Molecular Modeling of Saccharides, 7[◇]

The Conformation of Sucrose in Water: A Molecular Dynamics Approach

Stefan Immel and Frieder W. Lichtenthaler*

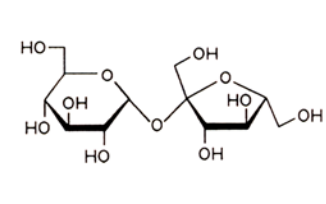
Institut für Organische Chemie der Technischen Hochschule Darmstadt,
Petersenstraße 22, D-64287 Darmstadt, Germany

Received June 2, 1995

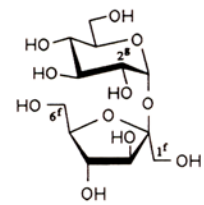
Key Words: Sucrose / Molecular mechanics and dynamics / Conformational properties / Hydration / Free energy calculations

A molecular mechanics analysis of the conformational properties of sucrose in vacuo in terms of the intersaccharidic torsion angles Φ and Ψ, revealed three energy minima. The geometry of the global minimum-energy closely resembles the solid-state structure. Most notably, the interresidue hydrogen bonding interaction 2^g-O···HO-1^f present in the crystal, is retained under vacuum boundary conditions, indicating the molecular geometries adopted in the crystal lattice and in vacuo to be similar. For aqueous solutions, detailed molecular dynamics simulations of sucrose "soaked" with 571 water molecules in a periodic box (truncated octahedron), revealed this direct H-bond interaction to be replaced by an indirect, water-mentioned one: an interresidue water-bridge of the 2^g-O···H_2O···HO-1^f type prevailed with a high significance and a long life-time. This means the linkage geometry of sucrose in water – despite the absence of direct interresidue hydrogen bonds – again closely resembles the solid-state and in vacuo geometry in terms of the orientation of the glucose and the fructose unit relative to one another. The solution dynamics of, and the hydration around sucrose were analyzed in terms of pair distribution functions. These indicate strong hydrogen bonding between all sucrose hydroxyls (as donors and acceptors) and water within a first, well-defined hydration layer (hydroxyl-oxygen – water distances 1.8–3.5 Å), whereas the acetalic oxygens are engaged to a lesser extent as H-bond acceptors. The second hydration shell (>4 Å) is rather diffuse and less pronounced, indicating those water molecules to be in a disordered state. The implications of the hydration shell and the water bridge on the crystallization process of sucrose and on binding towards transporter proteins, and the sweet-taste receptor, are discussed. Other sucrose conformations that may conceivably exist in aqueous solution, may have eluded the MD simulation search. The umbrella sampling technique was applied for establishing the free energy profile as a function of the intersaccharidic torsion angles. The resulting concise picture of the dynamics of sucrose in aqueous solution, encompassing the entire conformational space available, revealed only two energy minima. Of these, the by far, most populated global minimum structure corresponded to the most stable solution geometry, as found by molecular dynamics.

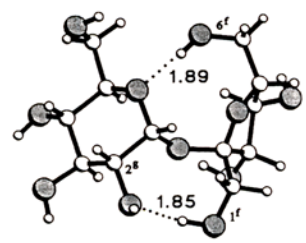
Sucrose, "the royal carbohydrate"[2], has for centuries been the world's most abundantly produced organic compound, and this in unparalleled purity. However, it was not until about 1930 – forty years after the unravelment of the sugar configurations by Emil Fischer[3] – that its chemical structure as depicted in **1a** was finally established[4]. Although this formula and its conformationally more refined graphic representation, **1b**, were sufficient to illustrate the chemistry of sucrose[5,6], they added little to the understanding of its chemical and biological properties. A more "realistic" molecular picture of sucrose, albeit for the solid state, emerged from neutron diffraction[7] and X-ray analysis[8]: the glucose portion adopts a $^4C_1$ conformation, the fructofuranose residue is in the $^4T_3$ twist form, and the overall structure (cf. ball-and-stick model **1c** and its graphic translation **1d**) is determined by two interresidue intramolecular hydrogen bonds[9], namely, 2^g-O···HO-1^f and 5^g-O···HO-6^f, both comparatively short as evidenced by O···H distances of 1.85 and 1.89 Å, respectively.



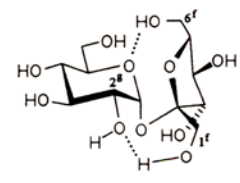
1a



1b



1c



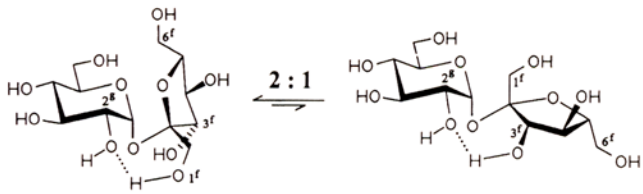
1d

The solution conformation of sucrose is surmised to strongly depend on the nature of the solvent, as the prob-

[◇] Part 6: Ref.[1].

  0947–3440/95/1111–1925 $ 10.00+.25/0

ability to disrupt one or both of the intramolecular hydrogen bonds present in the crystal is considerably higher in water than in polar aprotic solvents such as dimethyl sulfoxide (DMSO) or *N*,*N*-dimethyl formamide, which due to their hydrogen bond acceptor properties are apt to form different solvation shells. For DMSO solutions of sucrose, SIMPLE ^{1}H-NMR spectroscopic data[10] suggested the occurrence of two forms with competitive intramolecular hydrogen bond interactions, namely 2^g-O···HO-1^f (**1e**) and 2^g-O···HO-3^f (**1f**), the equilibrium favoring the former by about 2:1.



1e **1f**

In *N*,*N*-dimethyl formamide, the solvation shell around sucrose appears to be quite similar, inasmuch as deprotonation by chemical means (NaH)[11,12] or electrochemically[13] generates the 2^g-*O*-alkoxide with high preference, obviously due to the persistence of an intense 2^g-O···HO-1^f (or ···HO-3^f) intramolecular hydrogen bond, its back-donation effect making the glucosyl-2-OH the most acidic, and hence, the most readily removable by base.

The conformation of sucrose in aqueous solution has been under scrutiny by a variety of experimental probes, which have failed, as of now, though to provide an unequivocal picture. Raman[14] and solution X-ray diffraction[15] data have been interpreted to show the conformation of aqueous sucrose to be concentration-dependent, such that in dilute aqueous solutions (<0.7 M) no intramolecular hydrogen bonds are present, while at higher concentration, H-bonds are successfully formed resembling those in the crystal structure. However, subsequent detailed ^{13}C-NMR spin-lattice relaxation measurements[16,17] were not in accord with these conclusions as they seemed to support the contention that sucrose, in water, adopts a form represented by formula **1e**, in which the 2^g-O···HO-1^f hydrogen bond is maintained even in dilute solution. Similarly, the optical rotation of sucrose in water has been interpreted to indicate an equilibrium mixture of the two linkage conformers **1e** and **1f** with the former predominating[18]. More penetrating NMR investigations[19–22] have increasingly accumulated evidence as to "the conformational flexibility"[19] of sucrose in aqueous solution, and the "non-importance"[20], the "transient existence"[21], or altogether "non-existence"[22] of intramolecular hydrogen bonds. Most convincing in this context appears the evidence stemming from the temperature- and magnetic field-dependence of interglycosidic proton-proton NOE contacts[21] and from measurements of the hydroxyl proton exchange rates[22], which are virtually the same for all OH groups; most notably, the 1^f-OH exchange with water is as fast as that of the other two primary hydroxyls, that is, the 6^g-OH and 6^f-OH.

Conformational analysis of sucrose by computational methods has been carried out by rigid-residue calculations using the HSEA[16] and PFOS[20] methodology, and more recently, by relaxed-residue energy mapping[23–25] which provided energy minima smoother in contours and broader in shape. Aside from some differences in the hydrogen bond patterns emerging from the different force-fields used (CHARMM[23], PIMM88[24], and MM3[25]), the bulk of evidence uniformly points to three low-energy domains, of which the two major ones represent geometries corresponding closely to those depicted by formula **1e** and **1f**.

These molecular mechanics simulations refer to the conformations of sucrose adopted in vacuo, in other words, in the gas phase at zero pressure. The vacuum boundary conditions, however, may distort the shape of a molecule substantially, as they tend to minimize the surface area via back-folding. Thus, it cannot be expected a priori, that the in vacuo molecular geometries survive when surrounded with a sphere of water molecules, that is, in a sucrose-water interface.

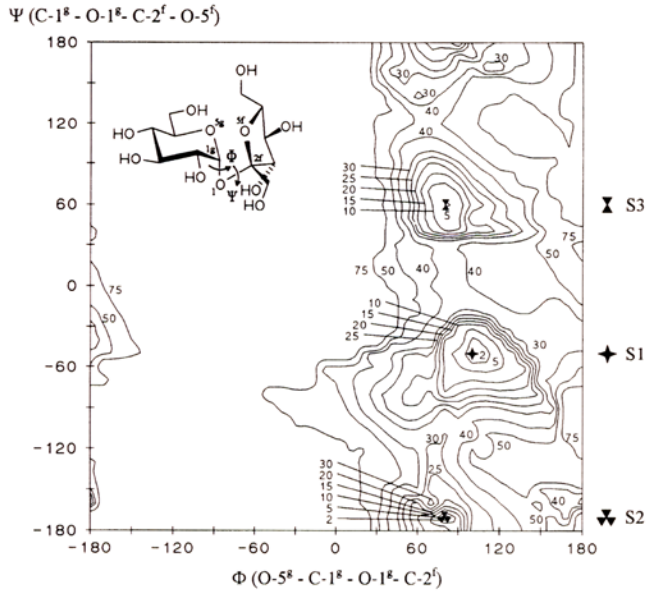
Our interest in unraveling the molecular features responsible for sweet taste elicitation[26,27], for which the sucrose linkage geometry in hydrophilic (e.g. water) and lipophilic surroundings appears imperative, has led us to further probe into this problem. As a result, here we detail and refine our previous modelings of sucrose[24] based on the PIMM88 force-field program, and report on GROMOS-based molecular dynamics (MD) simulations of sucrose in a box of 571 water molecules.

1. Sucrose in a Vacuum: Molecular Mechanics Calculations

The relaxed-residue energy potential surface of sucrose as generated with the PIMM88 force-field program[28] is shown in Figure 1, clearly revealing a limited conformational flexibility of the Φ torsion angle, that is, the rotation about the C-1^g–O-1^g bond, as compared to the Ψ torsion which covers the entire angular range available. Three low-energy regions are found on this map, designated S1, S2, and S3, with the corresponding molecular geometries depicted in Figure 2, and a calculated percentage distribution of 71:21:8 (Figure 3). Based on the Cremer-Pople parameters[29] as conformational descriptors of pyranoid and furanoid ring systems[30] (cf. Table 1), the glucose residue adopts a $^4C_1$ chair, while the fructofuranose unit invariably falls into a nearly perfect $^4T_3$ geometry.

