ELSEVIER

www.elsevier.nl/locate/carres

Carbohydrate Research 326 (2000) 305–322

CARBOHYDRATE RESEARCH

A QM/MM analysis of the conformations of crystalline sucrose moieties

Alfred D. French [a,*1], Anne-Marie Kelterer [b,*2], Christopher J. Cramer [c], Glenn P. Johnson [a], Michael K. Dowd [a]

[a] *Southern Regional Research Center, Agricultural Research Service, US Department of Agriculture, 1100 Robert E. Lee Boulevard, PO Box 19687, New Orleans, LA 70179-0687, USA*

[b] *Institut für Physikalische und Theoretische Chemie, Technische Universität Graz, Technikerstraße 4, A-8010 Graz, Austria*

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

Received 30 December 1999; accepted 4 February 2000

Abstract

Both ab initio quantum mechanics (QM) and molecular mechanics (MM) were used to produce a hybrid energy surface for sucrose that simultaneously provides low energies for conformations that are observed in crystal structures and high energies for most unobserved structures. HF/6-31G* QM energies were calculated for an analogue based on tetrahydropyran (THP) and tetrahydrofuran (THF). Remaining contributions to the potential energy of sucrose were calculated with MM. To do this, the MM surface for the analogue was subtracted from the MM surface for the disaccharide, and the QM surface for the analogue was added. Prediction of the distribution of observable geometries was enhanced by reducing the strength of the hydrogen bonding. Reduced hydrogen-bonding strength is probably useful because many crystalline sucrose moieties do not have intramolecular hydrogen bonds between the fructose and glucose residues. Therefore, hydrogen bonding does not play a large role in determining the molecular conformation. On the hybrid energy surface that was constructed with a dielectric constant of 3.5, the average potential energy of 23 sucrose moieties from crystal structures is 1.16 kcal/mol, and the population of observed structures drops off exponentially as the energy increases.

Keywords: Carbohydrates; exo-Anomeric effect; Sugar; HF/6-31G*; MM3; Modeling; Non-additivity; 2-Methoxytetrahydropyran; 2-Methoxytetrahydrofuran; Force field

1. Introduction

The chemical and physical properties of the disaccharide sucrose are of substantial interest because of its biological and commercial roles. Crops that produce table sugar are agricultural mainstays in many nations. Sucrose is a biosynthetic feedstock for cellulose, the polysaccharide that is the most prevalent biomolecule on the planet as well as the major component of cotton fiber. Sucrose esters are used as no-calorie fat substitutes, and sucrose could become an important industrial raw material [1] for many non-food products such as adhesives [2] or even medicines [3].

Sucrose conformations have been studied extensively because its molecular shapes are

[1] *Corresponding author. Tel.: +1-504-2864410; fax: +1-504-2864217; e-mail: afrench@nola.srrc.usda.gov

[2] *Corresponding author. Tel.: +43-316-8738223; fax: +43-316-8738225; e-mail: kelterer@ptc.tu-graz.ac.at

0008-6215/00/$ - see front matter
PII: S0008-6215(00)00051-3

thought to help determine its properties. Besides, even the relationship between its sweetness and its structure is not completely understood. An understanding of its conformational behavior is necessary for many sorts of investigations ranging from understanding interactions with proteins to computer-aided development of new materials. The basic conformation is usually expressed by the values of the torsion angles ϕ and ψ (Fig. 1) that describe the rotational orientations of the monosaccharide residues about their C-1–O-1 and O-1–C-2′ bonds to the glycosidic oxygen atom. Values of ψ for crystalline sucrose moieties range over almost 180°, while ϕ takes values over a more limited span of 30°. Thus, some flexibility exists.

Computer models have been used extensively to study sucrose. Ramachandran plots [4] depict how the calculated energy changes when ϕ and ψ are varied. The approximate relative populations of the various conformations in equilibria can be inferred from these energy surfaces, as can the heights of barriers between them. Such surfaces can be used to validate the method of computing the energy by comparison with observed structures. Conversely, energy surfaces indicate the extents of distortion for well-determined experimental geometries in particular environments. They are also important for the interpretation of less definitive experiments, and they indicate the likelihood of proposed structures.

Energies for disaccharide surfaces can be calculated in several ways, but at present, they are usually based on empirical (or molecular mechanics, MM) force fields. Some time ago, we reported [5–7] that the MM program MM3 [8] does not account very well for all of the structures of sucrose moieties that are observed with diffraction crystallography despite its apparent success with other disaccharides. On the deficient MM3 surface and those made with other force fields [9,10], some observed sucrose conformations corresponded to high energies. Also, parts of the low-energy region were not populated by observed structures. We refer to this situation as 'the sucrose prob-

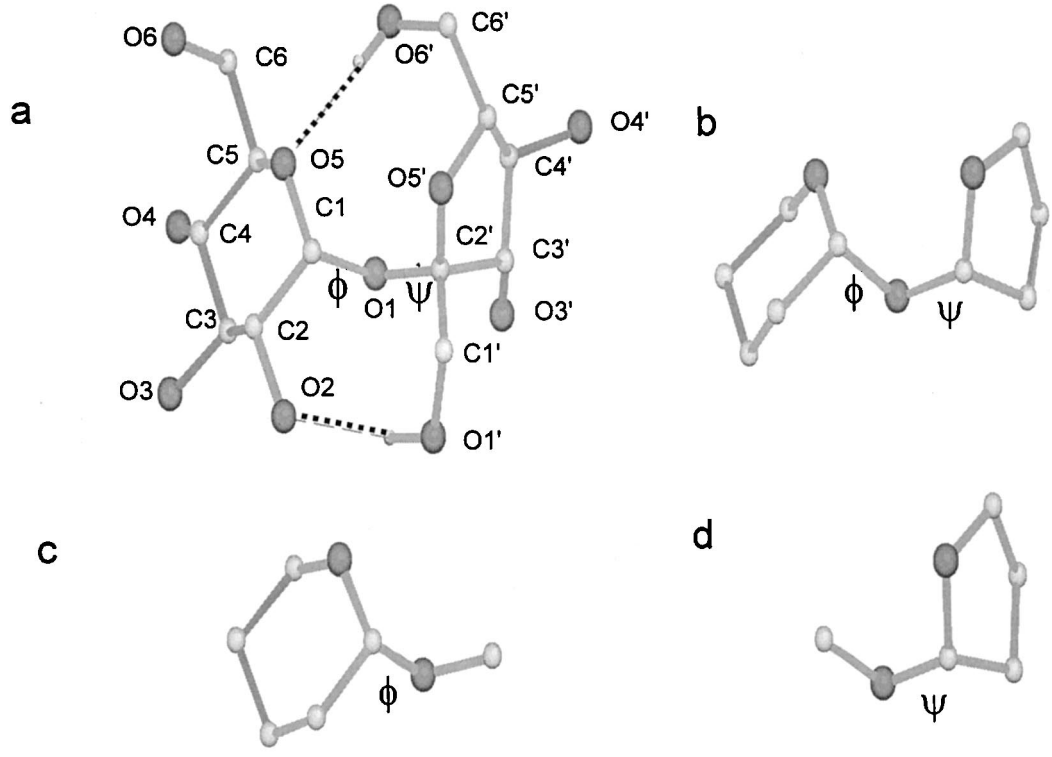


Fig. 1. Compounds used in this work. (a) Sucrose with most hydrogen atoms removed. (b) 1,2′-O-linked THP–O–THF analogue of sucrose (c) Axial 2-Me–O–THP. (d) 2-Me–O–THF. ϕ for sucrose is based on the O-5–C-1–O-1–C-2′ sequence and ψ is based on O-5′–C-2′–O-1–C-1. The torsion angles for the other species are defined analogously. Only hydrogen atoms involved in hydrogen bonds are shown.

lem'. Subsequently, energy surfaces were proposed for sucrose that were based on molecular dynamics simulations with explicit water [11,12], or on a model having united hydroxyl groups [13]. Those proposals were not entirely satisfying, and the purpose of the present paper is to describe a new and hopefully more convincing, potential-energy surface for sucrose. We then undertake a detailed analysis of the agreement of this surface with a variety of linkage conformations that are observed in crystal structures. In the process we introduce a statistical comparison based on a Boltzmann-like, exponential decay distribution of the conformational energies that appears to be useful in testing the validity (or at least the utility) of proposed potential-energy functions.

We had earlier attributed the problems with the MM3 surface for sucrose to failure to account for an 'overlapping exo-anomeric effect'. The normal exo-anomeric effect, as found in reducing disaccharides such as cellobiose [14–16] is reasonably well understood [17,18]. It leads to gauche preferences for O–C–O–C sequences because of delocalization of the lone-pair electrons on the glycosidic oxygen atom. This effect has been mimicked in MM force fields [8,19–23]. In the non-reducing sucrose molecule, however, two anomeric centers are linked to the same oxygen atom, and the energy may not be correctly calculated with current MM potential functions [24].

We have worked around this apparent problem in the present paper by using ab initio quantum mechanics (QM) for the critical part of the molecule and MM for the rest of the structure. In this sense, our method is a QM/MM hybrid [25], but it is simpler than the methods usually described by that phrase. In our method, the critical part is itself a complete molecule. It is an analogue of sucrose (Fig. 1(b)), based on the tetrahydropyran (THP) and tetrahydrofuran (THF) rings of the sucrose backbone and the linkage oxygen atom. We propose that this choice of the critical part provides the necessary atoms to allow for a reasonable approximation of the varying electronic structures that contribute to the torsional energies as the values of ϕ and ψ are changed.

