case for the $^1C_4$ conformers. Table 7 provides a listing of all group torsional angles for the $^1C_4$ conformers at several theoretical levels and the data for the $^4C_1$ conformers at two levels of theory.

Fig. 1 illustrates that the preferred orientation of the O-3 hydroxyl group in **d** minimizes unfavorable steric interactions from eclipsing one of the bonds of the six-membered ring, but does so at the expense of losing a hydrogen bond to the O-1 hydroxyl group. As discussed above, MM3 does not explicitly include any information about the stereochemical arrangement of the oxygen lone pairs. Therefore, the MM3 hydrogen bond strength(s) to the O-3 hydroxyl when serving as an acceptor are unaffected by the torsion angle. At ab initio levels of theory with small basis sets, on the other hand, Table 7 indicates that the O-3 hydroxyl rotates to permit formation of a hydrogen bond to O-1. It is likely that the favorable interactions associated with that hydrogen bond are probably overestimated at those levels. When larger, polarized basis sets are used, the O-3 hydroxyl returns to an orientation away from the ring. When correlation effects are added, however, the repulsive effects associated with eclipsing the C-2–C-3 ring bond are reduced, and the hydrogen bond to O-1 is present for all correlated levels. These results suggest that the two effects, hydrogen bonding and repulsive eclipsing interactions, counterbalance each other, and as a result the O-3 hydroxyl probably plays only a small role in stabilizing **d′**. Some of these same effects are observed to a lesser degree for the O-1, O-2, and O-4 hydroxyl groups.

Figs 1 and 2 and Table 7 illustrate that **c** has torsion angles that differ significantly from those of **c′** for all hydroxyl groups except HO-2. Again, the tendency at the MM3 level seems to be to minimize eclipsing of ring bonds. The ab initio QM structures at lower levels of theory are all reasonably similar to **c′**. Optimization of the MP2/cc-pVDZ structure with MM3(94) in this instance proved difficult since the MM3 potential energy surface is quite flat (see Section 6). In particular, the O-4 and O-6 torsional angles may change by $\pm 20°$ at little energetic cost.

Just as **c** and **d** failed to be stationary with respect to hydroxyl group torsional angles at quantum mechanical levels of theory, deviations of up to 66° in hydroxyl group torsion angles were observed for MM3(94) optimizations starting from **c′** and **d′**. Moreover, neither of the resulting conformations corresponded to **c″**, the deepest $^1C_4$ local minimum from the MM3(94) search of the hypersurface. MM3(94) structure **a″**, on the other hand, has hydroxyl group torsion angles within 10° of the best QM values. In energetic terms, structures derived from **c′** and **d′** dropped by more than 10 kcal in energy during the MM3(94) optimization, while structures derived from **a′** and **b′** each dropped only 4.5 kcal. In the $^1C_4$ cases, the improved energies were found primarily to result from changes in the torsional and hydrogen bonding terms, while in the $^4C_1$ cases changes were in the bond length and bond angle terms. Since the latter terms involve much larger force constants, the associated geometrical changes were indeed quite small.

As discussed above, the $^1C_4$ conformers showed more variation in ring puckering as a function of theoretical level than did the $^4C_1$ conformers. Table 8 lists the Cremer–Pople ring puckering parameters for {**a**} and {**c**} at several levels of theory. Although all {**a**}

Table 8
Cremer–Pople ring puckering parameters for the $^4C_1$ G^+ and $^1C_4$ G^+ conformers of β-D-glucose for geometries optimized at different theoretical levels [a]

Optimized geometry [b]	$^4C_1$ G^+ {a}		$^1C_4$ G^+ {c}		
	q	θ	q	θ	ϕ [c]
MM3(94)/$\epsilon = 1.5$	0.609	2	0.489	176	
AM1	0.576	0	0.439	173	325
PM3	0.574	0	0.473	164	18
HF/STO-3G	0.590	2	0.570	169	311
HF/3-21G	0.586	2	0.528	173	301
HF/6-31G*	0.568	1	0.505	174	290
HF/cc-pVDZ	0.565	2	0.506	174	292
MP2/3-21G	0.614	2	0.553	173	290
MP2/6-31G*	0.590	3	0.527	174	278
MP2/cc-pVDZ	0.589	3	0.537	173	284

[a] Amplitudes q in Å and values of θ and ϕ in degrees.
[b] Hydroxyl group torsional angles are found in Table 7.
[c] Not reported for values of θ below 5° or above 175° since ϕ values are not meaningful in this region.

structures are essentially perfect chairs, i.e., θ is close to 0°, the {c} structures are more distorted, i.e., θ is farther from 180°. The prediction of very small puckering amplitudes for the $^1C_4$ conformers at the AM1 and PM3 levels of theory is consistent with the well-known tendency [104] of semiempirical models to overemphasize intramolecular van der Waals repulsions.