The global energy minimum S1 is characterized by intersaccharidic torsion angles Φ/Ψ ≈ 105°/−48° (cf. Table 1), that differ little from those found in the crystal (+107.9° and −44.8°, respectively). Accordingly, the molecular geometry adopted by the major conformer S1 – as represented by the ball-and-stick model in Figure 2 – corresponds closely to that prevailing in the crystal; only the intramolecular 5^g-O···HO-6^f hydrogen bond has been released on proceeding from the sucrose packing in the crystal lattice to a single molecule in a vacuum environment.

Figure 1. Fully relaxed energy potential surface of sucrose as a function of the intersaccharidic torsion angles Φ and Ψ, corresponding to the rotation about the C-1^g–O-1^g (Φ) and O-1^g–C-2^f (Ψ) bonds, respectively. Energy contours are given in kJ/mol relative to the global minimum at Φ ≈ +110°/Ψ ≈ −50° (designated as S1). The molecular geometry for S1 (cf. ball-and-stick models of Figure 2) closely conforms with the conformation depicted in formula **1e**, that of S2 with formula **1f**



The second low-energy conformer S2 displays quite different intersaccharidic torsion angles (Φ ≈ +80°, Ψ ≈ −170°, cf. Table 1), obviously induced by the embodiment of a 2^g-O···HO-3^f hydrogen bond. Thus, the relative orientation of the fructofuranose and glucopyranose rings in S2 (Figure 2) resembles closely that depicted in the conventional formula drawing **1f**. The third conformer S3 is characterized by an alternative 5^g-O···HO-3^f H-bond with the glucopyranosyl ring oxygen being involved as acceptor.

In addition to the ball-and-stick models in Figure 2, for all sucrose conformers the MOLCAD-program[31] generated contact surfaces[32] in respect to a water molecule (i.e. "how water sees the molecule") are shown in dotted form. The spatial volume included by these surfaces was calculated to approximately 340–345 Å^3, and does not change significantly with the molecular conformation (solid state structure: 339.3, conformer S1: 342.2, S2: 343.8, and S3: 340.9 Å^3). The close agreement of the computed data with the apparent molar volume of $^{\Phi}V$ ≈ 203–212 cm^3/mol ≈ 337–351 Å^3 demanded by sucrose in different solvents such as H_2O, DMSO, and DMF[33] – that is accessible through simple density measurements on solutions – demonstrates the contact surfaces to properly describe the steric demands of the sucrose molecules.

Accordingly, there is a close similarity between the sucrose conformations calculated for vacuum boundary conditions and those detectable in DMSO solution by SIMPLE ^{1}H NMR measurements. Not only do the molecular geometries **1e** and **1f** emerge as the prevailing ones from both methodologies, but their apparent population is a 2:1 equilibrium **1e** ↔ **1f** in DMSO[10] versus a 71:21:8 probability

Figure 2. Ball-and-stick model representation, also comprising the contact surface in dotted form, of the three low-energy conformers (S1–S3) of sucrose that emerge from the PIMM88 force-field calculations of Figure 1. The major conformer S1 displays intersaccharidic torsion angles Φ and Ψ close to those observed in the crystal (cf. Table 1); its overall molecular geometry comprising the 2^g-O···HO-1^f interresidue hydrogen bond resembles largely that implied by the conventional formula drawing **1e**

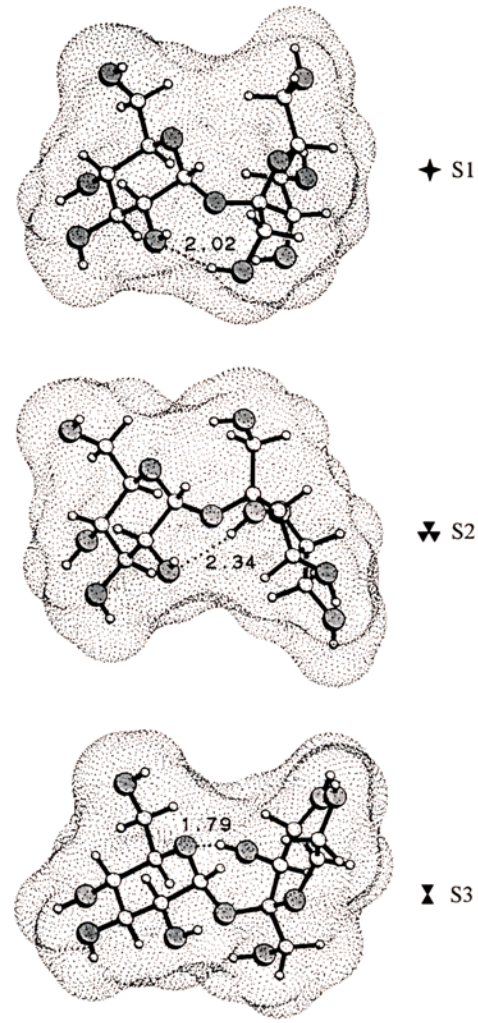


Figure 3. 3D-Plot of the percentage distribution of sucrose conformers as a function of the Φ/Ψ torsion angles, calculated from the PIMM88 energy potential surface (Figure 1) according to the Boltzmann equation for T = 300 K

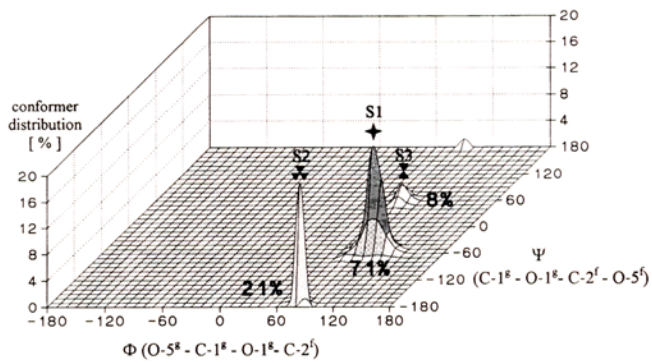


Table 1. Molecular parameters for the three low-energy conformers S1, S2, and S3 obtained from PIMM modelings (Figure 1), as compared to the conformational data for the crystal

Molecular parameters	Sucrose conformation S1	S2	S3	Crystal[a]	
ΔH_F^{298} [kJ/mol]	-1741.8	-1742.2	-1737.3	–	
Φ [°]	+105.4	+80.4	+80.1	+107.9	
Ψ [°]	-47.9	-170.0	+60.0	-44.8	
Hydrogen bonds:	2^g-O···HO-1^f	2^g-O···HO-3^f	5^g-O···HO-3^f	2^g-O···HO-1^f	5^g-O···HO-6^f
d (O···H) [Å]	2.02	2.34	1.79	1.85	1.89
φ (O···H-O) [°]	145.7	163.2	155.4	158.6	158.6
Calc. distribution [%]	71	21	8	–	
Glucose Cremer-Pople parameters:					
Q [Å]	0.555	0.559	0.550	0.556	
θ [°]	5.7	6.0	3.3	5.1	
ϕ[b] [°]	88.6	88.6	50.9	183.8	
conformation	$^4C_1$	$^4C_1$	$^4C_1$	$^4C_1$	
Fructose Cremer-Pople parameters:					
q [Å]	0.411	0.433	0.352	0.352	
ϕ [°]	261.0	261.3	279.8	265.2	
conformation	$^4T_3$ (→ E_3)	$^4T_3$ (→ E_3)	^4E (→ $^4T_3$)	$^4T_3$ (→ E_3)	

[a] Two intramolecular hydrogen bonds. – [b] For $\Theta \rightarrow 0°$, ϕ becomes meaningless, since $^4C_1$ conformations are identical to $^2C_5$ and $^0C_3$.

distribution of conformations S1 (≡ **1e**):S2 (≡ **1f**):S3 in vacuo. This situation entails the notion that the glucosyl-2-oxygen in sucrose has a pronounced tendency to engage in intramolecular hydrogen bonding in vacuo, in DMSO, and in the crystal. At least, a DMSO solvation shell around sucrose, in which the DMSO-oxygen presumably acts as H-bond acceptor for all available sucrose hydroxyl protons, does not intercept the interresidue hydrogen bond to the glucosyl-2-oxygen. This rationalization conceivably extends to all other aprotic solvents.

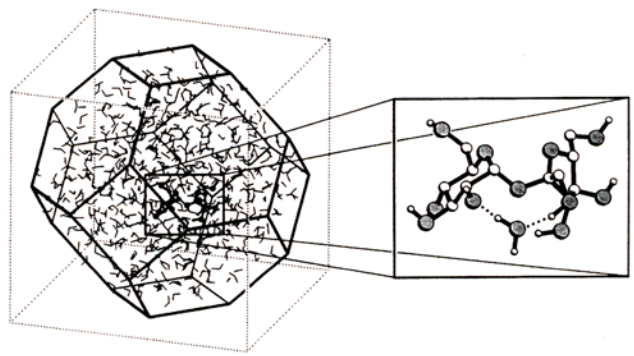
2. Sucrose in Water: Molecular Dynamics Simulations

While the conformation of sucrose in vacuo, in DMSO and in the crystal is dominated by an interresidue hydrogen bond between the glucosyl-2-O (as the acceptor) and the fructosyl-1-OH, in aqueous solution, this bond is obviously intercepted by solvation with water. At least, there is ample NMR evidence[19–22] as to the conformational flexibility of sucrose in an aqueous environment, and to the nonexistence of intramolecular hydrogen bonds altogether. If indeed so, this picture should also emerge from molecular dynamics (MD) simulations of sucrose "soaked" with water molecules. Accordingly, by using the GROMOS[34,35] force-field, MD simulations were carried out for sucrose in the center of a truncated octahedron of pregiven dimensions (box-size approximately 32.9 Å, $V \approx 17800$ Å^3), for which 571 water molecules were needed to fill the remaining empty space (Figure 4).

