Evolution of the Structural Representation of Sucrose [1]

Frieder W. Lichtenthaler, Stefan Immel, and Uwe Kreis, Darmstadt (Germany)

Following a historical account on the establishment of the constitutional formula of sucrose and of its conformational features, the present possibilities for interactive graphics display of its molecular geometry, based on X-ray structural data, are given. In addition, the MOLCAD program-computed contact surface is presented as well as – in a 16 color code ranging from violet ro red – the electrostatic potential and, most relevant for structure-sweetness relationship considerations, ist hydrophobicity potential profile on the contact surface. Finally, an attempt is made towards a tentative assessment of the computer-generated distribution of hydrophilic and hydrophobic regions over the surface of the sucrose molecule in terms of the "sweetness triangle" AH-B-X concept. Also given are the graphic displays for the contact surfaces and hydrophobicity profiles of leucrose, isomaltulose and its reduction products α-glucosyl-mannitol and -sorbitol.

Evolution der strukturellen Darstellung der Saccharose. Nach einem kurzen Abriß der Historie, die zur korrekten Konstitutions- und Konformationsformel der Saccharose führte, werden die derzeitigen Möglichkeiten präsentiert, die auf Röntgenstrukturdaten basierende molekulare Geometrie der Saccharose graphisch darzustellen. Darüberhinaus wird die mit dem Computer-Programm MOLCAD generierte Kontaktoberfläche der Saccharose erstmals bildlich dargestellt, ebenso wie – in einem von violett bis rot reichenden 16 Farben-Code – die Verteilung des elektrostatischen Potentials und der Hydrophobizität auf diese Oberfläche, was für das Verständnis von Struktur-Süßkraft-Beziehungen von grundlegender Bedeutung ist. Schließlich wird versucht, anhand des computergenerierten Hydrophobizitätsprofils der Saccharose eine vorläufige Bewertung des als „Süßkraft-Dreieck" apostrophierten AH-B-X Konzepts zu geben. Ebenso werden die Kontaktoberflächen und Hydrophobizitätsprofile von Leucrose, Isomaltulose sowie deren Reduktionsprodukte α-Glucosyl-mannitol und -sorbitol abgebildet.

A picture may instantly present what a book could set forth only in a hundred pages.
Ivan Sergeyevich Turgenev

1 Introduction

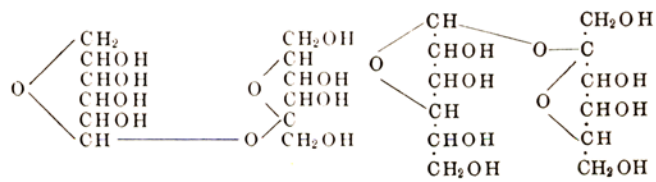
Knowledge of the structure of a molecule is of fundamental importance, inasmuch as its *composition* (molecular or sum formula), *constitution* (type of bonding between elements), *conformation* (three dimensional arrangement of elements) and its *dynamic stereochemistry* (time and solvent dependent aspects of conformation) not only determine the physical, chemical, and biological properties, but – ultimately – the industrial applications as well. This most pertinently applies to the fairly complex molecule of sucrose, which with an annual production of over a 100 million tons is certainly the most readily available, lowest cost, crystalline organic compound.

Sucrose being a leading world commodity for centuries, it is not surprising that the first efforts to unravel its structure reach back to the very beginnings of organic chemistry as it gradually unfolded with the advent of reasonably exact elemental analyses. The first, albeit not very accurate combustion analyses by *Prout* in 1827 [2], and their interpretations by *Liebig* [3], *Peligot* [4], *Berzelius* [5], and *Dubrunfaut* [6] eventually led to the correct molecular formula $C_{12}H_{22}O_{11}$. Its expansion into a satisfactory constitutional formula mystified a number of great organic chemists and proceeded in a number of distinct phases, each reflecting the progress in experimental methodology: methylation and periodation studies leading to the correct ring structures [7], investigations on the chemical and enzymatic hydrolysis of sucrose establishing the anomeric configurations about the intersaccharidic junction [7], X-ray- and neutron diffraction-based structural data unravelling ring conformations and non-bonded interactions between the saccharide portions [8] and NMR studies providing first insights into the dynamic stereochemistry of surcose in solution [8]. The newest phase of structural visualization of the sucrose molecule appears to have been launched by the availability of computer-based graphic and computational tools that allow molecular modeling with the potential for eventually understanding structure-activity (sweetness) relationships. In the sequel, an account is given on the gradual evolvement of the structure of sucrose from its molecular formula $C_{12}H_{22}O_{11}$ to the color-coded visualization of its hydrophobicity potential profile with particular emphasis to the current possibilities in computer-aided molecular modelling.

2 The Constitutional Formula of Sucrose

As early as 1883, *Tollens* [9] advanced the glucoseptanosyl fructofuranoside formula **1** for sucrose, followed, 10 years later, by *Emil Fischer*'s proposition [10] that sucrose rather is a glucofuranosyl fructofuranoside (**2**).

