9590 *J. Am. Chem. Soc.* **1994,** *116,* 9590–9595

Ab Initio-MIA and Molecular Mechanics Studies of the Distorted Sucrose Linkage of Raffinose

Christian Van Alsenoy,[†] Alfred D. French,*[,‡] Ming Cao,[§] Susan Q. Newton,[⊥] and Lothar Schäfer[§]

Contribution from the Department of Chemistry, University of Antwerp, B2610 Wilrijk, Belgium, Southern Regional Research Center, U.S. Department of Agriculture, P.O. Box 19687, New Orleans, Louisiana 70179, Department of Chemistry, University of Arkansas, Fayetteville, Arkansas 72701, and John Brown University, Siloam Springs, Arkansas 72761

Received August 2, 1993. Revised Manuscript Received June 20, 1994[⊗]

Abstract: A conformational energy map for a tetrahydropyran–tetrahydrofuran analogue of sucrose was calculated with ab initio quantum mechanics. Geometries were optimized at the HF/4-21G level, using the MIA approximation. The highest energy (with the exception of trichlorogalactosucrose) corresponding to a conformation of an observed crystal structure was for the sucrose moiety in crystalline raffinose. It was less than 1.5 kcal/mol above the local minimum occupied by sucrose, and calculations with larger HF basis sets converged at an energy difference of 2.9 kcal/mol. On the other hand, MM3(92) energies for the raffinose conformation are improbably high, 7 or 8 kcal/mol above the global minimum, whether or not the galactose residue is included in the calculation. Intra- and intermolecular MM3 forces in a miniature model of crystalline raffinose pentahydrate were unable to account for the observed conformation. Also, other observed crystalline sucrose conformations correspond poorly to low-energy regions on MM2, MM3, and other molecular mechanics surfaces. About 6 kcal/mol of the 8 kcal/mol difference in conformational energies comes from the torsion angles involving the anomeric center on the furanose ring. This suggests that even MM3, which has been very successful in modeling other disaccharides, miscalculates the energies of compounds that have overlapping anomeric sequences.

1. Introduction

Sucrose has been studied extensively because of its practical importance. Not only significant as a disaccharide, it is also the root molecule for raffinose (6-*O*-galactosyl-β-D-glucopyranoside-1↔2-β-D-fructofuranoside) and its extensions,[1] as well as polysaccharide molecules such as the fructans. Structural studies have used high-precision X-ray[2] and neutron[3] diffraction, NMR,[4–7] optical rotation,[8] and potential energy calculations.[4,9] The traditional variables of disaccharide analysis, the torsion angles ϕ and ψ that describe the conformation at the interresidue linkage (Figure 1), are among the points of interest in these studies.

As discussed in an early NMR and modeling study,[4] the interresidue linkage of sucrose was thought to be rigid, based partly on information from crystal structure determinations that showed two interresidue hydrogen bonds. However, their models incorporated rigid geometries of the individual glucose and fructose residues. That rigidity, useful in pioneering studies of disaccharides composed of pyranoid rings, may be less appropriate for disaccharides containing the more flexible furanoid rings.

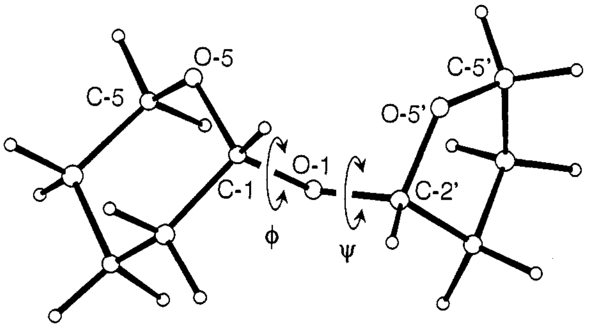

Figure 1. The tetrahydropyran–tetrahydrofuran analogue of sucrose, numbered as if it were sucrose. The torsion angles are ϕ = O(5)–C(1)–O(1)–C(2′) and ψ = O(5′)–C(2′)–O(1)–C(1).

Subsequent experimental work showed the sucrose linkage to be nearly as flexible as the linkage in another disaccharide, maltose. A survey of all the sucrose moieties in crystalline oligosaccharides[9–21] yields a 90° range in the ψ torsion angle, as well as a 40° range in ϕ. Newer NMR studies gave further evidence for a flexible sucrose linkage as well. Two conformations in 2:1 equilibrium dominate in DMSO solution,[5] and averaging over several conformations allowed a better reconciliation of NMR

[†] University of Antwerp.

[‡] U.S. Department of Agriculture.

[§] University of Arkansas.

[⊥] John Brown University.

[⊗] Abstract published in *Advance ACS Abstracts*, October 1, 1994.

(1) French, D. *Adv. Carbohydr. Chem.* **1954**, *9*, 149–184.

(2) Hanson, J. C.; Sieker, L. C.; Jensen, L. H. *Acta Crystallogr. Sect. B* **1973**, *29*, 797–808.

(3) Brown, G. M.; Levy, H. A. *Acta Crystallogr. Sect. B* **1973**, *29*, 790–797.

(4) Lemieux, R. U.; Bock, K.; Delbaere, L. T. J.; Koto, S.; Rao, V. S. R. *Can. J. Chem.* **1980**, *58*, 631–653.

(5) Christophodes, C.; Davies, D. B. *J. Chem. Soc., Chem. Commun.* **1985**, 1533–1534.

(6) du Penhoat, C. H.; Imberty, A.; Roques, N.; Michon, V.; Mentech, J.; Descotes, G.; Pérez, S. *J. Am. Chem. Soc.* **1991**, *113*, 3720–3727.

(7) Poppe, L.; van Halbeek, H. *J. Am. Chem. Soc.* **1992**, *114*, 1092–1094.

