

3

Sucrose : Generation of Molecular Electrostatic and Lipophilic Profiles and their Implications on Hydroxyl Group Reactivities and Sweetness Elicitation

Abstract: The mode of generation of **m**olecular **e**lectrostatic **p**otential (**MEP**'s) profiles and **m**olecular **l**ipophilicity **p**atterns (**MLP**'s) is described in detail, including the presently available possibilities for their graphical representation. Calculation of physico-chemical properties on molecular surfaces and their color-coded visualization provides new insights in the chemical and biological behavior of molecules. As exemplified for sucrose, accurate molecular models were obtained from conformational analysis in terms of the Φ/Ψ -intersaccharidic torsion angles using *in vacuo* molecular mechanics and molecular dynamics in aqueous solution, including free energy calculations following the "umbrella sampling" strategy. The relative reactivities of hydroxyl groups of sucrose after chemical or electrochemical deprotonation can be deduced from the corresponding MEP's, showing the glucosyl-2-OH group to be the most reactive one, followed by the 1'- and 3'-hydroxyl groups. Considering both MEP's and MLP's and their consequences on the hydrogen bonding ability of sucrose and its hydrophobic characteristics entails a substantial modification of prevailing sweetness concepts. Qualitative and quantitative structure-sweetness relationships on the basis of the MLP's lead to a comprehensive understanding of the relative sweetness of a large number of deoxy-halo-sucrose derivatives.

Sucrose and other sugars are over-functionalized with hydroxyl groups of almost identical properties, such that predictions of their relative chemical reactivities are highly speculative. Especially difficult is the assignment of relevant local electrostatic and hydrophobic molecular properties which determine their chemical and biological behavior. In view of the availability of modern molecular modeling techniques^[114], the MOLCAD-program^[48] in particular, it deemed opportune to probe this methodology with a structure as complex as sucrose. In this chapter, the computational basics for generating **m**olecular **e**lectrostatic **p**otential profiles and **l**ipophilicity **p**atterns (**MEP**'s and **MLP**'s) are detailed and the far-reaching chemical implications and biological significance is discussed for sucrose with the aim to gain further insights into hydroxyl group reactivities and the mechanism by which sweetness is elicited.

MEP's and MLP's are visual representations of molecular properties and interactions which may be of fundamental importance for molecular recognition processes. Attraction and repulsion between bioactive molecules are controlled by a variety of weak interactions: the van der Waals surfaces (or better "contact" surfaces^[46]) should

be complementary in the contact area, the local charge distributions should be of complementary sign close to the binding site in order to maximize Coulomb attraction (important for salt bridges and H-bonds), and finally hydrophobic parts of the molecules should be brought into close contact for an effective hydrophobic interaction^[115]. The repulsive interaction as a consequence of the intermolecular deformation and the overlap of two electron distributions and the Coulomb interaction as an electrostatic phenomenon are well understood today and can be adequately modeled by a local model approach. However, this is not true for the hydrophobic effect or the hydrophobic interaction^[116,117]. There are some attempts to establish a molecular lipophilicity potential based on phenomenological arguments, but there is still no unique concept available^[58,118,119]. In this chapter the argumentation considers a new approach to the analysis and display of local lipophilicity / hydrophobicity mapped on molecular surfaces which was published recently by Heiden *et al.*^[58].

Solid State and Solution Conformation of Sucrose

In order to reliably generate the electrostatic potential profiles and lipophilic patterns of sucrose, a detailed knowledge of its conformational properties is an essential prerequisite. The common chemical depiction used for sucrose (Fig. 3-1 A) neither gives any three-dimensional information about the actually predominating molecular conformation nor can it predict chemical properties. A more "realistic" molecular picture of sucrose, albeit for the solid state, can be derived from neutron diffraction^[25] and from X-ray analysis^[26] (Fig. 3-1 B), showing the glucose and fructose moiety fixed in their relative orientation by two interresidue hydrogen bonds between 5-O_g ... HO-6^f and 2-O_g ... HO-1^f.

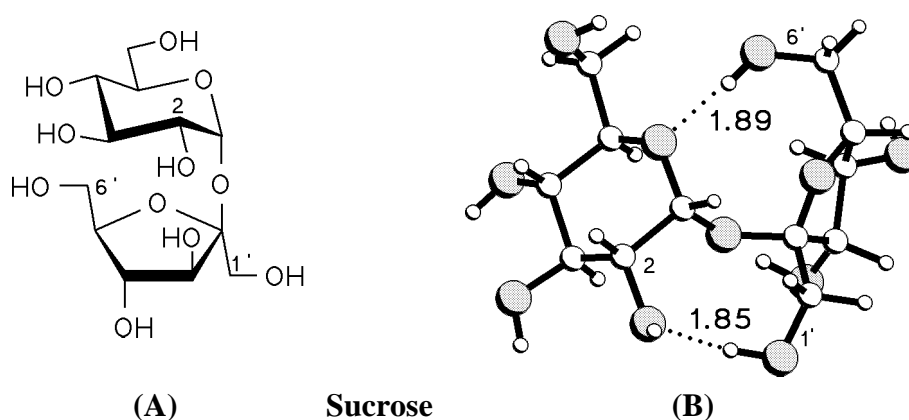


Fig. 3-1. Representation of the common chemical formula of sucrose (A) versus a ball and stick model of the solid state conformation (B) derived from neutron diffraction data^[25].

The solution conformation and flexibility of sucrose has been investigated by spectroscopical methods^[120,121], NMR-experiments^[27-37], force field based analysis^[38-44], and combined techniques^[39,40], since the solid state conformation cannot account for all the NMR-data^[34,37,39]. In solution, particularly in water, it is unlikely that both of the intramolecular hydrogen bonds present in the crystalline state are retained. Indeed, extensive X-ray^[121], ¹H- and ¹³C-NMR investigations^[29-32] strongly attest to the disintegration of the 5-O^g ... HO-6^f hydrogen bond by solvation, yet there is one NMR-study which even questions the existence of any stable intramolecular hydrogen bond in sucrose for aqueous solutions^[36]. For polar aprotic solvents like dimethyl sulfoxide, a competitive equilibrium between forms **A** and **B** (Fig. 3-2), with the former predominating in a 2 : 1 ratio, has been inferred^[30].

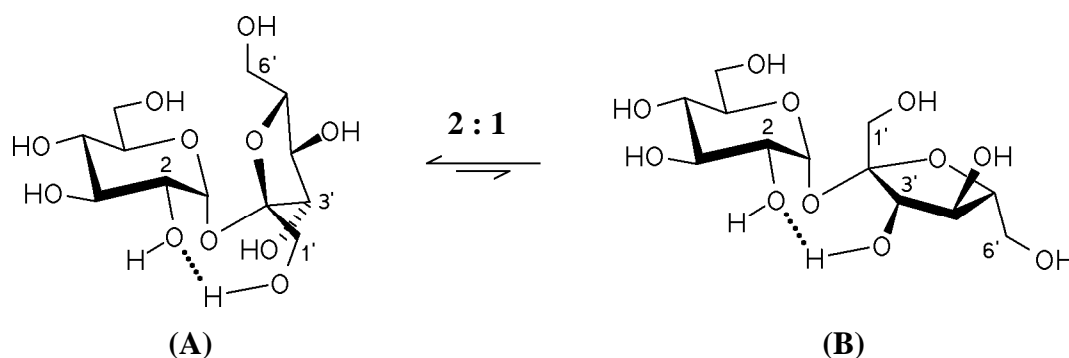


Fig. 3-2. Conformations of sucrose in solution: in DMSO a competitive equilibrium between **A** and **B** has been proposed^[30].

In conformer **A**, the glucose and fructose parts are fixed in their relative orientation by the same interresidue hydrogen bond as observed in the crystal lattice, thus the solid state conformation and **A** must correspond closely. Several calculations of the energy potential surface of sucrose^[29,38-44] have provided additional indications as to the relevance of the two forms.

Conformational Properties of the Isolated Sucrose Molecule

PIMM88^[45] force field based computation of the fully relaxed Φ ($O_5-C_1-O_1-C_2$) / Ψ ($C_1-O_1-C_2-O_5$) energy potential surface of sucrose (Fig. 3-3, for computational details see Appendix I) manifested three local energy-minima, differing mainly by the Ψ -torsion angle. The energy minimum at $\Phi \approx +100 - +110^\circ$ / $\Psi \approx -40 - -50^\circ$ corresponds closely to the crystal conformation of sucrose. The molecular geometries of sucrose associated with the observed energy minima are shown in Fig. 3-4, their molecular parameters are listed in Table 3-1 including those obtained from solid state structure analysis^[25].

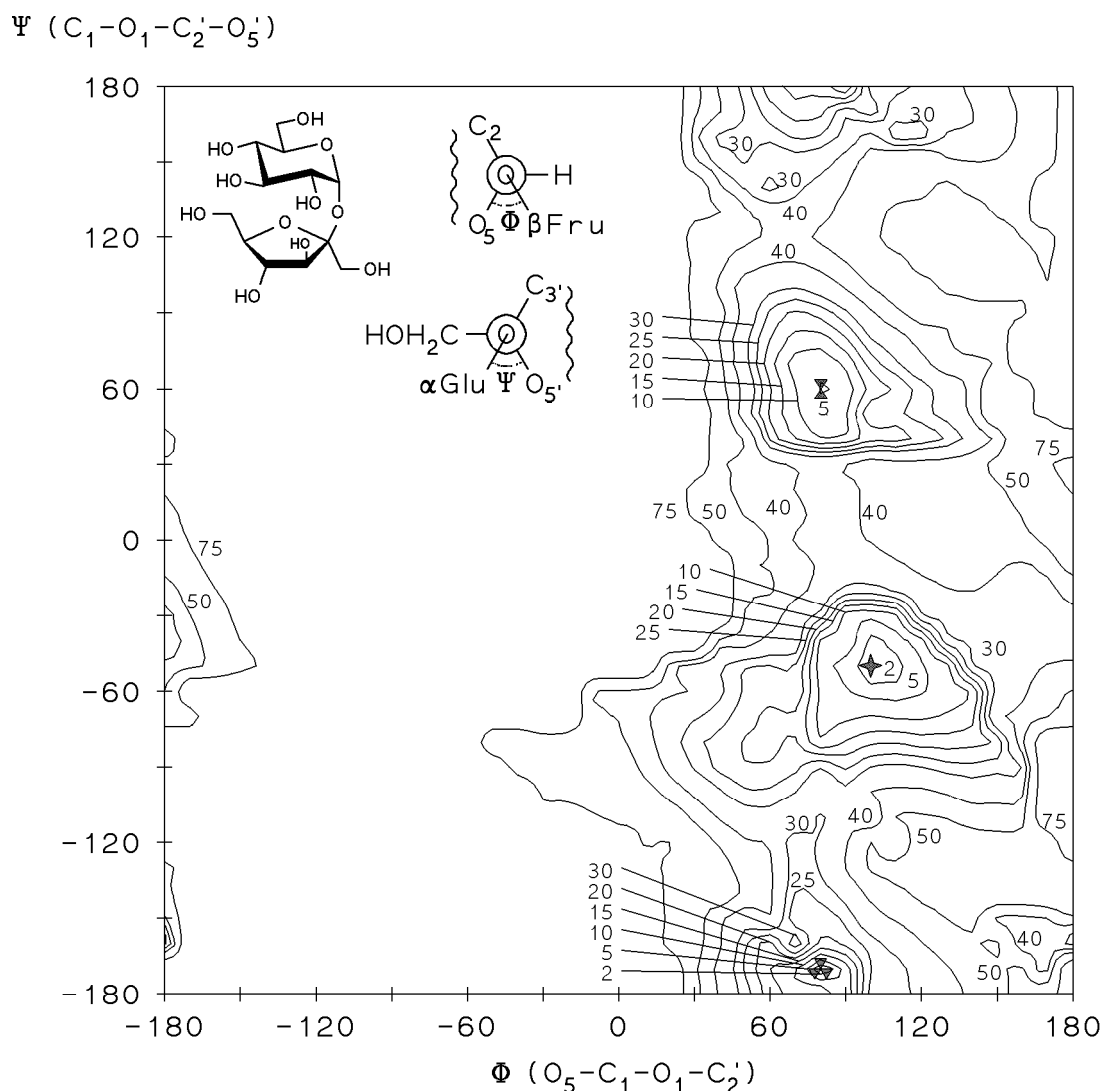


Fig. 3-3. Fully relaxed PIMM88^[45] energy potential surface of sucrose as a function of the intersaccharidic torsion angles Φ ($O_5-C_1-O_1-C_2'$) and Ψ ($C_1-O_1-C_2'-O_5'$), energy contours are given in kJ/mol relative to the global minimum.

As evidenced by the Cremer-Pople parameters^[122] as conformational descriptors of furanoid^[123] and pyranoid^[124] ring systems (cf. Table 3-1), the glucose residue adopts a 4C_1 -conformation, while the fructose unit invariably exhibits an almost ideal 4T_3 -geometry. Conformer **A** is characterized by nearly identical Φ / Ψ -torsion angles as observed in the solid state and the same interresidue hydrogen bond $2-O_g \cdots HO-1^f$, while the $5-O_g \cdots HO-6^f$ H-bond is missing. Moreover, both calculated sucrose geometries **A** and **B** of Fig. 3-4 are characterized by intramolecular $2-O_g \cdots HO-1^f$ (conformer **A**) and $2-O_g \cdots HO-3^f$ (form **B**) hydrogen bonds identical to those proposed on the basis of SIMPLE- 1H -NMR-experiments for DMSO-solutions^[30] (Fig. 3-2). Optical rotation data of sucrose also points towards an equilibrium between two different linkage conformers similar to **A** and **B**, with the former predominating^[41].

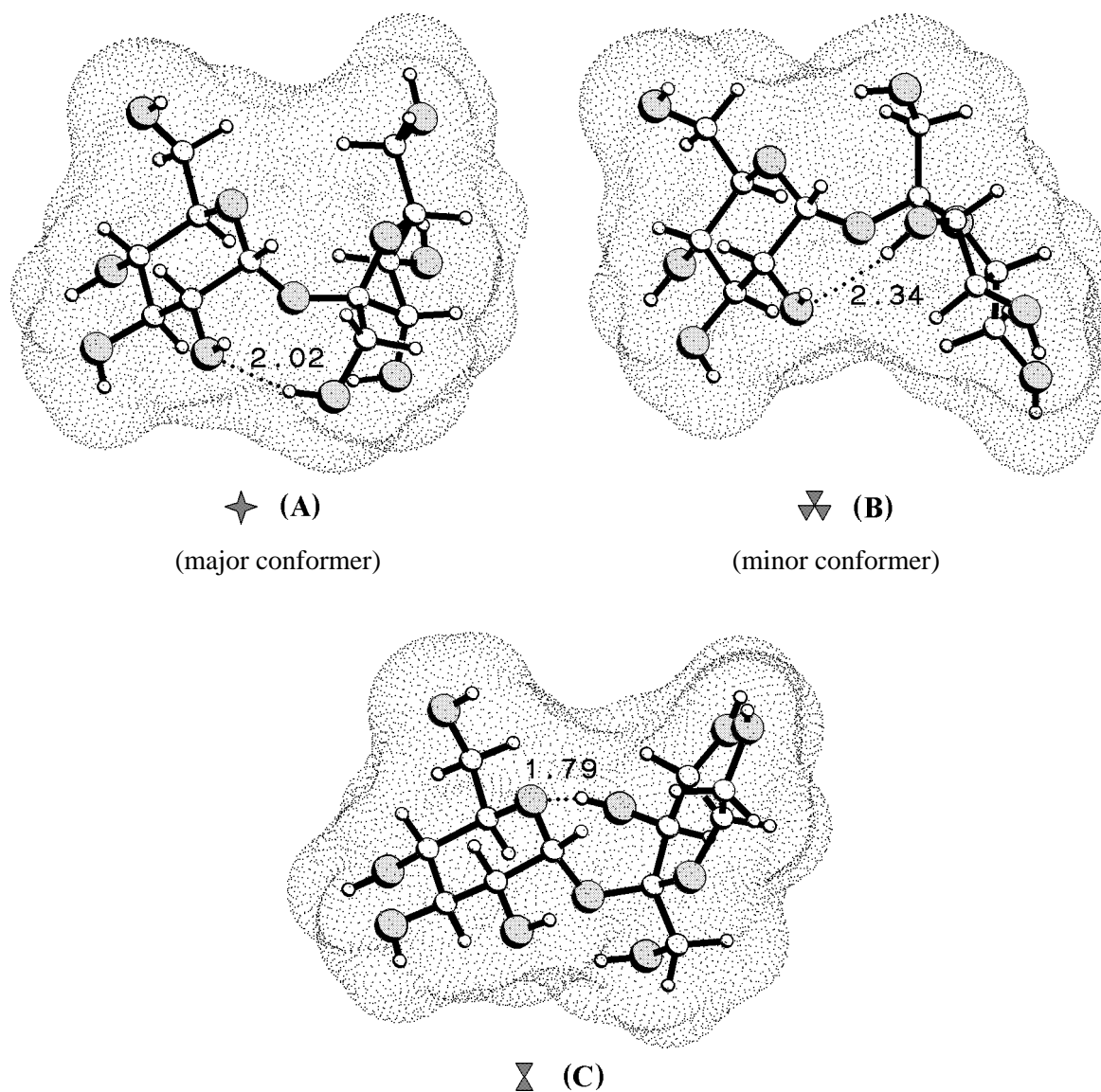


Fig. 3-4. Ball and stick model representation including the contact surface in dotted form of the low energy conformers of sucrose (cf. Fig. 3-3). Conformers **A** and **B** are likely to represent the conformations prevailing in solution^[30] (cf. Fig. 3-2), the third conformer **C** was not detected experimentally.