Backbone Conformation

For calculation of the averaged backbone conformation of aqueous sucrose, the PIMM-derived conformations S1, S2, and S3 were used as starting geometries for the MD

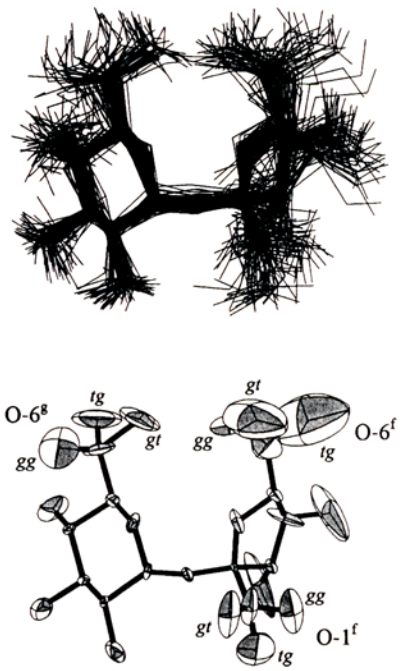
Figure 4. Snapshot of a MD simulation of sucrose surrounded by 571 water molecules (box-size of the truncated octahedron approximately 32.9 Å, volume ca. 17800 Å^3) with one water molecule hydrogen-bonded between 2^g-O and 1^f-O



simulations, which all relaxed within 50 ps of simulation time into a single energy minimum closely resembling in its intersaccharidic torsion angles Φ and Ψ the molecular geometry S1 (cf. ball-and-stick model in Figure 2, or conventional formula **1e**). No further conformational transitions occurred, albeit dynamic fluctuations of Φ and Ψ of ±30° around their mean values were observed. The conformational space occupied by this minimum structure and its flexibility are amply illustrated by superimposition of 100 snap-shot geometries taken in 5 ps intervals from a 500 ps MD simulation.

The three primary hydroxymethyl groups exhibit multiple conformational changes in the 50–75 ps range, the *gauche-gauche* (*gg*) and *gauche-trans* (*gt*) forms invariably being the most populated rotamers. As clearly evident from the fluctuations depicted in Figure 5, the $^4C_1$ chair conformation of the pyranoid ring of the glucose portion varies within

Figure 5. Superimposition of 100 sucrose snapshot-geometries taken in 5 ps intervals from a 500 ps molecular dynamics simulation on an assembly comprising sucrose in a box of 571 water molecules. For clarity, hydrogen atoms and water molecules are omitted. Least-squares fitting was performed by rigid body translation and rotation of the molecules, considering only the tetrahydropyranyloxy-tetrahydrofuran backbone. In the lower plot, the atomic mean positions and anisotropic thermal probability ellipsoids at the σ-level as obtained from the MD are shown. The thermal ellipsoids for the three staggered arrangements for each of the hydroxymethyl groups (*gg*, *gt*, and *tg* forms, respectively) are indicated separately; they signify their individual flexibility, not their relative populations, as adoption of the *gg* and *gt* forms is highly preferred. During the entire simulation only one transition of the glucose-6-CH_2OH to the *tg* conformation was observed

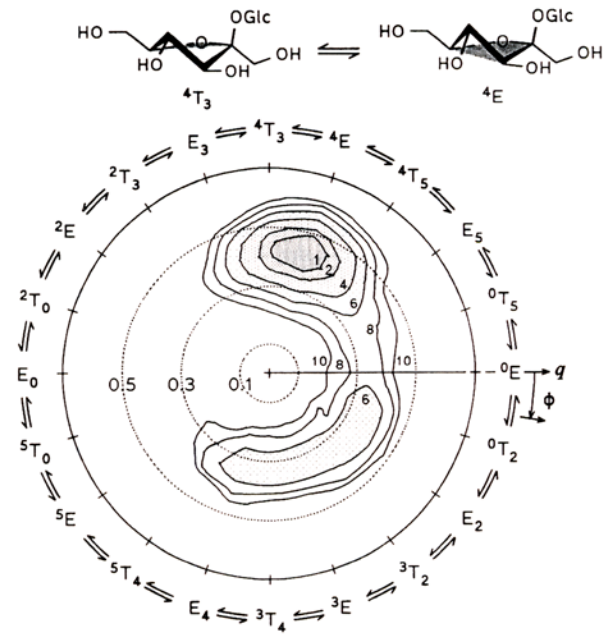


relatively small extensions, as opposed by the furanoid fructose part that is considerably more flexible. The analysis of distributions sampled during the MD-trajectories according to Boltzmann's law (T = 300 K), allows for calculation of the molecular energy of sucrose as a function of the Cremer-Pople[29] ring puckering parameters q and ϕ, which serve as conformational descriptors for five membered rings. The corresponding free energy potential surface (Figure 6) exhibits two clearly separated energy minima: the global minimum at $q \approx 0.40-0.45$ Å and $\phi \approx 270-290°$ corresponds to nothern $^4T_3 \leftrightarrow {}^4E$ ring conformations. The extended southern minimum (+4 to +6 kJ/mol, $q \approx 0.30-0.40$ Å and $\phi \approx 10-100°$) correlates with less favored $E_2 \leftrightarrow {}^3T_2 \leftrightarrow {}^3E \leftrightarrow {}^3T_4$ ring geometries. As evidenced by statistical crystal structure analysis and NMR data, the former $^4T_3$ type conformations are characteristic for β-D-fructofuranose derivatives not only in the solid state, but also in solution as well[36]. Apparently, pseudorotation via eastern pathways is preferred over western transitions.

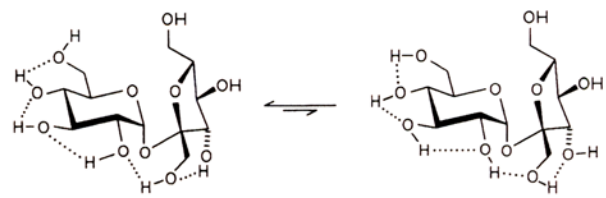
Hydrogen Bonding Patterns

An MD simulation of sucrose in vacuo exhibited an intramolecular hydrogen bond network stable during 2000 ps,

Figure 6. Polar coordinate contour plot of the β-D-fructofuranose pseudorotational energy potential surface as a function of the Cremer-Pople puckering parameters q and ϕ (500 ps MD simulation of sucrose including 571 water molecules). Energy contours are given in kJ/mol relative to the global minimum, broken lines indicate iso-contour lines of the puckering amplitude at q = 0.1, 0.3, and 0.5 Å. The global energy minimum of the furanose ring around $^4T_3 \leftrightarrow {}^4E$ conformations is accentuated by gray shading



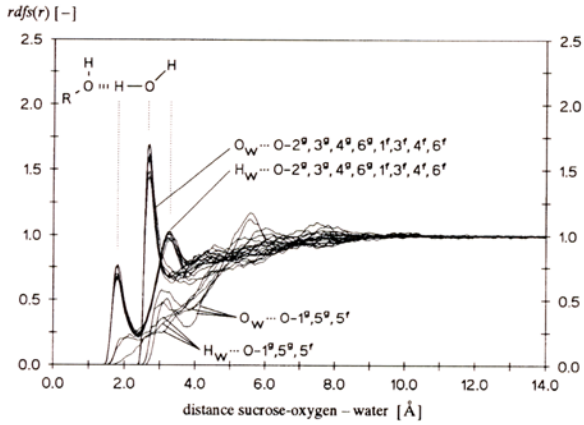
in the course of which (after ca. 680 ps) only *one* highly cooperative 30 ps flip-over from the clockwise arrangement (**1g**) to the alternative anti-clockwise disposition (**1h**) was observed.



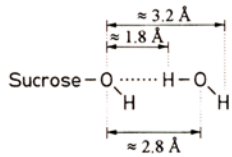
However, when performing the MD simulations of sucrose in the water box, all of these hydrogen bonds are entirely disintegrated, since the hydroxyl groups obviously satisfy their hydrogen bonding requirements by interaction with water molecules of the first hydration layer. Even the most favorable 2^g-O···HO-1^f hydrogen bond (vide infra) is intercepted.

This picture unambiguously emerges from a detailed computerized analysis of the radial pair distribution functions (*rdfs*)[37] derived from about 10000 sucrose geometries in the water box; each of the eight sucrose hydroxyls was inspected separately with respect to the probability with which a hydrogen or, alternately, an oxygen atom of a water molecule is being found at a certain distance from a given sucrose-oxygen. The results were transferred to 22 probability curves as depicted in Figure 7, all showing an essentially identical course for the first hydration sphere around

Figure 7. Plot of the radial pair distribution function *rdfs*(*r*) of water protons (H_W) and oxygens (O_W) around sucrose hydroxyl (2^g-, 3^g-, 4^g-, 6^g-, 1^f-, 3^f-, 4^f-, and 6^f-O) and acetalic (1^g-, 5^g-, and 5^f-O) oxygens (data obtained from 500 ps MD simulation of sucrose in a box of 571 H_2O, $2 \cdot 11$ individual *rdfs*). The curves indicate the relative distribution probability of water to occupy positions at a certain distance from sucrose oxygens; the peaks clearly denote a highly preferred occurrence of water molecules hydrogen-bonded to all hydroxyl groups (distances sucrose hydroxyl-O···$H_W \approx 1.8$ and 3.2 Å, and O···$O_W \approx 2.8$ Å) within a sharply defined first hydration shell. The outer second sphere hydration (O···$H_W \approx 4-6$ Å) is much less prominent but rather diffuse and weak. As evident from the almost uniform peak location and height, the strength of hydration does not vary to a significant extent for the individual hydroxyl groups. In contrast to the hydroxyl oxygens, the acetalic ring oxygens (5^g-O and 5^f-O) as well as the intersaccharidic linkage oxygen 1^g-O are less influenced by hydration: small peaks of the respective pair distribution functions shifted to higher distances demonstrate only weak hydrogen bonding interactions with water (O···$O_W \approx 3.2$ Å). Due to van der Waals repulsions, the curves *rdfs*(*r*) become zero for O···H-distances below ≈ 1.7 Å



the individual hydroxyl groups: two well-defined peaks at 1.8 and 3.2 Å, respectively; thereby, the smaller distance corresponds to a water proton hydrogen bonding to a sucrose-oxygen, while the peak at 3.2 Å is caused by the second non-hydrogen-bonded water proton that points towards the bulk water phase.

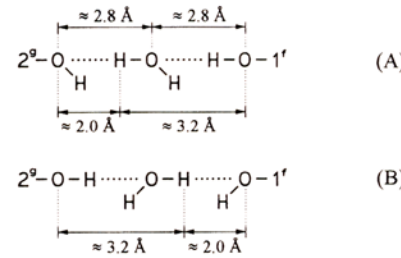


These findings are even further corroborated by the respective sucrose-oxygen–water-oxygen distances, the narrow peak obtained at 2.8 Å (cf. Figure 7) being clear evidence for single water molecules engaged in a donor-type hydrogen bonding interaction with each of the eight sucrose hydroxyls.

Another notable feature of these probability curves is that there are no significant effects beyond a 4 Å distance, indicating that the second hydration layer around sucrose is little influenced by the sucrose hydroxyls – a conclusion that has also been reached on the basis of density measurements of sucrose solutions[38]. The three acetalic oxygens of sucrose, that is, the intersaccharidic 1^g-O and the ring oxygens 5^g-O and 5^f-O, are considerably less effected by interaction with water molecules than the sucrose hydroxyls: the two peaks around 2.0 and 3.0 Å (cf. Figure 7) are broad and rather diffuse. Thus, these oxygens have no bearing on the water molecules in the first hydration sphere around sucrose.