(8) Stevens, E. S. *Biopolymers* **1992**, *32*, 1571–1579.

(9) Ferretti, V.; Bertolasi, V.; Gilli, G. *Acta Crystallogr. Sect. C* **1984**, *40*, 531–535.

(10) Berman, H. M. *Acta Crystallogr. Sect. B* **1970**, *26*, 290–299.

(11) Jeffrey, G. A.; Park, Y. J. *Acta Crystallogr. Sect. B* **1972**, *28*, 257–267.

(12) Rohrer, D. C. *Acta Crystallogr. Sect. B* **1972**, *28*, 425–433.

(13) Avenel, D.; Neuman, A.; Gillier-Pandraud, H. *Acta Crystallogr. Sect. B* **1976**, *32*, 2590–2605.

(14) Becquart, J.; Neuman, A.; Gillier-Pandraud, H. *Carbohydr. Res.* **1982**, *111*, 9–21.

(15) Accorsi, C. A.; Bertolasi, B.; Ferretti, V.; Gilli, G. *Carbohydr. Res.* **1989**, *191*, 91–104.

(16) Accorsi, C. A.; Bellucci, F.; Bertolasi, B.; Ferretti, V. *Carbohydr. Res.* **1989**, *191*, 105–116.

(17) Jeffrey, G. A.; Huang, D.-b. *Carbohydr. Res.* **1990**, *206*, 173–182.

(18) Jeffrey, G. A.; Huang, D.-b. *Carbohydr. Res.* **1991**, *210*, 89–104.

(19) Taga, T.; Inagaki, E.; Fujimori, Y.; Nakamura, S. *Carbohydr. Res.* **1993**, *240*, 39–45.

(20) Okuyama, K.; Noguchi, K.; Saitoh, M.; Ohno, S.; Fujii, S.; Tskuda, M.; Takeda, H.; Hidano, T. *Bull. Chem. Soc. Jpn.* **1993**, *66*, 374–379.

(21) Jeffrey, G. A.; Huang, D.-b. *Carbohydr. Res.* **1993**, *247*, 37–50.

0002-7863/94/1516-9590$04.50/0

coupling constants and nuclear Overhauser effects for sucrose in aqueous solution.[6]

In crystal structure and modeling studies of the sucrose linkage in 6-kestose,[9] Ferretti *et al.* used a simple energy function with rigid residues. On their energy map, the conformation of the sucrose moiety in crystalline raffinose corresponds to an energy of nearly 100 kcal/mol! Tran and Brady used the CHARMM program and a potential function for carbohydrates[22] to study the sucrose energy surface by both flexible-residue conformational searching[23] and dynamics simulation.[24] The energy of the raffinose geometry on their surface is about 8 kcal/mol, confirming the predictions of Ferretti *et al.* that relaxation of residue geometry would be important in energy calculations.[9] A subsequent energy map made with MM3(90) showed the raffinose geometry at an energy of 5.5 kcal/mol.[25] The relaxed-residue energy map by Lichtenthaler *et al.*, made with PIMM88, places the raffinose geometry at nearly 10 kcal/mol,[26] comparable to a result by Hooft, who used a version of MM2 modified for hydrogen bonding.[27] (Although the relative energies for the raffinose conformation are similar on these various energy maps, the maps do have considerable differences.)

Raffinose is not the only molecule with a conformation of its sucrose moiety that corresponds to a molecular mechanics energy that we consider to be improbably high. Nystose,[20,21] a tetrasaccharide, and the trisaccharide planteose[12] also had corresponding MM3 energies above 3.0 kcal/mol.[28] Even though these energies are much lower than those in the rigid-residue 6-kestose work, all are higher than those found for any other disaccharide by MM3(90)[29–32] or MM3(92).[33] In studies of maltose linkages, the highest energy observed for 39 crystalline examples was less than 3.0 kcal/mol,[33] and MM3 studies of other sugar linkages have discrepancies no more than 2 kcal/mol from the minimum energy. This was true regardless of substitution or complex formation.

The abnormally high energies in the various relaxed-residue molecular mechanics studies of sucrose, then, lead to several possible conclusions. First, all of the potential energy functions may contain substantial error, at least for the sucrose linkage. Because this difficulty with the energies is associated with rotation about the bonds of the glycosidic linkages, one must consider the possibility of an anomeric effect.[34] Initially, the anomeric effect was defined as the unexpected importance for axial substituents at the glycosidic carbon of sugar rings, compared to cyclohexane. That is, the torsion angle C(5)–O(5)–C(1)–O(1) (Figure 1) preferred values of about 60° instead of 180°. Later, the C–O bond lengths at these anomeric centers were found to vary. Next, an exo-anomeric effect was proposed that leads to a preference for positions of the methyl glycoside to be *gauche* (e.g. O(5)–C(1)–O(1)–C(2′) = ±60°) to the ring oxygen atom. From the viewpoint of molecular mechanics, the parametrization of the C–O–C–O sequence should, therefore, account for both the anomeric and exo-anomeric effects. In the development of MM3, ab initio calculations at the HF/4-21G level were used to parametrize the calculated energies of the C–O–C–O sequence. Also, corrections are introduced for the changes in bond lengths and angles as the sequence takes different torsion angles. In the case of sucrose, a possible cause for error is the extended sequence of C–O bonds that creates the possibility of an overlapping exo-anomeric effect that has not been considered during the parametrization of molecular mechanics force fields. This effect would require torsional energies to be determined by the geometry of a string of atoms longer than the four used to define a torsion angle.