The surprisingly close agreement between the calculated data and the experimental observations is impressively emphasized by the 3D-plot of the percentage distribution shown in Fig. 3-5: the calculated ratio of 71 : 21 : 8 (**A** : **B** : **C**) for the three sucrose conformers is close to the experimentally observed proportions of 67 : 33 : 0 in DMSO-solution^[30].

Table 3-1. Molecular parameters and percentage distribution of sucrose conformers as obtained from Fig. 3-2 and 3-3.

Conformer	ΔH_F^{298} [kJ/mol]	Φ [°] (O ₅ -C ₁ -O ₁ -C ₂)	Ψ [°] (C ₁ -O ₁ -C ₂ -O ₅)	H-Bonds	d [Å] (O \cdots H)	ϕ [°] (O \cdots H-O)	distribution [%] calc.	exp. ^{e)}
A ^{a)}	-1741.8	+105.4	-47.9	2-O ^g \cdots HO-1 ^f	2.02	145.7	71	67
B ^{b)}	-1742.2	+80.4	-170.0	2-O ^g \cdots HO-3 ^f	2.34	163.2	21	33
C	-1737.3	+80.1	+60.0	5-O ^g \cdots HO-3 ^f	1.79	155.4	8	– ^{d)}
crystal ^[25]	–	+107.9	-44.8	2-O ^g \cdots HO-1 ^f	1.85	158.6	–	–
				5-O ^g \cdots HO-6 ^f	1.89	167.1	–	–

Conformer	Cremer-Pople parameters ^[122]						
	Q	– glucose residue –			– fructose residue –		
		θ	$\phi^e)$	conformation	q	ϕ	conformation
A ^{a)}	0.555	5.7	88.6	⁴ C ₁	0.411	261.0	⁴ T ₃ (\rightarrow E ₃)
B ^{b)}	0.559	6.0	88.6	⁴ C ₁	0.433	261.3	⁴ T ₃ (\rightarrow E ₃)
C	0.550	3.3	50.9	⁴ C ₁	0.352	279.8	⁴ E (\rightarrow ⁴ T ₃)
crystal ^[25]	0.556	5.1	183.8	⁴ C ₁	0.352	265.2	⁴ T ₃ (\rightarrow E ₃)

a) corresponding to conformer **A** in Fig. 3-2. – b) conformer **B** (Fig. 3-2). – c) in DMSO^[30]. – d) not detected. – e) ϕ is meaningless for $\theta \rightarrow 0^\circ$ (⁴C₁-conformations are identical to ²C₅ and ⁰C₃).

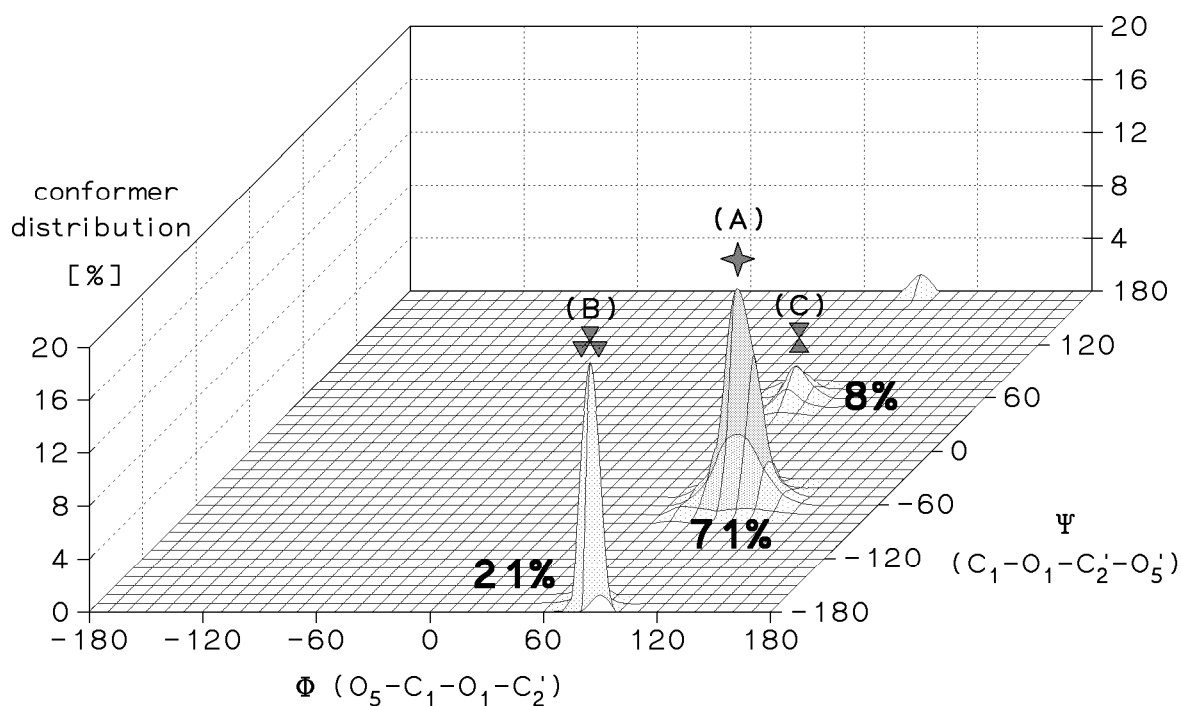


Fig. 3-5. 3D-Plot of the percentage distribution of sucrose conformers as a function of the Φ/Ψ -torsion angles, calculated from the energy potential surface (Fig. 3-3) according to the Boltzmann eq. for $T = 300$ K. SIMPLE-¹H-NMR measurements in DMSO^[30] indicate a 67 : 33 ratio of conformers **A** and **B**, without detecting a third one.

Despite the wide diversity of computational methods and force fields used for theoretical treatment of sucrose (HSEA^[29], CHARMM^[38], PFOS^[39,40] with and without H-bond energy terms, MM3^[40], MM3(92)^[43], MM2HB^[44], PIMM88, PCILO^[40], and 4-21/G *ab initio* calculations on the model compound tetrahydro-2-[(tetrahydro-2-furanyl)oxy])-2-*H*-pyran^[42]) and the fact that the respective locations and relative energies of the minima and barriers differ to quite some extent, some general trends are independently reproduced by all studies: the Φ -torsion around the axial C₁-O₁-bond is more restricted as compared to the Ψ -angle (pseudoaxial O₁-C₂-bond). In all cases, three main families of sucrose conformers are found with Φ / Ψ -angles of approx. +70 – +110° / -40 – -90° (**A**), +80 – +100° / +170 – -170° (**B**), +90 – +120° / +50 – +70° (**C**), which approximately coincide with the energy minima found here. The global energy minimum (**A**) invariably resembles closely to the solid state conformation, while the minima **B** and **C** exhibit almost equally higher energies (+8 – +25 kJ/mol). From PCILO calculations^[40] it was concluded that the latter might contribute substantially to the conformational equilibrium in different solvents, with the intramolecular H-bonds playing a minor role in determining the conformation of isolated sucrose molecules in solution^[39]. On the contrary, this study here clearly favors conformer **B** over **C**.

More recent 4-21/G* *ab initio* energy surface calculations of the tetrahydropyran-tetrahydrofuran analog around the sucrose crystallographic zone revealed the fundamental conformation determining information to rest in this molecular frame^[125]. While the MM3 map differs therefrom quite substantially, the location of the steeper shaped main minimum on the PIMM88 map coincides with the *ab initio* results and crystal structure derived informations on α -glucopyranosyl- β -(1→2)-fructofuranose linkages^[125]. Seemingly, the PIMM88 force field is better suited for treatment of the overlapping anomeric effects in sucrose than the MM3 parameters are.

However, the small differences in Φ / Ψ and energies found for the energy minima resemble uncertainties of energy calculations inherent in the different force fields and computational strategies. Relaxed energy potential surfaces (CHARMM, MM2-3, PIMM88) must be considered more valuable than rigid residues ones (HSEA, PFOS): in particular the fructofuranoside ring system is highly flexible and can adopt many different geometries to release energy strain, e.g. from interresidue steric repulsions. This leads to a broadening and smoothing of the energy minima and the contours, giving a more realistic picture of the molecular properties as already have been demonstrated for disaccharides^[43,126]. Due to the exclusion of vibrational effects, the extraction of energy barriers involved in conformational transitions might even be tenuous and misleading as evidenced by the wide range of values proposed for sucrose (**A** → **B**: 15 – 40 kJ/mol, **A** → **C** and **B** → **C**: 20 – >40 kJ/mol).

Molecular Dynamics of Sucrose in Aqueous Solution

In order to get an anticipation if the results obtained for the isolated molecule also apply to aqueous solutions, extensive molecular dynamics (MD) simulations of sucrose with explicit incorporation of surrounding water molecules were carried out. For this task it was necessary to use the simpler force field GROMOS^[127] (Fig. 3-6, for details see Appendix II).

The PIMM88 structures **A** – **C** as different starting geometries all relaxed within the first 100ps of simulation time to a single energy minimum substate closely resembling the most stable PIMM88 conformer (**A**) and the solid state conformer in terms of its Φ / Ψ -torsion angles. No further transitions occurred, but considerable flexibility of the intersaccharidic linkage was observed by vibration of the glucose and fructose unit against each other. As clearly evident, the Ψ -torsion and the fructosyl residue is more flexible than the Φ -angle and the glucose moiety (Fig. 3-7).

The hydroxymethyl groups exhibit multiple conformational changes in the 50 – 75ps range, and the fructofuranose ring undergoes fast pseudorotation with many transitions between distinctive conformational regions. The analysis of distributions sampled during the MD-trajectories according to Boltzmann's law (300K), allows for calculation of the molecular energy of sucrose as a function of the Cremer-Pople^[122] ring puckering parameters q and ϕ , which serve as conformational descriptors for five membered rings^[123]. The corresponding fully relaxed energy potential surface (Fig. 3-8) exhibits two clearly separated energy minima: the northern global minimum at $q = 0.40 - 0.45$ and $\phi = 270 - 290^\circ$ corresponds to ${}^4T_3 \leftrightarrow {}^4E$ ring conformations. The extended southern minimum (+4 – +6kJ/mol, $q = 0.30 - 0.40$ and $\phi = 10 - 100^\circ$) 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 (this work, Chapter 5) the former 4T_3 type conformations are characteristic for β -D-fructofuranose derivatives not only in the solid state, but also in solution. Apparently, pseudorotation via eastern pathways is preferred over western transitions.

Intramolecular hydrogen bonds proved to be very stable during a 2000ps simulation for vacuum conditions ($\epsilon = 1$). During this calculation only one 30ps flip-over of the "clockwise" hydrogen bond network (3'-OH ... 1'-OH ... 2-OH ... 3-OH ... 4-OH ... 6-OH) to the alternative "anti-clockwise" configuration (6-OH ... 4-OH ... 3-OH ... 2-OH ... 1'-OH ... 3'-OH) was observed after approx. 680ps. In the solvent-including MD's these H-bonds are entirely disintegrated since the hydroxyl groups can satisfy their hydrogen bonding requirements by interaction with water molecules in the first hydration layer of the solute. Even the most favorable 2-O^g ... HO-1^f hydrogen bond (*vide infra*) is interrupted to a large extent by the solvent.

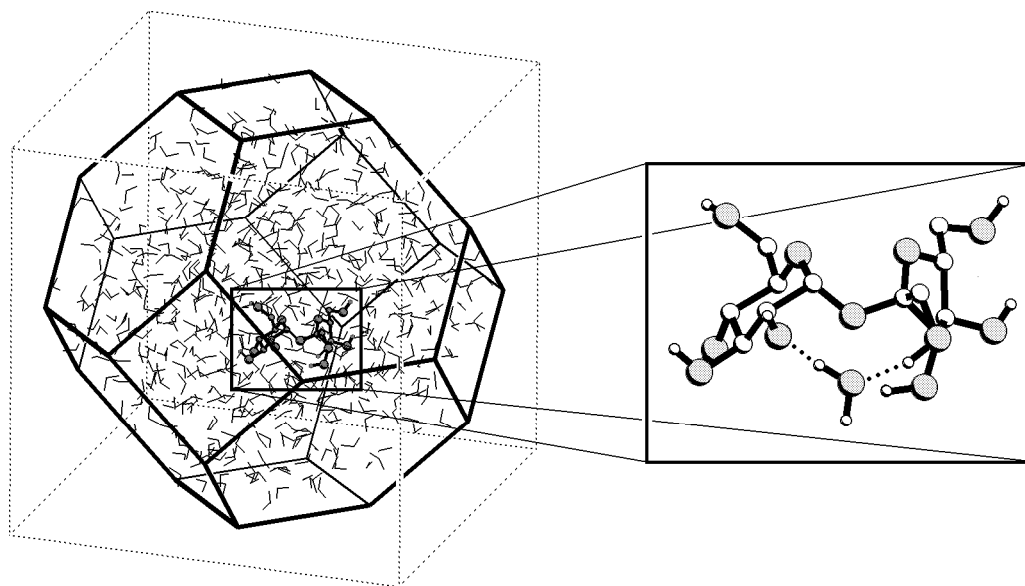


Fig. 3-6. Snapshot of a MD simulation of sucrose surrounded by 571 water molecules (boxsize of the truncated octahedron 32.9 Å) with a water molecule hydrogen bonded to O-2^g and O-1^f.