Semiempirical and minimum-basis-set QM methods.—It remains a subject of active debate how appropriate semiempirical QM levels of theory are for modeling hydrogen bonding [105–109], non-bonded interactions [110,111], and five- and six-membered rings [112–114]. Although the data in Table 7 indicate that AM1 and PM3 optimized structures are intermediate in quality between MM3(92) and better ab initio QM levels of theory, Table 9 illustrates that the relative energies are less trustworthy. In particular, the semiempirical QM models AM1 and PM3 are of little value for predicting the relative energetics of the $^4C_1$ and $^1C_4$ chairs. Indeed, a cursory search of the PM3 hypersurface located a $^1C_4$ structure (see Fig. 5 and Table 7) that is 3.1 kcal lower in energy than the structure derived from **b**. This $^1C_4$ structure places O-6 in the G^- position with a hydrogen bond of only 1.36 Å between HO-6 and O-1. This structure, moreover, has an unusual arrangement of the O-2 and O-4 hydroxyl groups—their hydrogen atoms point almost directly at each other with an H–H non-bonded distance of 1.44 Å. This latter observation is probably a manifestation of the unphysical shape of the H–H core–core repulsion function known to exist in the PM3 Hamiltonian [111,115]. On the other hand, a previous PM3 survey [12] identified **b** as having the lowest energy of 81 $^4C_1$ structures, illustrating that the semiempirical models may be more successful at predicting energy differences associated with other kinds of conformational isomerism. (We emphasize in this context, moreover, that the semiempirical solvation models employed in this and earlier [12] studies do not depend on the ability of the gas-phase Hamiltonians to accurately predict solute energies. Rather the earlier solvation studies depend on the quality of the predicted electronic charge distribution, and for that

Table 9
Calculated relative energies (kcal/mol) of D-glucose conformers in the gas phase [a]

Level of theory	$^4C_1$ G^+ [b] {a}	$^4C_1$ T {b}	$^1C_4$ G^+ {c}	$^1C_4$ G^- {d}
Quantum mechanical Hamiltonians				
AM1	0.00	0.10	0.10	1.00
PM3	0.00	−1.50	−0.90	2.30
HF/STO-3G	0.00	−1.31	0.38	−1.60
HF/3-21G	0.00	−2.08	−6.55	−7.68
HF/6-31G*	0.00	−0.15	6.73	6.76
HF/cc-pVDZ	0.00	−0.09	6.23	6.34
HF/cc-pVTZ//HF/cc-pVDZ	0.00	0.35	9.78	9.87
HF/cc-pVDZ//MP2/cc-pVDZ	0.00	−0.04	7.94	11.23
HF/cc-pVTZ//MP2/cc-pVDZ	0.00	0.40	12.21	15.73
HF/cc-p^{T}VQZ//MP2/cc-pVDZ	0.00	0.48	13.07	16.78
MP2/3-21G	0.00	−2.56	−17.28	−17.66
MP2/6-31G*	0.00	−0.45	0.66	−0.60
MP2/cc-pVDZ	0.00	−0.50	−0.11	−0.81
MP2/cc-pVTZ//MP2/cc-pVDZ	0.00	0.07	4.13	3.65
CCSD/6-31G*//MP2/6-31G*	0.00	−0.34	2.07	1.68
Composite E [c]	0.00	0.27	6.41	6.99
ZPVE [d]	0.00	0.18	1.05	1.09
ΔH_{298}(rot-vib) [d]	0.00	0.08	0.64	0.69
$-T\Delta S_{298}$(rot-vib) [d]	0.00	0.23	1.19	1.12
ΔG_{298}(rot-vib) [d]	0.00	0.31	1.83	1.81
Composite G_{298} [e]	0.00	0.58	8.24	8.80
Molecular mechanics Hamiltonians				
MM3(94)/$\epsilon = 1.5$	0.00	0.21	11.19	12.23
ZPVE [f]	0.00	0.33	1.20	1.46
ΔH_{298}(rot-vib) [f]	0.00	0.17	0.76	0.57
$-T\Delta S_{298}$(rot-vib) [f]	0.00	0.38	1.15	2.04
ΔG_{298}(rot-vib) [f]	0.00	0.56	1.91	2.61
G_{298} [g]	0.00	0.77	13.10	14.84

[a] See text after Eq. (2) for $w/x//y/z$ notation.
[b] Absolute QM energies (E_h) for this column are −105.046 78; −97.719 54; −674.480 55; −679.552 50; −683.331 95; −683.406 95; −683.617 64; −683.399 05; −683.606 17; −683.650 29; −680.789 73; −685.176 03; −685.341 07; −686.050 46; −685.247 89; −686.166 44; 0.214 59; 0.224 75; −0.029 63; 0.195 12; −685.971 32.
[c] From Eq. (2).
[d] Zero-point vibrational energy (ZPVE) and thermal rotational–vibrational enthalpy and entropy terms calculated from unscaled HF/6-31G* frequencies.
[e] Composite E + ZPVE + thermal rotational–vibrational enthalpy and entropy terms, latter contributors again at HF/6-31G* level.
[f] Calculated using MM3(94)/$\epsilon = 1.5$ structures.
[g] MM3(94) steric energy + ZPVE + thermal rotational–vibrational enthalpy and entropy terms.

purpose the semiempirical models have been shown to be in better agreement with higher level calculations [116–120]. The present (SM4) solvation calculations rely even less on the semiempirical methods, using them for the solute reorganization energy and as a zero-order operand for a CM1 [75] type mapping.)

Table 9 shows that the minimum-basis-set HF/STO-3G level, like the AM1 and PM3 semiempirical levels, predicts the range in relative energies spanned by the four

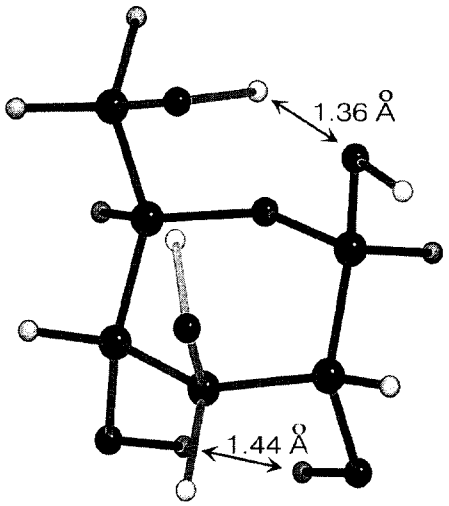


Fig. 5. A $^1C_4$ structure predicted to be lower in energy than any located $^4C_1$ structure at the PM3 level.

conformers to be fairly small. This level predicts a structure derived from **d** to be lower in energy than structures derived from either **a** or **b**.