The method requires three separate sets (or maps) of ϕ, ψ and energy values: $E_{\phi,\psi}$ for the disaccharide as computed by MM, $E_{\phi,\psi}$ for the analogue, computed by MM, and $E_{\phi,\psi}$ for the analogue computed by QM. At each ϕ, ψ point, the hybrid energy for sucrose is then defined by:

$$E_{\phi,\psi}(\text{hybrid}) = E_{\phi,\psi}(\text{QM, analogue}) + E_{\phi,\psi}(\text{MM, whole molecule}) - E_{\phi,\psi}(\text{MM, analogue})$$

This expression depicts the hybrid energies as being the result of starting with the QM energies for the backbone and then adding the net effect of the presence of the hydroxyl and hydroxymethylene groups from the MM calculations. Alternatively, the hybrid energy could be computed by adding the difference between the QM and MM calculations for the analogue to the MM energies for the disaccharide. That would depict the hybrid energy as being the MM energy for the disaccharide, with a correction to the MM values for the torsional energy. Both descriptions of the hybrid energy are equivalent.

The QM calculations allow us to treat directly the apparently different exo-anomeric effects in sucrose and also to deal directly with the potential issue of conformationally dependent atomic charges. The MM calculations for the less important parts of the molecule reduce the size of the QM calculations, which would otherwise be impractically time consuming. MM also allows us to deal rapidly with the manifold hydroxyl and primary alcohol group orientations.

Compared with reducing disaccharides, sucrose presents an additional challenge with its more flexible furanosyl ring. It also has the problems that are encountered when modeling 'normal' disaccharides. When predicting condensed phase shapes with isolated molecules, appropriate accounting for hydrogen bonding is difficult. Some intramolecular hydrogen bonds are observed in crystals, as in Fig. 1(a). More often, however, hydrogen bonds are intermolecular. Isolated models have no opportunity for intramolecular hydrogen bonds, and such models will have the lowest energy

when the strongest intramolecular hydrogen-bonding system is formed. This leads to high relative energies for structures that cannot form the intramolecular hydrogen bonds, even though they could form good intermolecular hydrogen bonds and otherwise have low energy. The stabilizing energy of a full-strength hydroxyl–hydroxyl hydrogen bond is around 5 kcal (MM3 gives 5.3 kcal {all energies herein are molar} lower energy for an isolated methanol dimer than for the two molecules at infinite separation [26]). Therefore, shapes of sucrose that permit both of these hydrogen bonds could be stabilized by 10 kcal. This overwhelming computed advantage over other structures that are otherwise equivalent in energy is an artifact of predicting condensed phase structures with isolated models. The MM part of our hybrid method allows us to give quickly (although crudely) varying strength to the hydrogen bonding through alterations of the dielectric constant.

2. Methods

Molecules used in this study are shown in Fig. 1, including sucrose (Fig. 1(a)), and its analogue based on tetrahydropyran and tetrahydrofuran, THP–O–THF (Fig. 1(b)). Also included are axial 2-methoxytetrahydropyran, Me–O–THP (Fig. 1(c)), and 2-methoxytetrahydrofuran, Me–O–THF (Fig. 1(d)). The driven linkage torsion angles are ϕ (O-5–C-1–O-1–C-2′) and ψ (O-5′–C-2′–O-1–C-1). Further discussion of the modeling strategies is presented elsewhere [27].

The QM calculations were carried out with GAMESS [28] and GAUSSIAN94 [29]. The QM surfaces were calculated at the RHF/6-31G(d) level of theory, which has been used to study anomeric effects in molecules that are smaller than our analogues [30]. Because of fortuitous cancellation of errors, HF/6-31G* energies for glucopyranose are similar to relative energies produced by advanced electronic structure theory [31,32]. Also, HF/6-31G* conformational energies are reasonably comparable to energies from higher levels of theory for dimethoxymethane [33,34], an important fragment of our analogs.

The MM calculations used MM3(96) [8]. (MM3 is available to academic users from the Quantum Chemistry Program Exchange, Department of Chemistry, Indiana University, Bloomington, IN. Others may obtain it from Tripos Associates, St. Louis, MO.) Calculations for the analogue used the default dielectric constant of 1.5, as intended for comparisons with QM calculations. For sucrose we used dielectric constants (ε) of 1.5, 3.5 (our choice for modeling condensed phases), and 7.5 (further de-emphasis of hydrogen bonding). The force field was altered by using the equilibrium distance (1.82 Å) and force constant (3.30 kcal) for H···O hydrogen bonding from the 1992 version of MM3. Our experience has been that those shorter or stronger parameters (compared with the 1996 values of 2.02 Å and 3.00 kcal) are better for modeling condensed phase systems. In particular, lattice dimensions for modeled crystals are closer to the experimental values, as are the geometries of individual hydrogen bonds.

Default criteria for termination of minimization were used in the respective programs. All internal coordinates were allowed to vary during minimization except the linkage torsions, ϕ and ψ. Although Tvaroška and Carver [35] had previously carried out the calculations for Me–O–THP, we repeated them at 20° increments, the same as for the other molecules in the present work. The QM calculations for the *analogue of sucrose* started with crystalline disaccharide geometry [36], after the exocyclic groups were replaced by hydrogen. The 'northern' starting conformation of its furanosyl ring was also modified to make starting geometries with 'southern' forms for both the MM3 and QM calculations. 'Northern' refers to furanoid geometries in the upper half of the conformational wheel, and 'southern' to those in the lower half [37,38]. For fructofuranose, the northernmost structure is $^4T_3$, and the southernmost is $^3T_4$. Starting the calculations with both northern and southern structures for the furanose ring was a way to make sure that the lowest-energy ring shape was used at each ϕ, ψ point. This simple strategy was confirmed by computing the entire pseudorotational itinerary at each of the three major minima.

The final MM3 calculations for the THP–O–THF analogue were started from the QM structures. For the *sucrose molecule*, 48 different starting geometries, with different combinations of exocyclic substituent orientations, were used at each ϕ, ψ point. Those structures, all started as northern forms, were sketched with CHEM-X [39]. These combinations included clockwise and counterclockwise orientations of the secondary OH groups on glucose and fructose. A second counterclockwise combination on glucose had O-2–H oriented trans to the C-2–H bond. The *tg* rotamer of the glucose C-6–OH group was omitted from the starting models because it rarely occurs in crystal structures. Two orientations of each primary alcohol group on the furanose ring, as observed in crystal structures, were included. Each minimization began with rigid rotations to the ϕ, ψ point in question from one of the starting structures. The lowest energy obtained at any ϕ, ψ point, regardless of input geometry, was used to construct each energy surface. Six $^{\text{Six}}\phi$, ψ points had starting structures that were so entangled that they could not be minimized, and those points were not included in the data.

Crystal structures [3,40–56] (Table 1) were taken from the Cambridge Structural Database (CSD) [57]. A sarcosine–sucrose adduct [58] and two molecules in the crystalline NaI–sucrose complex [59] were not in the CSD, but were taken from the original literature. The CSD search was based on molecules having the THP–O–THF backbone. All structures were included unless there were additional covalent bonds between the THP and THF rings. SURFER (version 7.00) [60] was used to construct the contour surfaces and to create the difference surfaces. Its optional multiquadric radial basis function method for interpolating the data was used to produce files with 1° intervals, and interpolated data were adjusted for the zero of relative energy. The corresponding crystal structure energies were derived from these grid files with the calculate residuals function of Surfer. We observed some minor variation in the reported energy values of the crystalline conformations compared to a previous version of SURFER.

3. Results and discussion

Components of the hybrid sucrose surface.— Fig. 2(a) shows the MM3 energy surface for sucrose calculated at $\varepsilon = 3.5$ and Fig. 2(b) and (c) depict the MM3 and QM surfaces for the analogue, similar to maps published previously [6]. The 23 linkage geometries from small-molecule crystal structures are indicated by symbols. SUC, RAF and TGS notations point, respectively, to the conformations in crystal structures of sucrose [55]; the trisaccharide, raffinose [53]; and 4,1′,6′-trichlorogalactosucrose [45].

On Fig. 2(a), some observed conformations of sucrose moieties have high energies, especially the RAF conformation with a corresponding energy of 8.26 kcal. Previously we showed that neither the galactose residue nor neighboring water or raffinose molecules would be likely to stabilize the RAF conformation if the high MM3 energy were correct [6]. Immel and Lichtenthaler also agree that the intramolecular interactions from the galactose residue are not important in determining the conformation of the sucrose moiety [12]. Only five of the 23 observed conformations are within the 1 kcal contour of Fig. 2(a), and none are very near the global minimum. Some of the moderately high-energy parts of the surface are occupied while others are not. These observations suggest that MM3 does not calculate the energy correctly for the sucrose linkage, as mentioned in Section 1. It would be desirable to have a more predictive surface than Fig. 2(a).