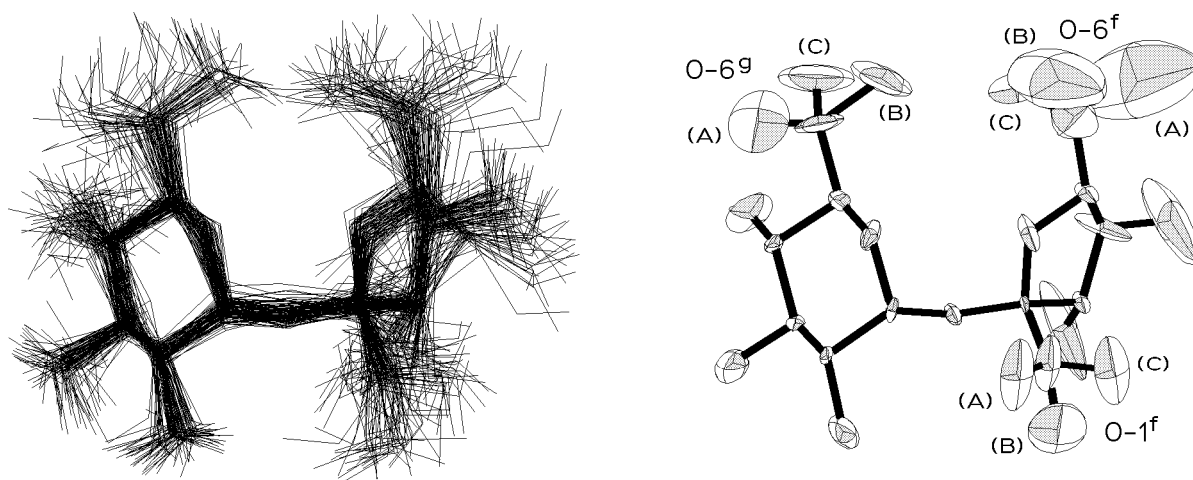


Fig. 3-7. Superimposition of 100 sucrose snapshot-geometries (*left* plot) taken in 5 ps intervals from a 500 ps MD simulation including 571 H₂O. Least-squares fitting was performed by rigid body translation and rotation^[128] of the molecules by considering the tetrahydro-2-[(tetrahydro-2-furanyl)oxy]-2*H*-pyran backbone of sucrose only, CH-hydrogen atoms were omitted for clearness. In the *right* plot, the atomic mean positions and anisotropic thermal probability ellipsoids at the σ -level as obtained from the MD are shown. The three staggered conformers for each of the hydroxymethyl units^[66,102] are indicated separately (a rare transition of the glucose CH₂OH-unit to the *tg* conformations was also observed!), hydrogen atoms were omitted. The thermal flexibility of the fructose residue is considerable larger than of the glucose moiety.

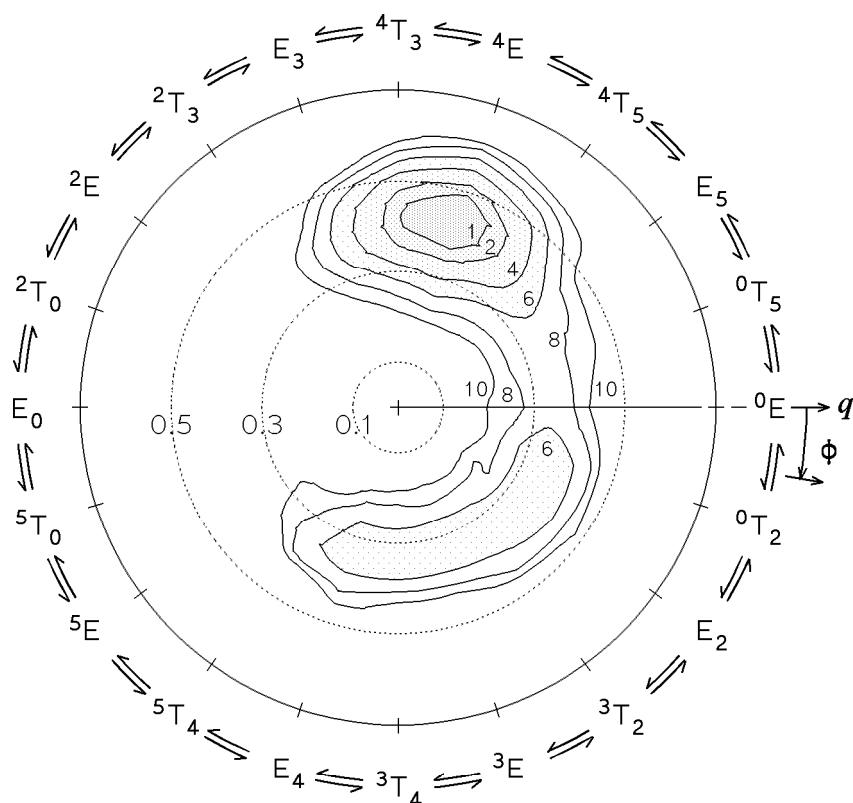


Fig. 3-8. Polar coordinate contour plot of the β -D-fructofuranose pseudorotational energy potential surface as a function of Cremer-Pople puckering parameters q and ϕ (500 ps MD simulation of sucrose including 571 water molecules). Energy contours are given in kJ/mol relative to the global minimum, dashed lines indicate iso-contour lines of the puckering amplitude at $q = 0.1, 0.3,$ and 0.5). The global energy minimum of the furanose ring conformations around ${}^4T_3 \leftrightarrow {}^4E$ is accentuated by grey shading.

The radial pair-distribution functions* of water protons and oxygen atoms in relation to the hydroxyl groups (500 ps simulation of sucrose + 571 H_2O molecules) show a strong first hydration sphere, and a weaker second one. The two-dimensional function calculated for water protons around the $O-2^g$ and $O-1^f$ atoms (Fig. 3-9) indicates the interresidue H-bond being replaced by a single water molecule, forming simultaneous interactions with both hydroxyl groups (cf. Fig. 3-6). In addition, trajectory analysis showed variable water molecules to be hydrogen bonded concurrently with both OH-groups, and thus confirms this effect to be related not only to statistical properties of the pair-distribution functions. The high preference of the $2-O^g \cdots H_2O \cdots HO-1^f$ water bridge directionality over the $2-OH^g \cdots OH_2 \cdots O-1^f$ configuration strongly reminds to the H-bond situation observed in the crystal structure

* Radial pair-distribution functions give the probability of solvent atoms (e.g. water protons) to occupy positions at a certain distance around solute atoms in relation to a purely random distribution, they are used to estimate solvation spheres.

and DMSO-solutions, with the water molecule being stripped off in non-aqueous and less polar environments and replaced by the direct 2-O^g ... HO-1^f interaction. Most notably, the water mediated interresidue bridge is consistent with NMR observations of the absence of any stable intramolecular hydrogen bonds in sucrose in aqueous solutions^[36].

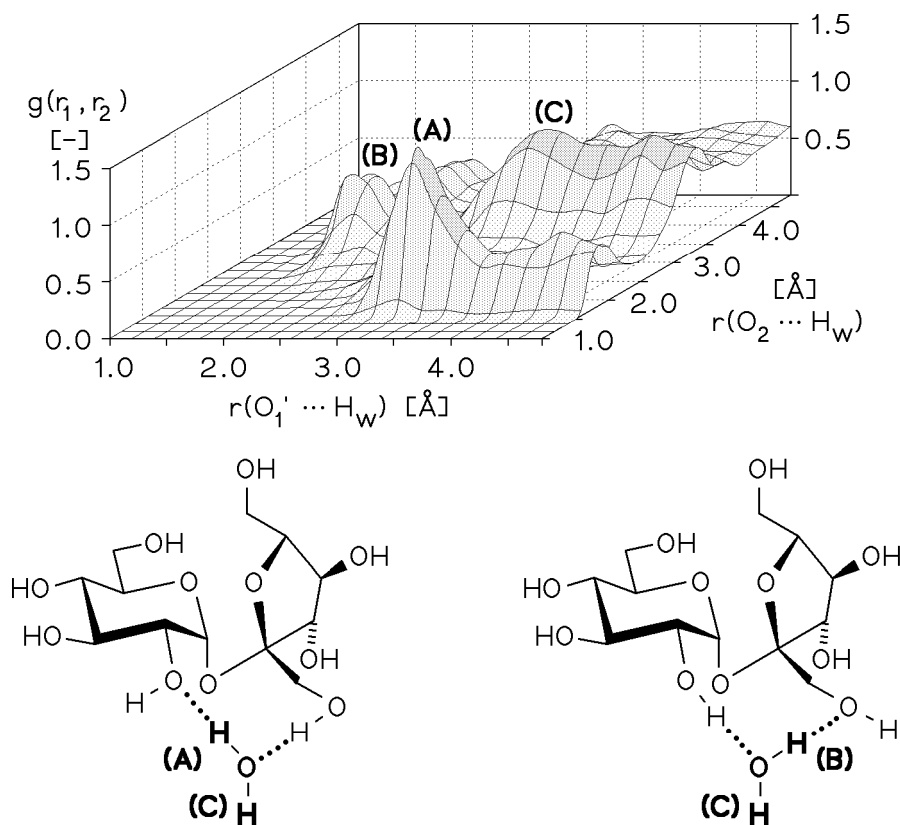


Fig. 3-9. Two dimensional pair-distribution function $g(r_1, r_2)$ of water protons around O-1^f (r_1) and O-2^g (r_2) of sucrose. The three observed peaks (A – C) can be traced back to water molecules forming simultaneously hydrogen bonds to the 1^f- and 2-OH groups differing in their directionality (*lower* formulas). The main peak (A) corresponds to the preferred configuration with a water proton H-bonded to 2-OH^g ($r_1 \approx 3.0 \text{ \AA}$, $r_2 \approx 2.0 \text{ \AA}$, *left* formula), the minor peak (B) to the inverse situation with the proton bonded to 1-OH^f ($r_1 \approx 2.0 \text{ \AA}$, $r_2 \approx 3.0 \text{ \AA}$, *right* formula), and the maximum (C) to the non-bonded proton ($r_1 \approx r_2 \approx 3.0 \text{ \AA}$). Correspondingly, a maximum occurrence probability of water oxygen atoms around O-1^f and O-2^g (not shown here), was found for $r_1 \approx r_2 \approx 2.8 \text{ \AA}$.

Relative Stabilities of Sucrose Solution Conformations

The missing conformational transitions of the intersaccharidic torsion angles of sucrose during the MD simulations neither allow an assessment of the relative free energy differences of conformational substates, nor the calculation of their contribution to the solution equilibrium. Free energy calculations^[129] by

thermodynamic integration require very long simulation times to obtain precise values^[130] and proved insuitable for disaccharide investigations^[44]. The more efficient non-Boltzmann "*umbrella sampling*"^[131] technique was used here to establish the free energy profile of sucrose in aqueous solution as a function of the most flexible intersaccharidic torsion angle Ψ ($\rightarrow \Delta G(\Psi)$).

Umbrella Sampling Procedure

Calculation of free energy differences $\Delta G(\varphi)$ from MD-sampled distributions $P(\varphi)$ in respect to any reaction or conformational coordinate φ (eq. 3-1) would require impossible long simulation times to ensure proper sampling over the whole conformational space including all substates.

$$G(\varphi) = -RT \ln P(\varphi) \quad (3-1)$$

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

$$G_i(\varphi) = -RT \ln P_i^*(\varphi) - U_i^*(\varphi) + const \quad (3-2)$$

$$P_i(\varphi) = const' \cdot P_i^*(\varphi) \exp(U_i^*(\varphi) / RT) \quad (3-3)$$

$$U_i^*(\varphi) = K_i [1 + \cos(n_i \varphi - \Theta_i)] \quad (3-4)$$

In contrast to the "*adaptive umbrella sampling*" method, where the *umbrella potential* is successively modified to finally yield $U^*(\varphi) = -G(\varphi)$ and a uniform distribution $P^*(\varphi)$ ^[44] – which is certainly hard to sample properly for large rotating groups like two monosaccharide units – variable potentials $U_i^*(\varphi)$ in consecutive molecular dynamics simulations were used. As indicated by alternative indices i , independent distributions $P_i^*(\varphi)$ were sampled for different φ -regions (for details see Appendix III). The force constant K_i , the multiplicity n_i , and the phase shift Θ_i in $U_i^*(\varphi)$ were chosen appropriately to explore as much of φ -space as possible, and simultaneously providing sufficient overlap of different $P_i^*(\varphi)$ distributions. These overlap regions are used to assemble the corresponding incomplete $G(\varphi)$ or $P(\varphi)$ functions by a least squares fitting procedure weighted by $(P_i^*(\varphi) \cdot P_j^*(\varphi))^{1/2}$ to an

overall fitting curve covering the entire ϕ -space*. The hysteresis of the overall $G(\phi)$ function obtained from fitting of distributions in the same ϕ -region generated by different conformational pathways was used for error estimation.

In this work the conformation coordinate for sucrose was defined in terms of the intersaccharidic torsion angle as $\phi \equiv \Psi$ and an *umbrella potential* of the form of eq. 3-4 was added**.

Free Energy Profile of Sucrose in Aqueous Solution

In Fig. 3-10, the GROMOS derived conformational free energy profile is shown as a function of the intersaccharidic torsion angle Ψ ($C_1-O_1-C_2-O_5$). $\Delta G(\Psi)$ was elaborated as a 10-term Fourier transform fit (refined parameters given in Table 3-2) in regard to the global energy minimum. The inherent relative error was estimated to be $\sigma(\Delta G(\Psi)) \approx \pm 1.0$ kJ/mol. The different conformational sub-families were denoted **A** – **C** in analogy to the minima of the energy potential surface of Fig. 3-3. The global energy minimum (**A**, $\Psi \approx -60^\circ$) again closely resembles the conformation in the crystalline state, while the conformer **B** ($\Psi \approx -170^\circ$) is approx. +10.3 kJ/mol higher in energy. For **C** not an energy minimum but a saddle point ($\Delta G \approx +13$ kJ/mol) is found around $\Psi \approx +30^\circ$.

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 **A** into **B**, a barrier of +16.5 kJ/mol ($\Psi \approx -125^\circ$) is computed, while for the pathway **B** \rightarrow **A** only a small barrier of +6.3 kJ/mol is found. The full range rotation of the glucose and fructose in respect to each other involves the transition **A** \rightarrow **B** \rightarrow **C** \rightarrow **A**, for which a central hindrance of +34.8 kJ/mol ($\Psi \approx +105^\circ$) has to be surmounted.

* The weighting factor relates to an error of $\sigma(P_i^*(\Psi)) \sim N_i(\Psi)^{1/2}$ with $N_i(\Psi)$ being the number of events sampled for the respective class of Ψ . Fitting of the sub-curves can be carried out either in terms of free energy (adding an energy offset $const_{i,j}$, eq. 3-2) or in probabilities (multiplication with a normalization factor $const'_{i,j}$, eq. 3-3). The latter is most conveniently carried out on a logarithmic scale, thus both methods and the results obtained therefrom are essentially identical.

** Actually, the umbrella potential was applied as a function of Ψ' ($C_1-O_1-C_2-C_3$), since Ψ ($C_1-O_1-C_2-O_5$) does not occur explicitly in the GROMOS molecular topology definition for sucrose. The more convenient function $\Delta G(\Psi)$ can be obtained via the strong correlation of $\Delta G(\Psi) \approx \Delta G(\Psi' + 115.7^\circ)$ due to the rather stiff pyranose ring system. Equivalently, the probability distribution for each small ϕ -interval can be calculated from

$$P_i(\phi + \Delta\phi) \sim \sum \exp[U_i^*(\phi')/RT] \quad (3-3b)$$

as the sum over all structures within each class $\phi + \Delta\phi$ for every umbrella run i (cf. eq. 3-3), whereas ϕ and ϕ' are different molecular functions, e.g. $\phi \equiv \Psi$ and $\phi' \equiv \Psi'$.