Additional Hartree–Fock comparisons.—At the valence double-ζ HF/3-21G level, the energetic separation between the two chairs becomes more pronounced, but remarkably it is the $^1C_4$ chair that is predicted to be significantly lower in energy! Apparently, not only minimum-basis-set ab initio methods but even double-ζ calculations without polarization functions are no more reliable than the semiempirical ones.

Evidently the problems with semiempirical methods and ab initio calculations with unpolarized basis sets are primarily due to the lack of *d* functions on at least some of the heavy atoms (probably the oxygens), since HF calculations using the heavy-atom-polarized-valence-double-ζ 6-31G* and cc-pVDZ basis sets dramatically reverse the relative energies. This change may likely be traced to the tendency for small, unpolarized basis sets to overemphasize ionic aspects of hydrogen bonding [108]; H-bond distances are shorter for the $^1C_4$ conformers than for the $^4C_1$ conformers, and the former are thus preferentially (but erroneously) stabilized. At the HF level, the effect of placing *p* functions on hydrogen seems fairly small (cc-pVDZ has them while 6-31G* does not). However, expanding the valence space to doubly polarized, triple-ζ (cc-pVTZ) does increase the relative energy of the $^1C_4$ conformers by about 3.5 kcal. Further improvement of the valence space using the cc-p^{T}VQZ basis set provides yet another 1.0 kcal.

The relative energy of {**b**} is lower than that of {**a**} at all ab initio levels of theory unless basis sets of at least polarized-triple-ζ quality are used. The observation that the aqueous solution population of β-D-glucose does not include any *T* hydroxymethyl conformers [89] has often been assumed to derive from improved solvation of the G^+ and G^- conformers [6,9,12]. While that differential solvation may well exist, it is not required to overcome an intrinsic gas-phase preference for the *T* conformer, at least based on the minimum-energy structures considered here.

In addition, comparing lines 2 and 3 of Table 2 illustrates that more complete polarized basis sets tend to preferentially stabilize the $^4C_1$ chair over the $^1C_4$ chair. This is probably best understood in terms of hydrogen bonding. Table 1 shows that, on average, the hydrogen bonds in the $^4C_1$ chairs are longer and deviate considerably more from collinearity than do those in the $^1C_4$ chairs. Increasing the flexibility of the valence space and the number of polarization functions is thus more important for the hydrogen bonds in the $^4C_1$ chair than for those in the $^1C_4$ chair.

Finally, for the four isomers in Table 9, the agreement between the HF/6-31G* relative energies and those derived via Eq. (2) is good, and Tvaroska and Carver [34] have also recently obtained encouraging results for carbohydrate conformational analysis at this level. On the other hand, the 0.4 kcal error in the relative energies of {**a**} and {**b**} at this level is sufficient to reverse their stabilities.

Level of electron correlation treatment.—Correlation effects at the MP2 level stabilize the $^1C_4$ conformers by about 6 kcal relative to the $^4C_1$ conformers for the larger basis sets. Taking further account of correlation by carrying out CCSD calculations with the 6-31G* basis set removes 20–30% of this difference. Thus, not only is it important to use large basis sets, it is also important to go beyond second order in electron correlation.

MM energetics.—Comparing the MM3(94)/$\epsilon = 1.5$ to the ab initio-correlated composite results (Table 9) shows agreement in relative energies for the two $^4C_1$ chairs to be within 0.01 kcal. Although such close agreement is fortuitous, it is noteworthy again that **b**″ (not pictured, but very similar in structure to **b**′) is predicted to be higher in energy than **a**″. On the other hand, MM3(94)/$\epsilon = 1.5$ predicts the $^1C_4$ structures to be 4 to 6 kcal higher in relative energy than is found from the best QM results. The MM3(94)/$\epsilon = 1.5$ deepest local minimum for the $^1C_4$ chair, **c**″, is 10.3 kcal higher in free energy than **a**″. Although this is in better agreement with the best QM numbers, the overall structure is quite different so this is of questionable relevance.

Level dependence of vibrational free energies.—Zero-point and thermal contributions to the free energies at 298 K calculated at the HF/cc-pVDZ level for the two $^4C_1$ structures gave a net change of only a few hundredths of a kilocalorie compared with the HF/6-31G* level. Table 9 indicates that MM3(94)/$\epsilon = 1.5$ also predicts rovibrational free energies at 298 K that are of a similar magnitude to those calculated at the HF/6-31G* level.

6. Computational details

General.—The MM3(92) minimizations used a block diagonal Newton–Raphson routine and the default termination criterion of no more than 1.9 cal/mol energy change over the last five iterations. The MM3(94) minimizations used a full-matrix Newton–Raphson routine with a more stringent termination criterion of 0.2 cal/mol. As has been noted previously for MM3, manual rotation of a surface hydroxyl group in the miniature crystal was sometimes required in order for the minimization to terminate normally [121]. This reflects a flaw in the MM3 optimizer. Hydroxyl torsion angles fail to converge when the potential energy hypersurface is nearly flat along the torsional mode.