Secondly, intramolecular forces could result in the distorted conformation. They could arise from the galactose residue of raffinose or from the exact combination of primary and secondary hydroxyl group orientations found in crystalline raffinose. The explicit arrangements of side groups are usually not used when constructing energy maps.[35] Instead, lower energy schemes with clockwise and counterclockwise orientations of hydroxyl groups are used. A third possibility is that intermolecular forces in the raffinose crystal distort the molecule. Such distortions are potentially large, given that lattice energies for carbohydrates are as large as 38 kcal/mol of glucose residues.[33] Of course, any of these factors could combine to give the high calculated energies for the raffinose conformation.

Previously, we proposed[36] that an overlapping anomeric effect might cause incorrect molecular mechanics energies for sucrose linkages. In that work, an analogue of sucrose, tetrahydro-2-[(tetrahydro-2-furanyl)oxy]-2*H*-pyran (Figure 1), was studied by MM3(90) and six ab initio optimizations at the HF/4-21G level. Those calculations, while insufficient to establish an energy surface, showed that the relative energy for the raffinose conformation might be much smaller than found with molecular mechanics.

In the present work, we investigate further the above three possibilities. Intramolecular interactions were studied with MM3(92), which incorporates improved hydrogen bonding potential functions for new studies of the sucrose linkage, along with increasing the number of side group orientations that were considered. We also constructed energy maps for the sucrose linkage that explicitly include the side group orientations and the galactosyl residue of crystalline raffinose. Intermolecular forces were studied in raffinose minicrystals, and the analogue was studied anew with MM2(77), MM3(92), and ab initio calculations. The resulting relaxed-residue ab initio energy surface, using d-orbitals on the oxygen atoms and the 4-21G basis set, may be the first for a disaccharide analogue. Also, three conformations of the tetrahydropyran analogue of the disaccharide α,β-trehalose were optimized with the ab initio method to learn if the low energies for changes in sucrose are unique to nonreducing disaccharides that contain furanose rings. Finally, we obtained convergence in the relative energies of the four most important conformations with some advanced basis sets. This latter work was done with a parallelized version of the BRABO software for ab initio calculations, running on several workstations (see Computations).

2. Results

An adiabatic, relaxed-residue MM3 map for sucrose based on 48 starting conformations is shown in Figure 2. Figure 3 shows similar MM3 maps that have the starting geometry based on crystalline raffinose.[17] Figure 3a shows the energy surface for just the sucrose moiety, while Figure 3b is for variation of the sucrose moiety but with the galactose residue present. The similarity of parts a and b of Figure 3 shows that presence or absence of the galactose residue has a negligible effect on the

---

(22) Ha, S. N.; Madsen, L. J.; Brady, J. W. *Biopolymers* **1988**, *27*, 1927–1952.

(23) Tran, V. H.; Brady, J. W. *Biopolymers* **1990**, *29*, 961–976.

(24) Tran, V. H.; Brady, J. W. *Biopolymers* **1990**, *29*, 977–997.

(25) Pérez, S.; Meyer, C.; Imberty, A.; French, A. D. In *Sweet Taste Chemoreception*; Birch, G. G., Mathlouthi, M., Kanters, J., Eds.; Elsevier: New York, 1993; pp 55–72.

(26) Lichtenthaler, F. W.; Immel, S.; Kreis, U. *Stärke* **1992**, *43*, 121–132.

(27) Hooft, R. W. W. Ph.D. Thesis, University of Utrecht, Utrecht, The Netherlands, 1993.

(28) French, A. D.; Mouhous-Riou, N.; Pérez, S. *Carbohydr. Res.* **1993**, *247*, 51–62.

(29) Dowd, M. K.; Reilly, P. J.; French, A. D. *J. Comput. Chem.* **1992**, *13*, 102–114.

(30) Dowd, M. K.; French, A. D.; Reilly, P. J. *Carbohydr. Res.* **1992**, *233*, 15–34.

(31) Dowd, M. K.; Zeng, J.; French, A. D.; Reilly, P. J. *Carbohydr. Res.* **1992**, *230*, 223–244.

(32) Dowd, M. K.; Reilly, P. J.; French, A. D. *J. Carbohydr. Chem.* **1993**, *12*, 449–457.

(33) French, A. D.; Dowd, M. K. *J. Mol. Struct.* (*Theochem*) **1993**, *286*, 183–201.

(34) Tvaroška, I.; Bleha, T. *Adv. Carbohydr. Chem.* **1989**, *47*, 45–123.

(35) French, A. D.; Tran, V.; Pérez, S. *ACS Symp. Ser.* **1990**, *430*, 191–212.

(36) French, A. D.; Newton, S. Q.; Schäfer, L. *Carbohydr. Res.* **1993**, *239*, 51–60.

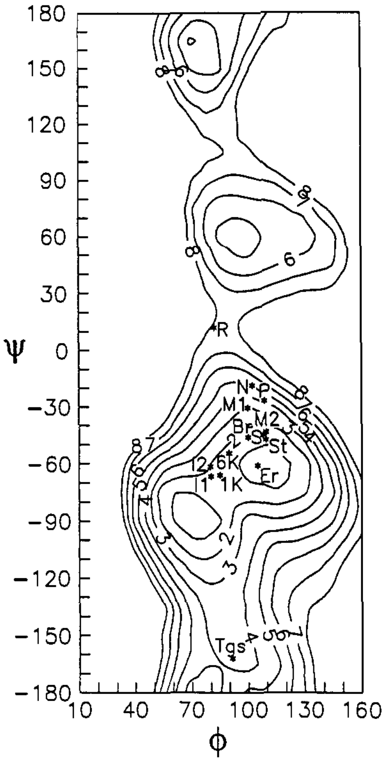

Figure 2. Energy surface for sucrose, computed with MM3(92), dielectric constant of 3, and 48 starting geometries that included various combinations of staggered orientations of the primary and secondary hydroxyl side groups. Plotted on the surface are conformations in single crystal structures. The letter codes correspond to the following: R, raffinose;[17] N, nystose;[21] P, planteose;[12] M1, melezitose 1;[13] M2, melezitose 2;[14] S, sucrose;[3] Br, NaBr sucrose complex;[16] St, stachyose;[18] 6K, 6-kestose;[9] Er, erlose;[19] 1K, 1-kestose;[11] I1 and I2, NaI sucrose complex;[15] and Tgs, Trichlorogalactosucrose.[39] Contours are at each kcal/mol above the global minimum.