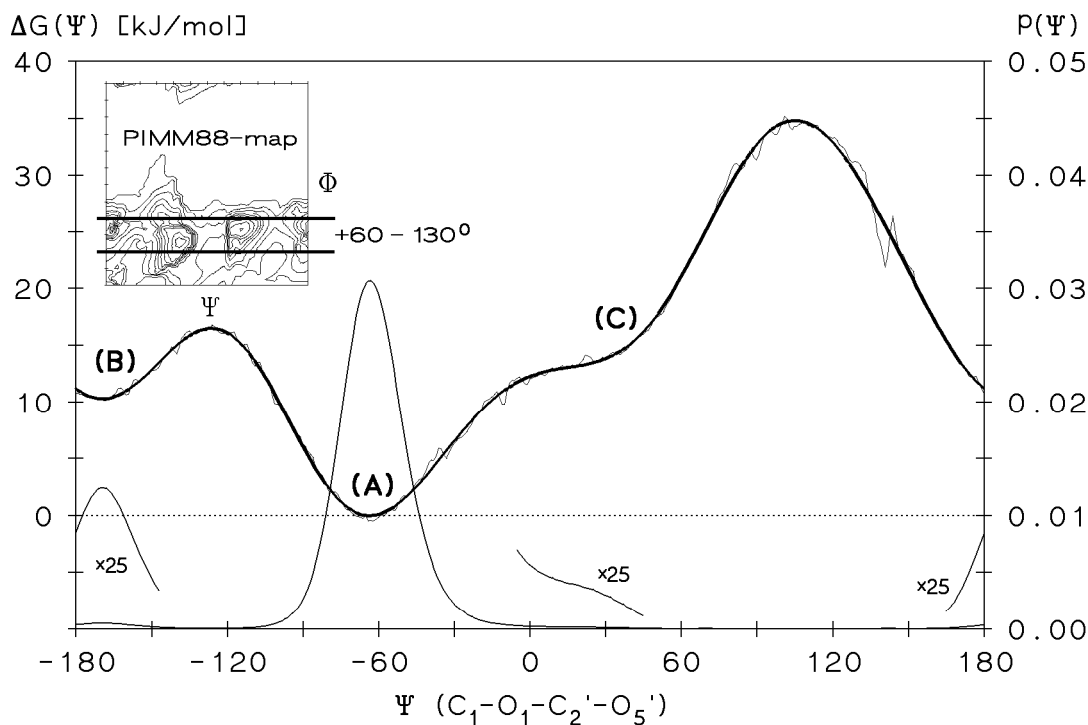


Fig. 3-10. Calculated free energy profile of sucrose conformations in aqueous solution as a function of the Ψ -torsion (*umbrella sampling* using the GROMOS^[127] force field, energies in kJ/mol relative to the global energy minimum), the thick line delineates the Fourier transform fit of the sampled ΔG -values (thin line) with conformational regions **A** – **C** marked, the estimated error is $\sigma(\Delta G(\Psi)) \approx \pm 1.0$ kJ/mol. The ΔG -function correlates with the cross section of the PIMM88 energy potential surface (Fig. 3-3) as indicated in the *upper left* pictogram ($\Phi \approx +60 - +130^\circ$, $\Psi = -180 - +180^\circ$). The Boltzmann distribution function $P(\Psi)$ (300 K, *right y-axis*, normalized to $\int P(\Psi) d\Psi = 1$) is shown in addition.

Table 3-2. 10-term Fourier transform fitting parameters (all in kJ/mol) of the free energy function $\Delta G(\Psi)$ for the conformations of sucrose in aqueous solution, higher order terms become insignificantly small.

$\Delta G(\Psi) = a_0 / 2 + \sum a_i \cos(i \Psi) + b_i \sin(i \Psi)$ 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 – 200ps MD simulations starting from different conformations (cf. above) are consistent with the absence of any energy barrier (**C** \rightarrow **A**) or a relatively small

hindrance of $\leq 2.5 RT$ (**B** \rightarrow **A**). The conformation **A** corresponds to the deep energy minimum and proved to be stable during all simulations (altogether approx. 1500ps).

The results correlate well with calculations of Hoofdt^[44], who on the basis of the same force field found three minima for **A**, **B**, and **C** with relative energies of ± 0.0 ($\Psi \approx -65^\circ$), $+9.7$ ($\Psi \approx -165^\circ$), and $+11.0$ kJ/mol ($\Psi \approx +40^\circ$), respectively. The corresponding transition barriers (**A** \rightarrow **B**: $+13.2$ kJ/mol, **A** \rightarrow **C**: $+12.4$ kJ/mol) are lowered to a little extent, while the central energy maximum of $+33.8$ kJ/mol ($\Psi \approx +115^\circ$) is of the same height and within the error margins as the barrier found here.

From the relative sucrose conformer distribution $P(\Psi)$, as shown additionally in Fig. 3-10, the equilibrium ratio of **A** : **B** is estimated as 98 : 2, the contribution of **C** is negligible small ($< 0.5\%$).

Free energy differences and transition barriers strongly depend on the definition of the reaction coordinate. The energy profile $\Delta G(\Psi)$ given in Fig. 3-10 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 $P_i(\Phi + \Delta\Phi, \Psi + \Delta\Psi) \sim \sum \exp[U_i^*(\Psi')/RT]$ for all MD structures within classes $\Phi + \Delta\Phi / \Psi + \Delta\Psi$ (cf. eq. 3-3b), results in partial energy potential surfaces. Subsequent fitting of the individual umbrella data sets as described above yields the section of GROMOS free energy potential surface shown in Fig. 3-11, which can be related to the PIMM88 surface. Calculating Φ / Ψ -probabilities increases the computational error to approx. $\sigma(\Delta G(\Phi, \Psi)) \approx \pm 2.0$ kJ/mol. The contours and the corresponding distributions in Fig. 3-11 become smoother and broader than in the PIMM88 map. Most notably, the minima of the conformational sub-families **A** and **B** of sucrose are located within similar regions ($\pm 15^\circ$) in both cases (**A**: $\Phi \approx +90^\circ$, $\Psi \approx -60^\circ$, and **B**: $\Phi \approx +90^\circ$, $\Psi \approx \pm 180^\circ$ for the GROMOS force field, for the PIMM88 data compare Table 3-1), but with different relative energies. The conformational family **A** invariably corresponds to the global energy minimum, while **B** is $\approx +12$ kJ/mol higher in energy on the GROMOS surface. The **C**-region even becomes a high-energy plateau ($\approx +15 - +20$ kJ/mol) rather than a minimum. The transition barriers (**A** \rightarrow **B**: $15 - 20$ kJ/mol, **A** \rightarrow **C** \rightarrow **B**: $35 - 40$ kJ/mol) were already estimated from the one-dimensional function $\Delta G(\Psi)$, small differences demonstrate the susceptibility of free energy analysis in respect to different definitions of the reaction coordinates.

However, it should be noted that these calculations fulfill van Gunsteren's criteria to exclude methodological errors in free energy calculations^[132]. First of all (and most

importantly) the free energy change for the closed thermodynamic cycle of the full range rotation around Ψ is close to zero ($< 2\text{kJ/mol}$), and second, the results are independent from the equilibration times applied for analysis of the MD-runs (for details see Appendix).

Since accurate sampling of the conformational space was established, the accuracy of the calculations itself depends solely on the GROMOS force field. From the 3D-plot of the probability distribution (Fig. 3-11) it becomes clear that in terms of the GROMOS force field only the conformational substate **A** – which again resembles closely to the solid state conformation – is to be considered relevant in solution.

Certainly, in relation to experimental data, the free energy differences between both conformers of **A** and **B** are being overestimated by the GROMOS force field, and correspondingly, the relative importance of **B** in the solution equilibrium is underestimated. However, all these results together clearly point towards the two PIMM88 models **A** and **B** to represent the major and minor component in the conformational equilibrium of sucrose in solution. The third conformer **C** was not detected by experimental methods and seems to have no physical relevance, therefore **C** is not being considered in the following discussions*. Both structures **A** and **B** have to be entered into more detailed molecular modeling studies to obtain not solely the basic (but itself useless) geometry information, but to gain even more insights in the chemical and biochemical behavior of this molecule.

In addition to the ball and stick models, the MOLCAD-generated^[48] contact surfaces^[46,47] in respect to a water molecule (i.e. "how water sees the molecule") are represented for both sucrose conformers in Fig. 3-4. The close agreement between the spatial volume included by these surfaces (solid state structure: 339.3, conformer **A**: 342.2, **B**: 343.8, and **C**: 340.9 Å³) and the apparent molar volume ϕV demanded by sucrose in different solvents like H₂O, DMSO, and DMF, for example ($\phi V \approx 203\text{-}212\text{cm}^3/\text{mol} \approx 337\text{-}351\text{Å}^3$)^[133,134], demonstrates the correct reproduction of molecular steric properties by the contact surfaces.

Calculation of closed surface models is achieved by triangulation^[135] of all dots, i.e. connecting all points in such a way that a closed net of triangles is obtained. As demonstrated in the sequel, for each of the surface points different physico-chemical molecular properties can be calculated, the respective numeric values may be transferred into a color-code for visualization on the closed surfaces^[48].

* MD simulations only yield averaged ("virtual") molecular conformations, which might have no physical meaning itself. MM structures are well-defined geometries, which can directly be used for more detailed molecular modeling studies.

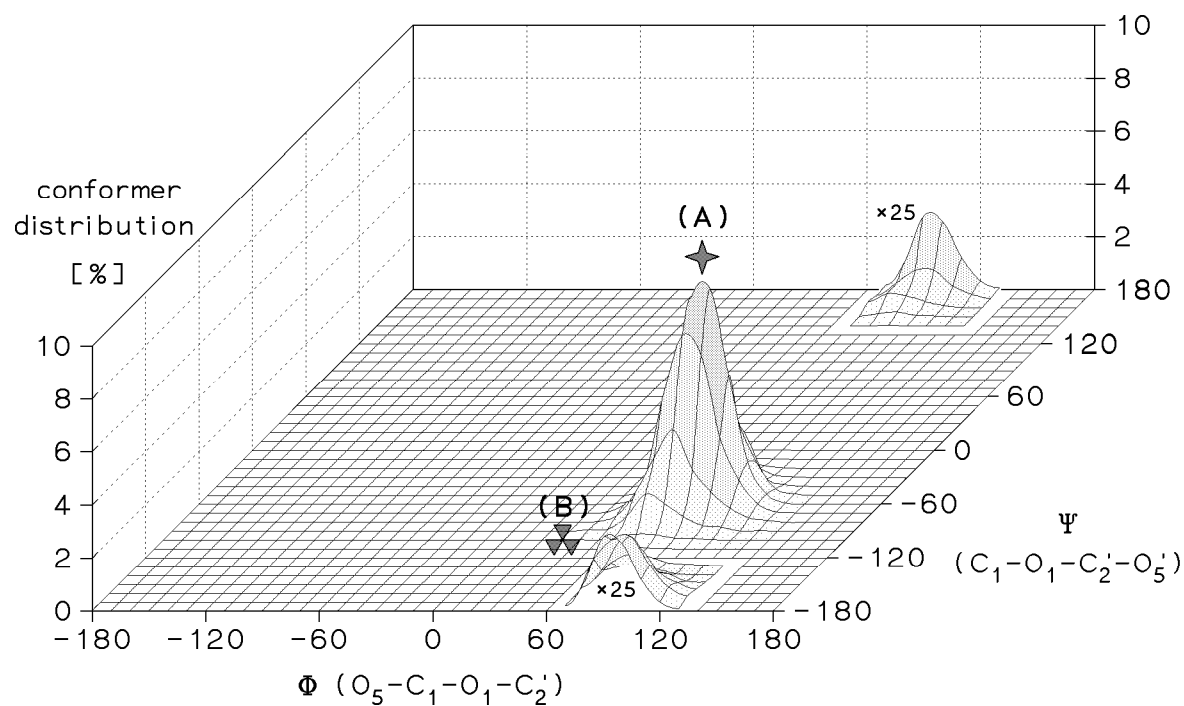
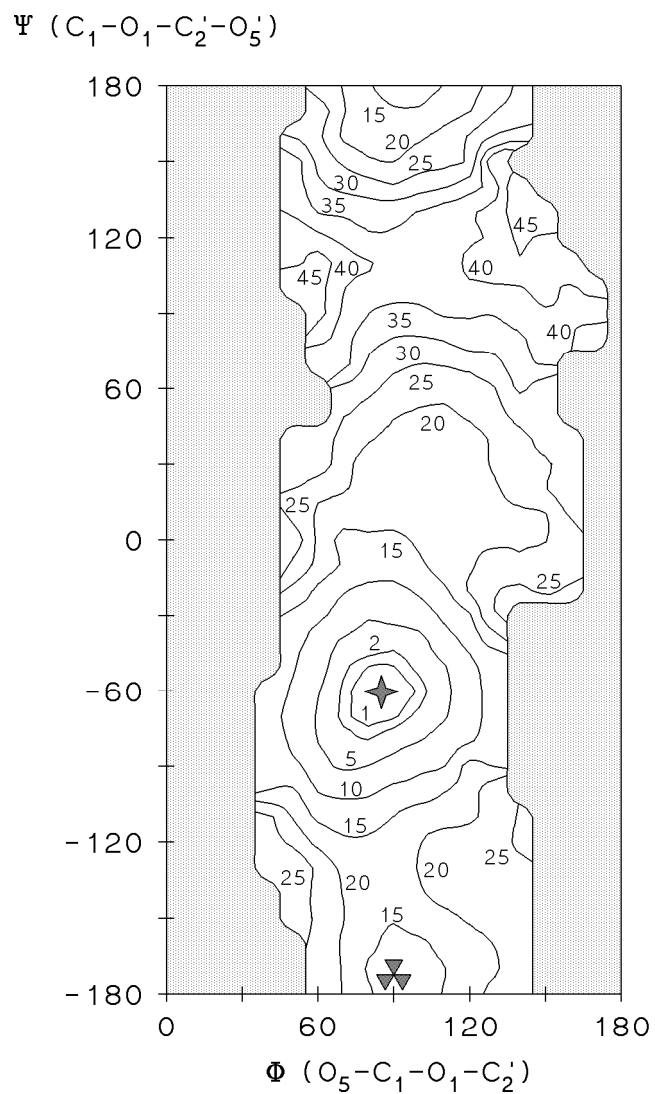


Fig. 3-11. Contour plot of the sucrose conformational free energy as function of the intersaccharidic torsion angles (*left plot* $\Delta G(\Phi, \Psi)$, energies in kJ/mol) as obtained from 2D-analysis of the *umbrella sampling* data (compare to Fig. 3-3, grey shading indicates insufficient data for contour plotting; estimated uncertainty $\sigma(\Delta G(\Phi, \Psi)) \approx \pm 2.0$ kJ/mol). The *upper* 3D-plot of the corresponding percentage distribution (cf. Fig. 3-5) indicates only one relevant conformational substate (\equiv **A**, note that the distribution of **B** is expanded vertically by a factor of 25).

The Molecular Electrostatic Potential (MEP) Profile of Sucrose

Computational Basics

Electrostatic interactions of molecules are long-ranged, strong effects which determine many properties of molecular systems^[136]. Starting from atomic partial charges q_i – which may be calculated by different methods including semiempirical or *ab initio* approaches – the molecular electrostatic potential profiles^[49] can be obtained by classical electrostatic laws (eq. 3-5),

$$MEP(\vec{r}) = \sum_{i=1}^n q_i / |\vec{r} - \vec{r}_i| \quad (3-5)$$

where n is the number of atoms in the molecule and r_i is the position vector for the i -th atom. There is no unique way to determine the effective partial atomic charges from the electronic charge distribution. In this work the AM1^[51]-atomic charges are used, computed with the MOPAC^[50]-program system.

Visualization of Data

For both relevant sucrose conformers (cf. Fig. 3-4) the MEP-profiles were calculated on the basis of eq. 3-5. In Fig. 1-3 (p. 4)*, the MEP patterns are represented on the solid surface models in a 16-color code, ranging from red (most positive electrostatic potential) to violet (area of highest electronegativity) as calculated with the MOLCAD module^[48].**

In each case, the color-code was adapted to the range of calculated MEP-values, a procedure referred to as individual mapping on each surface. The absolute energy-values (here approx. -25 to +25 kcal/mol) of the electrostatic potential profiles are sensitive to the type of atomic-charges used, which may vary significantly with the different calculation modes (*ab initio*, semiempirical, or simple force field charges). However, the relative potential distributions are much less susceptible to this.

* In Fig. 1-3 on p. 4, the MOLCAD-program^[48] generated molecular electrostatic potential (MEP) profiles^[49] are depicted for the sucrose conformers **A** (*left*) and **B** (*right*) (cf. Fig. 1-2 and 3-4) on the corresponding contact surfaces. The 16-color code ranging from violet (most negative potential) to red (most electropositive potential) was applied in relative terms for each molecule separately. To facilitate visualization, the front side-opened forms of the two conformers are also provided with a ball and stick model inserted. In either case, it is evident that the proton of the 2-OH group of the glucose part is characterized by a high positive electrostatic potential (red).