The standard state for all free energies is 298 K and 1 mol/L solute concentration both in the gas phase and in solution. At the QM level, rovibrational contributions to the relative free energy of each stereoisomer were calculated using standard statistical mechanical formulae and harmonic frequency calculations at the HF/6-31G* level [65]. The MM3(94) rovibrational free energies were calculated using the full-matrix optimization option. Translational free energies were not considered because they do not distinguish between stereoisomers.

Software.—Stereoisomer geometries for each chair were generated using the program CHEM-X [122]. MM3(92) and MM3(94) force-field calculations used the MM3 program [40,123]. Ab initio calculations were carried out using the Gaussian 92/DFT suite of programs [124]. Semiempirical calculations used the AMSOL [125] program with local modifications to include the SM4 solvation models.

AM1-SM4-SRP model for sugars.—The notation SM4 is used generically for any solvation model based on CM1 charges and bond-order-dependent CDS parameters [47]. In a general sense we reserve the unadorned name SM4 for a general parameterization, i.e., one based on a broad range of solutes, and we use the notation SM4-SRP for specific-range-parameters, i.e., parameters based on a limited (specific) range of solutes. In the present work we use the general SM4 model [48] for *n*-hexadecane, but since such a general parameterization is unavailable for water, we developed an SM4-SRP model specifically for this paper. (For brevity though, having stated this, we drop the SRP in the bulk of the text.) In this section we present the details of the AM1-SM4-SRP parameterization for sugars in water.

The SM4 models for the two solvents calculate $\Delta E_{\text{internal}}$ and ΔG_{P} in the same manner, but using different optimized parameters for the latter and different dielectric constants of 78.3 (water) and 2.06 (*n*-hexadecane), and they differ in the way in which G^0_{CDS} is treated.

In the SM4–hexadecane model, every atom is assigned two surface tensions that depend on its atomic number and bond orders [126] to other atoms. Each atom contributes to G^0_{CDS} based on two calculated solvent-accessible surface areas (SASA) [127]. One surface tension accounts for short-range interactions, and the corresponding SASA is calculated using an effective radius of 2.0 Å for the solvent. The other surface tension accounts for longer range interactions and the corresponding SASA is calculated using an effective radius of 4.9 Å for the solvent. The SM4–water model, on the other hand, has only a single atom-specific surface tension, and SASA is calculated using a radius of 1.4 Å for the solvent (i.e. a single surface tension accounts for both short-range and long-range effects since the water molecule is small).

For the SM4–water solvation model, the C and O parameters and the radius parameters for H were taken from a previous (less general) SM4-SRP parameterization based on hydrocarbons, ethers, and aldehydes [47]. Since glucose has H atoms bonded to both C and O, we require surface tension parameters, called $\sigma^{\text{CDS}}_{\text{CH}}$ and $\sigma^{\text{CDS}}_{\text{OH}}$, respectively [48], for these two situations. The value of $\sigma^{\text{CDS}}_{\text{CH}}$ was taken as 9.09 cal/mol/Å^2 which, reproduces the previous SM4-SRP parameterization described above [47] (which differed somewhat in form since hydrogenic surface tension did not depend on the bond order of hydrogen to heavy atoms). The value of $\sigma^{\text{CDS}}_{\text{OH}}$ was taken as −188 cal/mol/Å^2, which, with all other parameters held fixed, minimizes the

root-mean-square (rms) error in the calculated free energies of solvation for methanol, ethanol, 1-propanol, 2-propanol, 2-methyl-2-propanol, 2-methoxyethanol, phenol, and 1,2-ethanediol relative to experiment (with a resulting rms error of 0.88 kcal/mol). All experimental values are from Cabani et al. [128], with the exception of that for 1,2-ethanediol, which is taken from Suleiman and Eckert [129].

As stated above, the main motivation for developing this SRP model is that a general solvation model for water based on CM1 charges is not completed yet. Nevertheless, the present SRP model may be of use for sugar modeling, even after the more general model is done in that, being based on a more limited solute set, it may actually be empirically more accurate for that kind of solute than will be the general set, which will necessarily involve compromises to fit a broader range of solutes.

7. Conclusions

Convergence at the ab initio level of relative energies of chair configurations to within 1 kcal, or relative energies of the hydroxymethyl conformers within each chair to within 0.1 kcal, requires basis sets of at least triple-ζ quality and an accounting for correlation at levels beyond second-order many-body perturbation theory. For both high-level ab initio QM calculations and MM3(94)/$\epsilon = 1.5$ molecular mechanics calculations, the optimized gas-phase geometries of β-D-glucose in the $^4C_1$ G^+ conformation agree well with X-ray crystal structural data with respect to bond lengths, valence angles, and torsion angles. Errors of up to 0.03 Å in bond lengths at the anomeric center are noted for MM3(94), and both the QM and MM levels underestimate the ring puckering parameter θ. For other β-D-glucose conformers for which crystal data are not available, agreement between these same levels of theory remains good in general, but hydroxyl group torsional angles can vary widely depending on the theoretical level at which they are optimized. The forms of the MM3 potential functions for hydrogen bonding permit orientations of hydroxyl groups serving as hydrogen-bond acceptors that differ from those predicted at the best QM level, especially in the $^1C_4$ structures. The hydrogen bonding is more favorable in the $^1C_4$ conformers than in the $^4C_1$ conformers, but that is insufficient to overcome the 1,3-*trans* diaxial steric repulsions. Thus, the relative gas-phase free energies of the $^1C_4$ structures remain more than 8 kcal above the $^4C_1$ structures, which is consistent with the $^1C_4$ chair never having been observed experimentally.