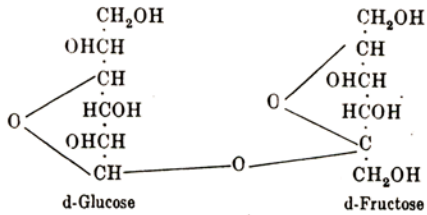
Neither formulations contained configurational implications, nor were their ring structures adequately supported by experimental evidence of which *Fischer* was well aware: "*Ich halte die schon von Tollens vor längerer Zeit aufgestellte, allerdings nur ungenügend begründete Structurformel mit einer kleinen von mir vorgenommenen Änderung im Wesentlichen für richtig*" [10].



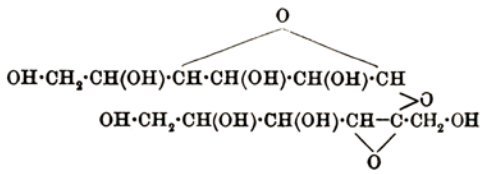
1 B. Tollens, 1883 **2** E. Fischer, 1893

Around 1915 structural representations of type **3** [11], **4** [12], and **5** [13] appeared to be the state of the art. Thereby it is worthy

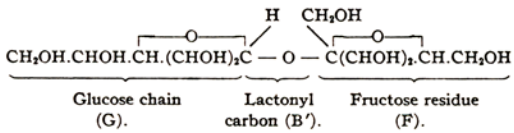
to note that *Tollens'* formulation **3** contains the configurational features for the glucose and fructose portions as established by *E. Fischer* in 1891 [14], whereas *Haworth* [12] and *Hudson* [13], obviously, were not prepared yet to accept *Fischer*'s projection formula representation advanced 25 years earlier.



3
B. Tollens, 1914

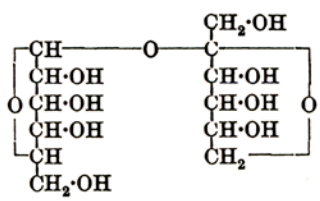


4
W. N. Haworth, 1916

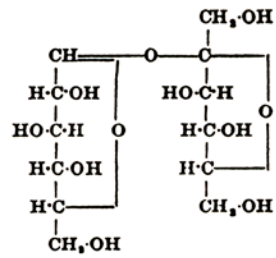


5
C. S. Hudson, 1916

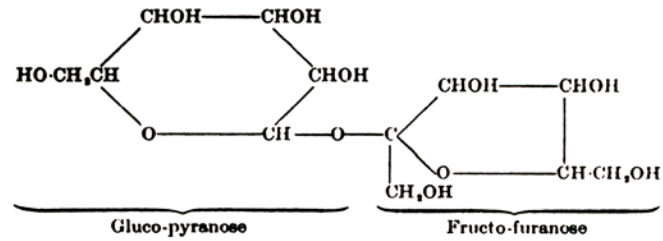
It was not until 1926 – a hundred years after *Prout*'s combustion analyses [2] – that *Haworth* and his coworkers [15], on the basis of their classical methylation studies, deduced the correct ring structures for sucrose as depicted in the representations **6** [15] (without configurational assignments), **7** [16] (implementing configurational relationships), and **8** [16]. The latter, in fact, appears to be a first, cautious formulation at what was to become the "*Haworth* perspective formula", that in the ensuing years found general acceptance in the form propagated by *Pigman* [17], *Purves* [18], and others. The particularly lucid, didactically auspicious formulation **10** used by *Morrison* and *Boyd* [19] in the first edition of their textbook on Organic Chemistry (1959) appears to be the first instance where conformational concepts were incorporated.



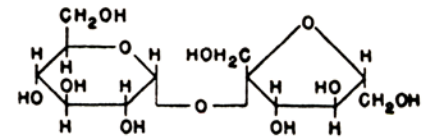
6
W. N. Haworth, 1926



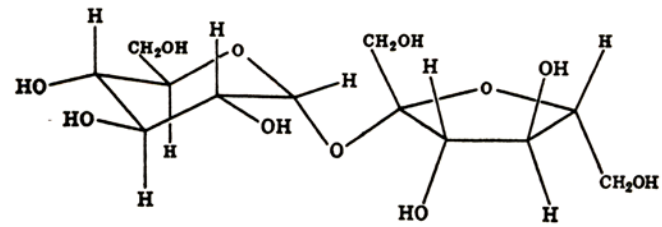
7
W. N. Haworth, 1929



8
W. N. Haworth, 1929



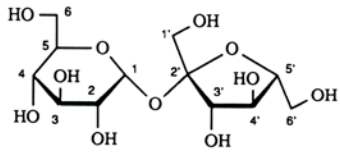
9
W. W. Pigman, 1948



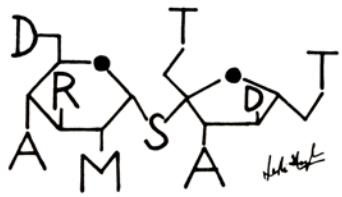
(+)-Sucrose
α-D-Glucopyranosyl β-D-fructofuranoside
β-D-Fructofuranosyl α-D-glucopyranoside
(no anomers; *non-mutarotating*)