** In contrast to earlier publications^[17-19] a color-inverted code was used here for the MEP profiles for reasons of standardization^[48].

Accordingly, I refrain from relying on absolute values and want to discuss the implications emerging from relative effects only.

As evident from Fig. 1-3, each of the two relevant sucrose conformers has the red, i.e. most electropositive area, centered around the glucosyl-2-OH, undoubtedly caused by the cooperativity^[52-54] of the hydrogen bond directed towards the oxygen of this hydroxyl group. The electron-withdrawing effect of the adjacent anomeric center can contribute thereto only to a minor extent, since this would also have to be observed for the 1'- and 3'-OH groups in the fructose moiety.

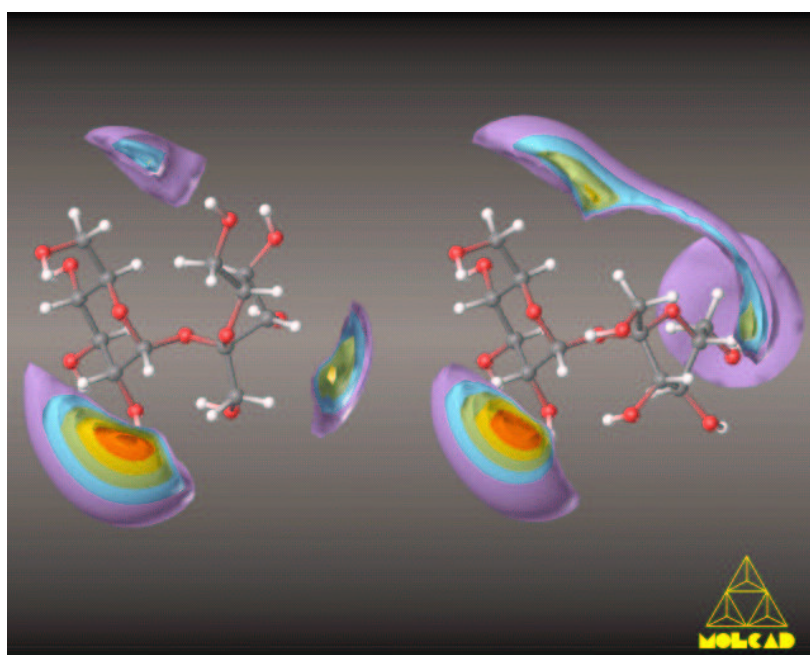


Fig. 3-12. Iso-energy contour plots for the interactions of a negatively charged probe sphere with the two sucrose conformers of Fig. 3-4. Contours are drawn at levels of +2, +4, and +6kcal/mol above the respective global energy minimum. The high positivation of the 2-OH of the glucose units points towards an enhanced acidity of this hydroxyl group over the others, as well as to its energetically favorable hydrogen bond donor capabilities.

The strong positivation of the 2-OH group becomes particularly intelligible when calculating iso-energy contour surfaces^[48] for the interaction of a negatively charged probe sphere with the sucrose molecule. At negative energy levels, these iso-energy plots delineate spatial volumes of energetically favorable interactions with nucleophiles. In Fig. 3-12, the energy contours for sucrose conformers **A** and **B** are given at levels of +2 – +6kcal/mol above the respective global energy minimum of about -10.5kcal/mol. In either case, the lowest and thus most electrophilic regions are invariably centered around the 2-hydroxyl group of the glucose moiety.

Chemical Implications of the MEP Profiles

The data presented in Figs. 1-3 and 3-12 necessarily entail the notion that the strong positivation of the glucosyl-2-OH must lead to enhanced acidity of this hydroxyl group over the others, i.e. should be deprotonated first on base treatment. There is ample experimental evidence that this, indeed, is the fact.

Benylation of sucrose under basic conditions (benzyl bromide / Ag₂O in DMF) yields a 86 : 10 : 3 : 1 – mixture of monobenzyl ethers, from which a reactivity of the OH-groups in the order 2-OH > 3'-OH > 1'-OH > 3-OH has been deduced^[137]. More recent findings^[138] demonstrated that 2-*O*-acyl and 2-*O*-(*N*-carbamoyl)-derivatives of sucrose are obtained in useful yields by NaH-deprotonation of sucrose in pyridine and reaction with 3-acylthiazolidinethiones. Studies on this NaH-induced deprotonation of sucrose and subsequent exposure of the mono-anion to benzyl bromide opened up a large scale adaptable route towards 2-*O*-benzyl-sucrose that proved to be an ideal starting material for the synthesis of 2-*O*-modified derivatives such as 2-keto- or 2-deoxy-sucrose^[139].

A remarkably high regioselectivity favors the secondary 2- and 3'-hydroxyl groups over the generally more reactive primary positions. These experimental effects, incomprehensible on the basis of common formula writings, may be understood on the basis of the MEP profiles (Fig. 1-3): from this it is concluded that the 2-OH group must be easiest activated by deprotonation. But also, the anion at this position seems to be the best stabilized one.

Statistical crystal structure analysis^[52,140] and a force field based calculation of different standard types of intramolecular hydrogen bonding geometries in carbohydrates was carried out. The results summarized in Fig. 3-13 – despite being crude approximations due to the mode of calculation and the negligence of various hydrogen bond directionalities – allow clear conclusions:

- ◆ Intramolecular H-bonds between vicinal hydroxyl groups cannot contribute effectively to stabilization of anions due to their rather long O ... H distances^[140] of 2.6 – 2.8 Å and the small O ... HO angles^[140] of $\varphi \approx 90 - 100^\circ$.
- ◆ Interactions between hydroxyl and exocyclic hydroxymethyl groups^[140], despite their shorter distances and larger angles ($d(\text{O} \cdots \text{H}) \approx 2.0 - 2.6 \text{ \AA}$, $\varphi(\text{O} \cdots \text{HO}) \approx 110 - 130^\circ$), are unfavorable in solution due to 1,3-diaxial like repulsions between the negatively charged oxygen atoms^[53,67-73].

In disaccharides, hydrogen bonds between the two sugar residues^[140] are stereochemically less restricted and much more favorable, and therefore will contribute to a large extent to anion stabilization, in sucrose most pronouncedly the H-bonds

between 1'-OH and 2-O (conformer **A** in Fig. 3-4 with $\varphi \approx 145^\circ$) and 3'-OH \cdots 2-O (conformer **B**, $\varphi \approx 165^\circ$), both being comparatively short (2.0 – 2.3 Å, cf. Table 3-1).

The interresidue 2-O^g \cdots HO-1^f and 2-O^g \cdots HO-3^f hydrogen bonds in sucrose obviously not only determine the MEP profile, i.e. the electropositive 2-OH proton, but are also responsible for the stabilization of the anion at these positions.

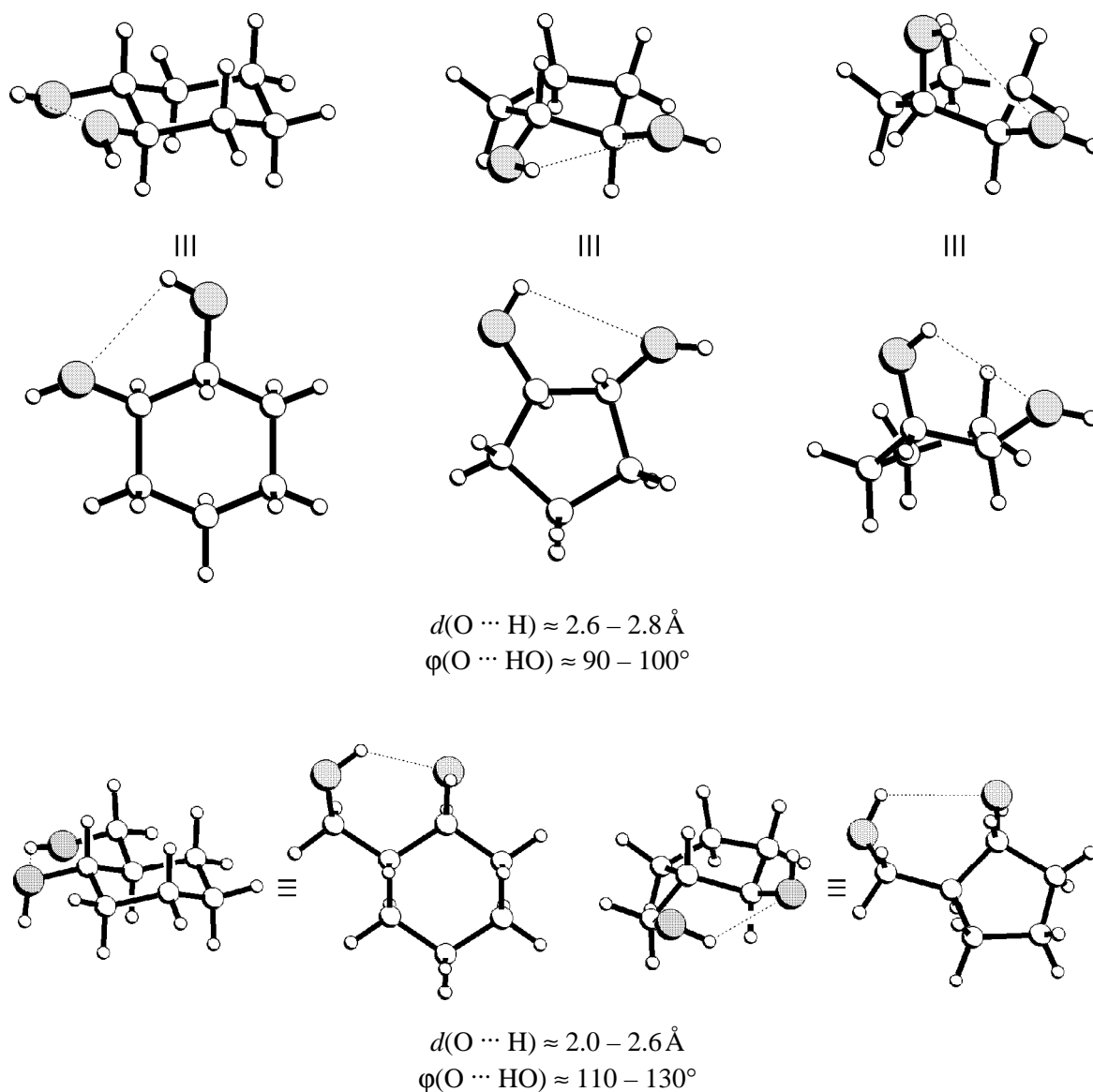
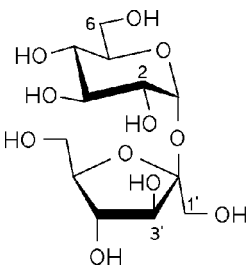


Fig. 3-13. A selection of typical hydrogen bond geometries (PIMM88 [45]) between vicinal hydroxyl groups (*top row*) and hydroxyl- and CH₂OH-groups (*lower row*) in cyclohexane- and cyclopentane-type model compounds; the approximate O \cdots H distances and O \cdots HO angles are given (in all cases, a projection of the molecule perpendicular to the O \cdots HO-plane is depicted in addition, oxygen atoms are shaded).

The results are additionally supported by certain electrochemical data: The electroreductive deprotonation^[55] of alcohols within the cathodic compartment of an electrolysis cell is initiated by a single electron transfer (SET) at the electrode to the alcohol, the instable radical anion formed thereby losing a hydrogen atom ($\rightarrow \text{H}_2$)^[56] with generation of the respective alkoxide, which is subsequently trapped by alkylating or acylating reagents. In the case of sucrose, electrolysis in DMF in the presence of LiClO_4 and subsequent alkylation generates 2-*O*-alkyl derivatives with high selectivity (cf. Table 3-3), while benzylation preferentially leads to the 1'-*O*-substituted products^[57].

Table 3-3. Product distribution of the etherification and esterification of the electrochemically generated sucrose mono-anion with various reagents^[57].

	reagent*	product distribution [%]				reaction conditions
		O-2	O-6	O-1'	O-3'	
	BnCl	49	–	41	10	<i>ex situ</i>
	BnBr	48	–	39	13	<i>ex situ</i>
	MeI	54	–	27	19	<i>ex situ</i>
	AllBr	39	–	61	–	<i>ex situ</i>
	BzCl	–	–	100	–	<i>ex situ</i>
	BzOMe	–	50	50	–	<i>in situ</i>

* Bn = $\text{CH}_2\text{C}_6\text{H}_5$, Me = CH_3 , All = $\text{CH}_2\text{CH}=\text{CH}_2$, Bz = COC_6H_5 .

On the basis of the MEP profiles of Fig. 1-3, these selectivities can be rationalized as follows: the sucrose molecule is oriented in the electric field in such a way that the glucosyl-2-OH side is directed towards, and hence, arrives at the cathode, thus eliciting the single electron transfer at that very position. Thereby, the 2-OH is converted into the alkoxide. There is an alternative possibility though: SET-induced deprotonation occurs at any other hydroxyl group of sucrose, e.g. at one of the sterically more accessible primary hydroxyls (6-OH, 1'-OH or 6'-OH), and the resulting anion undergoes equilibration via very fast intramolecular proton shifts between the various OH-groups elaborating the thermodynamically most stable anion. The time scales for the formation of the radical anion, for releasing a hydrogen atom, equilibration, and the subsequent trapping reactions of the anions (even when carried out *in situ*) differ by several orders of magnitude, and therefore, the observed product distribution upon trapping the anions by alkylation or acylation may, but does not necessarily have to be the result of the regioselectivity in the initial electrochemically induced deprotonation.

Although being preparatively impractical, the product distributions observed after quenching electrolyzed DMF-solutions of sucrose with alkylating agents such as benzyl chloride, benzyl bromide, or methyl iodide (cf. Table 3-3) correlate very well with the ones after standard base-induced deprotonation^[137-139]; in all cases the 2-*O*-alkylated compounds are obtained as the major products, followed by the 1'-*O*- and 3'-*O*-derivatives. That benzylation of the electrochemically generated sucrose mono-anion gives the 1'-*O*-benzoate preferentially (Table 3-3), may conceivably be attributed to the well-established fact that acyl groups migrate under basic conditions from the place of the first chemical attack to thermodynamically more stable positions^[141].

MEP Pattern of Sucrose and Hydrogen Bonding

The sugar-substrate specificity^[142] of many enzymes must be attributed to characteristic hydrogen bonding patterns which are sensitive to the respective stereochemistry of different carbohydrates. From the MEP-profiles (Fig. 1-3) and the iso-energy contour plots (Fig. 3-12) of sucrose, it can be concluded that the strong positivation of the glucose-2-OH group makes it an energetically favorable hydrogen bond donor for interaction with an oppositely polarized region of a receptor protein, and thus an important unit for biochemical recognition processes of sucrose. Indeed, from examination of the substrate-specificity of a sucrose transporter complex in plant membranes, it was deduced that the 2-OH group of the glucose moiety as well as the 3- and 4-OH groups are essentially required for molecular transport^[143]. Since the conformational properties of the 2-, 3-, and 4-deoxy-sucroses used in these investigations were shown to be similar to those of sucrose^[143], their substantially altered biological properties must be directly related to changes in hydrogen bondings with the transporter proteins.

The Molecular Lipophilicity Pattern (MLP) of Sucrose

The interactions of substrates in general (and of carbohydrates in particular) with receptor proteins are not only governed by electrostatic effects – including hydrogen bonding phenomena – but also by dispersion forces (van der Waals contacts) and hydrophobic effects^[142]. Even when looking at CPK-models, the assessment of this kind of interaction is not always straight forward. Utilizing the MOLCAD-program^[48] as in the case of the molecular electrostatic potential (MEP) profiles, it is also possible to calculate molecular lipophilicity patterns (MLP's) on the molecular contact surfaces.