At our best level of theory, as well as at the MM3(94)/$\epsilon = 1.5$ level, the $^4C_1$ G^+ hydroxymethyl conformer **{a}** is lower in potential energy than the $^4C_1$ T conformer **{b}**.

A systematic examination of a series of calculations in which one or more of the factors contributing to these energy differences are treated less completely provides new insight into the physical effects controlling these conformational energy differences. Especially interesting is that we find the interplay between hydrogen bonding and van der Waals repulsion to be finely balanced in the gas-phase $^1C_4$ conformers. When the theoretical level overemphasizes ionic contributions to hydrogen bonding, the $^1C_4$ conformers are inaccurately predicted to be more stable than the $^4C_1$.

Hartree–Fock level calculations tend to overestimate nonbonded repulsion as they are incapable of accounting for the (attractive) intramolecular London dispersive interac-

tions that are expected to be more important for the sterically congested $^1C_4$ conformers. Including electron correlation effects ameliorates this deficiency. Correlation effects evaluated at the MP2 level stabilize the $^1C_4$ conformers by about 5–7 kcal relative to the $^4C_1$ structures.

In addition to potential energy differences of 6 kcal, rovibrational free-energy contributions destabilize the $^1C_4$ conformers relative to the $^4C_1$ by an additional 2 kcal. The shorter hydrogen bonds in the $^1C_4$ structures cause the corresponding vibrational modes to be stiffer in these conformers than in the $^4C_1$ cases. As a result both the relative zero-point vibrational energies and the relative vibrational free energies of the former isomers are raised by about 1 kcal each. A similar argument applies to the relative free energy of the $^4C_1$ *T* conformer {**b**}, with its nearly ideal O-6 H $\cdots$ O-4 hydrogen bond, compared to $^4C_1$ G^+ {**a**}.

The SM4 series of semiempirical generalized Born/surface tension models predicts solvation in both water and *n*-hexadecane to further destabilize the $^1C_4$ conformers relative to the $^4C_1$. In solution, polarization favors the $^4C_1$ conformers. The differential polarization is about four times larger in water than in *n*-hexadecane. In water, favorable hydrogen-bonding interactions with the first solvation shell are also larger for the $^4C_1$ chairs, so that overall solvation favors them over the $^1C_4$ chairs by about 5 to 9 kcal. In *n*-hexadecane, the larger solvent accessible surface areas of the $^4C_1$ conformers provides an additional increase in favorable dispersion interactions, so that overall the $^4C_1$ structures are better solvated by about 2–4 kcal.

We will report presently on an investigation of the anomeric equilibrium for D-glucopyranose carried out at levels of theory similar to those used here.

Acknowledgements

The authors are grateful to David Giesen, Gregory Hawkins, and Joey Storer for contributions to the locally modified version of AMSOL used for these calculations. We thank Professor S.J. Angyal for a summary of his research in progress and Professor N.L. Allinger for a preprint of ref. [55] and a pre-release version of the MM3(94) program that included the MM3(94) force field. This work was supported in part by the University of Minnesota Supercomputer Institute and the National Science Foundation.

References

[1] J. Lechner and F. Wieland, *Annu. Rev. Biochem.*, 58 (1989) 173–194.
[2] J.V. Hunt, M.A. Bottoms, and M.J. Mitchinson, *Biochem. J.*, 291 (1993) 529–535.
[3] R.A. Dwek, C.J. Edge, D.J. Harvey, and M.R. Wormald, *Annu. Rev. Biochem.*, 62 (1993) 65–100.
[4] S.J. Angyal, *Aust. J. Chem.*, 21 (1968) 2737–2746.
[5] S.N. Ha, A. Giammona, M. Field, and J.W. Brady, *Carbohydr. Res.*, 180 (1988) 207–221.
[6] J.W. Brady, *J. Am. Chem. Soc.*, 111 (1989) 5155–5165.
[7] L.M.J. Kroon-Batenburg and J. Kroon, *Biopolymers*, 29 (1990) 1243–1248.
[8] B.P. van Eijck, L.M.J. Kroon-Batenburg, and J. Kroon, *J. Mol. Struct.*, 237 (1990) 315–325.
[9] S. Ha, J. Gao, B. Tidor, J.W. Brady, and M. Karplus, *J. Am. Chem. Soc.*, 113 (1991) 1553–1557.
[10] R.G. Zhbankov, *J. Mol. Struct.*, 275 (1992) 645–684.
[11] P.L. Polavarapu and C.S. Ewig, *J. Comput. Chem.*, 13 (1992) 1255–1261.
[12] C.J. Cramer and D.G. Truhlar, *J. Am. Chem. Soc.*, 115 (1993) 5745–5753.