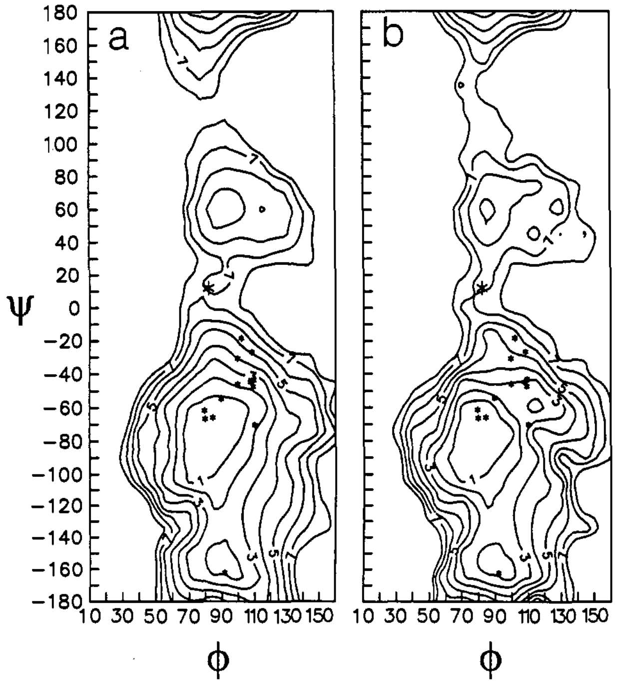

Figure 3. MM3 energies based on the geometry found in crystalline raffinose. (a) Energy of the sucrose moiety. A hydroxyl hydrogen atom was added to O(6) of the glucose moiety after the galactose–glucose bond was broken and the galactose residue removed. (b) Energies for the raffinose trisaccharide, with only ϕ and ψ contrained. Contours at 1 kcal/mol.

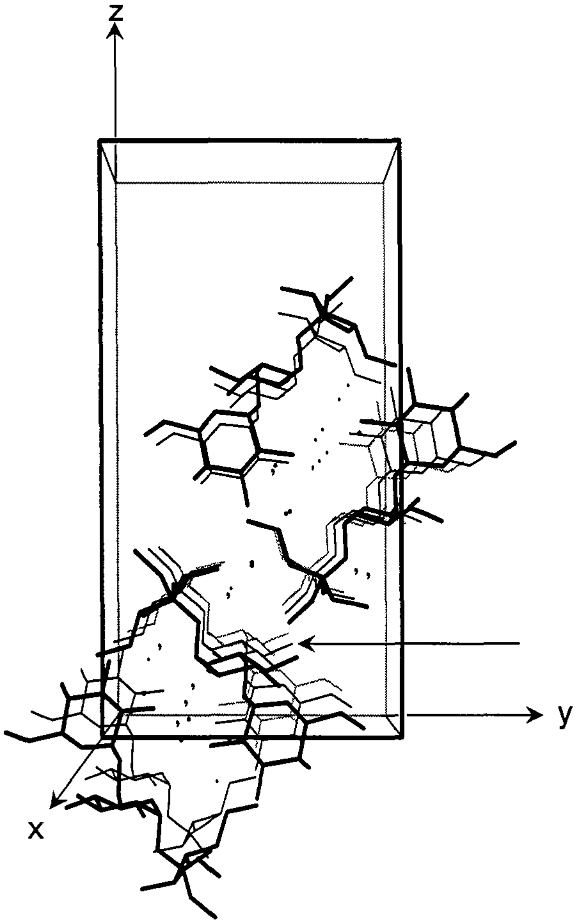

Figure 4. Miniature crystal model of raffinose pentahydrate, drawn in perspective with CHEM-X. Ten raffinose and 30 water molecules are present, all generated from the published coordinates by $P2_12_12_1$ symmetry operators. The dielectric constant was 3.5. Compared to other oligomeric carbohydrate crystal structures, the channel of water molecules is unusual. As shown in Table 1, the average movement from these original positions was fairly large, even for the "central" molecule, shown with the arrow.

Table 1. Mean Movement (Å) of Atoms in Minicrystals during Optimization

	movement
11-raffinose model, no water	
overall	0.643
central molecule	0.545
10-raffinose model, 30 waters	
overall	0.434
central molecule	0.396

conformational energy of the sucrose linkage. Also, the similarities of Figures 2 and 3 show that even the molecular mechanics model of the exact intramolecular interactions found in the crystal cannot account for the observed conformation. Sucrose maps (not shown) calculated with various dielectric constants (ϵ) from 1.0 to 8.0 were very similar to Figures 2 and 3, except for one computed at $\epsilon = 1$. That map corresponded to the observed crystal structures more poorly than the others.