Localization of Hydrophobicity

Hydrophobicity, the hydrophobic effect and hydrophobic interactions are terms used for the discussion of the specific interaction of a molecule with water and for the molecular interaction in aqueous solution. Nevertheless, neither the hydrophobic effect^[116,117] nor the hydrophobic interaction is fully understood today^[117]. Phenomenological, the hydrophobic effect may be measured in terms of the partition coefficient of a substance in the water/*n*-octanol system. The corresponding log *P*-values are related to a free energy change $\Delta G_{\text{transfer}}$ for the transfer of the solute from one solvent to another, but they only can be regarded as an overall property, which does not allow to discriminate between hydrophilic and hydrophobic molecular regions. There have been some attempts to formulate the hydrophobic effect in analogy to the electrostatic interaction in terms of a local lipophilic "potential"^[58,118,119], based on empirical findings such as the fact that the hydrophobicity of a molecule (measured, for example, by its log *P*-value, where *P* is the partition coefficient between two solvents) can be regarded as a sum of increments related to the lipophilicity of its fragments, f_i :

$$\log P = \sum_{i=1}^n f_i \quad (3-02)$$

Heiden *et al.*^[58] have demonstrated that early approaches^[144] in which a molecular lipophilicity "potential" (MLP) was introduced in the form

$$MLP(\vec{r}) = \sum_{i=1}^n f_i g(d_i) \quad ; \quad d_i = |\vec{r} - \vec{r}_i| \quad (3-6)$$

with atomic increment values f_i and atomic positions r_i only lead to reasonable hydrophobicity profiles, when the distance functions $g(d)$ are chosen properly. Here, the f_i -values of Ghose, Crippen and others^[145,146] and the distance functions

$$g(d) = N \left\{ \exp[a(d - d_{\text{cut-off}})] + 1 \right\}^{-1} \quad (3-7)$$

with the "normalization"

$$N^{-1} = \sum_{i=1}^n g(d_i) \quad (3-8)$$

where n is the number of atomic increments and $d_{\text{cut-off}} = 4 \text{ \AA}$, $a = 1 \text{ \AA}^{-1}$ have been chosen^[58].

The approach by Heiden *et al.*^[58] does not really lead to a lipophilic "potential" like the electrostatic potential, but it represents a method for a weighted projection of Crippen's f_i -values onto the contact surface of the molecule, and so gives the possibility to visualize hydrophobicity in a qualitative way with a certain color coding with the aid of computer graphics techniques^[48]. Throughout this context relative coding of hydrophobicity was used, i.e. the most hydrophilic values are visualized in blue, while the most hydrophobic values are shown in yellow-brown. The visualizations have been performed with texture mapping strategies^[59].

Application of the Hydrophobicity Mapping

Applying the methodology to sucrose^[17-19], the MLP's for both relevant conformers were generated and transferred into a color-code (Fig. 1-4 on p. 6)*, ranging from blue for hydrophilic regions (indicating its relationship to water) to an "oily" yellow-brown for designating hydrophobic areas.

From Fig. 1-4 it is obvious that the back side portion of the fructose moiety is decisively hydrophobic, and hence, must contribute most to the hydrophobic interaction with the sweetness receptor. The most hydrophilic regions, likely to be involved in hydrogen bonding, are located on the opposite molecular side. A more quantitative relationship between local hydrophobicity and the corresponding surface area is presented in Fig. 3-14. The relative surface hydrophobicity is given in arbitrary units, which may be interpreted as relative local contributions to the overall partition coefficient. Both, the glucosyl and fructosyl residue contribute in nearly equal proportions of approx. 150 \AA^2 each** to the total surface area of sucrose ($\approx 300 \text{ \AA}^2$). Yet, the hydrophobicity distribution within the glucose residue is considerably shifted towards higher hydrophilicity when compared to the fructose region.

* In Fig. 1-4 on p. 6, the MOLCAD-program^[48] generated molecular lipophilicity patterns (MLP's)^[58] are shown for the two sucrose conformers of Fig. 1-2 and 3-4. For visualization by texture mapping^[59], a two-color code graded into 32 shades is used. The color-coding was adapted to the range of relative hydrophobicity calculated for each molecule, using 16 colors ranging from dark-blue (most hydrophilic surface areas) over light blue to full yellow (most hydrophobic regions) for mapping the computed values onto the surface. The remaining 16 color shades (blue to brown) were used to indicate iso-contour lines in between former color scale, allowing for a more quantitative assessment of relative hydrophobicity on different surface regions. For both sucrose conformers, the entire "backside" of the fructose moiety is decisively hydrophobic.

** Mathematically, the surface consists of small triangles, the dots in Fig. 3-4 indicating the respective corners. Here, each triangle was assigned to the closest neighbor atom and accordingly used for surface calculation.

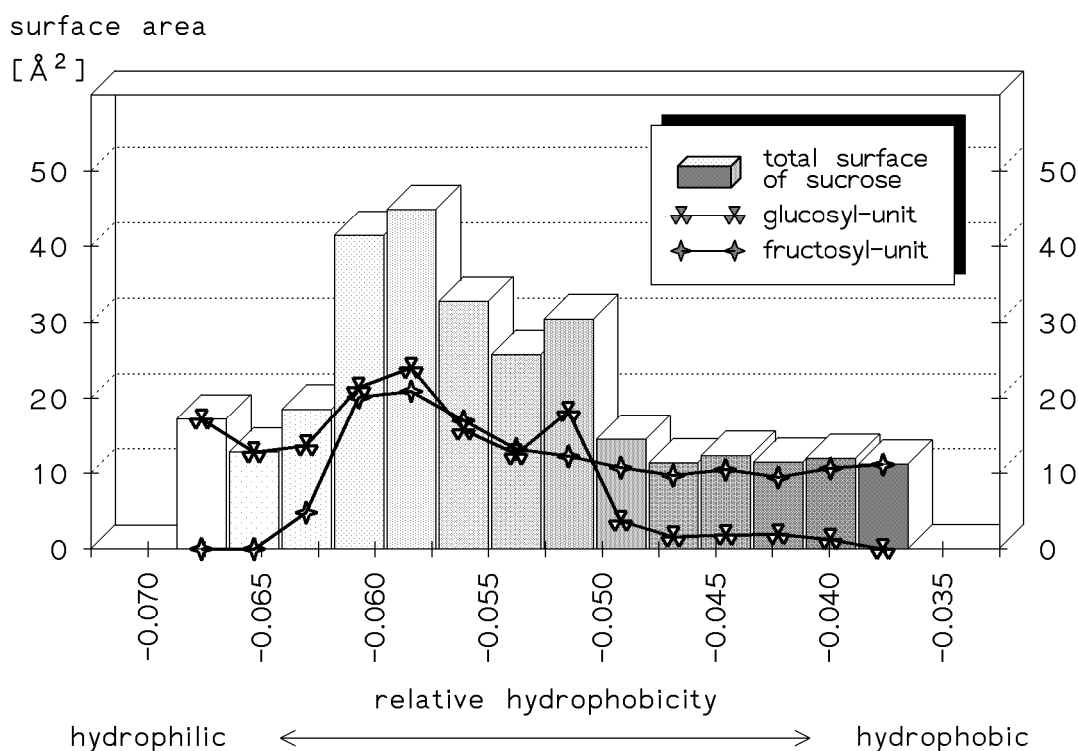


Fig. 3-14. Surface area of sucrose (major conformer) in relation to relative hydrophobicity (in arbitrary units). The solid lines indicate the contributions of the glucosyl and fructosyl residue, respectively (total surface area approx. 300Å^2). The glucose portion contributes to hydrophilic surface areas to about the same extent ($\approx 150\text{Å}^2$) as the fructose part contributes to hydrophobicity.

Biological Significance of the Molecular Lipophilicity Profiles

Studies on the binding characteristics of sucrose to a plant transporter protein^[143,147] indeed showed a large contribution of hydrophobic interactions to molecular recognition. As expected from the MLP's presented here, of the topography of sucrose amenable for interaction with the carrier, especially the fructosyl residue was shown to contribute most to the hydrophobic interaction with the protein, while the glucose unit provides the hydroxyl groups for the polar reciprocal action^[147].

Undoubtedly, the energetical contribution of hydrophobic substrate-receptor interactions not only depends on the chemical structures involved in binding, but is also directly related to the size of the molecular contact regions^[148-150]. Thus, the MLP of sucrose does not only allow an unequivocal assignment of hydrophobic and hydrophilic molecular regions, but necessarily must also lead to revision of the common AH-B-X-concept of sweetness.

The Modified AH-B-X-Concept of Structure-Sweetness-Relationships

Present structure-sweetness relationships^[4,151-154] rest on ideas of Shallenberger^[1,2] and Kier^[3], such that all sweet tasting molecules have a common tripartite structural element: a hydrogen bond donor (AH-portion), a hydrogen bond acceptor functionality (B-unit), and an additional hydrophobic binding point X. From the lipophilicity patterns presented in Fig. 1-4, it is clear that not a single binding point, but an entire extended molecular region must be responsible for the hydrophobic interaction with the sweetness receptor (Fig. 3-15). Moreover, the hydrogen bonding AH-B-couple must be contained in an opposite-side located, more hydrophilic region of the sweetener.

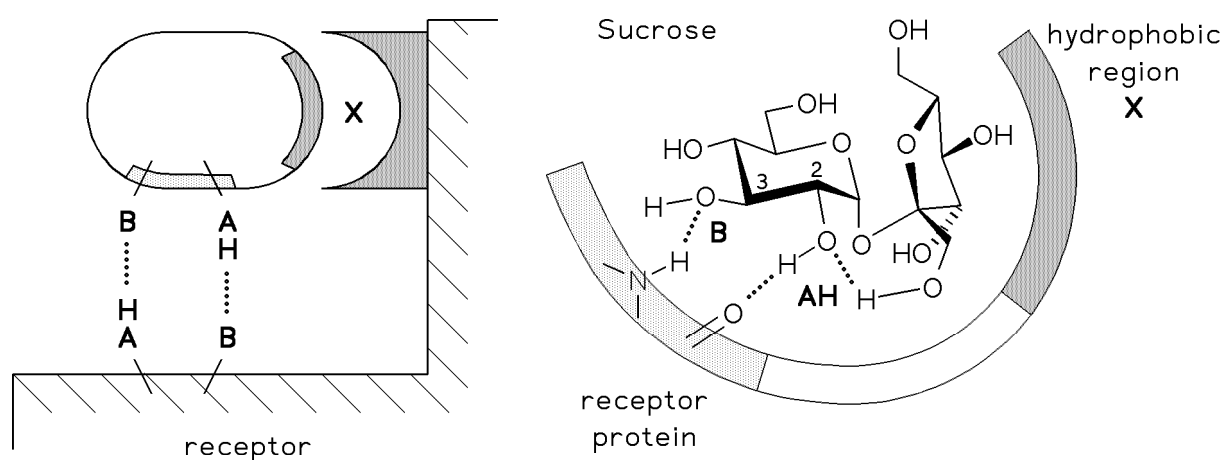


Fig. 3-15. Modified AH-B-X-concept of sweetness (*left*): an extended hydrophobic region rather than a specific binding point, and the opposite-side located, hydrophilic AH-B-couple are the basic structural characteristics of all sweeteners. In sucrose (*right*), the AH-B-couple is represented by the glucosyl-2- and 3-OH groups, while the entire backside of the fructose contributes to the hydrophobic interaction with the sweetness receptor.

On the basis of the MEP profiles of Fig. 1-3 it is clear that the distinctly electropositive glucosyl-2-OH group of sucrose is the energetically most favored to engage in hydrogen bonding as a donor, and correspondingly may be assigned as the AH-unit. The hydrogen bond acceptor (B-unit) must be located in its direct vicinity, clearly pointing towards the glucosyl-3-OH as a likely candidate (Fig. 3-15). This assessment is further substantiated by the MLP's: the H-bond donor as well as the acceptor functionality are located within the most hydrophilic molecular region. On the other hand, the entire "backside" of the fructose moiety must be responsible for the hydrophobic driving force of binding to the receptor. In an extension of this model, Jeffrey^[155] suggested that water mediated cooperative H-bonds in the solvation sphere

may even govern the long range approximation of the sweetness receptor by the sucrose molecule, before the water is successively being stripped off while entering the receptor site.

Quantitatively, the AH-B-X-structural elements can be located by calculation of iso-potential cuts on the molecular surface, where only those parts of the surface are depicted, which carry a surface quantity (electrostatic potential, hydrophobicity, etc.) which is higher (or lower) than a given contour value. Since these surfaces are bounded by iso-values of the considered quantity they are termed **iso-contour surfaces (ICS)***. In Fig. 3-16, surface areas of high and low electrostatic potentials (preferred hydrogen bond donor and acceptor functionalities), and regions of high hydrophobicity have been cut off from the MEP's (Fig. 1-3 on p. 4) and MLP's (Fig. 1-4 on p. 6) respectively, and are depicted simultaneously.

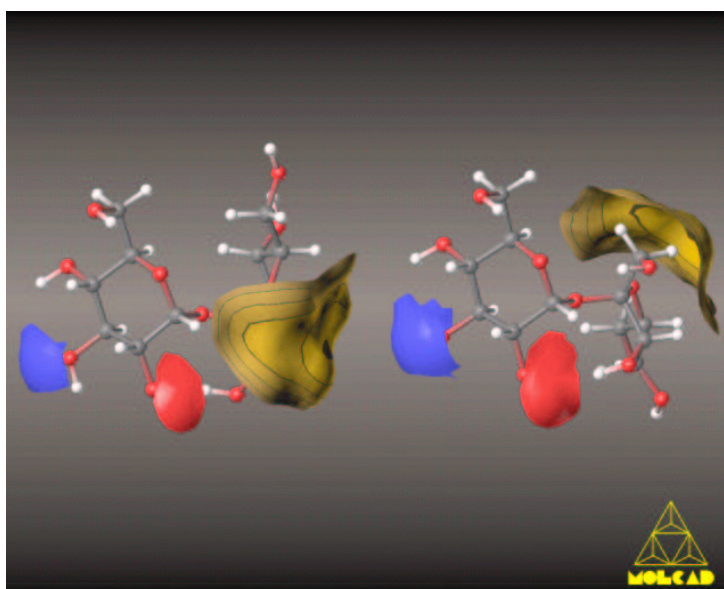


Fig. 3-16. Iso-potential cuts of the molecular surface of the relevant sucrose conformers. The surfaces (cf. Fig. 1-3 on p. 4 and 1-4 on p. 6) are sliced along constant potentials in order to identify the hydrogen bond donor (AH) unit (MEP > +18.0 kcal/mol), the H-bond acceptor (B) portion (MEP < -8.0 kcal/mol), and the hydrophobic binding region X (MLP > -0.042, $\approx 30 \text{ \AA}^2$ each) as the structural essentials required for sweetness (for assignment of the B-unit, only the relevant surface parts cut off are shown).

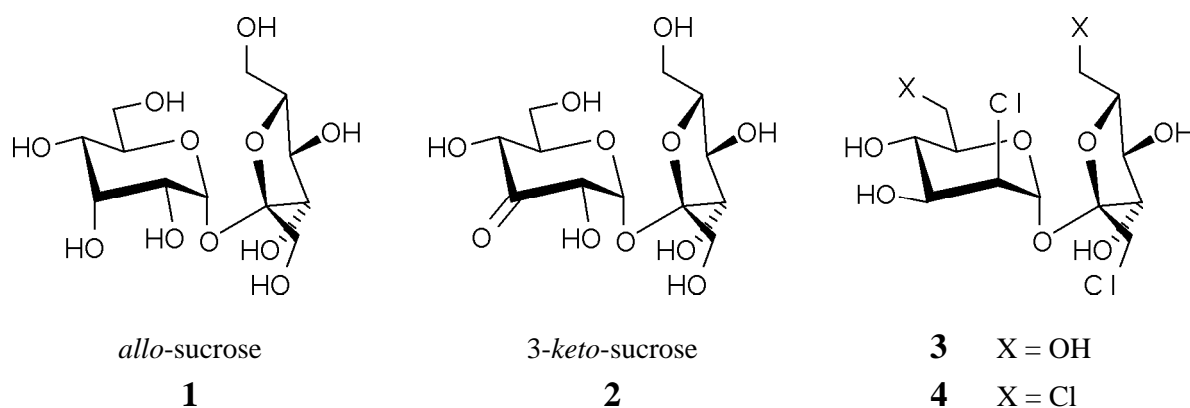
The hydrophobic areas and the regions of high electropositive potential are shown as they emerge from the calculation. For assessment of the B-unit, the electronegative surface regions not located in direct vicinity to the AH-portion were removed. Thereby, for the two relevant solution conformers of sucrose, the molecular geometry

* In contrast to the iso-energy surfaces of Fig. 3-12 which designate three-dimensional volumes of low potentials (energetically favorable interactions), the iso-contour surfaces describe two-dimensional areas of high or low electrostatic or lipophilic potentials.

and the steric AH-B-X-structural requirements for binding to the sweetness receptor are readily delineated, confirming the proposals made above. In contrast to the AH- and B-units, which are clearly associated with the 2- and 3-OH groups, the opposite-side located, large hydrophobic region comprises the whole fructosyl-backside.