The one-dimensional probability distribution of water molecules around sucrose, as depicted in Figure 7, provided ample evidence that each of the eight sucrose hydroxyls is engaged in hydrogen bonding to water in very much the same way. Furthermore, inspection of analogous pair distribution functions of water oxygens and protons around the sucrose hydroxyl-protons attests to the fact, that each hydroxyl group not only acts as a strong hydrogen bond acceptor for water, but also as a favorable H-bond donor towards the solvent molecules within the first hydration shell. However, the data in Figure 7 have no bearing on the obvious possibility that *one* water molecule is engaged in hydrogen bonding to two sucrose oxygen simultaneously. To evaluate such a contingency, the occurrence probability of water oxygens and hydrogens relative to two sucrose oxygens was determined simultaneously; the respective two-dimensional pair distribution functions are graphically represented in Figure 8.

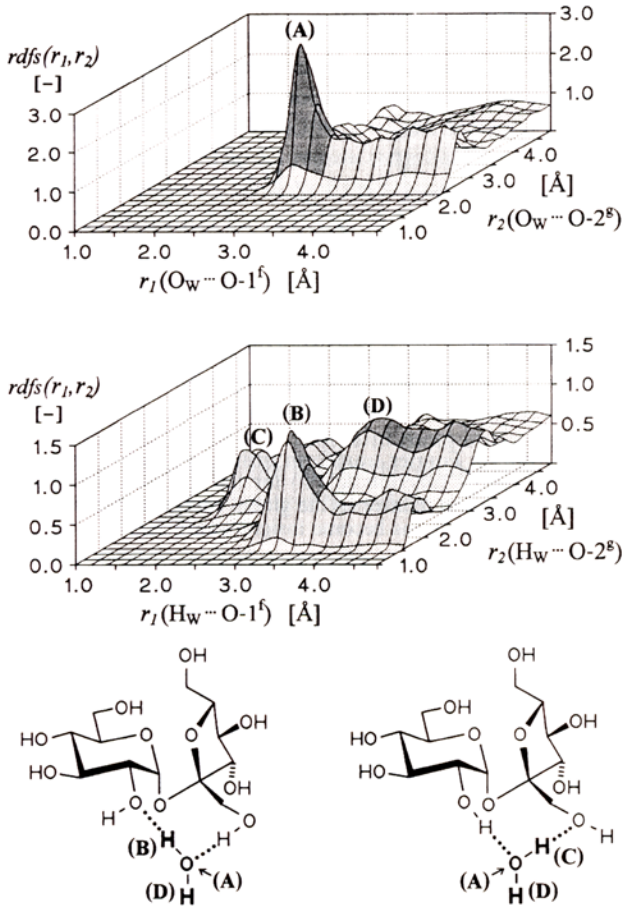
The *rdfs* of Figure 8 reveal a high occurrence probability of water oxygen atoms at 2.8 Å around 2^g-O *and* 1^f-O (Figure 8, peak (A) in the upper 3D-plot), indicating water molecules to be engaged in hydrogen bonding interactions with both hydroxyl groups simultaneously. The corresponding *rdfs* of water protons around 2^g-O and 1^f-O (Figure 8, lower diagram) displays two distinct peaks, (B) and (C), respectively, and a third, substantially broader one (D). The main peak (B) denotes the high probability with which a hydrogen atom of a solvating water molecule is found 2.0 and 3.2 Å away from the two sucrose oxygens 2^g-O and 1^f-O, that are engaged in an interaction of the type 2^g-O···H_2O···HO-1^f.



The reverse directionality 2^g-OH···OH_2···O-1^f of the hydrogen in this "water bridge" gives rise to peak (C), its lower intensity indicating the former configuration to be substantially preferred over the latter type of interaction. The third, comparatively broad-spread region marked D (ca. 3.2 Å away from 2^g-O as well as 1^f-O) derives from the second hydrogen of the bridging water molecule, being more flexible due to its non-involvement in direct hydrogen bonding with the solute.

On the basis of geometrical grounds, an analogous interresidue water bridge might be expected to prevail between the primary hydroxymethyl groupings 6^g-OH and 6^f-OH of the glucose and the fructose moiety. Indeed, MD trajectory

Figure 8. Two-dimensional pair distribution function $rdfs(r_1,r_2)$ of water oxygens (upper diagram) and protons (lower 3D-plot) around 1^f-O (r_1) and 2^g-O (r_2) of sucrose. The maximum occurrence probability of water oxygens was found for $r_1 \approx r_2 \approx 2.8$ Å (peak A in the upper plot), that indicates a single water molecule to be hydrogen-bonded simultaneously to both the 1^f-OH and 2^g-OH groups. The three observed peaks (B–D) in the distribution of water protons (lower diagram) around 1^f-O and 2^g-O can also be traced back a water mediated bridge of hydrogen bonds with two alternative configurations that differ in their H-bond directionality (cf. formula drawings). The main peak (B) corresponds to the preferred configuration with a water proton H-bonded to 2^g-OH ($r_1 \approx 3.2$ Å, $r_2 \approx 2.0$ Å, left formula), and the minor peak (C) to the inverse situation with the proton bonded to 1^f-OH ($r_1 \approx 2.0$ Å, $r_2 \approx 3.2$ Å, right formula). The broad maximum (D) is attributed to the remaining non-hydrogen-bonded water proton ($r_1 \approx r_2 \approx 3.2$ Å)



analysis in respect to this indicated the presence of single configurations with a 6^g-O···H_2O···O-6^f bridge with different H-bond directionalities. However, the stability, the occurrence probability as well as the average life-time of this alternative water bridge is by a factor of 4–5 lower than that for the 2^g-O···H_2O···O-1^f interaction. The faster fluctuations express themselves in broader and less pronounced peaks in the corresponding two-dimensional *rdfs*. Obviously, the increased rotameric flexibility of the two primary 6^g-OH and 6^f-OH hydroxymethyl groups and the unfavorable entropic factor emerging therefrom, prevents the formation of a water bridge with a long lifetime. The much higher significance of the indirect interresidue hydrogen bonding interaction between 2^g-O and 1^f-O must be attributed to the restricted flexibility of the former secondary hydroxyl group that is part of the rather rigid glucose unit.

These *rdfs*-based MD analysis establish that the state of sucrose in water is characterized by encapsulation in a first, well-defined shell of water molecules engaged in hydrogen bonding to each of the eight hydroxyl groups. These results are fully in accord with quite elaborate NMR measurements[19–22], most notably the hydroxyl proton exchange rates, which are virtually the same for all OH groups of sucrose[22]. There are definitely no direct intramolecular hydrogen bonds between the glucose and fructose portions of sucrose in water. However, the calculatory data clearly reveal the existence of an *indirect* one: a water molecule links the glucosyl-2-O with the fructosyl-1-O, an *interresidue water bridge*, so to say. Obviously, it is this water bridge, that largely determines the linkage geometry in aqueous solution, such that the intersaccharidic torsion angles vary only within a comparatively narrow range around $\Phi/\Psi \approx +90°/-60°$ and, thus, closely resemble the linkage geometries found in the crystal (**1c** and **1d**), in DMSO solution (**1e**), and in vacuo (S1 in Figure 2) – except for the direct interresidue hydrogen bonds realized there.

Interresidue water bridges of the type found here for aqueous solution are not uncommon in disaccharides that crystallize from water as hydrates. As evidenced by the crystal structures, α,α-trehalose dihydrate[39], melibiose hydrate[40], and sophorose hydrate[41], each contain a water molecule linking the respective monosaccharide portions with distances and geometries nearly identical to those found for the water bridge in sucrose. Various other examples of water bridges are known to occur in the solid state structures of cyclodextrins[42], and in the dihydrates of disaccharide alcohols, namely, glucosyl-α(1→1)-mannitol[43] and its 2-amino analog[44]. In the two latter crystal structures even a "double" water bridge is realized, that is, two water molecules linking hydroxyl groups along the alditol chain. Since these water bridges receive further stabilization from the cooperativity of hydrogen bonds within chains and networks[45], it may be surmised that these hydrates, proven to exist in the solid state, are also present in aqueous solutions from which these compounds crystallize.

The presence of an interresidue water bridge in aqueous sucrose solutions which has emerged from these data also has a bearing on the dissolution of sucrose in water as well as on its crystallization from aqueous solutions. The dissolution process is rationalized as involving disintegration of the two intramolecular hydrogen bonds present in the crystal structure (cf. formula **1c**): one being "solvated off" entirely, and the other, shorter and more intense one, being intercepted by a water molecule to form an indirect hydrogen bond, thus providing the molecule higher flexibility. The reverse process, assembly of sucrose molecules out of aqueous solution into the anhydrous crystal lattice, must necessarily comprise the "stripping off" of all water molecules. The one incorporated in the interresidue water bridge undoubtedly is the most strongly bound and the most difficult to remove, and hence, the last one to leave before nucleation occurs. This situation conceivably ac-

counts for the unusually broad oversaturation range[46] of aqueous sucrose solutions before eventually crystallization is induced.

3. Search for other Sucrose Conformations through Umbrella Sampling

The molecular dynamics methodology used so far has provided information only about a rather narrow set of intersaccharidic torsion angles for the sucrose conformation in water, namely, those corresponding to molecular geometry S1 (or its chemical formula drawing **1e**), comprising variations around $\Phi/\Psi \approx 105°/-48°$ with comparatively small fluctuations in the orientations of the glucose and fructose portions relative towards each other. Alternative starting geometries (i.e. the PIMM88 structures S2 or S3, respectively) – altogether six independent MD runs of 100–200 ps length each – all relaxed very fast within the first 50 ps of simulation time to the single S1 energy minimum. The absence of multiple conformational transitions back and forth between the substates did not reveal if other conformations are energetically accessible in solution, nor did it provide any information about the relative stabilities of substates or their contributions to a conformational equilibrium in solution.

To cover the entire conformational space of possible Φ/Ψ torsion angle combinations, in other words, to locate any other conformational energy minima that may exist, we determined the free energy profile $G(\Psi)$ of sucrose in aqueous solution as a function of the more flexible intersaccharidic torsion angle Ψ, in order to search for alternative geometries. As free energy calculations[47] by thermodynamic integration require very long simulation times to obtain precise values[48], the more efficient non-Boltzmann *"umbrella sampling"*[49] technique was used for evaluating the $G(\Psi)$ function; the computational basics and details of this procedure are given in the Experimental section.