[13] B.P. van Eijck, R.W.W. Hooft, and J. Kroon, *J. Phys. Chem.*, 97 (1993) 12093–12099.
[14] T.M. Glennon, Y.J. Zheng, S.M. Le Grand, B.A. Shutzberg, and K.M. Merz, Jr, *J. Comput. Chem.*, 15 (1994) 1019–1040.
[15] M.K. Dowd, A.D. French, and P.J. Reilly, *Carbohydr. Res.*, 264 (1994) 1–19.
[16] S.J. Angyal, *Angew. Chem., Int. Ed. Engl.*, 8 (1969) 157–166.
[17] C.G. Pearson and O. Rumquist, *J. Org. Chem.*, 33 (1968) 2572–2574.
[18] G.A. Jeffrey, J.A. Pople, and L. Radom, *Carbohydr. Res.*, 25 (1972) 117–131.
[19] G.A. Jeffrey, J.A. Pople, and L. Radom, *Carbohydr. Res.*, 38 (1974) 81–95.
[20] S. Wolfe, M.-H. Whangbo, and D.J. Mitchell, *Carbohydr. Res.*, 69 (1979) 1–26.
[21] G.A. Jeffrey and J.H. Yates, *Carbohydr. Res.*, 96 (1981) 205–213.
[22] A.J. Kirby, *The Anomeric Effect and Related Stereoelectronic Effects at Oxygen*, Springer-Verlag, Berlin, 1983.
[23] R.W. Franck, *Tetrahedron*, 39 (1983) 3251–3252.
[24] C.W. Andrews, J.P. Bowen, and B. Fraser-Reid, *J. Chem. Soc., Chem. Commun.*, (1989) 1913–1916.
[25] M.C. Krol, C.J.M. Huige, and C. Altona, *J. Comput. Chem.*, 11 (1990) 765–790.
[26] C.W. Andrews, B. Fraser-Reid, and J.P. Bowen, *J. Am. Chem. Soc.*, 113 (1991) 8293–8298.
[27] O. Kysel and P. Mach, *J. Mol. Struct. (Theochem)*, 227 (1991) 285–293.
[28] R. Montagnani and J. Tomasi, *Int. J. Quant. Chem.*, 39 (1991) 851–870.
[29] C.J. Cramer, *J. Org. Chem.*, 57 (1992) 7034–7043.
[30] E. Juaristi and G. Cuevas, *Tetrahedron*, 48 (1992) 5019–5087.
[31] C.L. Perrin, K.B. Armstrong, and M.A. Fabian, *J. Am. Chem. Soc.*, 116 (1994) 715–722.
[32] K.B. Wiberg and M. Marquez, *J. Am. Chem. Soc.*, 116 (1994) 2197–2198.
[33] W.L. Jorgensen, P.I.M. Detirado, and D.L. Severance, *J. Am. Chem. Soc.*, 116 (1994) 2199–2200.
[34] I. Tvaroska and J.P. Carver, *J. Phys. Chem.*, 98 (1994) 9477–9485.
[35] J.-P. Praly and R.U. Lemieux, *Can. J. Chem.*, 65 (1987) 213–223.
[36] E.G. Cox and G.A. Jeffrey, *Nature*, 143 (1939) 894–895.
[37] R.E. Reeves, *J. Am. Chem. Soc.*, 72 (1950) 1499–1506.
[38] E.L. Eliel and S.H. Wilen, *Stereochemistry of Organic Compounds*, John Wiley and Sons, New York, 1994, pp 686–720.
[39] J.A. Hirsch, *Top. Stereochem.*, 1 (1967) 199–222.
[40] N.L. Allinger, X.F. Zhou, and J. Bergsma, *J. Mol. Struct. (Theochem)*, 118 (1994) 69–83.
[41] N.L. Allinger, Y.H. Yuh, and J.H. Lii, *J. Am. Chem. Soc.*, 111 (1989) 8551–8566.
[42] J.H. Lii and N.L. Allinger, *J. Am. Chem. Soc.*, 111 (1989) 8566–8575.
[43] N.L. Allinger, M. Rahman, and J.H. Lii, *J. Am. Chem. Soc.*, 112 (1990) 8293–8307.
[44] A.D. French, L. Schäfer, and S.Q. Newton, *Carbohydr. Res.*, 239 (1993) 51–60.
[45] C.J. Cramer, G.D. Hawkins, and D.G. Truhlar, *J. Chem. Soc., Faraday Trans.*, 90 (1994) 1802–1804.
[46] Y.-J. Zheng, S.M. Le Grand, and K.M. Merz, Jr, *J. Comput. Chem.*, 13 (1992) 772–791.
[47] J.W. Storer, D.J. Giesen, G.D. Hawkins, G.C. Lynch, C.J. Cramer, D.G. Truhlar, and D.A. Liotard, in C.J. Cramer and D.G. Truhlar (Eds), *Structure and Reactivity in Aqueous Solution*, ACS Symposium Series 568, American Chemical Society, Washington, DC, 1994, pp 24–49.
[48] D.J. Giesen, J.W. Storer, C.J. Cramer, and D.G. Truhlar, *J. Am. Chem. Soc.*, 117 (1995) 1057–1068.
[49] D.J. Giesen, C.J. Cramer, and D.G. Truhlar, *J. Phys. Chem.*, 99 (1995) 7137–7146.
[50] A. Finkelstein, *J. Gen. Physiol.*, 68 (1976) 127–135.
[51] N.P. Franks and W.R. Lieb, *Nature*, 274 (1978) 339–342.
[52] M.H. Abraham, G.S. Whiting, R. Fuchs, and E.J. Chambers, *J. Chem. Soc., Perkin Trans. 2*, (1990) 291–300.
[53] R.M. Venable, Y. Zhang, B.J. Hardy, and R.W. Pastor, *Science*, 262 (1993) 223–226.
[54] N.L. Allinger and L. Yan, *J. Am. Chem. Soc.*, 115 (1993) 11918–11925.
[55] I.J.-H. Lii and N.L. Allinger, *J. Phys. Org. Chem.*, 7 (1994) 591–609.
[56] S.F. Boys and F. Bernardi, *Mol. Phys.*, 19 (1970) 553–566.
[57] D.W. Schwenke and D.G. Truhlar, *J. Chem. Phys.*, 82 (1985) 2418–2426; 86 (1987) 3760 (E).
[58] S.M. Cybulski and G. Chalasinski, *Chem. Phys. Lett.*, 197 (1992) 591–598.
[59] E.R. Davidson and S.J. Chakravorty, *Chem. Phys. Lett.*, 217 (1994) 48–54.
[60] N.L. Allinger, M. Rahman, and J.-H. Lii, *J. Am. Chem. Soc.*, 112 (1990) 8293–8307.