Assessment of the Modified AH-B-X-Concept with Sucrose Derivatives

If these rather theoretical conclusions are correct, any chemical modification within the AH-B-couple, as well as changes of the hydrophobicity of the fructose portion must lead to significantly altered sweetness characteristics. The following brief evaluation of sucrose analogs as to their sweetness appears to substantiate the conclusions reached.

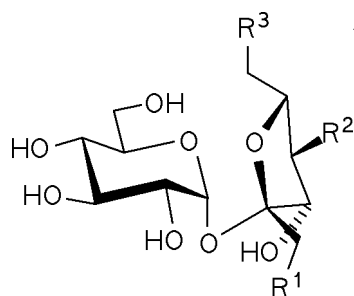


Inversion of configuration at C-3 of the glucosyl-unit leads to *allo*-sucrose (**1**), which is tasteless^[65] obviously due to the fact that an axially disposed 3-OH is not capable to function properly as the hydrogen bond-accepting B-component. On the other hand, *3-keto*-sucrose (**2**) is sweet^[65], i.e. the 3-carbonyl function retains the hydrogen bond acceptor capabilities. As exemplified by two 2-chloro-2-deoxy-*manno*-sucrose derivatives **3** and **4**, chemical substitution of the glucosyl-2-OH (with inversion) by chlorine has a dramatic effect: **3** is "not sweet"^[81,82], **4** is even "extremely bitter"^[85].

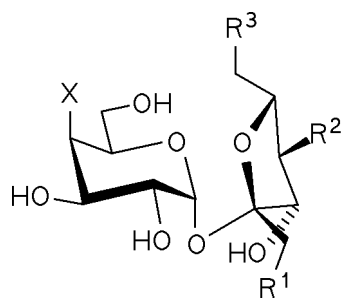
The sweetness characteristics of a variety of deoxy-halo-sucroses^[21-23,62,64,74-85] also appear to corroborate the modified AH-B-X-concept outlined above. *De facto* it implies that the hydrophobic cleft of the sweetness receptor protein corresponds to the hydrophobic region on the "backside" of the fructose unit. Indeed, as evidenced by the data listed in Table 3-4, substitution of the hydroxyl groups at the 1'-, 4'-, and / or 6'-positions invariably leads to sweeteners **5** – **10** more potent than sucrose.

Table 3-4. Relative sweetness of deoxy-halo-sucroses modified in the fructose portion.

	R ¹	R ²	R ³	relative	
				sweetness	Refs
sucrose	OH	OH	OH	1	–
5	Cl	OH	OH	20	21,62,64,80
6	OH	OH	Cl	20	21,62,64
7	Cl	Cl	OH	30	81
8	Cl	OH	Cl	80	21,62
9	Br	OH	Br	80	84
10	Cl	Cl	Cl	100	81

**Table 3-5.** Relative sweetness (sucrose = 1) of 4-deoxy-4-halo-*galacto*-sucroses.

	Glucose	- Fructose -			relative	Refs
	X	R ¹	R ²	R ³		
sucrose	OH _{eq}	OH	OH	OH	1	–
11	Cl	OH	OH	OH	5	21,64,80
12	Cl	OH	OH	Cl	50	84
13	Cl	Cl	OH	OH	120	21,62
14	Cl	OH	Cl	Cl	160	81
15	Cl	Cl	Cl	OH	220	81
16*	Cl	Cl	OH	Cl	650	78
17	Cl	Br	OH	Br	800	84
18	Cl	Cl	F	Cl	1000	81
19	Cl	Cl	Cl	Cl	2200	79,82
20	Cl	Cl	Br	Cl	3000	81
21	Cl	Cl	I	Cl	3500	78,81
22	Cl	Br	Br	Br	7000	81
23	F	OH	OH	F	4	84
24	F	F	OH	F	40	78
25	F	Cl	Cl	Cl	200	81
26	Br	Cl	OH	Cl	375	84
27	Br	Br	OH	Br	800	84
28	Br	Br	Br	Br	7500	78,81
29	I	I	OH	I	120	78,84



* Sucralose

The same trend is observed in the case of a large number of 4-deoxy-4-halo-*galacto*-sucrose derivatives (Table 3-5). The 4-chloro-*galacto*-sucrose **11** is 5 times sweeter than sucrose, thus indicating that the glucose unit contributes less to the hydrophobic interaction with the sweetness receptor than the fructose moiety does: the respective 1'- and 6'-mono-chloro derivatives **5** and **6** are four times as sweet as **11** (cf. Table 3-4). More pronounced effects are observed – not unexpectedly – on successive replacement of the fructose hydroxyl groups with halogen substituents: sweetness enhancements of several orders of magnitude are observed, reaching a solitary maximum in the bromo-derivatives **22** and **28**, which are 7000 – 7500 times sweeter than sucrose.

Especially informative is a series of 4,1',6'-trihalo compounds: going from the fluoro- (**24**) to the chloro- (**16**) and bromo-analogs (**27**), sweetness power increases continuously from 40 to 650 and 800 times that of sucrose. While fluorine seems to be too small to properly fit into the receptor, iodine appears to be too big: the sweetness of the triiodo-compound (**29**) again drops to a value of only 120.

Hydrophobicity Pattern of Sucralose

Of the halo-substituted sucrose analogs mentioned above, 4,1',6'-trichloro-*galacto*-sucrose **16** (Sucralose®), recently approved in Canada as a non-caloric, high-potency sweetener for food uses^[86], was chosen to further probe the assignments of the glucophoric unit in sucrose. In Fig. 2-9 on p. 33, the X-ray-derived solid state structure^[87] and the lowest energy conformation emerging from PIMM88^[45] force field calculations are depicted as ball and stick models, including the contact surfaces in dotted form. Both conformers exhibit essentially the same intersaccharidic torsion angles, thus having the same relative orientation of the glucose and the fructose moieties, representing one low energy conformation on the energy potential surface of sucralose^[44,156-158]. Interestingly, the directionality of the interresidue hydrogen bond observed in the crystalline state (2-OH^g ... O-3^f) is reversed in the calculated structure (2-O^g ... HO-3^f), resembling more closely the situation proved to prevail in solution^[88].

The hydrophobicity patterns of both sucralose conformers are shown in Fig. 2-11 (p. 35), expectedly showing the two chlorine atoms in the fructose portion to represent the hydrophobic center (X-region), which is extended over the entire "backside" of the fructose-residue. The third chlorine, imbedded into the more hydrophilic glucose portion is considerably less hydrophobic than the other two. The MLP's of sucralose resemble closely those of sucrose (cf. Fig. 1-4), particularly, the opposite-side location of hydrophobic and hydrophilic regions is preserved in each case, and the AH-B-couple made up by the glucosyl-2- and 3-OH groups is located in the most hydrophilic molecular region.

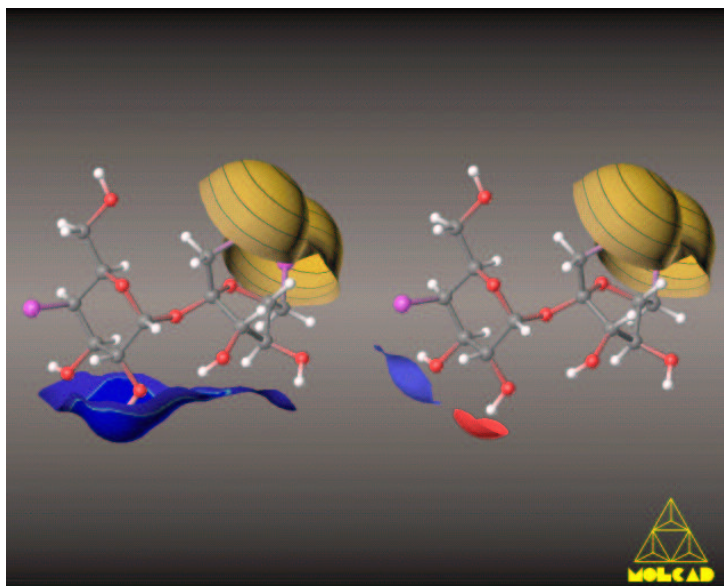


Fig. 3-17. Iso-potential cuts of the molecular surface of sucralose (*right* conformer of Fig. 2-9 – 2-11, p. 33–35). On the *left* side the opposite-side location of most hydrophilic (light shading, $MLP < -0.002$) and hydrophobic (dark shading, $MLP > +0.040$) surface regions becomes obvious. On the *right* side the AH-B-X-system could be identified by the MEP and MLP surface sections (cf. Fig. 3-16, $MEP > +12.0 \text{ kcal/mol}$ (AH-unit), $MEP < -12.0 \text{ kcal/mol}$ (B-portion), and the hydrophobic binding region X ($\approx 65 \text{ \AA}^2$) with $MLP > +0.040$).

As outlined above for sucrose (Fig. 3-16), the identification of the AH-B-X-unit of sucralose can also be made on the basis of iso-potential surface cuts of the corresponding MEP's and MLP's. From Fig. 3-17, the essential opposite-side location of hydrophilic and hydrophobic regions becomes evident, the backside of the fructose moiety including the 1'- and 6'-chlorine atoms contributing to the most hydrophobic surface parts. The AH-B-couple as characterized by high or low electrostatic surface potentials respectively, again is located within the most hydrophilic molecular region. Of course, in terms of the absolute partition coefficients ($\log P$ -values), sucralose is much more hydrophobic than sucrose. The colors in Fig. 1-4 and 2-11 are given in relative terms for each molecule (individual mapping), no absolute values are represented.

In order to get an anticipation of the absolute hydrophobicity differences of sucrose and sucralose, the correlation of surface areas and the relative hydrophobicities is shown in Fig. 3-18 (compare also to Fig. 3-14) for both compounds, as well as for 1'-chloro-1'-deoxy-sucrose (**5**) and 1',6'-dichloro-1',6'-dideoxy-sucrose (**8**).

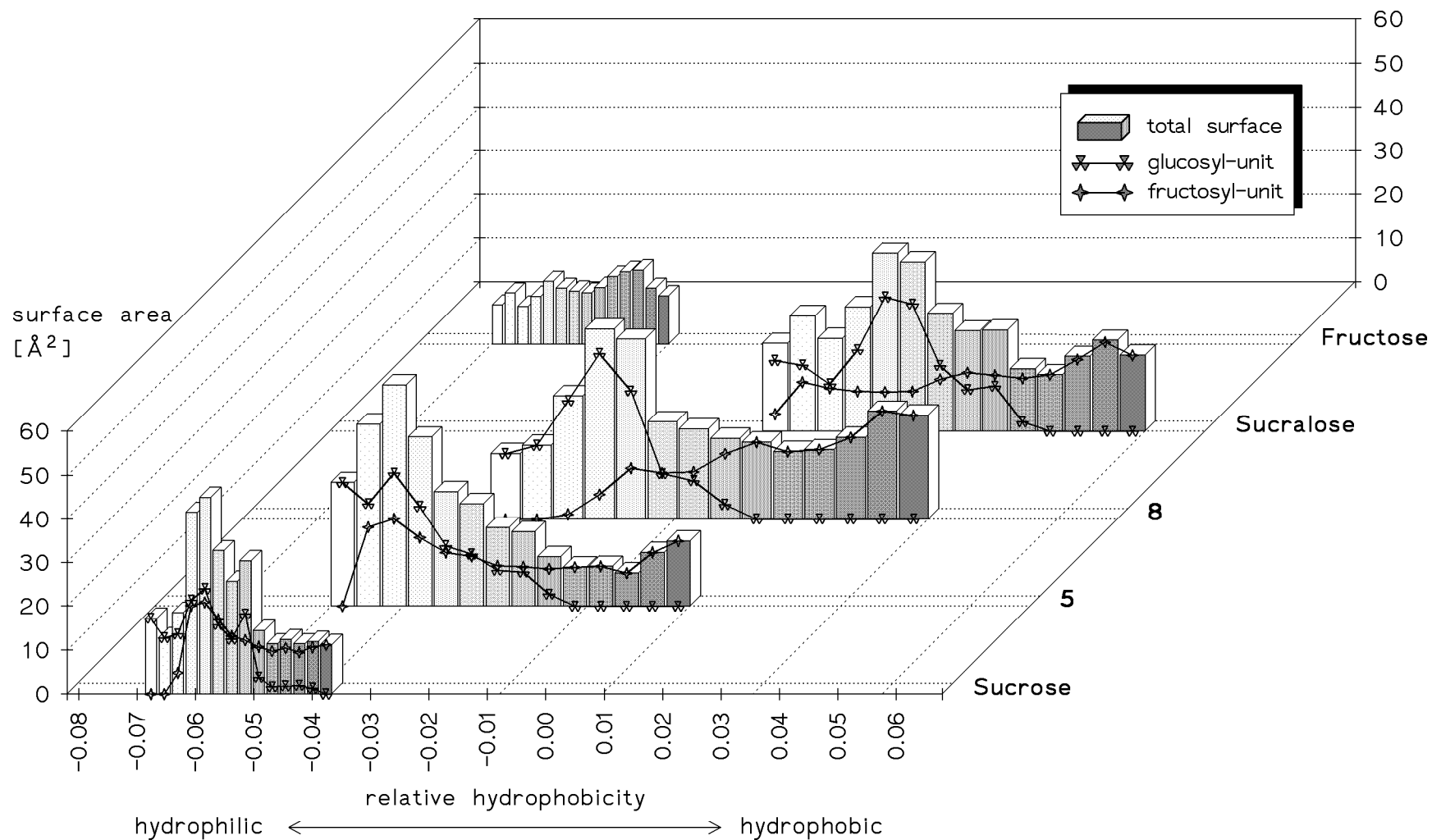


Fig. 3-18. Correlation of surface area and relative hydrophobicity for 1'-chloro-sucrose (5), 1',6'-dichloro-sucrose (8), and sucralose as compared to sucrose (compare with Fig. 3-14). As expected, increasing chloro-substitution leads to an enhancement of hydrophobicity (shift of the diagrams to the *right*), and, in consequence, to higher sweetness (cf. Table 3-4). In all cases, the fructosyl-residue is more hydrophobic than the glucose-unit. For comparison, the data for β -D-fructopyranose has been included.

Clearly, the absolute hydrophobicity (local log *P*-contributions on the X-axis) continuously increase going from sucrose to the mono- (**5**), di- (**8**), and trichloro-analogs (sucralose), with sweetness potency increasing simultaneously (sucrose: 1x, **5**: 20x, **8**: 80x, and sucralose: 650x, cf. Table 3-4 and 3-5). As indicated by the solid-line curves, in all cases the fructose unit contributes more to the hydrophobic surface regions than the glucose unit does, invariably the glucose residue being more hydrophilic.

The fact that the hydrophobic binding region of the sweetness receptor is flexible and capable to accommodate not only sucrose, but its halogen substituted derivatives as well, is even more emphasized by the taste characteristics of L-sucrose: the mirror image sugar was reported to exhibit a taste indistinguishable from the natural occurring enantiomer^[159]. Obviously, the hydrophobic backbone of the L-fructofuranosyl unit is still able to fit into the binding site of the sweetness receptor.