Fig. 2(b) shows that the observed conformations are also poorly distributed on the MM3 energy surface for the THF–O–THP analog of sucrose. The discrepancy between the surface and the observed structures is very similar to the discrepancy in Fig. 2(a). This suggests that a method that gives an improved calculation of the energy for the analogue could lead to a more predictive surface for the disaccharide. The biggest contribution to the difference between the energy of the conformation at the minimum in MM3 energy and the RAF shape of this simplified model is from the torsional energy term [6]. Even without the partitioning of the energy given by the MM3 report, it appears that substantial varia-

Table 1
Crystal structures, their conformations, their corresponding energies on the $\varepsilon = 3.5$ hybrid surface and their hydrogen bonds

Molecule name	Ref.	Ref. code	ϕ, ψ (°)	kcal	Hydrogen bonding [a]	
					$O\text{-}6_{fru}\cdots O\text{-}5_{glu}$	$O\text{-}1_{fru}\cdots O\text{-}2_{glu}$
6-Kestose	[40]	CELGIJ	89.58, −54.50	1.63	NA	
NaBr–sucrose dihydrate	[41]	DINYOO10	99.84, −46.06	0.41	weak *	
Erlose monohydrate	[42]	HAHXUJ	104.54, −32.56	0.52		
Xylo-sucrose hemihydrate	[43]	HAHYUK	108.25, −46.90	0.02	*	
Xylo-sucrose hemihydrate	[43]	HAHYUK	96.95, −37.18	0.72		
Erlose trihydrate	[44]	HEGXOG	98.00, −55.71	0.75	weak *	
Erlose trihydrate	[44]	HEGXOG	109.90, −39.76	0.15		
Trichlorogalactosucrose	[45]	KANJOY	91.43, −162.21	4.00	NA	$O\text{-}3_{fru}\cdots O\text{-}2_{glu}$
1-Kestose	[46]	KESTOS	84.64, −65.90	2.00		weak *, R
Potassium sucrose octasulfate	[47]	KSCOSF	107.25, −41.52	0.10	NA	NA
Melezitose I	[48]	MELEZT01	99.77, −30.61	0.74	*	
Melezitose II	[49]	MELEZT02	109.58, −43.42	0.04		
6,6′-Dichloro-6,6′-dideoxy-2,3,4,3′,4′-penta-*O*-benzylsucrose	[50]	NIVPIR	105.79, −52.96	0.18	NA	*, R
Nystose trihydrate	[51]	PEKHES01	102.33, −18.64	1.38		R
Di-iodoplatinum monohydrate of 6,6 diamino hexa-*O*-methylsucrose H_2O	[3]	PELWEI	110.32, −70.29	1.34	NA	NA
Planteose	[52]	PLANTE10	108.21, −26.18	1.01	NA	
Raffinose	[53]	RAFINO01	82.08, 11.98	3.88		
Stachyose	[54]	STACHY01	109.32, −47.31	0.01	*	*
Sucrose	[55]	SUCROS11	108.20, −45.09	0.03	*	*
Sucrose octaacetate	[56]	ZZZSTI01	93.22, −21.79	1.50	NA	NA
3 NaI·2 sucrose·3 H_2O complex	[59]	not in CSD	79.90, −66.60	2.19		
3 NaI·2 sucrose·3 H_2O complex	[59]	not in CSD	79.60, −61.60	2.56		
Sarcosine–sucrose adduct	[58]	not in CSD	94.1, −20.60	1.48		

[a] Criteria for hydrogen bonding were: $d_{H\cdots O} < 3.0$ Å, $\angle_{O\text{–}H\cdots O} > 90°$. * means that bond was observed. NA means that substitution on that molecule precludes hydrogen bonding. Weak means that $2.5 < d_{H\cdots O} < 3.0$ Å. R means that only the $O\text{-}2_{glu}$–H can be a donor; $O\text{-}1_{fru}$ is in the linkage to another monomer and cannot donate.

tion in MM3 conformational energy in the area of the crystal structures arises from the torsional energies and not from forces such as hydrogen bonding, etc. This is because the analogue has no capability for hydrogen bonding, and the nonbonded interactions would be quite different for sucrose and this analogue.

The QM surface for the same analogue (Fig. 2(c)) is quite different from the MM3 analogue surface (Fig. 2(b)) in the general region of the crystal structures, although it is otherwise similar, even in areas of high energy. The crystal structures are clustered about a secondary minimum of 0.68 kcal near $\phi = 99°$ and $\psi = -54°$. While Fig. 2(c) also does not provide an even distribution of the conformations from crystal structures about the global minimum, all of the linkages from these crystals have corresponding energies of < 3 kcal. Fig. 2(d) shows the difference that was computed by subtracting the relative QM energy for the analogue from the relative MM3 energy for the analogue at each ϕ, ψ point. Some crystal conformations occur where the two methods give the same relative energy, but the RAF conformation in particular falls where the differences are among the greatest.

The discrepancy between MM3 and QM maps for the sucrose analogue was pointed out earlier [6]. The improvements here are that the new QM surface covers the full range of ϕ and ψ. This was important because it shows substantial agreement between the QM and MM3 energies in other regions of ϕ, ψ space. Also, the new map is based on the more complete HF/6-31G* theory instead of HF/4-

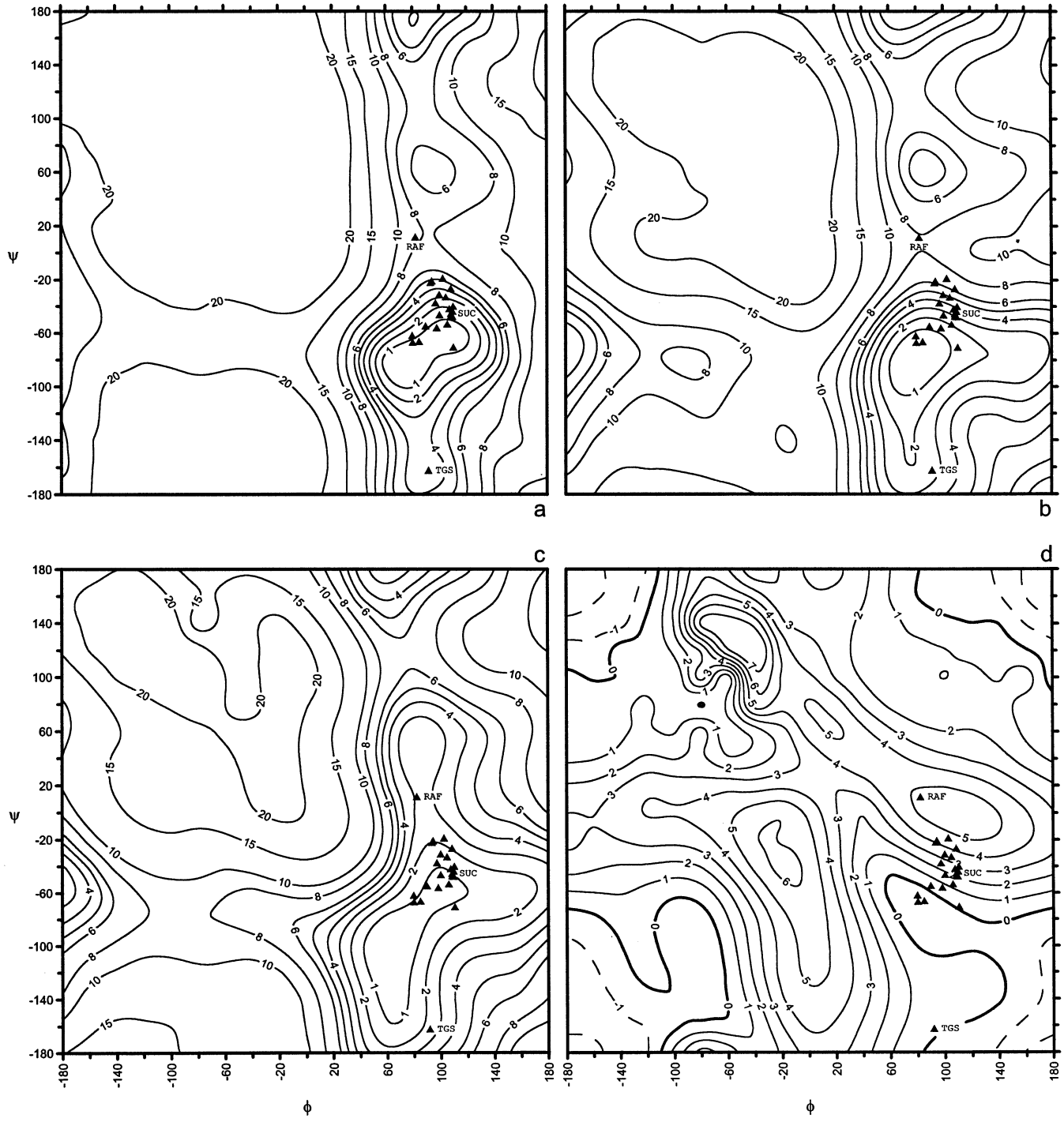
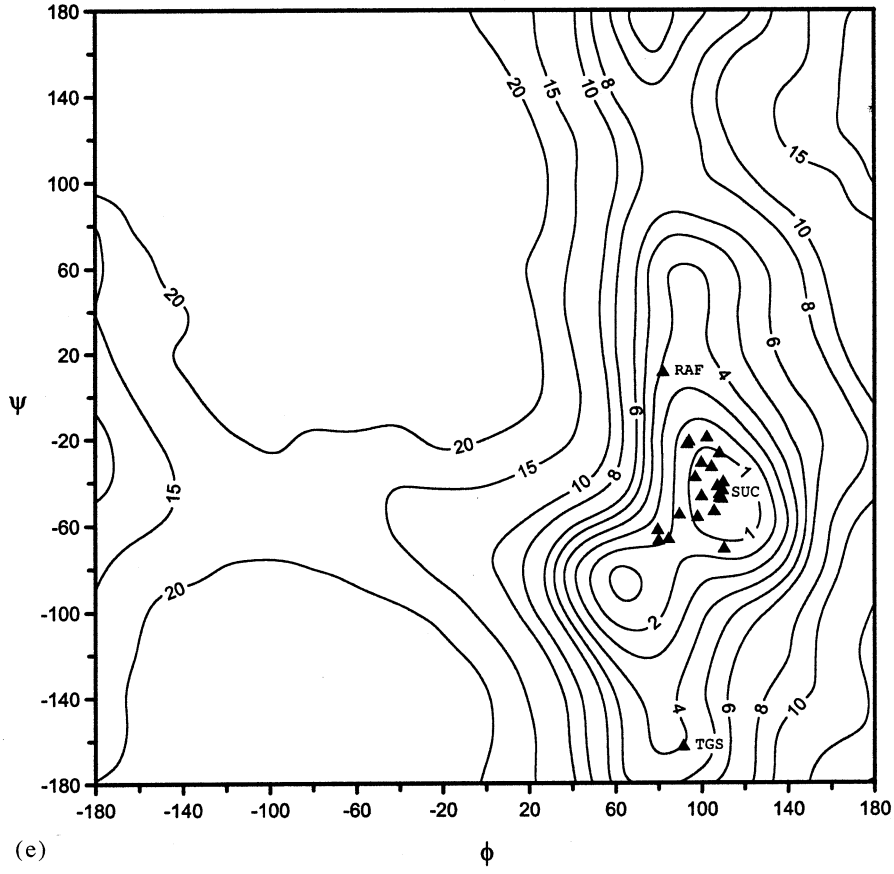
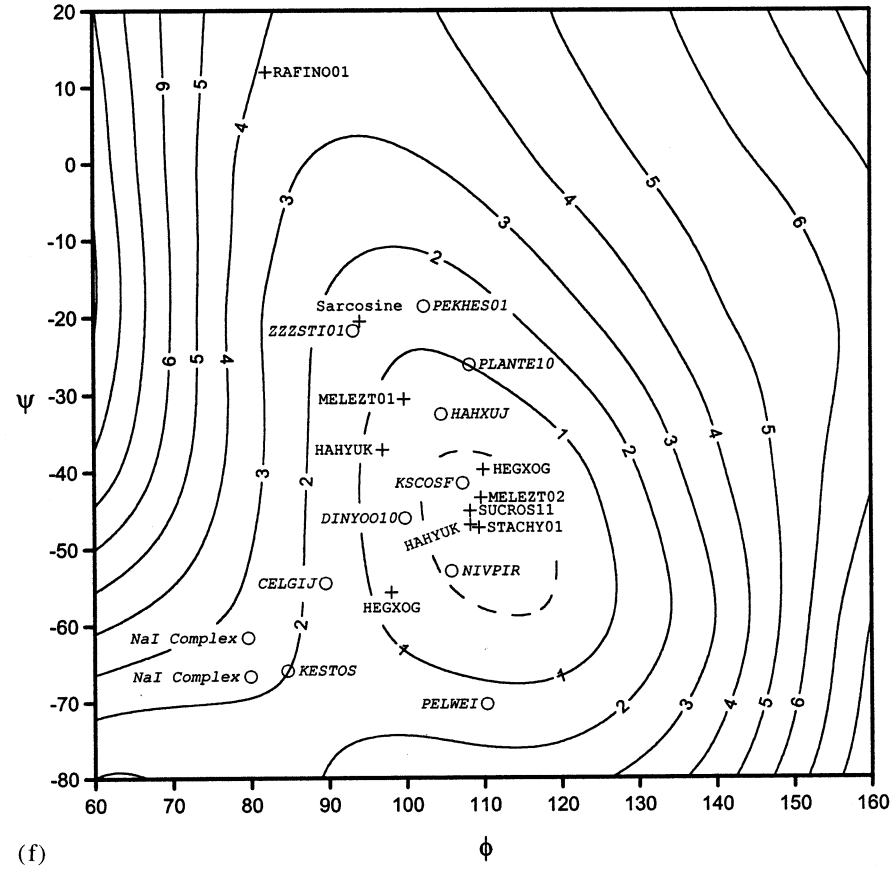