[61] A.D. French and M.K. Dowd, *J. Comput. Chem.*, 15 (1994) 561–570.
[62] C. Van Alsenoy, A.D. French, M. Cao, S.Q. Newton, and L. Schäfer, *J. Am. Chem. Soc.*, 116 (1994) 9590–9595.
[63] M.J.S. Dewar, E.G. Zoebisch, E.F. Healy, and J.J.P. Stewart, *J. Am. Chem. Soc.*, 107 (1985) 3902–3909.
[64] J.J.P. Stewart, *J. Comput. Chem.*, 10 (1989) 221–264.
[65] W.J. Hehre, L. Radom, P.v.R. Schleyer, and J.A. Pople, *Ab Initio Molecular Orbital Theory*, Wiley, New York, 1986.
[66] W.J. Hehre, R.F. Stewart, and J.A. Pople, *J. Chem. Phys.*, 51 (1969) 2657–2664.
[67] J.S. Binkley, J.A. Pople, and W.J. Hehre, *J. Am. Chem. Soc.*, 102 (1980) 939–947.
[68] R. Ditchfield, W.J. Hehre, and J.A. Pople, *J. Chem. Phys.*, 54 (1971) 724–728.
[69] W.J. Hehre, R. Ditchfield, and J.A. Pople, *J. Chem. Phys.*, 56 (1972) 2257–2261.
[70] P.C. Hariharan and J.A. Pople, *Theor. Chim. Acta*, 28 (1973) 213–222.
[71] T.H. Dunning, *J. Chem. Phys.*, 90 (1989) 1007–1023.
[72] J. Cizek, *Adv. Chem. Phys.*, 14 (1969) 35–89.
[73] G.D. Purvis and R.J. Bartlett, *J. Chem. Phys.*, 76 (1982) 1910–1918.
[74] C.J. Cramer and D.G. Truhlar, *J. Am. Chem. Soc.*, 113 (1991) 8305–8311; 9901.
[75] J.W. Storer, D.J. Giesen, C.J. Cramer, and D.G. Truhlar, *J. Comput. Aid. Mol. Des.*, 9 (1995) 87–110.
[76] C.J. Cramer and D.G. Truhlar, *J. Am. Chem. Soc.*, 116 (1994) 3892–3900.
[77] C.J. Cramer and D.G. Truhlar, in K.B. Lipkowitz and D.B. Boyd (Eds), *Reviews in Computational Chemistry*, Vol. 6, VCH, New York, 1995, pp 1–72.
[78] C.J. Cramer and D.G. Truhlar, in P. Politzer and J.S. Murray (Eds), *Quantitative Treatments of Solute / Solvent Interactions, Theoretical and Computational Chemistry*, Vol. 1, Elsevier, Amsterdam, 1994, pp 9–54.
[79] B. Lee and F.M. Richards, *J. Mol. Biol.*, 55 (1971) 379–400.
[80] R.B. Hermann, *J. Phys. Chem.*, 76 (1972) 2754–2759.
[81] C.J. Cramer and D.G. Truhlar, *Chem. Phys. Lett.*, 198 (1992) 74–80; 202 (1993) 1567(E).
[82] P.R. Muddasani, E. Bozo, B. Bernet, and A. Vasella, *Helv. Chim. Acta*, 77 (1994) 257–290.
[83] J.C. Christofides, D.B. Davies, J.A. Martin, and E.B. Rathbone, *J. Am. Chem. Soc.*, 108 (1986) 5738–5743.
[84] S.J. Angyal, personal communication.
[85] B.R. Leefland, J.F.G. Vliegenthart, L.M.J. Kroon-Batenburg, B.P. van Eijck, and J. Kroon, *Carbohydr. Res.*, 230 (1992) 41–61.
[86] L. Poppe and H. Van Halbeek, *J. Am. Chem. Soc.*, 114 (1992) 1092–1094.
[87] B. Adams and L. Lerner, *J. Am. Chem. Soc.*, 114 (1992) 4827–4829.
[88] S.A. Stortz and A.S. Cerezo, *J. Carb. Chem.*, 13 (1994) 235–247.
[89] Y. Nishida, H. Ohrui, and H. Meguro, *Tetrahedron Lett.*, 25 (1984) 1575–1578.
[90] S.S.C. Chu and G.A. Jeffrey, *Acta Crystallogr.*, B24 (1968) 830–838.
[91] G.R.J. Thatcher, Ed., *The Anomeric Effect and Associated Stereoelectronic Effects*, ACS Symposium Series 539, American Chemical Society, Washington DC, 1993.
[92] P.P. Graczyk and M. Mikolajczyk, in E.L. Eliel and S.H. Wilen (Eds), *Topics in Stereochemistry*, Vol. 21, John Wiley and Sons, New York, 1994, pp 159–349.
[93] A. Abe, K. Inomata, E. Tanisawa, and I. Ando, *J. Mol. Struct.*, 238 (1990) 315–323.
[94] U. Salzner and P.v.R. Schleyer, *J. Am. Chem. Soc.*, 115 (1993) 10231–10236.
[95] I. Tvaroska and J.P. Carver, *J. Phys. Chem.*, 98 (1994) 6452–6458.
[96] A.D. French, R.S. Rowland, and N.L. Allinger, in A.D. French and J.W. Brady (Eds), *Computer Modeling of Carbohydrate Molecules*, ACS Symposium Series 430, American Chemical Society, Washington, DC, 1990 pp 120–140.
[97] D. Cremer and J.A. Pople, *J. Am. Chem. Soc.*, 97 (1975) 1354–1358.
[98] G.A. Jeffrey and J.H. Yates, *Carbohydr. Res.*, 74 (1979) 319–322.
[99] J.T. Ham and D.G. Williams, *Acta Crystallogr.*, B26 (1970) 1373–1383.
[100] K. Hirotsu and A. Shimada, *Bull. Chem. Soc. Jpn*, 47 (1974) 1872–1879.
[101] H. Takeda, N. Yasuoda, and N. Kasai, *Carbohydr. Res.*, 53 (1977) 137–152.
[102] C.C. Rohrer, A. Sarko, T.L. Bluhm, and Y.N. Lee, *Acta Crystallogr.*, B36 (1980) 650–654.