So far, the ideas and computational data on structure-sweetness relationships point towards the general validity of the modified AH-B-X-concept, considering the hydrophobic part, interacting with a matching hydrophobic section in the receptor protein, to be an entire molecular region rather than a distinct corner of an AH-B-X-tripartite system ("sweetness triangle") as originally suggested by Shallenberger and Kier (cf. Fig. 2-2 on p. 19, upper left). Despite the fact, that the initial model has been used extensively, it was outlined by Leel^[4] in his 1987 review on this topic that it is devoid of "any predictive value". Due to the uncertain assignment of the individual corners of the sweetness triangle to distinct structural elements of any sweet carbohydrate, for sucrose many different concepts have been advanced (cf. Fig. 2-2)^[20-23], so far none of them was able to explain the sweetness characteristics of all sucrose derivatives. The molecular modelings presented here provide overwhelming evidence for the necessity of modifying the AH-B-X-concept as indicated in Fig. 3-15 to obtain more reliable predictions in future.

Quantitative Sweetness-Hydrophobicity Relationships for Halo-Sucroses

From the relative hydrophobicity patterns presented, the characteristic binding sites of sucrose and its derivatives can be evaluated. The model developed therefrom and the AH-B-X-assignments made are able to explain the sweetness potency of a large number of compounds. General sweetness-lipophilicity correlations^[160] proved hydrophobic interactions to represent a crucial factor guiding a sweet substrate into the receptor taste bud, and locking it into the correct position for eliciting the sweetness response. Therefore, it was interesting to correlate the hydrophobicity of the halo-sucroses listed above with their sweetness to obtain a more quantitative relationship.

The correlation between the relative sweetness intensities (sucrose = 1) and calculated water / *n*-octanol partition coefficients (log *P*-values)^[145] for sucrose and the compounds **5** – **29** is shown in Fig. 3-19. Least squares fit^[161] of the data (considering the standard deviation $\sigma_{\log P} = \pm 0.5$ ^[145] and a factor of 2 as the error in sweetness^[158]) to a straight line $y = a + b \cdot x$ was performed by iteratively minimizing the χ^2 -function given below (eq. 3-9).

$$y(x) = a + b x \quad \chi^2(a, b) = \sum_{i=1}^N \frac{(y_i - a - b x_i)^2}{(\sigma_{y_i}^2 + b^2 \sigma_{x_i}^2)} \quad (3-9)$$

$$Q(a, x) = Q\left(\frac{N-2}{2}, \frac{\chi^2}{2}\right) = 1 - P\left(\frac{N-2}{2}, \frac{\chi^2}{2}\right) = \frac{1}{\Gamma(a)} \int_x^{\infty} e^{-t} t^{a-1} dt \quad (3-10)$$

$$\Gamma(a) = \int_0^{\infty} e^{-t} t^{a-1} dt$$

$$r_{xy} = \frac{\sum (x_i - \bar{x})^2 \cdot (y_i - \bar{y})^2}{\sum (x_i - \bar{x})^2 \cdot \sum (y_i - \bar{y})^2} \quad (3-11)$$

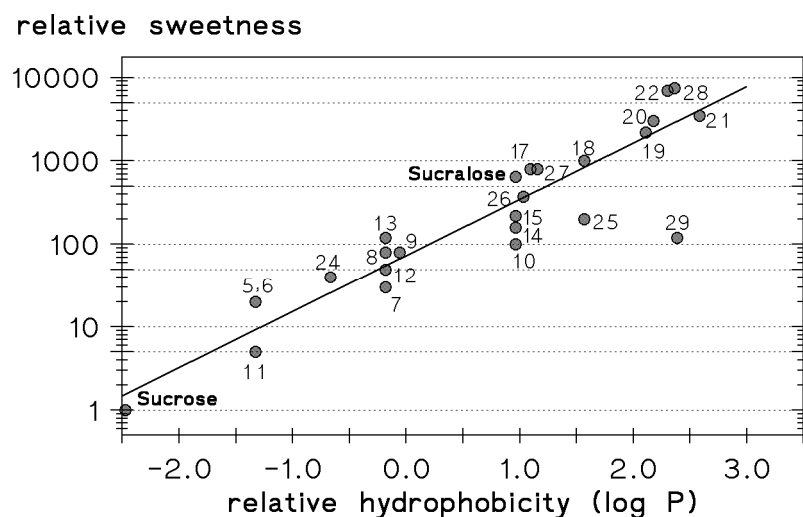


Fig. 3-19. Relationship between calculated partition coefficients in the water / *n*-octanol system (log *P*-values) and relative sweetness intensities (sucrose = 1) of halogen substituted sucrose derivatives (compound numbers correlate with Table 3-4 and 3-5). Compound **29** was not considered for fitting the data (see text for details).

The formulas for the goodness-of-fit estimation Q and the correlation coefficient r_{xy} are given in eqs. 3-10 and 3-11^{[161],*}. From the primary data set, compound **29** was removed, since the comparatively low sweetness of this triiodo-compound can be attributed to a steric misfit introduced by the three iodine atoms, and hence, may have no relation to molecular hydrophobicity. Recalculation of the fitting parameters a (intercept) and b (slope) finally yielded the relation

$$\log_{10}(\text{sweetness}) = (1.87 \pm 0.11) + (0.73 \pm 0.07) \cdot \log P$$

with a very good goodness-of-fit probability $Q = 0.993$ and a correlation coefficient of $r_{xy} = -0.955$.

Recently, the relative sweetness of various halo-sucroses has been correlated with their internal conformational energy (MM2HB molecular energy for a conformation at the torsion angles $\Phi = 95^\circ$ ($O_5-C_1-O_1-C_2$) and $\Psi = 75^\circ$ ($C_1-O_1-C_2-O_5$) relative to the global minimum of the Φ / Ψ -energy potential surface)^[158]. The linear relation of $\log_{10}(\text{sweetness}) = (7.1 \pm 1.3) - (1.0 \pm 0.2) \cdot \text{energy}(\text{kcal/mol})$ derived therefrom suffers from relatively large standard deviations of the parameters a and b , from a poor goodness-of-fit probability ($Q = 0.005$), and a low correlation coefficient ($r_{xy} = -0.615$), indicating that the proposed correlation "may be acceptable"^[161] only, if the respective standard deviations (errors given $\sigma_{\text{energy}} = \pm 1 \text{ kcal/mol}$, and a factor of 2 in relative sweetness) have been underestimated. However, this approach is only valid, if all molecules adopt the same conformation in binding to the receptor, and if the energy of interaction is constant in all cases – conditions which are unlikely to prevail. Comparison of quantitative structure-sweetness-relationships derived from the free energy of conformation obtained from MD simulations in water (GROMOS)^[44] and the sweetness-hydrophobicity correlation presented here even points towards the hydrophobic character of the receptor cavity with conformational effects playing a secondary role only.

The partition coefficients ($\log P$ values) used in this work correspond to a free energy contribution of hydrophobic binding to the receptor. When translated into energies, the error $\sigma_{\text{energy}} = \pm 0.7 \text{ kcal/mol}$ ($= \sigma_{\log P} = \pm 0.5$, at 300K) assumed here is lower than the one in the conformational energy correlation. The clear relationship between hydrophobicity and sweetness furthermore indicates that all sucrose analogs conceivably bind in a very similar fashion to the same sweetness receptor rather to several different ones as proposed by some authors^[7,8]. This appears to suggest the

* Fitting of the data set weighted by two-dimensional errors in X ($\log P$) and Y (rel. sweetness) yields a slightly different and more accurate relationship than simple straight line fitting to data points without error consideration does.

view that the modified AH-B-X-glucophore outlined above is the same at least for all of the sucrose derivatives discussed here.

Conclusions

Modern molecular modeling techniques such as the MOLCAD-program open up the possibility to obtain more "realistic" ideas on the physico-chemical properties of molecules, and consequently their chemical behavior and biological activity. As exemplified for sucrose, the incorporation of the three-dimensional shape of this molecule, of its contact surface, and the MEP-profiles provide an unambiguous explanation of the regioselectivities observed in trapping electrochemically (or simple base-induced) generated sucrose-alkoxide. The assessment of hydrogen bond donor and acceptor functionalities on the basis of these MEP patterns leads to new insights into the biological importance – particularly in view of sweetness – of individual hydroxyl groups in sucrose, results which are furthermore supported by the MLP's. As a major consequence, the model generally used for structure-sweetness relationships has been revised in order to increase its predictive value. The correlations presented above provide a comprehensive explanation of the relative sweetness potencies of a large number of sucrose analogs.

In future, **quantitative structure-activity relationships (QSAR)** may be developed from more detailed considerations of molecular properties, including fuzzy-logic supported recognition algorithms of possible molecular regions involved in binding to the sweetness-receptor. The models presented point towards modeling the complementary receptor side in future, too, using analogous molecular modeling techniques.

Appendix – Computational Methods

I. Φ/Ψ – Energy Potential Surfaces and Contour Plots. Molecular mechanics calculations were carried out using the PIMM88^[45] force field for vacuum conditions ($\epsilon = 1$). Different starting structures were generated from the solid state geometry by variation of the torsions involved in the three hydroxymethyl groupings. The Φ/Ψ -torsion angles were driven in steps of 10° each and fixed, whilst all other molecular parameters were fully optimized in each step. New geometries were generated from conformationally adjacent, already optimized conformers, the self-consistency of the map was counter checked by double calculation of structures on same Φ/Ψ -grid points, generated by different conformational pathways. Hydrogen bonds were treated as described in Ref. [45], no additional energy terms have been used. The PIMM88 force field was shown to properly reproduce the anomeric and

exoanomeric effects in acyclic and cyclic acetal structures such as dimethoxy methane, dihydroxy methane, and tetrahydro-pyranosides without any application of additional potentials^[162]. Contour plots were computed using cubic regression formulas^[163].

II. Molecular Dynamics Simulations. The GROMOS^[127] force field with the united atoms model for CH- and CH₂-groups and SPC-type water^[164] was used for all MD simulations. Standard GROMOS parameters of CS1, CS2 (CH- and CH₂-groups), CB (C-2'), OA-HO (OH-groups), and OS (acetal oxygens) were used (total 31 atoms for sucrose), with charge groups^[165] defined as C_n-O_n-H_n ($n = 2, 3, 4, 6, 1', 3', 4'$, and 6' with charges +0.150, -0.548, and +0.398) and C₅-O₅-C₁-O₁-C₂'-O₅'-C₅' (charges C-5 and C-5' +0.160, C-1 and C-2' +0.380, all oxygens -0.360). All sucrose starting geometries were placed into a pre-equilibrated box of water. Prior to the MD simulations the systems were first geometry optimized with fixed atomic positions of the solute (relaxation of the solvent phase only), followed by an unrestrained full energy minimization of the entire system. Periodic boundaries (truncated octahedron) and isothermal and isobaric conditions ($T = 300\text{K}$, $p = 1\text{bar}$) were applied, with a temperature relaxation time of 100 fs, a pressure relaxation time of 500 fs (during the first 1 ps of equilibration $\tau_T = 10\text{fs}$ and $\tau_p = 50\text{fs}$ were used)^[166], and a cut-off radius of 8.5 Å for the long-range electrostatic interactions. 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^[167]. Coordinates were saved every 25 steps (50 fs) for analysis purposes. The crystal structure^[25] and PIMM88 models were used as starting geometries (X-ray: +211 H₂O, $r_{\text{box}} \approx 23.85(9)\text{Å}$, $t = 100\text{ps}$; **A**: 571 H₂O, 32.90(8) Å, 500 ps; **B**: 265 H₂O, 25.64(8) Å, four runs with 100-200 ps; **C**: 256 H₂O, 25.34(8) Å, 2 runs of 100 ps each), the first 25 ps of every simulation were discarded for equilibration.

The force field was checked by 200 ps MD simulation of $2 \cdot 2 \cdot 3$ unit cells (a total of 32 sucrose molecules) of the crystal structure (monoclinic space group P2₁, $a = 10.86\text{Å}$, $b = 8.71\text{Å}$, $c = 7.76\text{Å}$, and $\beta = 102.95^\circ$, $Z = 2$) with periodic boundary conditions. The molecular geometry and all intra- and intermolecular hydrogen bonds were reproduced with satisfactory accuracy.

The fructofuranose pseudorotational energy potential surface as a function of the Cremer-Pople^[122] puckering parameters q and ϕ (Fig. 3-8) was derived from the 500 ps MD simulation including 571 H₂O molecules according to the Boltzmann statistics ($T = 300\text{K}$, width of classes $\Delta q = 0.05$ and $\Delta\phi = 18^\circ$), contours were plotted in q/ϕ -polar coordinates^[163].

III. Free Energy Calculations using Umbrella Sampling. An umbrella potential $U_i^*(\Psi')$ according to eq. 3-4 with $\phi \equiv \Psi'$ (C₁-O₁-C₂'-C₃') was used for the free energy

calculation, dividing Ψ' into 360 classes of 1° width each. Probability distributions $P_i^*(\Psi')$ were calculated as a fraction of sampled points per class for each trajectory. Starting from the 100ps pre-equilibrated sucrose MD including 211 H₂O (simulation parameters as described above, $r_{\text{box}} \approx 23.85(9)\text{\AA}$, $\Psi' \approx \pm 180^\circ$), two independent molecular dynamics series of 13 consecutive simulations each (all 150ps, incl. 25ps of equilibration, altogether 3900ps!) were initiated. The phase angle Θ_i was varied from $0^\circ \rightarrow -180^\circ$ (series I) and $0^\circ \rightarrow +180^\circ$ (series II) in steps of 20° and / or the force constant K_i was changed in the range of 35 – 300kJ/mol ($n_i = 1$ for all MD runs) between individual simulations, final ("pre-equilibrated") structures 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 $\Psi' \approx -15^\circ$ (series I: **A** \rightarrow **B** \rightarrow central barrier, series II: **A** \rightarrow **C** \rightarrow central barrier, with $\Psi'(\mathbf{A}) \approx \pm 180^\circ$, $\Psi'(\mathbf{B}) \approx +70^\circ$, and $\Psi'(\mathbf{C}) \approx -90^\circ$). Each distribution $P_i^*(\Psi')$ covered a Ψ' -space of approx. $70 \pm 15^\circ$, with overlap regions of $\approx 55 \pm 10^\circ$ between neighbored distributions. Least squares fitting (weighted by $(P_i^*(\varphi) \cdot P_j^*(\varphi))^{1/2}$, points differing by more than 2σ were rejected) of the $\Delta G(\varphi)$ sub-curves for series I and II yielded the overall free energy function. The Ψ' -distribution obtained from the 500ps simulation of sucrose incl. 571 water molecules (cf. Appendix II) was used as a common starting reference for both MD series I and II.

The relative error estimated from the hysteresis on top of the central barrier (+34.8kJ/mol) was found to be $\sigma(\Delta G) \approx \pm 1.0\text{kJ/mol}$.

Essentially the same results were obtained when applying different weighting schemes $(P_i^*(\varphi) \cdot P_j^*(\varphi))^n$ with $n = 0 - 3$, and / or when discarding up to 125ps from each 150ps trajectory for equilibration purposes, leaving only the rest for calculation of the distribution functions $P_i^*(\varphi)$, thus indicating the excellent statistics obtained during all simulations.

Two-dimensional analysis of the umbrella samplings (Fig. 3-11) was performed using classes of $\Delta\Phi = \Delta\Psi = 10^\circ$ width and application of the same weighting and fitting scheme as above. The error was estimated to $\sigma(\Delta G(\Phi, \Psi)) \approx \pm 2.0\text{kJ/mol}$.

Acknowledgement: I am grateful to Prof. Dr. W. F. VAN GUNSTEREN, Eidgenössische Technische Hochschule, Zürich, for using the GROMOS molecular dynamics software package and granting me access to his computational facilities. I am also indebted to Prof. Dr. A. D. FRENCH, Southern Regional Research Center, New Orleans, for providing an advance plot of his *ab initio* results on the sucrose analog.