The results from a miniature crystal (Figure 4) study are given in Table 1. The presence of water molecules in the miniature crystal model decreased the intramolecular changes in the carbohydrate during minimization. However, even with the water molecules, the miniature crystal model did not fix the distorted linkage conformation, unlike the results of similar studies on maltose[33] and gentiobiose.[37] The tabulated overall mean atomic movements are substantially higher than values around 0.2 Å seen for other carbohydrates.[33,38] Unlike those studies, the movement in the central raffinose molecule was nearly equal to that of the surface molecules. In similar work on maltose, the mean atomic movement in the central molecule was much lower, about 0.06 Å. After raffinose, the next largest movement in minicrystals was for nystose.[28] In both cases, the ψ torsion angles

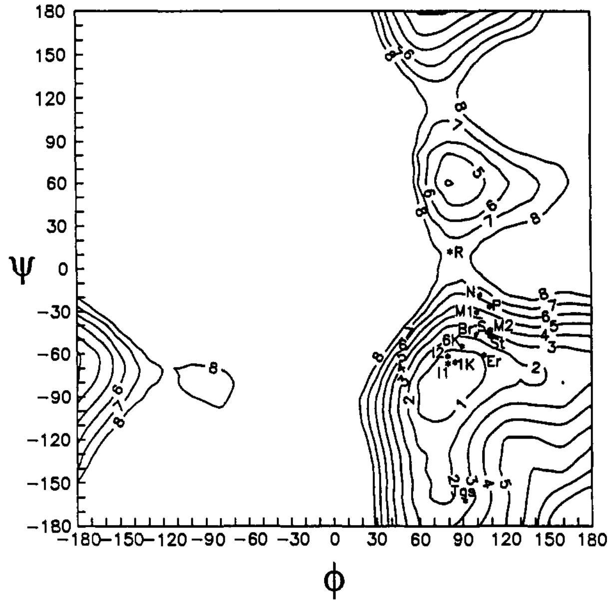

Figure 5. MM3 energies for the sucrose analogue shown in Figure 1. The value of ε = 1.5 was used to best mimic the ab initio calculations. Letter codes are in the caption of Figure 2.

Table 2. Components of MM3(92) Energy for the Sucrose Analogue (kcal/mol)

overall							
φ	ψ	ε	total	vdW 1,4	vdW other	dipole	torsion
90	10	4.0	28.61	10.20	0.16	-2.04	10.76
84	-69	4.0	19.58	10.41	-0.79	-1.55	4.09
90	10	1.5	25.21	10.30	0.30	-5.48	10.84
84	-69	1.5	16.97	10.47	-0.80	-4.20	4.16

Torsional Energies Differing by More than 0.4 kcal/mol at ε = 1.5

	φ = 84.2, ψ = -69.0	φ = 90.0, ψ = 10.0
C(1)-O(1)-C(2′)-O(5′)	-1.719	1.944
C(1)-O(1)-C(2′)-C(3′)	0.005	0.890
C(1)-O(1)-C(2′)-H(C(2′))	0.014	0.619
C(5′)-O(5′)-C(2′)-O(1)	-1.874	-1.303
C(4′)-C(3′)-C(2′)-O(5′)	0.862	1.358
sums	-2.712	3.508
difference of sums	6.220	

optimized toward the value found for the isolated sucrose molecule. Work with a larger minicrystal of raffinose hydrate (1976 atoms) did restrict the atomic movement during energy minimization to a greater extent, but the movement of the central molecule was still greater than in the small minicrystals of other carbohydrates.

Figure 5 shows the MM3 relaxed energy map for the analogue in Figure 1. Except for the extension of its low-energy area from the global minimum in the direction of increased φ, it is similar to Figure 2. The dissimilar area arises because the analogue lacks the C(1′) primary methylol group and O(2) hydroxyl group of sucrose. A map made with MM2(77) for the analogue (not shown) gives a relative energy for raffinose of 6 kcal/mol, somewhat lower than the MM3 map.

Table 2 breaks down the MM3 total steric energy for a conformation near the minimum of Figure 5, and for a point near the raffinose linkage conformation. The energies of the analogue are somewhat sensitive to dielectric constant, even though the analogue lacks hydroxyl groups. Despite differences in "dipole–dipole" and "van der Waals other" energies for the two conformers, the torsion angle terms cause the main energy difference between the two conformations. The lower part of Table 2 shows the torsional energies that vary the most. All individual energy differences greater than 0.4 kcal/mol arise from torsion angles that contain the anomeric carbon on the furanose ring, C(2′).

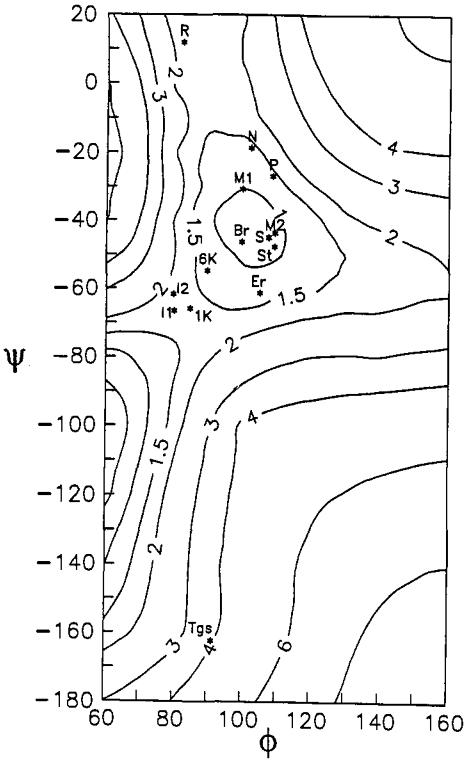

Figure 6. Ab initio energy map for the sucrose analogue, with the observed crystal structures. The HF/4-21G basis set was used, with d-functions for the oxygen atoms. Increments were 20°, and contours are at 0.5, 1, 1.5, 2, 3, 4, 6, and 8 kcal/mol. Letter codes are defined in the caption of Figure 2.