[103] A.D. French and D.P. Miller, in D.A. Smith (Ed), *Modeling the Hydrogen Bond*, ACS Symposium Series 569, American Chemical Society, Washington DC, 1994, pp 235–251.
[104] J.J.P. Stewart, in K.B. Lipkowitz and D.B. Boyd (Eds), *Reviews in Computational Chemistry*, Vol. 1, VCH, New York, 1989, pp 45–81.
[105] M. Khalil, R.J. Woods, D.F. Weaver, and V.H. Smith, *J. Comput. Chem.*, 12 (1991) 584–593.
[106] A. Gonzalez-Lafont, T.N. Truong, and D.G. Truhlar, *J. Phys. Chem.*, 95 (1991) 4618–4627.
[107] M.A. Ríos and J. Rodríguez, *J. Comput. Chem.*, 13 (1992) 860–866.
[108] Y.-J. Zheng and K.M. Merz, Jr, *J. Comput. Chem.*, 13 (1992) 1151–1169.
[109] M.W. Jurema and G.C. Shields, *J. Comput. Chem.*, 14 (1993) 89–104.
[110] W. Thiel, *Tetrahedron*, 44 (1988) 7393–7408.
[111] G.I. Csonka, *J. Comput. Chem.*, 14 (1993) 895–898.
[112] I. Tvaroska and J.P. Carver, *J. Chem. Res. (S)*, (1991) 6–7.
[113] K. Gundertofte, J. Palm, I. Petterseon, and A. Stamvik, *J. Comput. Chem.*, 12 (1991) 200–208.
[114] D.M. Ferguson, I.R. Gould, W.A. Glauser, S. Schroeder, and P.A. Kollman, *J. Comput. Chem.*, 13 (1992) 525–532.
[115] V. Buss, J. Messinger, and N. Heuser, *QCPE Bull.*, 11 (1991) 5–6.
[116] M.N. Ramos and B.d.B. Neto, *Chem. Phys. Lett.*, 199 (1992) 482–486.
[117] C. Alemán, F.J. Luque, and M. Orozco, *J. Comput. Chem.*, 14 (1993) 799–808.
[118] G. Rauhut and T. Clark, *J. Comput. Chem.*, 14 (1993) 503–509.
[119] G.P. Ford and B. Wang, *J. Comput. Chem.*, 14 (1993) 1101–1111.
[120] B. Wang and G.P. Ford, *J. Comput. Chem.*, 15 (1994) 200–207.
[121] A.D. French, D.P. Miller, and A. Aabloo, *Int. J. Biol. Macromol.*, 15 (1993) 30–36.
[122] CHEM-X, Chemical Design, Ltd., United Kingdom, 1994.
[123] MM3 Program, Technical Utilization Corporation, Inc., 235 Glen Village Court, Powell, OH 43065 (all users); Tripos Associates, 1699 South Hanley Rd., St. Louis, MO 63144 (commercial users); and QCPE, Indiana University, Bloomington, IN 47405 (academic users).
[124] M.J. Frisch, G.W. Trucks, H.B. Schlegel, P.M.W. Gill, B.G. Johnson, M.W. Wong, J.B. Foresman, M.A. Robb, M. Head-Gordon, E.S. Replogle, R. Gomperts, J.L. Andres, K. Raghavachari, J.S. Binkley, C. Gonzalez, R.L. Martin, D.J. Fox, D.J. Defrees, J. Baker, J.J.P. Stewart, and J.A. Pople, *Gaussian 92/DFT, Revision G.1*, Gaussian, Inc., Pittsburgh, PA, 1993.
[125] C.J. Cramer, G.D. Hawkins, G.C. Lynch, D.J. Giesen, D.G. Truhlar, and D.A. Liotard, *QCPE Bull.*, 14 (1994) 55–57.
[126] D.R. Armstrong, P.G. Perkins, and J.J.P. Stewart, *J. Chem. Soc., Dalton Trans.*, (1973) 838–840.
[127] F.M. Richards, *Annu. Rev. Biophys. Bioeng.*, 6 (1977) 151–176.
[128] S. Cabani, P. Gianni, V. Mollica, and L. Lepori, *J. Solution Chem.*, 10 (1981) 563–595.
[129] D. Suleiman and C.A. Eckert, *J. Chem. Eng. Data*, 39 (1994) 692–696.