In Figure 9, the GROMOS-derived free energy profile for sucrose conformations in aqueous solution is shown as a function of the intersaccharidic torsion angle Ψ representing the rotation around the O-1^g–C-2^f glycosidic bond. Thereby, $\Delta G(\Psi)$ was elaborated as a 10-term Fourier transform fit (refined parameters cf. Table 2) with respect to the global energy minimum. The inherent relative error in computed energies was estimated at $\sigma[\Delta G(\Psi)] \approx \pm 1.0$ kJ/mol. The different conformational sub-families were denoted S1–S3 in analogy to the minima of the PIMM88 energy potential surface of Figure 1. The global energy minimum (S1, $\Psi \approx -60°$) again closely resembles the conformation in the crystalline state, while the conformer S2 ($\Psi \approx -170°$) is approximately +10.3 kJ/mol higher in energy. S3 is found around $\Psi \approx +30°$ not as an energy minimum but as a saddle point ($\Delta G \approx +13$ kJ/mol).

In contrast to the energy potential maps, the free energy profile allows the direct estimation of physically relevant transition barriers. For the conformational change of S1 into S2, a barrier of +16.5 kJ/mol ($\Psi \approx -125°$) is computed, while for the way back (S2 → S1), only a small barrier of +6.3 kJ/mol is found. The full range rotation of the glucose and fructose with respect to each other involves the transition S1 → S2 → S3 → S1, for which a central hindrance of +34.8 kJ/mol ($\Psi \approx +105$) has to be surmounted.

Figure 9. Calculated free energy profile $G(\Psi)$ of sucrose conformations in aqueous solution as a function of the Ψ-torsion (*umbrella sampling* using the GROMOS[34] force-field, energies in kJ/mol relative to the global energy minimum). The bold line delineates the Fourier transform fit ot the sampled G-values (thin line) with conformational regions marked S1–S3; the estimated error is $\sigma(\Delta G(\Psi)) \approx \pm 1.0$ kJ/mol. The ΔG-function correlates surprisingly well with a cross section of the PIMM88 energy potential surface around $\Phi \approx +60 - +130°$ and $\Psi = -180 - +180°$ (cf. Figure 1). The $G(\Psi)$-derived Boltzmann distribution function $P(\Psi)$ ($T = 300$ K, right y-axis, normalized to $\int P(\Psi)\, d\Psi = 1$) is also shown by the dashed line

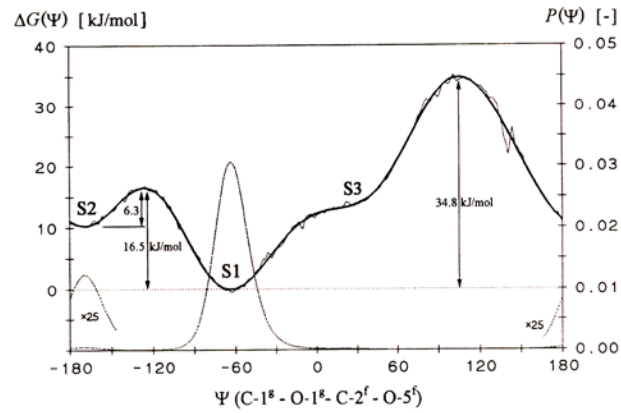


Table 2. 10-term Fourier transform fitting parameters (in kJ/mol) of the free energy function $\Delta G(\Psi)$ for the sucrose conformations in aqueous solution relative to the global energy minimum; higher order terms become insignificantly small

$\Delta G(\Psi) = a_0/2 + \Sigma\, a_i \cos(i\Psi) + b_i \sin(i\Psi)$	Parameters [kJ/mol]					
i:	0	1	2	3	4	5
a	+32.2	-4.20	-3.84	+4.82	-0.04	-0.12
b	–	+11.23	+0.37	-1.76	-0.82	+0.33

It is important to note that the conformational transitions observed during 100–200 ps MD simulations starting from different conformations (cf. above) are consistent with the absence of any energy barrier (S3 → S1) and a relatively small hindrance of $\leqslant 2.5\ RT$ (S2 → S1). The conformation S1 corresponds to the deep energy minimum and proved to be stable during all simulations of altogether about 1500 ps. In terms of the relative percentage distribution $P(\Psi)$ of these conformers, as depicted additionally in Figure 9 (dashed line), the equilibrium ratio of S1:S2 is estimated as 98:2; the contribution of S3 is negligible small (<0.5%).

Free energy differences and transition barriers strongly depend on the definition of the coordinate of reaction or conformation. The free energy profile $\Delta G(\Psi)$ given in Figure 9 implies averaging of all internal coordinates perpendicular to Ψ (including the second intersaccharidic torsion angle Φ) and cannot be related to a single vertical cross-cut of the Φ/Ψ-energy potential surface. A more detailed two-dimensional analysis of the umbrella samplings in terms of Φ *and* Ψ according to $P_i(\Phi + \Delta\Phi, \Psi + \Delta\Psi) \sim \Sigma \exp[U_i^*$

$(\Psi')/RT]$ for all MD structures within classes $\Phi + \Delta\Phi/\Psi + \Delta\Psi$ (cf. equations 1–5 in the Experimental), results in partial free energy potential surfaces. Subsequent fitting of the individual umbrella data sets, as described in the Experimental, yields the section of GROMOS free energy potential surface shown in Figure 10, which can be related to the PIMM88 surface. Calculating Φ/Ψ probabilities increases the computational error to approximately $\sigma[\Delta G(\Phi,\Psi)] \approx \pm 2.0$ kJ/mol. The contours and the corresponding distributions in Figure 10 become smoother and broader than in the PIMM88 map. Most notably, the minima of the conformational sub-families S1 and S2 of sucrose are located within similar regions (±15°) in both cases (S1: $\Phi \approx +90°$, $\Psi \approx -60°$, and S2: $\Phi \approx +90°$, $\Psi \approx \pm 180°$ for the GROMOS force-field, for the PIMM88 data compare Table 1), but with different relative energies. The conformational family S1 invariably corresponds to the global energy minimum, while S2 is $\approx +12$ kJ/mol higher in energy on the GROMOS surface. The S3 region even becomes a high-energy plateau ($\approx +15$ to $+20$ kJ/mol) rather than a minimum. The transition barriers (S1 → S2: 15–20 kJ/mol, S1 → S3 → S2: 35–40 kJ/mol) were already estimated from the one-dimensional function $G(\Psi)$; the small differences demonstrate the susceptibility of free energy analysis with respect to different definitions of the reaction coordinates.

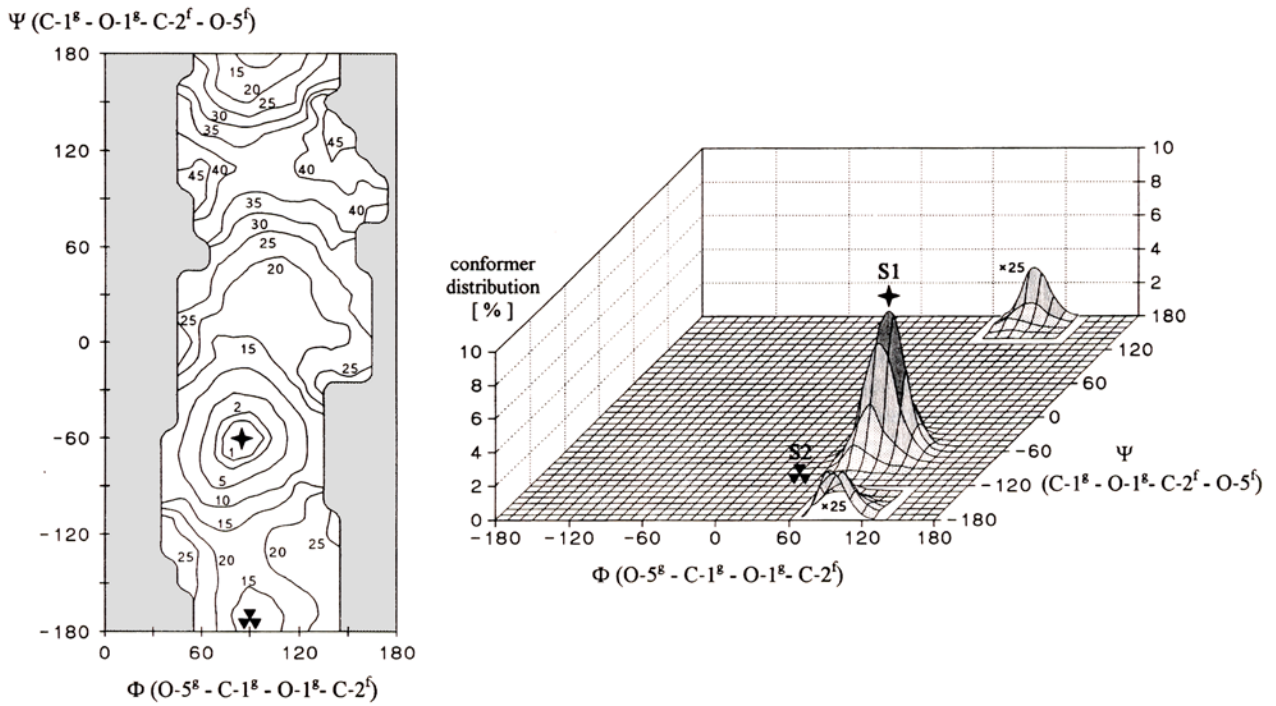
It is to be stressed, that our calculations fulfill van Gunsteren's criteria for exclusion of methodological errors in free energy calculations[50]: first, and most importantly, the free energy change for the closed thermodynamic cycle of the full range rotation around Ψ is close to zero (<1–2 kJ/mol), and second, the results are independent of the equilibration times applied for analysis of the MD-runs. Since accurate sampling of the conformational space was established, the accuracy of the calculations itself depends solely on the GROMOS force-field. It may be that the free energy differences between both conformational families S1 and S2 are overestimated by GROMOS, entailing an underestimation of the relative importance of S2 in the solution equilibrium.

4. Conclusion: Chemical and Biological Implications

The bulk of MD calculatory data, corroborated by the umbrella sampling strategy covering the entire conformational space accessible to sucrose, clearly points towards an equilibrium composition of only two alternative conformations in aqueous solution, that is, S1 and S2. Although these exhibit a considerable degree of flexibility, the major conformer closely resembles the geometry adopted in the solid state in terms of the intersaccharidic torsion angles Φ and Ψ, despite the fact that there are no interresidue hydrogen bonds because all sucrose hydroxyls satisfy their H-bonding requirements through bonding with the solvent. The reason for this is to be found in the elaboration of an indirect, water-mediated hydrogen bond bridge linking the glucose and fructose portions, thus keeping the Φ and Ψ angles within a geometry frame more flexible, yet nonetheless similar to that observed in the crystal and, calculationwise, for vacuum boundary conditions.