Further, at ε = 1.5, the sum of those 5 torsional energies corresponds to 6.22 kcal of the total difference of 8.23 kcal/mol. (The torsional energies were about the same at either dielectric constant; 1.5 was chosen for similarity to the ab initio results.)

Figure 6 shows the ab initio energy surface for the sucrose analogue, covering the area wherein crystalline sucrose moieties are found. It is very different in character from the molecular mechanics maps. Use of d-functions in the present work changed the location of the minimum on the ab initio map relative to the minima reported earlier[36] at φ = 84.2°, ψ = -69.2° and at φ = 129°, ψ = -70°. The new minimum (φ = 102°, ψ = -42°) corresponds closely to the conformation observed for crystalline sucrose and is about 0.5 kcal/mol above the overall minimum. Except for trichlorogalactosucrose[39] (Tgs), the crystal structure conformations plotted on this map are all within 1.5 kcal of the local minimum.

Compared to the molecular mechanics studies, this ab initio relaxed map predicts a much lower energy for the distortion of the sucrose linkage in crystalline raffinose. Evidence that the ab initio values are more correct than the MM3 values comes from the distortion energy. It seems that the raffinose crystal would not nucleate and grow, given an energy preference for a staggered conformation of 8 kcal/mol. The fits of the models can also be judged on the basis of the root mean square (rms) predicted energies of experimental structures. On the MM3 sucrose surface, the 13 observed linkages (Tgs excepted) have an rms energy of 3.2 kcal/mol. This contrasts sharply with the rms MM3 energy of 1.19 kcal/mol[33] for 39 maltose linkages, a number that we consider satisfactory. On the ab initio surface for the analogue, the rms energy (relative to the local minimum and excepting Tgs) is just 0.82 kcal/mol. Our approach implicitly assumes that such distortions by crystal fields will, given a sufficient number

---

[37] Dowd, M. K.; Reilly, P. J.; French, A. D. *Biopolymers* **1994**, *34*, 625–638.

[38] French, A. D.; Aabloo, A.; Miller, D. P. *Int. J. Biol. Macromol.* **1993**, *15*, 30–36.

[39] Kanters, J. A.; Scherrenberg, R. L.; Leeflang, B. R.; Kroon, J.; Mathlouthi, M. *Carbohydr. Res.* **1988**, *180*, 175–182.

Table 3. Conformational Energies for Three Conformations of the Tetrahydropyran Analogue of α,β-Trehalose (4-21G Basis Set)

ϕ	ψ	rel energy (kcal/mol)
84	12	3.037
84	−57	2.182
57	−97	0.00 (freely optimized)

of crystal structures, be randomly distributed. The structure having the minimum in energy as an isolated molecule (with a dielectric constant appropriate for a condensed phase) would be most probable. Thus, it is also important to examine the distribution of conformations of relevant molecules on a conformational energy surface instead of looking only at a specific molecule. The crystal structures occur in a random manner about the local minimum on the ab initio map but are all on one side of the MM3 maps.

Ab initio calculations on the tetrahydropyran analogue of α,β-trehalose are summarized in Table 3. There is less than 1 kcal/mol difference in the energies of the conformations that mimic the second NaI–sucrose and raffinose conformations. This is quite parallel to the results in Figure 6 for the sucrose analogue, and it indicates that the energies for overlapping anomeric sequences of pyranose–furanose and pyranose–pyranose linkages are not very different from each other. The conformation from free optimization is almost the same as the overall minimum on Figure 6, albeit at slightly lower relative energy.

From Figure 6, we deduced that four conformations (ϕ, ψ = 82, 12; 102, −44; 83, −68; 60, −100) would be of special interest for calculations with advanced basis sets. The first and third points correspond to the experimental raffinose conformation and the barrier between the global and local minima, and their linkage torsion angles were held fixed during optimization. The second and fourth points correspond to the local and global minima, and those conformations were allowed to optimize fully. Table 4 shows that the split valence basis sets gave energies for the raffinose conformation about 2.1 kcal/mol above the local minimum and that the triple-ζ basis sets gave slightly higher energies, about 2.9 kcal/mol. Although the location of the local minimum changed, from ϕ, ψ = 102, −44 to 101, −58, its position is still consistent with the observed distribution of crystalline conformations. The global minimum was stationary.

3. Discussion

Application of MM3 to carbohydrates has routinely given reasonable results. Therefore, apparently erroneous results, like those obtained in the present MM3 analysis of sucrose, are of general interest because they show that our knowledge of intramolecular interactions, as expressed by the potential energy functions of the model, is incomplete. In the present case, the MM3 torsional parameters for the atoms in non-reducing, pseudoequatorial glycosidic linkages from furanose rings almost certainly yield errors totalling several kilocalories per mole. The conclusion that the MM3 results are erroneous is based on both the high calculated energy of several observed conformations and the non-random distribution of observed crystalline conformations on the MM3 energy surface. Other explanations for the majority of the high MM3 energy of the raffinose conformation such as specific intramolecular effects and crystal field effects are rejected, at least within the ability of MM3 to model them. Because the MM3 energies for the torsional component of the conformational energy at the anomeric center of fructose are incorrect, even though they appear to be correct for many other molecules, including most disaccharides, we again propose that the discrepancy arises from an "overlapping anomeric effect". In the parlance of molecular mechanics, this means that the conformational energies in long sequences of alternating carbon and oxygen atoms may depend on more than just the four atoms currently used to define torsional energies.