Fig. 2. Potential-energy surfaces involved in the preparation of the hybrid surface for sucrose and the hybrid surface. The symbols show the locations of conformations observed in small-molecule crystals. Contours are at levels of 1, 2, 3, 4, 5, 6, 8, 10, 15, and 20 kcal. The RAF, SUC and TGS notations show the locations of structures discussed in the text for raffinose pentahydrate [53], surface [53], sucrose and trichlorogalactosucrose [45]. (a) The MM3 map for the full sucrose disaccharide with $\varepsilon = 3.5$. (b) The MM3 map for the THF–O–THP sucrose analogue at $\varepsilon = 1.5$. (c) The HF/6-31G*//HF/6-31G* map for the THP–O–THF analogue. (d) The map of the difference between the previous two surfaces (MM3 − HF/6-31G*). The dashed contours represent areas where the relative QM energies were higher than the MM3 energies. (e) The hybrid surface obtained by subtracting the MM3 analogue surface (2b) from the MM3 disaccharide surface (2a) and adding the QM surface (2c). (f) A magnified view of the hybrid surface (2e) showing the region of crystal structures (except TGS). The labels are the CSD Ref. codes in Table 1. The + symbols indicate the molecules that are most likely to have linkage conformations similar to sucrose, while the ○ symbols and italicized labels show the structures that have substitution in the region of the glycosidic linkage or are ionic complexes.



(e)



(f)

Fig. 2. (*Continued*)

21G (d on O). (The 4-21G partial surface was similar to the corresponding part of the 6-31G* surface.) Because of the lack of observed linkage geometries in the region of the global minimum on both the current and previous QM analogue maps, we were still concerned about their applicability to sucrose. That concern is answered by our hybrid surface. We have also created a HF/6-31G* surface for the C-2′-methylated THP–O–THF analogue of sucrose. The global minimum on that map coincides with most of the crystal structures [61].

The hybrid surface.—Fig. 2(e) shows the hybrid MM3–QM surface, which is the result of subtracting Fig. 2(d) from Fig. 2(a). Fig. 2(f) shows a closeup view of the region of all crystal structures except TGS. Table 1 shows the details of the fit of the crystal structures to the hybrid energy surface. On the hybrid surface, the observed structures are scattered randomly about the global minimum, with none corresponding to high (>4.0 kcal) energies. There are no large, low-energy areas that are unpopulated. There is an exception to this intuitive accounting. An unpopulated local minimum of just under 1 kcal exists near $\phi = 65°$, $\psi = -85°$. That local minimum is separated from the global minimum by a low (<3 kcal) barrier and is in the region of the global minimum of the QM THF–O–THP map. Crystalline sucrose itself has an energy of 0.03 kcal on our hybrid surface (Table 1), compared with 1.51 kcal on the MM3 surface, and the energy of the RAF conformation was reduced by 4.38 kcal from its value on the MM3 map to 3.88 kcal on the hybrid map.

Other observed molecules with conformations that are found inside the 0.25 kcal contour of Fig. 2(f) include stachyose [54], one of the erlose trihydrate molecules [44], one of the xylo-sucrose molecules [43], the potassium salt of sucrose octasulfate [47], melezitose II [49], and 6,6′-dichloro-6,6′-dideoxy-2,3,4,3′,4′-penta-*O*-benzylsucrose [50]. Despite very similar ϕ and ψ values, the inter-residue, intramolecular hydrogen-bonding schemes in these seven molecules vary. No hydrogen bonds are possible for the octasulfate, and none occur for this melezitose form, although one occurs for a different polymorph, melezitose I. None occur in this particular erlose molecule either, although one of the two other erlose molecules in the database does have one intramolecular hydrogen bond. Dichloropentabenzylsucrose does form an intramolecular hydrogen bond with its only hydroxyl group. Sucrose and stachyose have similar O-1_{fru}···O-2_{glu} and O-6_{fru}···O-5_{glu} links, but the bonds in sucrose are shorter. Xylo-sucrose forms an O-1_{fru}···O-2_{glu} link in the lower energy molecule in its two-molecule asymmetric unit.

The ψ torsion angle of the TGS molecule (Fig. 2(e), Table 1), which has a non-exo-anomeric (trans) geometry and a high corresponding energy, is stabilized by an O-3_{fru}···O-2_{glu} bond of very good geometry. The hydrogen bonds favored by sucrose and stachyose are not possible in TGS because O-1_{fru} and O-6_{fru} have been replaced by chlorine. Of the 23 sucrose moieties in our set, 16 of the O-6_{fru}, O-1_{fru} and O-2_{glu} sites are substituted, precluding their serving as hydrogen-bond donors. At the remaining 30 unblocked hydrogen bonding sites, nine intramolecular hydrogen bonds are formed. These observations imply that intramolecular hydrogen bonds do not have a large role in determining linkage conformations in carbohydrate crystal structures. In contrast to the decreased corresponding energy of the RAF geometry, the relative energy for the TGS linkage geometry was raised 0.62 kcal on the $\varepsilon = 3.5$ hybrid map compared to the MM3 map. Both RAF and TGS depended on the dielectric constant, being over 8 kcal on the hybrid $\varepsilon = 1.5$ map.

Except for the TGS structure, the crystal structures with the most negative ψ values include 1-kestose [46] (KESTOS) and a heavily substituted derivative (PELWEI [3]), as well as the two molecules in the asymmetric unit of the NaI complex [60]. The furanose rings in the sucrose moieties in KESTOS and PELWEI have southwestern conformations (E_4 and ^5E, respectively) [36]. The change from northern to southern conformation entails a change of the disposition of the glycosidic bond from the furanose ring. That change, from pseudo-axial to pseudo-equatorial, changes the O-1–C-2_{fru}–O-5_{fru}–C-5_{fru} torsion angle and might change the

C-1_{glu}–O-1–C-2_{fru}–O-5_{fru} torsional energy. The corresponding energies on Fig. 2(f) of the KESTOS and PELWEI conformations, 2.00 and 1.34 kcal, are above average but still within the range for structures having northern ring forms. The calculated energies for sucrose structures having the conformations found in the NaI complex are higher, at 2.19 and 2.56 kcal. Tvaroška and Carver [62] showed that the anomeric effects were altered in the presence of ions, so the actual energy of distortion might be less. Alternatively, the structures of the two sucrose molecules in the NaI complex may be distorted beyond the extent normally found in purely molecular crystals. The observed shapes permit nearly total coordination of the hydroxyl groups and ring and linkage oxygen atoms with the three Na^+ and I^- ions. On the other hand, the conformation of sucrose in the 1:1 complex with NaBr (DINYOO10) is similar to that of sucrose itself, with a calculated distortion energy of 0.41 kcal.