The proclivity of aqueous sucrose to elaborate an indirect, water-mediated hydrogen bond bridge between the

Figure 10. Contour plot of the sucrose conformational free energy as function of the intersaccharidic torsion angles (left plot $\Delta G(\Phi,\Psi)$, energies in kJ/mol relative to the global energy minimum) as obtained from 2D-analysis of the *umbrella sampling* data (compare to Figure 1 and 9, gray shading indicates insufficient data for contour plotting); the estimated uncertainty in energy is $\sigma[\Delta G(\Phi,\Psi)] \approx \pm 2.0$ kJ/mol. On the right, the 3D-plot of the corresponding Boltzmann-derived percentage distribution indicates only one relevant conformational substrate, that is S1. Note, that the distribution of the minor conformational family S2 is expanded by a factor of 25



two sugar residues not only has implications on its quite complex crystallization process: the broad oversaturation range prevailing before nucleation occurs is obviously caused by difficulties in expelling the water from the inter-residue water bridge. Moreover, it is likely that the conformational preferences found for sucrose also apply to various trisaccharides containing sucrosyl moieties. Crystallographic data on melezitose (3^f-*O*-glucosyl-sucrose)[51], planteose (6^f-*O*-galactosyl-sucrose)[52], and erlose (4^g-*O*-glucosyl-sucrose)[53] show sucrose linkage geometries ($\Phi \approx +100° - +110°$, $\Psi \approx -25° - -45°$) that are close to that observed for crystalline sucrose or its global energy minimum S1. The two kestoses which have been X-rayed, that is, sucroses carrying β-D-fructofuranosyl residues at the 1^f-*O*-[54] and 6^f-*O*-positions[55], feature intersaccharidic torsion angles of $\Phi \approx +85°$ and $+90°$ versus $\Psi \approx -65°$ and $-55°$ for their sucrose linkages, respectively. These minor deviations from the sucrose solid-state geometry (cf. graphic representation **1c**) are readily rationalized though on the basis that in each of the kestoses, one of the primary fructosyl-OH which engage in intramolecular hydrogen bonding in the sucrose crystal, is blocked. Only raffinose, a 6^g-*O*-galactosyl-sucrose crystallizing as a pentahydrate[56], shows a peculiar linkage geometry for the sucrose portion; it exhibits torsion angles of $\Phi \approx +82°$ and $\Psi \approx +11°$ that are significantly shifted from the S1 energy minimum towards S3. There appears to be no direct intramolecular repulsive interaction of the galactosyl residue pushing the fructosyl unit into a quite different orientation relative to the glucose portion. On the other hand, the tetrasaccharide stachyose, a raffinose galactosylated in its galactosyl residue, and also crystallizing as a pentahydrate[57], shows an essentially perfect sucrose linkage geometry ($\Phi \approx +109°$ and $\Psi \approx -48°$) indicating that the peculiar conformation of raffinose in the crystal is likely to be due to packing effects in the lattice. In aqueous solution, though, these peculiarities are surmised to disappear. This point is being pursued by us via MD simulations in water as well as through analysis of NOE contacts. At least, similar isotope effect characteristics derived from SIMPLE ^{1}H-NMR data suggest that in DMSO solutions, the equilibrium conformations of sucrose, raffinose, and stachyose do not differ to a significant extent[17b]. Moreover, the similarities observed for the long-range proton-carbon coupling constants in DMSO and D_2O indicate that the conformations about the C-1^g–O-1^g bond closely correspond to each other[58].

Some *chemical implications* may be derived from the conformational data presented. In polar aprotic solvents such as DMSO or DMF, direct interresidue hydrogen bonds 2^g-O···HO-1^f (or 2^g-O···HO-3^f) are formed, for which the PIMM88-derived models S1 and S2 represent reasonable "frozen" molecular pictures. This result, and particularly the generation – on the respective contact surfaces (cf. Figure 3) – of the molecular electrostatic potential (MEP) profiles[12] entailed the notion, that the glucosyl-2-OH is the most acidic of the eight sucrose hydroxyls, hence, the one most readily deprotonated by base, and, in turn, the first to be seized by alkylation and acylation agents. Induced by these considerations, a synthesis of 2^g-*O*-modified sucroses such as the 2-keto-, 2-deoxy- and 2-amino derivatives could be elaborated[12], based on the benzylation of sucrose in DMF with NaH/benzyl bromide occurring with unusually high selectivity (>80%). This, in turn, substantiates the retention of the intramolecular 2^g-O···HO-1^f hydrogen bond in aprotic solvents.

The *biological consequences* of these modelings on the molecular recognition mechanisms taking place in the active transport of sucrose in plants[59], and in particular, on the mechanism of sweet-taste elicitation in the taste-bud receptor[26] may be even more far-reaching. In both processes, proteinaceous binding sites are involved, which a priori are filled with water. In the docking procedure this water must be displaced by the sucrose molecule, which in turn must strip off its hydration shell. In view of the pronounced tendency of the glucosyl-2-OH of sucrose to engage in hydrogen bonding (to the 1^f-OH or to water), it is likely that it is also involved in binding to the receptor, conceivably to an amide group of the protein. In terms of the original AH-B-X structure-sweetness concept[60,61] and its recent modification incorporating the hydrophilic and hydrophobic regions of sucrose as essentials to make it predictive[26], the glucosyl-2-OH represents the AH-part of the tripartite contact. Accordingly, after docking of the sucrose molecule into the proteinaceous binding site – whether this is the global minimum conformation S1 or another one of higher energy – it is likely that the elaboration of a hydrogen bond of the glucosyl-2-OH is one of, or even the decisive step in launching the complex cascade of biochemical reactions that eventually transduce the sweet sensation to the brain.

We gratefully acknowledge financial support for this work by the *Südzucker AG*, Mannheim-Ochsenfurt, and the *Fond der Chemischen Industrie*, Frankfurt am Main. Our thanks are also due to Prof. Dr. *W. F. van Gunsteren*, Eidgenössische Technische Hochschule, Zürich, for kindly granting one of us (S.I.) access to the GROMOS molecular dynamics software package and his computational facilities.

Experimental

Φ/Ψ Energy Potential Surfaces and Contour Plots: Molecular mechanics calculations were carried out using the PIMM88[28] force-field for vacuum conditions ($\varepsilon = 1$). Different starting structures were generated from the solid-state geometry by variation of the torsions involved in the three hydroxymethyl groupings. The Φ (O-5^g–C-1^g–O-1^g–C-2^f) and Ψ (C-1^g–O-1^g–C-2^f–O-5^f) torsion angles were driven in steps of 10° each and fixed, while all other molecular parameters were fully optimized in each step. New geometries were generated from conformationally adjacent, preoptimized conformers, the self-consistency of the map was counterchecked by double calculation of structures on same Φ/Ψ-grid points generated by different conformational pathways. Hydrogen bonds were treated as described in ref.[28], without using additional energy terms. The PIMM88 force-field was shown to properly reproduce the anomeric and exoanomeric effects in acyclic and cyclic acetal structures such as dimethoxymethane, dihydroxymethane, tetrahydro-pyranosides and -furanosides without any application of additional potentials. Contour plots were computed using cubic regression formulas[62].

Molecular Dynamics Simulations: The GROMOS[34] force-field with the united atoms model for CH and CH_2 groups and SPC-type water[63] was used for all molecular dynamics (MD) simulations. Standard GROMOS parameters of "CS1", "CS2" (CH and CH_2 groups), "CB" (C-2^f), "OA"−"HO" (OH-grous), and "OS" (acetal oxygens) were used (a total of 31 atoms for sucrose), with charge groups[64] defined as $C_n-O_n-H_n$ (n = 2^g, 3^g, 4^g, 6^g, 1^f, 3^f, 4^f, and 6^f with charges +0.150, −0.548, and +0.398) and C-5^g−O-5^g−C-1^g−O-1^g−C-2^f−O-5^f−C-5^f (charges C-5^g and C-5^f +0.160, C-1^g and C-2^f +0.380, all oxygens −0.360). Periodic boundaries (truncated octahedron) and isothermal and isobaric conditions (T = 300 K, p = 1 bar) were applied with a temperature relaxation time of 100 fs, a pressure relaxation time of 500 fs (during the first picosecond of equilibration τ_T = 10 fs and τ_p = 50 fs were used)[65], and a cut-off radius of 8.5 Å for the long-range electrostatic interactions (ε = 1.0). The simulation time step was 2 fs, the non-bonded pair list was updated every 10 steps, and bond lengths were constrained using the SHAKE method[66]. Coordinates were saved every 25 steps (50 fs) for analysis purposes. The crystal structure[7] and PIMM88 models S1−S3 were used as starting geometries for independent MD runs [X-ray: +211 H_2O, $r_{box} \approx$ 23.85(9) Å, t = 100 ps; S1: 571 H_2O, 32.90(8) Å, 500 ps; S2: 265 H_2O, 25.64 (8) Å, four runs with 100−200 ps; S3: 256 H_2O, 25.34(8) Å, two runs of 100 ps each], whereby the first 25 ps of every simulation were discarded for equilibration.

The geometry S1 was used as a starting point for a 2000 ps MD of sucrose in vacuo (ε = 1.0) using the simulation parameters described above. The force-field was checked by 200 ps MD simulation of 2 · 2 · 3 unit cells (a total of 32 sucrose molecules) of the crystal structure (monoclinic space group $P2_1$, a = 10.86 Å, b = 8.71 Å, c = 7.76 Å, and β = 102.95°, Z = 2) with periodic boundary conditions. The molecular geometry of sucrose and all intra- and intermolecular hydrogen bonds were reproduced with satisfactory accuracy.

In general, for analysis of all MD trajectories the existence of hydrogen bonding interactions was assumed for O···H distances less than 2.4 Å and O···H−O angles greater than 120°. Least-squares fitting of the sucrose molecules was performed by rigid body translation and rotation, considering only the tetrahydropyranyloxy-tetrahydrofuran backbone without substituents and hydrogen atoms[67].

The fructofuranose pseudorotational energy potential surface as a function of the Cremer-Pople[29] puckering parameters q and ϕ (Figure 6) was derived from the 500 ps MD simulation including 571 H_2O molecules according to Boltzmann statistics (T = 300 K, width of classes Δq = 0.05 Å and $\Delta\phi$ = 18°, a correction was applied to account for different class volumes); contours were plotted in q/ϕ-polar coordinates[62].