Table 4. Advanced Basis Set Energies for the Sucrose Analogue[a]

basis set	ϕ	ψ	energy
SVP-d5	82.00	12.00	2.87
	101.27	−53.69	0.80
	83.00	−68.00	1.12
	60.25	−99.62	0.00
SVP-d6	82.00	12.00	2.95
	101.06	−54.30	0.79
	83.00	−68.00	1.07
	60.80	−99.12	0.00
TZP-d5	82.00	12.00	3.49
	101.30	−57.80	0.64
	83.00	−68.00	0.98
	61.95	−98.83	0.00
TZP-d6	82.00	12.00	3.49
	101.28	−57.84	0.64
	83.00	−68.00	0.98
	62.00	−98.82	0.00
TZVP-d5	82.00	12.00	3.52
	101.33	−58.18	0.67
	83.00	−68.00	0.99
	62.13	−98.94	0.00

[a] Conformations with ϕ, ψ = 82, 12 and 83, −68 were fixed.

The conclusion that the molecular mechanics analyses are defective is strengthened by the ab initio relaxed map for the analogue, even though its underlying calculations were at a relatively low level of theory. This compromise, the use of an analogue, and restriction of the area of ϕ, ψ space were all required by limits of available computational resources. To further buttress our point, the four most important points on the ab initio map were computed with the best basis sets that we could afford (up to 336 basis functions for the four optimized structures). These calculations converged and suggested that the distortion of the sucrose linkage in raffinose is less than 3.0 kcal/mol, consistent with our molecular mechanics studies of other disaccharides.

The position of the apparent global minimum in energy is a remaining mystery. Its ϕ, ψ coordinates, as determined by all five higher levels of theory, correspond to very low (<1.0 kcal/mol) energies on all our relaxed maps, including the MM3 studies. Yet, no crystal structures fall in this area. Instead, all of the structures, with the exception of Tgs, are clustered about an adjacent local minimum on the ab initio map. It may be that the correct intrinsic, relative energy of the apparent global minimum is as high as 2.0 kcal/mol in condensed phases. That energy would suffice to lower the probability that crystal structures would be found in the region. There may, however, also be an overall structural reason that prevents crystal packing in this region, and it may even just be luck that no crystal structure has been found there. We have no strategy to resolve such a relatively small energy discrepancy based on the absence of observed structures.

Quantum mechanical and molecular mechanical energies will generally not be the same with available modeling systems. However, the current observed discrepancy is important because the anomeric torsional constants for MM3 were developed with ab initio calculations at the 4-21G level, the same as used for Figure 6 except that we used d-functions on oxygens. Regardless of the validity of the 4-21G energies, the disagreement of MM3 and 4-21G energies suggests that there can be important further interactions in larger molecules that are not incorporated in current force fields. A comparable situation exists for a new force field based on quantum mechanics. On the sucrose map calculated with their force field,[40] the raffinose conformation has a relative energy 6 kcal above the local minimum. Their HF/6-31G* calculations are roughly comparable to the TZP-d5 basis set used here, which showed a 2.9 kcal difference. Both admittedly imperfect analogies suggest that the difference in results between these molecular and quantum mechanics calculations arises because torsional energies in state-of-the-art molecular mechanics

---

(40) Ewig, C. S.; Hwang, M.-J. Biosym Technologies, Inc., Personnal communication.

programs depend only on the four atoms that define the torsion angles. It remains to be seen whether improved parameters for torsional energies based on four atoms can overcome the above discrepancies, or whether it is ultimately necessary to expand the molecular mechanics potential function.

Only the Tgs structure[39] has an unusually high energy ($\approx$3.9 kcal/mol) on the ab initio surface, and it may be that the analogue is a poor model for that molecule. Neither of the available molecular mechanics studies of Tgs show its crystalline conformation at the global minimum.[41,42] However, Tgs is near a local minimum in those studies and in the CHARMM,[23] PIMM88,[26] and present MM3 maps for sucrose.

The MM3(92) energy map for sucrose (Figure 2) shows a greater difference between the minimum energy and the energy for the raffinose conformation (8 kcal/mol) than its MM3(90) predecessor (5.5 kcal/mol) for several reasons. The new global minimum conformation has several intramolecular hydrogen bonds that are strengthened by improvements in the 1992 version of the program. Also, hydrogen bonding was strengthened by using a dielectric constant of 3.0, instead of the value of 4.0 used before. Finally, the starting geometry that yielded the minimum energy was not part of the original set. The part of the $\approx$2.5 kcal/mol increase in the relative energy that came about because of the new starting geometry is an ever-present danger of modeling studies that do not search all possible conformations exhaustively. The earlier map was based on 17 starting geometries, instead of the 48 used in this work. Both numbers are small fractions of the 177 147 possible combinations of staggered side-group orientations for the disaccharide, so we still cannot be absolutely certain that we have found the global minimum on the MM3 energy surface. However, a lower-energy global minimum would only serve to increase the relative energy of the raffinose conformation, since its exact orientations of primary and secondary hydroxyl groups were studied herein.

The use of 48 starting geometries in the MM3(92) study brought the sucrose map into closer conformity with the analogue map, compared to the MM3(90) map. Similarity between an analogue map computed with the semiempirical quantum chemical program PCILO[43] and the MM3(90) map for α,α-trehalose was noted earlier,[29] as were similarities between MM3 maps for glucose–glucose disaccharides and their respective tetrahydropyran analogues.[29–31] The MM3(92) map for the sucrose analogue in the present study is also very similar to the MM3(90) map for the analogue of α,β-trehalose.