Average energies of the crystal structures on the MM3 and QM surfaces for the analogues were 3.51 and 1.38 kcal, respectively. Table 2 shows the average energies of the crystal conformations on the MM3 and hybrid sucrose surfaces at the different dielectric constants. The average energies of the crystal structures are lower on the hybrid maps, but the higher dielectric constants also give lower energies. This should not be taken to mean that the hydrogen bonds in MM3 are intrinsically too strong. In fact, hydrogen bonds in MM3 have energies that are quite comparable with experimental results and with ab initio calculations on isolated molecules [36,63]. However, we think that this reduction in average energy with increased dielectric constant is meaningful. The distribution of observed structures on the energy surfaces is clearly predicted better by the surfaces with dielectric constants higher than 1.5 [27]. Some parts of the high-energy regions are populated, while some lower-energy regions are unpopulated on the $\varepsilon = 1.5$ surface. Besides the obvious pragmatism of using an elevated dielectric constant, numerous arguments can be advanced for its necessity:

(1) An elevated dielectric constant (4.0) is recommended by Allinger for MM3 models of condensed phase structures [64]. In MM3, the hydrogen-bonding potential function replaces the van der Waals interactions, and those energies are added to the dipole–dipole interactions. Both of these contributions depend on the dielectric constant.

(2) Our unpublished calculation with the MM3 force field used herein and $\varepsilon = 3.5$ gave a value of 40.1 kcal for the lattice energy of *myo*-inositol, a heavily hydrogen-bonded crystal structure. That value can be compared to the experimental and modeled estimations of the heat of sublimation ranging from 37.0 to 42.5 kcal [65]. Thus, the strength of hydrogen bonding used in our calculations may accurately reflect the forces on the molecule.

(3) According to complete basis set ab initio calculations at the MP2 level, the energy of the hydrogen bond in the isolated water dimer is only -2.82 kcal after correction for zero-point vibrational energy, in comparison with the difference in the electronic energy of -4.94 kcal [66]. That -2.82 kcal value is substantially different from the -5.3 kcal of steric energy given by MM3 for the isolated methanol dimer at $\varepsilon = 1.5$ [26]. The MM3 steric energy with 1992 parameters at $\varepsilon = 3.5$ is -2.2 kcal.

(4) Hydrogen-bonding schemes in crystals invariably involve all or almost all of the hydroxyl groups. Therefore, the total hydrogen-bonding energy for a given molecule in condensed phase should be nearly constant, regardless of conformation. Heats of combustion of isomeric solid saccharides are very uniform, supporting that idea. Therefore, hydrogen-bonding would have little influence on observed conformations. This is not the case when the molecule is modeled in isolation with full-strength hydrogen-bonding interactions, however. Therefore, it makes sense to use a reduced strength of hydrogen bonding so that there is not an undue influence of hydrogen bonding on the modeling results.

The MM3 potential for the C–O–C–O torsion angle was derived from experimental ΔH_f data from solutions [8]. No allowance was made for dependence of the anomeric energies on dielectric constant (Allinger, personal communication). Such a dependency has been proposed [15,24] for both single and overlapping

exo-anomeric effects. Nevertheless, we have, in effect, replaced the experimental, solution-based torsional parameters from MM3 with the net torsional parameters from gas-phase QM data. Even though our increased dielectric constants only affected dipole–dipole and hydrogen-bonding calculations in the full disaccharides, and not the exo-anomeric torsional energies, our results seem mostly satisfactory. This observation is similar to an experience with AMBER*. Its exo-anomeric torsional energies derived from QM calculations do not vary with the dielectric constant but still gave satisfactory results in solution studies on THP compounds [22].

Comparison with other surfaces.—When Engelsen et al. [11] studied sucrose by energy minimization in vacuum with a modified CHARMM force field, their fit of the crystal structures was roughly similar to the accounting on our MM3 surface (Fig. 2(a)). A different result was obtained by their molecular dynamics (MD) simulations of sucrose in water with the same force field. They presented a contoured surface based on the frequencies or probabilities (denoted by P) of the various conformations. Six of the crystal structure conformations of Table 1 are found within their contour of $P = 0.01$. This is equivalent to 2.75 kcal, assuming a Boltzmann distribution and room temperature. Three of the present structures are on the $P = 0.01$ line, and 12 more, including RAF, are found within the $P = 0.001$ contours (4.1 kcal). Their outer, 0.0001 contour (5.5 kcal) includes the PELWEI structure but does not provide a pathway to the TGS conformation or a value for its probability. The improved fit in their dynamics study was achieved partly because the RAF conformation is stabilized in the solution dynamics study by a bridging water molecule, even though it is not one of the liaisons in the raffinose pentahydrate crystal structure [60]. Also, the solvated dynamics method provides explicit opportunities for intermolecular hydrogen bonding so the calculated relative energies for any given ϕ, ψ conformation should not depend so much on whether there is an intramolecular hydrogen bond.

A solvated dynamics simulation by Immel and Lichtenthaler [12] was based on the 1987 GROMOS united-atom potential function [67]. Unlike the above simulation, this simulation has a path to the TGS structure with plausible energies for both the barrier (< 4.0 kcal) and the observed conformation (about 2.5 kcal). Also, the RAF conformation is accounted for, with an energy of about 4 kcal. Their 1 kJ contour (1 kJ = 0.24 kcal) surrounds the NaI sucrose and kestose structures, KESTOS and CELGIJ, while the cluster of structures that includes sucrose (SUCROS11) and stachyose (STACHY01) corresponds to energies between 1 and 2 kcal. In their work, the preferred solution conformation is stabilized by a bridging water molecule, which again is not part of the observed crystal structure. While these authors did not apparently attempt to predict crystal structures, theirs is the closest of the energy surfaces discussed herein to our hybrid surface.

A united-hydroxyl group version of the CHARMM force field, CHEAT95, was used to model sucrose [13]. Because all 14 of their small-molecule crystal structures had energies of less than 4 kcal, the authors concluded that their force field does not have our sucrose problem. Still, none of their structures (or our additional nine) corresponded to less than 2 kcal on their surface, and their average energy

Table 2
Average energy values (kcal) corresponding to observed crystal structures based on MM3 and hybrid energy surfaces for the full disaccharide

Number of linkages	MM3			Hybrid		
	$\varepsilon = 1.5$	$\varepsilon = 3.5$	$\varepsilon = 7.5$	$\varepsilon = 1.5$	$\varepsilon = 3.5$	$\varepsilon = 7.5$
23	3.93	2.51	2.00	3.06	1.16	0.93

would be similar to the value of 2.51 kcal on our $\varepsilon = 3.5$ MM3 map. Also, only a part of their surface with energies of about 3 kcal was occupied.

Rasmussen has published several energy surfaces for sucrose or its THF–O–THP analogue, exemplifying the difficulty in modeling sucrose. There are two energy surfaces for the THP–O–THF analogue in the first paper [68]. The first shows the RAF conformation to be nearly 10 kcal above the global minimum at $\phi = 63°$, $\psi = 175°$. On his improved analogue surface, the RAF conformation is similar in energy to the SUC conformation, which is about 4 kcal above the global minimum. The sucrose surface in a newer paper [69] shows a global minimum at $\psi = -175°$ for the sucrose molecule. In that work, several minimizations led to a variety of results under similar circumstances, and the energy surface is quite different at the -180 and $+180$ values of ψ where it should be identical. This indicates technical difficulties in producing the surface.

Exponential decay.—By virtue of the low average corresponding energy of the crystal structures, its pathway to the TGS structure, and the visual distribution of the crystal structures, our hybrid map is a better predictor of observed conformations than other published results. Its utility as a predictor of conformations in crystal structures can also be assessed by how well the distribution fits a Boltzmann-like exponential decay pattern. According to Bürgi and Dunitz [70] and Bye et al. [71], the most frequently observed conformations should have corresponding energies near the global minimum in energy. Further, the number of observations should fall off according to a pattern of exponential decay as the energy increases. This of course assumes that there is no particular packing mode that is favored by a large portion of the structures. Their main question regards the effective value of T (temperature) to be used in the Boltzmann equation when comparing potential-energy values with crystallographic data.

One way to determine whether the distribution of structures has a Boltzmann-like distribution would be to plot a histogram. The number of structures having energies in the range of 0.0–0.5 kcal, for example, could be plotted as a vertical bar, and bars for succeeding 0.5 kcal increments would also be plotted (see Table 1 for the energy value for each structure). The bar with the greatest number of structures would be scaled to have a value of 1.0, and the same scaling would be applied to the other bars. A curved line would then be fit to the tops of the bars and compared with the Boltzmann curve. Given the small number of crystal structures (23) that contain sucrose moieties, it is likely that such a plot would be rather crude. Instead, we used a 'cumulative frequency' technique in our analysis. The corresponding relative energies from all 23 linkages from small-molecule crystals were sorted from lowest to highest calculated energy. These sorted energies are plotted in Fig. 3(a), where the y-coordinates are based on the position in the list for a given energy value. The highest point (the one with the lowest energy) has a y-value of $1-(1/23)$, the second point is $1-(2/23)$ and so forth, until the last point is reached with a y-value of $1-(23/23)=0$. This graph, based on cumulative frequencies, is comparable to the histogram described above. However, there is no need to choose the bin size, and fitting a curve to the energy distribution is simplified because there is a value of x and y for each point. For reference, a Boltzmann distribution at room temperature is plotted. That distribution is based on $P_i = e^{-\Delta G_i/RT}$, where P_i is the probability or frequency of structures having the ith value of the energy and R is the universal gas constant.