Free Energy Calculations using Umbrella Sampling. General Procedure: Calculation of the free energy profile $G(\Psi)$ of sucrose as a function of the intersaccharide torsion angle Ψ(C-1^g−O-1^g−C-2^f−O-5^f) using probability distributions $P(\Psi)$ obtained from MD simulations (cf. equation 1, Boltzmann's law) requires infinite sampling times to ensure proper exploration of the entire conformational space.

$$G(\Psi) = -RT \ln P(\Psi) \quad (1)$$

To enhance monitoring in particular of high-energy regions, an arbitrary shaped additional restraining *"umbrella potential"* $U^*(\Psi)$ (which is essentially identical to a *"driving force"* shifting Ψ to different regions!) is imposed to the potential energy of the system, for which a modified probability distribution $P^*(\Psi)$ is now sampled. Apart from an undefined energy offset or a normalization factor, the free energy $G(\Psi)$ or probability distribution $P(\Psi)$ of the unperturbed system can be calculated from equation 2 or 3:

$$G_i(\Psi) = -RT \ln P_i^*(\Psi) - U_i^*(\Psi) + \text{const} \quad (2)$$

$$P_i(\Psi) = P_i^*(\Psi) \cdot \exp[U_i^*(\Psi)/RT] \cdot \text{const}' \quad (3)$$

In contrast to the *"adaptive umbrella sampling"* method, where the *umbrella potential* is successively modified to finally yield $U^*(\Psi) = -G(\Psi)$ and a uniform distribution $P^*(\Psi)$ − which is certainly hard to sample properly for large rotating groups like two monosaccharide units − we used variable potentials $U_i^*(\Psi)$ (cf. equation 4) in consecutive and independent molecular dynamics simulations; different MD runs are indicated by alternative indices i. The restraining potential $U_i^*(\Psi)$ was chosen appropriately to explore as much of Ψ-space as possible, and simultaneously providing sufficient overlap of different $P_i^*(\Psi)$ distributions. These overlap regions are required to assemble the corresponding incomplete $G(\Psi)$ or $P(\Psi)$ functions by using a least squares fitting procedure weighted by $[P_i^*(\Psi) \cdot P_j^*(\Psi)]^{1/2}$, to form an overall fitting curve covering the entire Ψ-space. The weighting scheme relates to an error of $\sigma[P_j^*(\Psi)] \sim N_i(\Psi)^{1/2}$ with $N_i(\Psi)$ being the number of events sampled for each class of Ψ. Fitting of the sub-curves can be carried out either in terms of free energy (adding an energy offset $const_{ij}$, equation 2) or in terms of probabilities (multiplication with a normalization factor $const'_{ij}$, equation 3); the latter is most conveniently carried out on a logarithmic scale, and thus both methods and the results obtained therefrom become essential identical. The hysteresis of the overall $G(\Psi)$ function obtained from fitting of distributions in the same Ψ-region that were generated by different conformational pathways is used for error estimation.

Umbrella Sampling. Application to Sucrose: Starting from a system including sucrose and 211 H_2O molecules that was pre-equilibrated by 100 ps unrestrained molecular dynamics (simulation parameters as described above, $r_{box} \approx$ 23.85(9) Å, conformational family S1) two independent umbrella sampling runs were initiated. Each run comprises 13 consecutive simulations of 150 ps length, respectively, including 25 ps of equilibration phase, thus amounting to a total simulation time of altogether 3900 ps. The umbrella potential [$U_i^*(\Psi')$, cf. equation 4] was applied as a function of the intersaccharidic torsion angle Ψ' (C-1^g−O-1^g−C-2^f−C-3^f), since Ψ (C-1^g−O-1^g−C-2^f−O-5^f) does not occur explicitly in the GROMOS molecular topology definition for sucrose.

$$U_i^*(\Psi') = K_i[1 + \cos(n_i\Psi' - \Theta_i)] \quad (4)$$

The phase angle Θ_i was varied from 0° → −180° (series I) and 0° → +180° (series II) in steps of 20° and/or the force constant K_i was changed in the range of 35−300 kJ/mol between individual simulations (multiplicity n_i = 1 for all MD runs). The final ("pre-equilibrated") structures of each umbrella run were used as starting points for the next trajectory.

The probability densities were shifted via two independent and opposite conformational pathways towards the central barrier at $\Psi' \approx -15°$ and $\Psi \approx +100°$ [series I: S1 → S2 → central barrier, series II: S1 → S3 → central barrier, with Ψ'(S1) ≈ ±180°, Ψ'(S2) ≈ +70°, and Ψ'(S3) ≈ −90°]. The full range of values accessible to the torsion Ψ' was divided into 360 classes of 1° width each, and probability distributions $P_i^*(\Psi')$ were calculated as a fraction of sampled points per class for each trajectory separately. Each distribution $P_i^*(\Psi')$ covered a Ψ'-space of approximately 70 ± 15°, with overlap regions of ≈55 ± 10° between neighbored distributions. Transformation of $P_i^*(\Psi')$ into the partial free energy functions $G_i(\Psi')$ of the unperturbed system (cf. equation 2) and subsequent least squares fitting (weighting scheme as outlined above, points differing by more than 2σ were rejected) yielded the

overall free energy function $G(\Psi')$. The probability distribution of Ψ' obtained from the 500 ps simulation of sucrose incl. 571 water molecules (cf. above) was used as a common starting reference for fitting both umbrella series I and II.

Due to the rather stiff pyranose ring system and the resulting strong correlation of Ψ and Ψ', the more convenient function $G(\Psi)$ can be obtained via $G(\Psi) \approx G(\Psi' + 115.7°)$. Equivalently, for each umbrella run i, we calculated the distribution $P_i(\Psi \pm \Delta\Psi/2)$ from equation 5, as the sum over the occurrence probabilities of all structures in the unperturbed system and within each class $\Psi \pm \Delta\Psi/2$.

$$P_i(\Psi \pm \Delta\Psi/2) \sim \Sigma \exp[U_i^*(\Psi')/RT] \quad (5)$$

Transformation into free energies (cf. equation 1) and fitting directly yielded the function $G(\Psi)$ that is depicted in Figure 9.

The relative computational error in energy $\sigma(\Delta G)$ was estimated from the hysteresis obtained for the two independent umbrella runs. On top of the central barrier ($\Delta G \approx +34.8$ kJ/mol) that was approached via opposite conformational pathways, the free energy at any Ψ-value differs by less than $\sigma(\Delta G) \approx \pm 1.0$ kJ/mol for both umbrella series. Essentially the same results were obtained when applying different weighting schemes $[P_i^*(\Psi) \cdot P_j^*(\Psi)]^n$ with $n = 0-3$, and/or when discarding up to 125 ps from each 150 ps trajectory for equilibration purposes, leaving only the rest for calculation of the distribution functions $P_i^*(\Psi)$, thus indicating excellent statistics that had been obtained during all simulations.

Two-dimensional analysis of the umbrella samplings (Figure 10) was performed cf. equation 5 using classes of $\Delta\Phi = \Delta\Psi = 10°$ width and application of the same weighting and fitting scheme as described above. The error was estimated to $\sigma[\Delta G(\Phi,\Psi)] \approx \pm 2.0$ kJ/mol.

Determination of the free energy profile $G(\Psi)$ of sucrose was also attempted through umbrella sampling using vacuum boundary conditions. For dielectric constants of $\varepsilon = 1.0$ and 10.0, as well as a distance dependent parameter $\varepsilon = r$ [Å], unrestrained starting trajectories of 2000 ps length each were generated. In all cases, two series with 12 umbrella runs of 150 ps each were initiated; application of the same weighting and fitting scheme as outlined above resulted in $G(\Psi)$ profiles for the vacuum environment, respectively. The limited flexibility of sucrose in vacuum caused by strong hydrogen bonding interactions leads to an increase of the hysteresis and the associated computational error in energy to about 5–10 kJ/mol. Furthermore, the energy differences between the conformational substates as well as the corresponding transition barriers become significantly larger by up to a factor of 2.0 when compared to the solution dynamics with explicit incorporation of the solvent. Undoubtedly, when using the GROMOS force-field, the vacuum dynamics proved unsatisfactory for revealing the conformational properties of sucrose.

[1] S. Immel, J. Brickmann, F. W. Lichtenthaler, *Liebigs Ann.* **1995**, 929–942.

[2] Introductory dedicational metaphor, in A. Hugill, *Sugar and all That. A History of Tate & Lyle*, Gentry Books, London, **1978**.

[3] [3a] E. Fischer, *Ber. Dtsch. Chem. Ges.* **1891**, *24*, 1836–1845; 2683–2687. – [3b] F. W. Lichtenthaler, *Angew. Chem.* **1992**, *104*, 1577–1593; *Angew. Chem. Int. Ed. Engl.* **1992**, *31*, 1541–1556.

[4] For a detailed account on the evolution of the structural formula of sucrose, see: F. W. Lichtenthaler, S. Immel, U. Kreis, in *Carbohydrates as Organic Raw Materials* (Ed.: F. W. Lichtenthaler), VCH, Weinheim/New York, **1991**, pp. 1–32; *Starch/Stärke* **1991**, *43*, 121–132.

[5] C. E. James, L. Hough, R. Khan, *Progr. Chem. Org. Nat. Prod.* **1989**, *55*, 117–184.

[6] F. W. Lichtenthaler, *Zuckerindustrie (Berlin)* **1991**, *116*, 701–712.

[7] G. M. Brown, H. A. Levy, *Science* **1963**, *141*, 921–923; *Acta Crystallogr., Sect. B*, **1973**, *29*, 790–797.

[8] J. C. Hanson, L. C. Sieker, L. H. Jensen, *Acta Crystallogr., Sect. B*, **1973**, *29*, 797–808.

[9] The atoms of the *gluco* and *fructo* residues of sucrose are denoted with "g" and "f" superscripts, respectively.

[10] J. C. Christofides, D. B. Davies, *J. Chem. Soc., Chem. Commun.* **1985**, 1533–1534.

[11] C. Chauvin, K. Baczko, D. Plusquellec, *J. Org. Chem.* **1993**, *58*, 2291–2295.

[12] F. W. Lichtenthaler, S. Immel, P. Pokinskyj, *Liebigs Ann.* **1995**, 1939–1947, following paper.

[13] C. H. Hamann, S. Fischer, H. Polligkeit, P. Wolf, *J. Carbohydr. Chem.* **1993**, *12*, 173–190.

[14] [14a] M. Mathlouthi, D. V. Luu, *Carbohydr. Res.* **1980**, *81*, 203–212. – [14b] M. Mathlouthi, C. Luu, A. M. Meffroy-Biget, D. V. Luu, *Carbohydr. Res.* **1980**, *81*, 213–223.

[15] M. Mathlouthi, *Carbohydr. Res.* **1981**, *91*, 113–123.

[16] K. Bock, R. U. Lemieux, *Carbohydr. Res.* **1982**, *100*, 63–74.