4. Conclusions

Relaxed-residue conformational mapping based on ab initio studies of a sucrose analogue at the 4-21G level accounted for all observed conformations of sucrose moieties in the various relevant crystal structures except Tgs, giving a rms energy of 0.82 kcal/mol above the local minimum. This gives a much better rationalization of the observed crystal structures than provided by previous work, all of which was based on molecular mechanics studies. For example, energies of the crystal structures had a rms MM3 energy of 3.2 kcal/mol. Further, according to MM3 and other force fields, relaxed-residue energies for the conformation of the sucrose moiety in crystalline raffinose are from 6 to 10 kcal/mol, much higher than seen for other types of glycosidic linkages. The same conformation had ab initio energies of 1.25 to 2.9 kcal. Not only were the energies of the crystal structures much lower on the ab initio surface, but it accommodated the distribution of observed structures very well. These results confirmed that the raffinose geometry does not correspond to a high potential energy. Finally, ab initio studies for a tetrahydropyran analogue of α,β-trehalose gave similar energy values to the sucrose analogue. Thus, most of the error in the relaxed-residue molecular mechanics energy predicted for torsional changes in sucrose molecules does not arise from the presence of a furanose ring but appears to arise from the absence of molecules with overlapping anomeric sequences in the set of structures used to parametrize molecular mechanics force fields.

5. Computations

The ab initio calculations were executed with the program BRABO[44] on IBM RS/6000 workstations. HF geometry optimizations were performed with the 4-21G basis set[45] with d-functions on the oxygen atoms only. In previous studies,[46] this basis set gave fairly accurate energies for compounds of this type. A total of 66 points (20° increments of ϕ and ψ) were energy minimized, constraining the linkage torsions at the increment values.

In addition to the HF/4-21G* calculations, more advanced RHF geometry optimizations were performed for several selected conformations (see above) using the TZVP, TZP, and SVP bases from Schäfer et al.[47] These basis sets consist of the following functions. TZVP: O, C, (11s,6p,1d)/[5s,3p,1d] with contraction scheme {62111/411/1}; H, (5s,1p)/[3s,1p] with contraction scheme {311/1}. TZP: O, C, (10s,6p,1d)/[6s,3p,1d] with contraction scheme {511111/311/1}; H, (5s,1p)/[3s,1p] with contraction scheme {311/1}. SVP: O, C, (7s,4p,1d)/[3s,2p,1d] contracted to {511/31/1}; H, (4s,1p)/[2s,1p] contracted to {31/1}.

The program BRABO was modified to allow for parallel calculations. Part of the ab initio calculations for this study were performed in parallel on three IBM RS/6000 (type 375, 365 and 320H) workstations connected with thin ethernet using the PVM software by Geist et al.[48]

Standard versions of MM2(77) and MM3(92)[49,50] ran on VAX computers. The sucrose maps were calculated at increments of 15° in ϕ and ψ, 20° for the analogue. Various dielectric constants were used as described, and plotted sucrose energies were the lowest of the 48 values for each ϕ,ψ pair. Energy minimization terminated at the default criterion of 0.00008 kcal per atom per iteration. Dihedral driver option 4 was used for MM3(92) conformational searching. That driver starts all optimizations with rigid rotations from the same initial structure. The miniature crystals of raffinose were constructed with CHEM-X[51] and optimized with MM3. Additional graphical manipulation of some of the surface hydroxyl hydrogen atoms was required to obtain a final result that conformed to both the energy-based termination criterion and our criteria of 0.00008 Å mean movement per iteration and 0.0018 Å maximum atomic movement. The overall mean atomic movements, calculated with CHEM-X, included the carbon and oxygen atoms of the carbohydrate and water molecules. Contour plots were constructed with SURFER.[52] Interpolated rms energy values were obtained with the "SLICE" utility in the SURFER package. This utility interpolates the dependent variable at points supplied by the user, such as the ϕ,ψ values of the crystal structures. It also yields an interpolated value at each intersection with a grid line.

Acknowledgment. C.V.A. thanks the Belgian National Science Foundation for an appointment as Onderzoeksleider. D. P. Miller devised the small miniature crystal, and C. J. Cramer and N. L. Allinger contributed helpful discussions.

---

(41) Hooft, R. W. W.; Kanters, J. A.; Kroon, J. *J. Comput. Chem.* **1991**, *12*, 943–947.

(42) Meyer, C.; Pérez, S.; du Penhoat, C. H.; Michon, V. *J. Am. Chem. Soc.* **1993**, *115*, 10300–10310.

(43) Tvaroška, I.; Václavík, L. *Carbohydr. Res.* **1987**, *160*, 137–149.

(44) (a) Van Alsenoy, C. *J. Comput. Chem.* **1988**, *9*, 620–626. (b) Van Alsenoy, C.; Peeters, A. *J. Mol. Struct.* **1993**, *286*, 19–34.

(45) Pulay, P.; Fogarasi, G.; Pang, F.; Boggs, J. E. *J. Am. Chem. Soc.* **1979**, *101*, 2550–2560.

(46) Frey, R. F.; Coffin, J.; Newton, S. Q.; Ramek, M.; Cheng, V. K. W.; Momany, F. A.; Schäfer, L. *J. Am. Chem. Soc.* **1992**, *114*, 5369–5377.

(47) Schäfer, A.; Horn, H.; Ahlrichs, R. *J. Chem. Phys.* **1992**, *97*, 2571–2577.

(48) Geist, A.; Beguelin, A.; Dongarra, J.; Jiang, W.; Manchek, R.; Sunderam, V. *PVM3 User's Guide and Reference Manual*; ORNL/TM 12818, Oak Ridge National Laboratory: Oak Ridge, TN, 1993.

(49) Allinger, N. L.; Yuh, Y. H.; Lii, J.-H. *J. Am. Chem. Soc.* **1989**, *111*, 8551–8566.

(50) Allinger, N. L.; Rahman, M.; Lii, J.-H. *J. Am. Chem. Soc.* **1990**, *112*, 8293–8307.

(51) CHEM-X is developed and distributed by Chemical Design, Ltd.: Oxfordshire, UK.

(52) SURFER, for IBM-PC compatibles, is available from Golden Software: 809 14th St., Golden, CO 80401-1866.