We will take our relative energies of distortion as ΔG, even though in one analysis they differ from the potential energy for different conformers by as much as 0.5 kcal [72]. In cases where the population of structures has Boltzmann-like exponential decay, equations of similar form, $P = e^{-\Delta E/\beta}$, apply and the value of β corresponds approximately to the average energy of the population. Table 2 has average values of 1.16 for the hybrid sucrose surface at a dielectric constant of 3.5 and 1.38 for the QM surface of the analogue, instead of 0.59 kcal (RT at room temperature.) In Fig. 3(a), a fitted exponential decay curve with $\beta = 1.12$ kcal is a good approximation to the observed distribution on the hybrid surface (Fig. 2(e)). In Fig. 3(b), the fitted decay curve

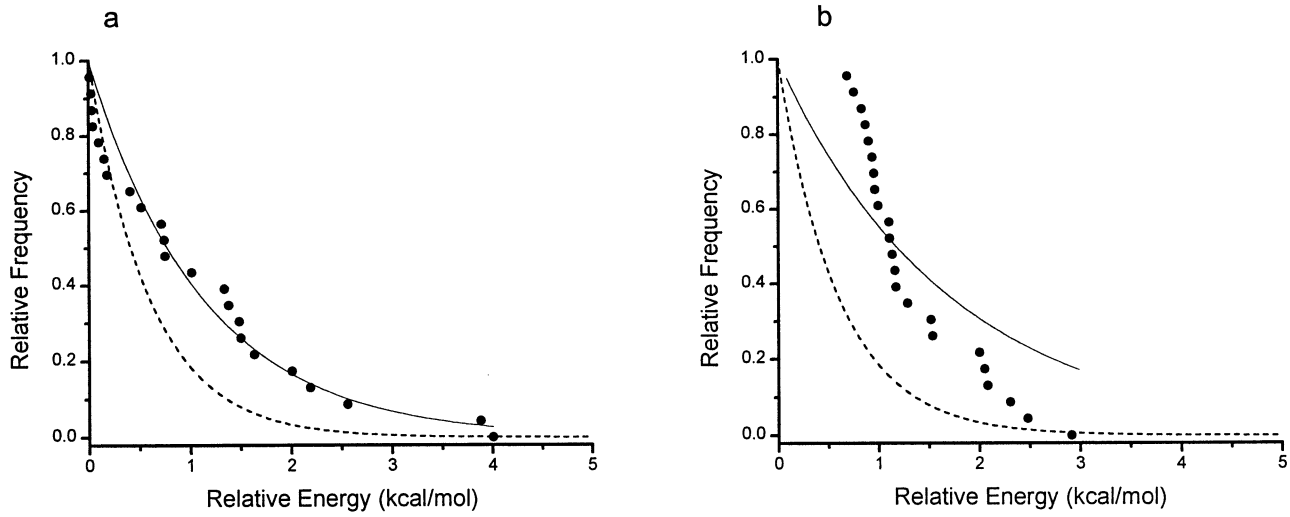


Fig. 3. Distributions of crystalline conformations against energy for (a) the hybrid sucrose surface (Fig. 2(e)) and (b) the QM analogue surface (Fig. 2(c)). The dashed lines show an ideal Boltzmann distribution, the ● show the observed points and the solid lines show the Boltzmann-like exponential curve fitted to the observed structures on the given surface (see text). In (a) the exponential line agrees fairly well with the observed data, while in (b) there is a substantial discrepancy. For (b), observations were clustered near a secondary minimum and no energies were < 0.69 kcal, while the exponential decay curve must take a value of 0 kcal at the maximum probability of 1.0.

with $\beta = 1.67$ is a poor fit to the observed distribution on the QM analogue surface (Fig. 2(c)). This shows a way to distinguish the quality of fits of the observed data to different potential-energy surfaces, despite similar values of the average energy. In this case, the better fit for the hybrid sucrose model is reasonable because both the substituents and the QM stereoelectronics are included. The poorer fit to the QM analogue surface results because of the missing substituents in the THP–O–THF model.

The question as to whether the distribution of structures should match the room-temperature Boltzmann equation remains unanswered. We anticipate that there is still substantial room for improvement in our calculation. The various simplifications, such as spherical atoms for the MM interactions, and the dielectric constant means of diminishing the strength of the electrostatic interactions, have substantial limitations. The level of QM theory used can only be regarded as a useful approximation. When electron correlation was included, either at the MP2/6-31G* or the B3LYP/6-31G* level, the differences between the energies of the secondary minimum near the SUC geometry and the saddlepoint near the RAF geometry (Fig. 2(c)) drop from 2.3 kcal (RHF/6-31G*) to 1.8 and 1.4 kcal, respectively. This flattening of the potential energy surface could significantly lower the average energies of the crystal structures.

The average of the crystal-structure energies might be higher than the theoretical value because we have considered all of the available conformations of structures that contain sucrose moieties regardless of chemical substitution or other factors. Many of the present crystal structures could be disqualified from inclusion in the analysis because of the presence of ions, or substituents that might alter the linkage conformation of the sucrose molecule. The remaining, 'best' analogues include HAHXUJ, HAHYUK, HEGXOG, MELEZT01, MELEZT02, RAFINO01, STACHY01, SUCROS11, and the sarcosine adduct. Raffinose, stachyose and erlose are substituted at the O-4 of the glucose moiety, as far as possible from the linkage. Xylo-sucrose is the same as sucrose except for the loss of the C-6 primary alcohol group on the glucose ring. Again, this is a change fairly distant from the linkage. The only sucrose moieties that are modified on their fructose residue that were included in this set of most analogous moieties comes from melezitose. There, the substitution is on C-3′. This inclusion is not as clear-cut as the other tri- and tetrasaccharides because the C-3′ hydroxyl is involved in a

hydrogen bond in TGS, although in a very different ϕ, ψ conformation. The average energy of these 11 structures on the $\varepsilon = 3.5$ surface is 0.76 kcal, with RAFINO01 contributing the most energy. Removal of that outlier gives ten structures with an average of 0.45 kcal.

The energy surfaces determined by other workers, based on united-atom or solvated dynamics methods as described above, have more structures at higher energies than at the lowest energies. They would not fit the exponential decay curves as well as the fit in Fig. 3(a). Data from our hybrid $\varepsilon = 1.5$ surface [27] also deviate from an exponential decay because none of the observed structures corresponds to an energy of less than 0.69 kcal on that surface. Our corresponding energies from the hybrid $\varepsilon = 7.5$ surface [27] fit an exponential decay curve well.

Because it has an average energy value closer to the ideal Boltzmann coefficient, the hybrid $\varepsilon = 7.5$ surface might be preferred. Of the 23 linkage conformations, 22 had lower energies on the hybrid $\varepsilon = 7.5$ surface. However, the lattice energy might then be underestimated, and the intramolecular hydrogen bonding geometry would not be very good. Also, our similarly prepared $\varepsilon = 3.5$ hybrid surfaces for other molecules gave an overall average, including sucrose, for the crystal structures of about 0.9 kcal [27]. Therefore, we anticipate that there is still some error in our calculations for sucrose. Also, the available crystal structures for sucrose moieties may be unusually distorted.

Thoughts on the overlapping exo-anomeric effect.—The phrase 'overlapping exo-anomeric effect' was first mentioned by Schleifer et al. [73], who were concerned about the abnormalities that it might introduce in their survey of the geometric features associated with the anomeric centers. The earlier analysis by Tvaroška and Václavík [24] suggested that an MM method could not properly calculate the torsional energies of molecules that have extended sequences of carbon and oxygen atoms such as trehalose or sucrose. This sequence covers C-5–O-5–C-1–O-1–C-2′–O-5′–C-5′, as shown in Fig. 1(a). The idea is that the delocalization of the lone-pair electrons is mostly responsible for the exo-anomeric effect. In the case of these longer sequences, the lone pairs have added opportunities for delocalization, and the delocalization of the lone pairs from the shared glycosidic oxygen atom will be dependent on the conformation of both the ϕ and ψ torsion angles. No MM force field that we know of provides torsion–torsion interactions. According to Hwang et al. [23], torsional terms are very important in determining the energy profiles of anomeric sequences, so incorrect torsional energies are a likely source of the difficulties in modeling sucrose with MM. It is attractive to imagine that the improvement in modeling sucrose that was effected in this work by direct QM calculation of the torsional energies is evidence of the importance of such an effect. More study is needed before such a conclusion is warranted, however.

For example, one reason why the HF/6-31G* and MM3 energies for the sucrose analogue do not agree is that they do not agree for the simple exo-anomeric torsions. Fig. 4 shows the HF/6-31G* and MM3 computed energies as the methoxyl groups undergo torsional rotation for 2-Me–O–THP (Fig. 4(a)) and 2-Me–O–THF (Fig. 4(b)). Their geometries were taken from crystalline sucrose. The THF moiety, if it were a pyranose ring instead of a furanose ring with a pseudo-axial glycosidic bond, would correspond to α-L-glucopyranose. There is some resemblance between the QM and MM3 curves, but there is a substantial difference for each molecule. The difference curves would be nearly identical to each other if one were plotted on a reversed ϕ-axis. From this it is understandable that the MM3 and HF/6-31G* contours for the THF–O–THP molecule (Fig. 2(b) and (c)) will differ from each other. Further manipulation of these data suggests, however, that even if the MM3 and QM data for these monomeric fragments matched, the aggregate correction would not be the same as the correction developed for THF–O–THP itself. This could be due to problems with the electrostatic terms in MM3. There is little variation in the MM3 dipole–dipole energies for the different THF–O–THP conformations, but they may not be correct. Electrostatics have been iden-

tified as important components of the exo-anomeric effect [74].

The discrepancies between MM3 calculations and HF/6-31G* calculations on dimethoxymethane are similar to those in Fig. 4 when one end of the C–O–C–O–C sequence is held in conformations that resemble the ring structures (results not shown). Overall, MM3 does not fit very well with the MP2/6-31G* or B3LYP/6-31G* data for a complete Ramachandran surface [34]. These observations suggest that improved fit of MM with the QM data on dimethoxymethane would be useful. When both the C–O–C–O and O–C–O–C sequences in dimethoxy methane can take the complete range of torsion angles, the overlapping exo-anomeric effect can also become evident. The longer sequence in sucrose and the trehaloses may not be needed to study the overlapping effect. Even so, the ultimate solution may require additions to the force field. Alternatively, the fairly close agreement of the solvated dynamics surface of Immel and Lichtenthaler and our hybrid surface suggests that it may be possible to achieve a good representation of the potential energy surface of sucrose with a relatively simple potential function. Ironically, however, modifications to the exo-anomeric torsional potential and other terms have recently been offered for the GROMOS87 potential function to make it more suitable for modeling cellulose and other carbohydrates [75].