[17] [17a] D. C. McCain, J. L. Markley, *Carbohydr. Res.* **1986**, *152*, 73–80; *J. Am. Chem. Soc.* **1986**, *108*, 4259–4264; *J. Magn. Reson.* **1987**, *73*, 244–251. – [17b] D. B. Davies, J. C. Christofides, *Carbohydr. Res.* **1987**, *163*, 269–274. – [17c] H. Kovacs, S. Bagley, J. Kowalewski, *J. Magn. Reson.* **1989**, *85*, 530–541.

[18] E. S. Stevens, C. A. Duda, *J. Am. Chem. Soc.* **1991**, *113*, 8622–8627.

[19] J. M. Duker, A. S. Serianni, *Carbohydr. Res.* **1993**, *249*, 281–303.

[20] C. Hervé du Penhoat, A. Imberty, N. Roques, V. Michon, J. Mentech, G. Descotes, S. Pérez, *J. Am. Chem. Soc.* **1991**, *113*, 3720–3727.

[21] L. Poppe, H. van Halbeek, *J. Am. Chem. Soc.* **1992**, *114*, 1092–1094.

[22] B. Adams, L. Lerner, *J. Am. Chem. Soc.* **1992**, *114*, 4827–4829.

[23] V. H. Tran, J. W. Brady, *Biopolymers* **1990**, *29*, 961–976; 977–997.

[24] F. W. Lichtenthaler, S. Immel, D. Martin, V. Müller, *Starch/Stärke* **1992**, *44*, 445–456.

[25] [25a] A. D. French, L. Schäfer, S. Q. Newton, *Carbohydr. Res.* **1993**, *239*, 51–60. – [25b] A. D. French, M. K. Dowd, *J. Mol. Struct. (Theochem)* **1993**, *286*, 183–201.

[26] F. W. Lichtenthaler, S. Immel, in *Sweet Taste Chemoreception* (Eds.: M. Mathlouthi, J. A. Kanters, G. G. Birch), Elsevier Appl. Science, London, **1993**, pp. 21–61.

[27] F. W. Lichtenthaler, S. Immel, *Internat. Sugar J.* **1995**, *97*, 12–22.

[28] [28a] H. J. Lindner, *Closed Shell PI-SCF-LCAO-MO Molecular Mechanics Program (PIMM88)*. Technische Hochschule Darmstadt, **1988**. – [28b] A. E. Smith, H. J. Lindner, *J. Comput.-Aided Mol. Des.* **1991**, *5*, 235–262.

[29] D. Cremer, J. A. Pople, *J. Am. Chem. Soc.* **1975**, *97*, 1354–1358.

[30] [30a] G. A. Jeffrey, J. H. Yates, *Carbohydr. Res.* **1979**, *74*, 319–322. – [30b] G. A. Jeffrey, R. Taylor, *Carbohydr. Res.* **1980**, *81*, 182–183.

[31] [31a] J. Brickmann, *Molecular Computer-Aided Design (MOLCAD)*, Technische Hochschule Darmstadt, **1992**. – [31b] J. Brickmann, T. Goetze, W. Heiden, G. Moeckel, S. Reiling, H. Vollhardt, C.-D. Zachmann, *Interactive Visualization of Molecular Scenarios with MOLCAD/SYBYL*, in *Data Visualization in Molecular Science – Tools for Insight and Innovation* (Ed.: J. E. Bowie), Addison-Wesley, Reading, Mass, **1995**, pp. 83–97.

[32] [32a] B. Lee, F. M. Richards, *J. Mol. Biol.* **1971**, *55*, 379–400. – [32b] F. M. Richards, *Ann. Rev. Biophys. Bioeng.* **1977**, *6*, 151–176. – [32c] M. L. Connolly, *J. Appl. Cryst.* **1983**, *16*, 548–558; *Science* **1983**, *221*, 709–713.

[33] [33a] P. J. Bernal, W. A. Van Hook, *J. Chem. Thermodyn.* **1986**, *18*, 955–968; 969–978. – [33b] M. V. Kaulgud, S. S. Dhondge, *Indian J. Chem. Sect. A*, **1988**, *27*, 6–11.

[34] W. F. van Gunsteren, H. J. C. Berendsen, *Groningen Molecular Simulation (GROMOS) Library Manual*, Biomos, Nijenborgh 16, Groningen, The Netherlands, **1987**.

[35] W. F. van Gunsteren, H. J. C. Berendsen, *Angew. Chem.* **1990**, *102*, 1020–1055; *Angew. Chem. Int. Ed. Engl.* **1990**, *29*, 992–1023.

[36] S. Immel, Doctoral Thesis entitled *Computersimulation of Chemical and Biological Properties of Saccharides: Sucrose, Fructose, the Cyclodextrins, and Starch*, Technische Hochschule Darmstadt, **1995**, pp. 109–146.

[37] M. P. Allen, D. J. Tildesly, *Computer Simulation of Liquids*, Clarendon Press, Oxford, **1987**.

[38] M. Eszterle, *Zuckerindustrie (Berlin)* **1993**, *118*, 459–464.

[39] G. M. Brown, D. C. Rohrer, B. Berling, C. A. Beevers, R. O. Gould, R. Simpson, *Acta Crystallogr. Sect. B*, **1972**, *28*, 3145–3158.
[40] [40a] K. Hirotsu, T. Higuchi, *Bull. Chem. Soc. Jpn.* **1976**, *49*, 1240–1245. – [40b] J. A. Kanters, G. Roelofson, H. M. Doesburg, T. Koops, *Acta Crystallogr. Sect. B*, **1976**, *32*, 2830–2837. – [40c] M. E. Gress, G. A. Jeffrey, D. C. Rohrer, *Acta Crystallogr. Sect. B*, **1978**, *34*, 508–512.
[41] J. Ohanessian, F. Longchambon, F. Arene, *Acta Crystallogr. Sect. B*, **1978**, *34*, 3666–3671.
[42] G. A. Jeffrey, W. Saenger, *Hydrogen Bonding in Biological Structures*, Springer Verlag, Berlin/Heidelberg, **1991**, pp. 309–350.
[43] H. J. Lindner, F. W. Lichtenthaler, *Carbohydr. Res.* **1981**, *93*, 135–140.
[44] M. Kunz in *Carbohydrates as Organic Raw Materials* (Ed.: F. W. Lichtenthaler), VCH, Weinheim/New York, **1991**, pp. 137ff.
[45] [45a] G. A. Jeffrey, *Carbohydr. Res.* **1973**, *28*, 233–241. – [45b] G. A. Jeffrey, M. E. Gress and S. Takagi, *J. Am. Chem. Soc.* **1977**, *99*, 609–611. – [45c] G. A. Jeffrey, S. Takagi, *Acc. Chem. Res.* **1978**, *11*, 264–270.
[46] A. van Hook, *Zuckerindustrie (Berlin)* **1987**, *112*, 587–600; C. Brennan, M. Mathlouthi, *ibid.* **1994**, *119*, 652–658.
[47] P. Kollman, *Chem. Rev.* **1993**, *93*, 2395–2417.
[48] M. J. Mitchel, A. McCammon, *J. Comput. Chem.* **1991**, *12*, 271–275.
[49] [49a] G. M. Torrie, J. P. Valleau, *Chem. Phys. Lett.* **1974**, *28*, 578–581; *J. Comput. Phys.* **1977**, *23*, 187–199; *J. Chem. Phys.* **1977**, *66*, 1402–1406. – [49b] M. Mezei, D. L. Beveridge, *Ann. N.Y. Acad. Sci.* **1986**, *482*, 1–23.
[50] A. E. Mark, S. P. van Helden, P. E. Smith, L. H. M. Janssen, W. F. van Gunsteren, *J. Am. Chem. Soc.* **1994**, *116*, 6293–6302.
[51] [51a] D. Avenel, A. Neuman, H. Gillier-Pandraud, *Acta Crystallogr., Sect.B*, **1976**, *32*, 2598–2605. – [51b] J. Becquart, A. Neuman, H. Gillier-Pandraud, *Carbohydr. Res.* **1982**, *111*, 9–21.
[52] D. C. Rohrer, *Acta Crystallogr., Sect. B*, **1972**, *28*, 425–433.
[54] G. A. Jeffrey, Y. J. Park, *Acta Crystallogr., Sect. B*, **1972**, *28*, 257–267.
[55] V. Ferretti, V. Bertolasi, G. Gilli, C. A. Accorsi, *Acta Crystallogr., Sect. C*, **1984**, *40*, 531–535.
[56] [56a] H. M. Berman, *Acta Crystallogr., Sect. B*, **1970**, *26*, 290–299. – [56b] G. A. Jeffrey, De-bin Huang, *Carbohydr. Res.* **1990**, *206*, 173–182.
[57] [57a] R. D. Gilardi, J. L. Flippen, *J. Am. Chem. Soc.* **1975**, *97*, 6264–6266. – [57b] R. Gilardi, J. L. Flippen-Anderson, *Acta Crystallogr., Sect. C*, **1987**, *43*, 806–808.
[58] B. Mulloy, T. A. Frenkiel, D. B. Davies, *Carbohydr. Res.* **1988**, *184*, 39–46.
[59] [59a] P. J. Card, W. D. Hitz, K. G. Ripp, *J. Am. Chem. Soc.* **1986**, *108*, 158–161. – [59b] W. D. Hitz, P. J. Card, K. G. Ripp, *J. Biol. Chem.* **1986**, *261*, 11986–11991. – [59c] N. Sauer, W. Tanner, *Bot. Acta* **1993**, *106*, 277–286.
[60] [60a] R. S. Shallenberger, *J. Food Sci.* **1963**, *28*, 584–589. – [60b] R. S. Shallenberger, T. E. Acree, *Nature (London)* **1967**, *216*, 480–482; *J. Agric. Food Chem.* **1969**, *17*, 701–703.
[61] L. B. Kier, *J. Pharm. Sci.* **1972**, *61*, 1394–1397.
[62] S. Immel, *Energy Contouring Program, Dissertation*, Technische Hochschule Darmstadt, **1995**.
[63] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, H. J. Hermans, *Intermolecular Forces* (Ed.: B. Pullmann), Reidel, Dordrecht, The Netherlands, **1981**, pp. 331–342; *Biopolymers* **1984**, *23*, 1513–1518.
[64] R. H. Stote, D. J. States, M. Karplus, *J. Chim. Phys.* **1991**, *88*, 2419–2433.
[65] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. Dinola, J. R. Haak, *J. Chem. Phys.* **1984**, *81*, 3684–3690.
[66] [66a] J.-P. Ryckaert, G. Cicotti, H. J. C. Berendsen, *J. Comput. Phys.* **1977**, *23*, 327–341. – [66b] W. J. van Gunsteren, H. J. C. Berendsen, *Mol. Phys.* **1977**, *34*, 1311–1327.
[67] D. J. Heisterberg, *QTRFIT – Rigid Body Rotation and Fitting Program*, The Ohio Supercomputer Center, Columbus, Ohio 43212, **1991**.

[95160]