4. Conclusions

In the first part of this paper, we proposed that a hybrid QM/MM energy surface for sucrose would provide a better prediction of conformations that are determined by crystallography than previous surfaces. The hypothesis was that the QM calculations would account for non-additivity of torsional energies that arises at the head-to-head glycosidic bonds in sucrose. A superior accounting for the sucrose conformations was indeed obtained by that method. This improvement came even though our QM surface for the analogue did not, by itself, account very well for the observed distribution of crystal structures. We also suggested that intramolecular hydrogen bonding would not be very important in determining the conformations that would be found in the crystal structure. Our lower average energy and improved visual fit of the structures showed that de-emphasis of the hydrogen bonding through an increased dielectric constant was important in isolated-molecule modeling for obtaining a good prediction of the distribution of crystal structures.

The crystal structures that contain sucrose moieties were divided into two categories. One was based on molecules that did not contain serious impediments to the sorts of intramolecular interactions that sucrose would

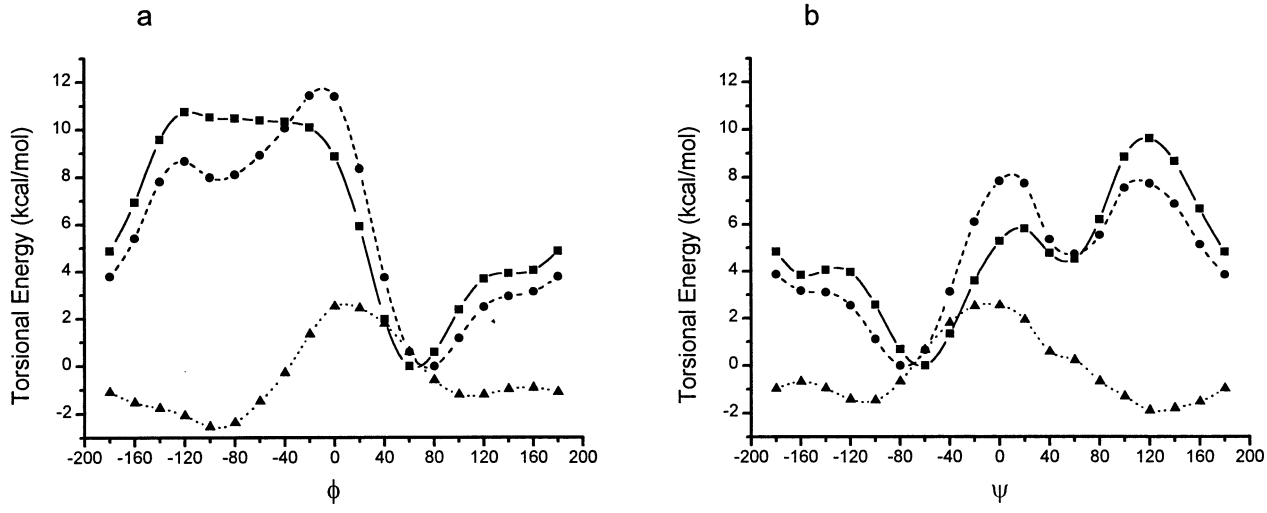


Fig. 4. HF/6-31G*, MM3(96) and the difference (MM − QM) torsional energies for the methoxylated THP and THF molecules. (a) axial 2-Me–O–THP (ϕ). (b) 2-Me–O–THF (ψ). On these drawings, the QM data are shown by –●–, the MM data by –■–, and the difference (MM − QM) by ···▲···.

be expected to experience. These structures included sucrose itself, sucrose in the sarcosine adduct, stachyose, erlose, xylosucrose, melezitose and raffinose. Of the eleven linkages available from these seven molecules, five are within 0.25 kcal of the global minimum and another four are still less than 1.0 kcal. Only the sacrosine (1.48 kcal) and raffinose (3.88 kcal) conformations correspond to energies greater than 1.0 kcal. The sucrose moieties in the second echelon were deemed less analogous because of impediments to normal intramolecular interactions. These impediments included substitution involving O-2 of the glucose ring or O-1′ or O-6′ of the fructose ring, and the presence of ions. Although there was some overlap of the individual energies and conformations of the two classes of structures, the average energy of the second group was 1.55 kcal, about twice as high as for the best 11.

We used a statistical technique to show that the distribution of linkage conformations that are observed in crystal structures is consistent with an exponential decrease in their population as the energy on our hybrid surface increases. Although it is not proven that conformational energies in crystals conform to a Boltzmann-like distribution, modeling methods that provide such a distribution of observed structures have maximal predictive utility.

The conclusions reached herein for sucrose are bolstered by analogous results for other disaccharides that will be reported elsewhere [76]. In particular, another hybrid map for sucrose [27] was very similar although a different analogue was used. It had a methyl group on the anomeric center of the glycosidic carbon of the furanoid ring. While it provided a hybrid surface that was comparable to the present one, the QM analogue map [62] itself accounted for the observed structures much better than the present QM map for the THP–O–THF analogue.

Modeling sucrose with energy minimization is particularly challenging, even compared to other disaccharides, in part because of its flexible furanose ring. For example, the ring shapes in the MM and QM analogue models are not necessarily the same as each other or in the sucrose models. Nevertheless, our procedure has shown that the observed conformations of sucrose are predictable without invoking an explicit environment. Also, it has pointed to areas for improvement, in empirical representations of molecular structure of molecules subject to anomeric effects.

Note added in proof

Two conformational analyses of molecules with the same backbone as sucrose have been reported. Conformations of sucrose octasulfate were studied with NMR and MM2* modeling [77]. A single, deep minimum was found, with the KSCOSF shape within the 10 kcal contour. Its furanose ring has a southwestern ${}^5T_4$ conformation. A map for TGS was based on MM2HB, giving (with our conventions) a global minimum at about $\phi = +100°$, $\psi = +50°$ [78]. The KANJOY crystal structure was less than 4 kcal higher in energy. On that surface, the SUCROS11 conformation falls between two other secondary minima and its energy is more than 2 kcal above the global minimum. Finally, further work on the hydration of sucrose in the context of solvated molecular dynamics calculations was recently published by Andersson and Engelsen [79]. They reiterate the importance of water for explaining the conformational and other properties of sucrose.

Acknowledgements

The authors thank Professor L. Schäfer for suggesting the collaboration between A.D.F. and A.M.K. We had fruitful discussions with Professor N.L. Allinger. The cumulative frequency statistical method was suggested by D. Boykin. Scans of the CSD were carried out by Dr D.P. Miller. Professor J.D. Dunitz corresponded regarding the distribution of conformations. Professor T.B. Grindley and Drs N. Sachinvala and W.E. Franklin commented on a presubmission draft of the paper. Professors C.F. Brewer and F. Tobiason commented on an early draft of this work.

References

[1] L. Hough, in F.W. Lichtenthaler (Ed.), *Carbohydrates as Organic Raw Materials*, VCH, Weinheim, 1991, pp. 33–55.
[2] N.D. Sachinvala, D.L. Winsor, R.K. Menescal, I. Ganjian, W.P. Niemczura, M.H. Litt, *J. Polym. Sci., Part A: Polym. Chem.*, 36 (1998) 2397–2413.
[3] N.D. Sachinvala, H. Chen, W.P. Niemczura, E. Furusawa, R.E. Cramer, J.J. Rup, I. Ganjian, *J. Med. Chem.*, 36 (1993) 1791–1795.
[4] (a) V. Sasisekharaan, *Collagen Proceedings Symposium, 1960*, Madras, India, 1962, pp. 39–78. (b) V.S.R. Rao, P.R. Sundararajan, C. Ramakrishnan, G.N. Ramachandran, *Conformation in Biopolymers*, Vol. 2, Academic Press, London, 1963.
[5] A.D. French, L. Schäfer, S.Q. Newton, *Carbohydr. Res.*, 239 (1993) 51–60.
[6] C. Van Alsenoy, A.D. French, M. Cao, S.Q. Newton, L. Schäfer, *J. Am. Chem. Soc.*, 116 (1994) 9590–9595.
[7] A.D. French, M.K. Dowd, *J. Mol. Struct. (Theochem.)*, 286 (1993) 183–201.
[8] N.L. Allinger, M. Rahman, J.-H. Lii, *J. Am. Chem. Soc.*, 112 (1990) 8293–8307.
[9] V.H. Tran, J.W. Brady, *Biopolymers*, 21 (1990) 961–976.
[10] F.W. Lichtenthaler, S. Immel, U. Kreis, *Stärke*, 43 (1992) 121–132.
[11] S.B. Engelsen, C. Hervé du Penhoat, S. Pérez, *J. Phys. Chem.*, 99 (1995) 13,334–13,351.
[12] S. Immel, F.W. Lichtenthaler, *Justin Liebigs Ann. Chem.*, 11 (1995) 1925–1937.
[13] M.L.C.E. Kouwijzer, P.D.J. Grootenhuis, *J. Phys. Chem.*, 99 (1995) 13,426–13,436.
[14] W.A. Szarek, D. Horton (Eds.), *The Anomeric Effect, Origin and Consequences*, ACS Symposium Series, Vol. 87, American Chemical Society, Washington, DC, 1979.
[15] I. Tvaroška, T. Bleha, *Adv. Carbohydr. Chem. Biochem.*, 47 (1989) 45–123.
[16] G.R.J. Thatcher (Ed.), *The Anomeric Effect and Associated Stereoelectronic Effects*, ACS Symposium Series, Vol. 539, American Chemical Society, Washington, DC, 1994.
[17] C.J. Cramer, *J. Mol. Struct. (Theochem.)*, 370 (1996) 135–146.
[18] C.J. Cramer, *J. Org. Chem.*, 57 (1992) 7034–7043.
[19] L. Nørskov-Lauritsen, N.L. Allinger, *J. Comput. Chem.*, 5 (1983) 326–335.
[20] R.J. Woods, R.A. Dwek, C.J. Edge, B. Fraser-Reid, *J. Phys. Chem.*, 99 (1995) 3832–3846.
[21] K.-H. Ott, B. Meyer, *Carbohydr. Res.*, 281 (1996) 11–34.
[22] H. Sendrowitz, C. Parish, W.C. Still, *J. Am. Chem. Soc.*, 118 (1996) 8985–8985.
[23] M.J. Hwang, X. Ni, M. Waldman, C.S. Ewig, A.T. Hagler, *Biopolymers*, 45 (1998) 435–468.
[24] I. Tvaroška, L. Václavík, *Carbohydr. Res.*, 160 (1987) 137–149.
[25] T. Matsubara, F. Maseras, N. Koga, K. Morakuma, *J. Phys. Chem.*, 100 (1996) 2573–2580.
[26] A.D. French, D.P. Miller, *ACS Symp. Ser.*, 569 (1994) 235–251.
[27] A.D. French, A.-M. Kelterer, G.P. Johnson, M.K. Dowd, C.J. Cramer, *J. Mol. Graph. Model.*, in press.
[28] M.W. Schmidt, K.K. Baldridge, J.A. Boatz, S.T. Elbert, M.S. Gordon, J.H. Jensen, S. Koseki, N. Matsunaga, K.A. Nguyen, S.J. Su, T.L. Windus, M. Dupuis, J.A. Montgomery, *J. Comput. Chem.*, 14 (1993) 1347–1363.
[29] M.J. Frisch, G.W. Trucks, H.B. Schlegel, P.M.W. Gill, B.G. Johnson, M.A. Robb, J.R. Cheeseman, T. Keith, G.A. Petersson, J.A. Montgomery, K. Raghavachari, M.A. Al-Laham, V.G. Zakrzewski, J.V. Oritz, J.B. Foresman, J. Cioslowski, B.B. Stefanov, A. Nanayakkara, M. Challacombe, C.Y. Peng, P.Y. Ayala, W. Chen, M.W. Wong, J.L. Andres, E.S. Replogle, R. Bomperts, R.L. Martin, D.J. Fox, J.S. Binkley, D.J. DeFrees, J. Baker, J.P. Stewart, M. Head-Gordon, C. Gonzalez, J.A. Pople, GAUSSIAN94, Gaussian Inc., Pittsburgh, PA, 1995.
[30] U. Salzner, P.v.R. Schleyer, *J. Org. Chem.*, 59 (1994) 2138–2155.
[31] S.E. Barrows, F.J. Dulles, C.J. Cramer, A.D. French, D.G. Truhlar, *Carbohydr. Res.*, 276 (1995) 219–251.
[32] S.E. Barrows, J.W. Storer, C.J. Cramer, D.G. Truhlar, A.D. French, *J. Comput. Chem.*, 19 (1998) 1111–1129.
[33] K.B. Wiberg, M.A. Murcko, *J. Am. Chem. Soc.*, 111 (1989) 482–4828.
[34] J.R. Kneisler, N.L. Allinger, *J. Comput. Chem.*, 17 (1996) 757–766.
[35] I. Tvaroška, J. Carver, *J. Phys. Chem.*, 98 (1994) 9477–9485.
[36] G.M. Brown, H.A. Levy, *Acta Crystallogr., Sect. B*, 29 (1973) 790–797.
[37] A.D. French, *Encyclo. Comput. Chemistry*, Wiley, New York, 1998, pp. 233–246.
[38] A.D. French, M.K. Dowd, P.J. Reilly, *J. Mol. Struct. (Theochem.)*, 395–396 (1997) 271–287.
[39] CHEM-X, Oxford Molecular Group Inc., Beaverton, OR.
[40] V. Ferretti, V. Bertolasi, G. Gilli, C.A. Accorsi, *Acta Crystallogr., Sect. C*, 40 (1984) 531–535.
[41] C.A. Accorsi, F. Bellucci, V. Bertolasi, V. Ferretti, G. Gilli, *Carbohydr. Res.*, 191 (1989) 105–116.
[42] T. Taga, E. Inagaki, Y. Fujimori, S. Nakamura, *Carbohydr. Res.*, 240 (1993) 39–45.
[43] T. Taga, E. Inagaki, Y. Fujimori, K. Fujita, K. Hara, *Carbohydr. Res.*, 241 (1993) 63–69.
[44] T. Taga, E. Inagaki, Y. Fujimori, S. Nakamura, *Carbohydr. Res.*, 251 (1994) 203–212.
[45] J.A. Kanters, R.L. Scherrenbert, B.R. Leeflang, J. Kroon, M. Mathlouthi, *Carbohydr. Res.*, 180 (1988) 175–182.
[46] G.A. Jeffrey, Y.J. Park, *Acta Crystallogr., Sect. B*, 28 (1972) 257–267.
[47] Y. Nawata, K. Ochi, M. Shiba, K. Morita, Y. Iitaka, *Acta Crystallogr., Sect. B*, 37 (1981) 246–249.
[48] D. Avnel, A. Neuman, H. Gillier-Pandraud, *Acta Crystallogr., Sect. B*, 32 (1976) 2598–2605.
[49] J. Becquart, A. Neuman, H. Gillier-Pandraud, *Carbohydr. Res.*, 111 (1982) 9–21.
[50] Z. Ciunik, S. Jarosz, *Pol. J. Chem.*, 71 (1997) 207–212.
[51] G.A. Jeffrey, D.-b. Huang, *Carbohydr. Res.*, 247 (1993) 37–50.
[52] D.C. Rohrer, *Acta Crystallogr., Sect. B*, 28 (1972) 425–433.
[53] G.A. Jeffrey, D.-b. Huang, *Carbohydr. Res.*, 210 (1991) 89–104.
[54] G.A. Jeffrey, D.-b. Huang, *Carbohydr. Res.*, 206 (1990) 173–182.
[55] J.C. Hanson, LC. Sieker, L.H. Jensen, *Acta Crystallogr., Sect. B*, 29 (1973) 797–808.
[56] J.D. Olivier, L.C. Strickland, *Acta Crystallogr., Sect. C*, 40 (1984) 820–824.
[57] F.H. Allen, O. Kennard, *Chem. Des. Autom. News*, 8 (1993) 1 and 31–37.

[58] R.V. Krishnakumar, S. Natarajan, *Carbohydr. Res.*, 296 (1996) C1–C1.
[59] C.A. Accorsi, B. Bertolasi, V. Ferretti, G. Gilli, *Carbohydr. Res.*, 191 (1989) 91–104.
[60] SURFER, Golden Software, Golden, CO, USA.
[61] A.D. French, A.-M. Kelterer, G.P. Johnson, M.K. Dowd, C.J. Cramer, *J. Comput. Chem.*, submitted for publication.
[62] I. Tvaroška, J. Carver, *J. Phys. Chem.*, 99 (1995) 6234–6241.
[63] J.-H. Lii, N.L. Allinger, *J. Phys. Org. Chem.*, 7 (1994) 591–609.
[64] A.D. French, R.S. Rowland, N.L. Allinger, *ACS Symp. Ser.*, 430 (1990) 120–140.
[65] G. Barone, G. Della Gatta, D. Ferro, V. Piacente, *J. Chem. Soc., Faraday Trans.*, 86 (1990) 75–79.
[66] S.S. Xantheas, *J. Chem. Phys.*, 104 (1996) 8821–8824.
[67] W.F. van Gunseren, H.J.C. Berendsen, GROMOS, Groningen Molecular Simulation, Library Manual, Bomos, Nijengorgh 15, Groningen, The Netherlands, 1987.
[68] Kj. Rasmussen, *J. Mol. Struct.* (*Theochem.*), 395–396 (1997) 91–106.
[69] Kj. Rasmussen, *J. Carbohydr. Chem.*, 18 (1999) 789–805.
[70] H.-B. Bürgi, J.D. Dunitz, *Acta Crystallogr., Sect. B*, 44 (1988) 445–448.
[71] E. Bye, B. Schweizer, J.D. Dunitz, *J. Am. Chem. Soc.*, 104 (1982) 5893–5898.
[72] S.B. Engelsen, Kj. Rasmussen, *J. Carbohydr. Chem.*, 16 (1997) 773–788.
[73] L. Schleifer, H. Senderowitz, P. Aped, E. Tartakovsky, B. Fuchs, *Carbohydr. Res.*, 206 (1990) 21–39.
[74] I. Tvaroška, T. Bleha, *Tetrahedron Lett.* (1975) 249–252.
[75] S.A.H. Spieser, J.A. van Kuik, L.M.J. Kroon-Batenburg, J. Kroon, *Carbohydr. Res.*, 322 (1999) 264–273.
[76] A.D. French, A.-M. Kelterer, G.P. Johnson, M.K. Dowd, C.J. Cramer, in preparation.
[77] J. Shen, L.E. Lerner, *Carbohydr. Res.*, 273 (1995) 115–127.
[78] R.W.W. Hooft, J.A. Kanters, J. Kroon, in M. Mathlouthi, J.A. Kanters, G.G. Birch (Eds.), *Sweet-Taste Chemoreception*, Elsevier, London, 1992, pp. 11–19.
[79] C. Andersson, S.B. Engelsen, *J. Mol. Graph. Model.*, 17 (1999) 101–105.