10
R. T. Morrison
R. N. Boyd, 1959

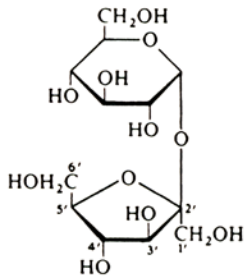
Except for writing the respective C-hydrogens, these formula representations are still used today, as for example, the formulation **11** favoured by *Hough* [8], which is sufficiently clear and simple to straightforwardedly develop the chemistry of sucrose and which, as evidenced by **12** [20], may even be subjected to unusual substitution to account for carbohydrate symposium meeting places.



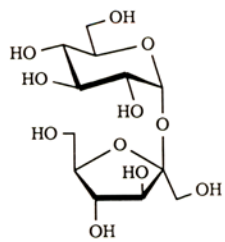
11



12



13

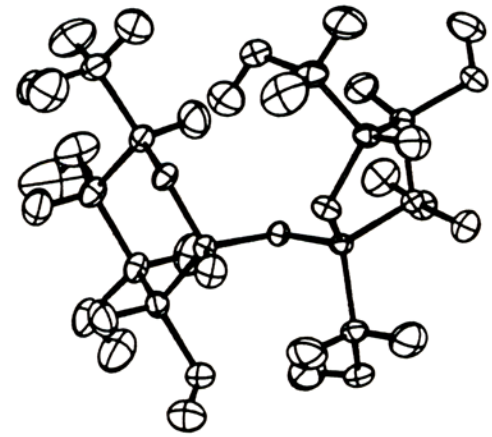


14

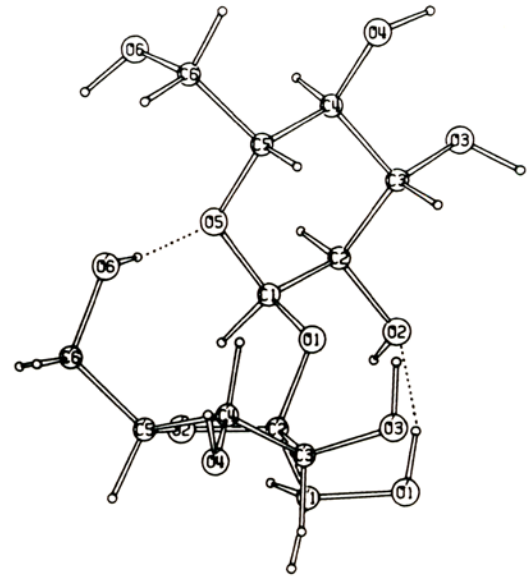
Another representation that is reasonably clear and economical in space are formulae **13** [21] and the Chem.Text. [22] graphics program-generated **14**. Both do not give, of course, the correct disposition of the glucose and fructose portions in the solid state or in solution, but have the didactic advantage that the configurational relationship of the two monosaccharide units, i.e. identical arrangements at C-3 to C-5 of pyranose and furanose rings, become particularly obvious.

3 The Conformation of Sucrose

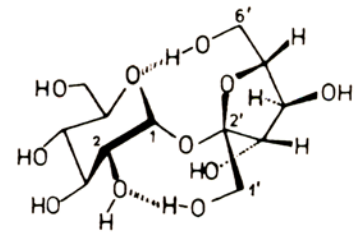
A new possibility to further probe into the structural subtleties of sucrose opened up by crystal structure determinations. The first measurements accurate enough to locate the positions of the hydrogen atoms, were based on neutron diffraction data by *Brown* and *Levy* in 1963 [23], which upon later refinement [24,25] resulted in structure **15**: the glucose portion adopts a $^4C_1$ conformation whereas that of the fructofuranose residue is a $^4T_3$ twist. The two rings are approximately at right angles, yet the overall structure, most notably, is determined by two strong intramolecular hydrogen bonds: one between the terminal 6'-OH of fructose to the pyranoid ring oxygen of 1.89 Å length, the other, even shorter, reaching from the primary 1'-OH of fructose to O-2 of glucose (1.85 Å), as depicted in formulae **16** [24], **17** [26], and **18** [27]. These hydrogen bonds serve to fix the molecule in a well-ordered, rigid conformation.



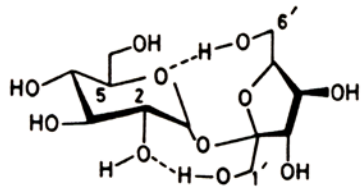
15



16

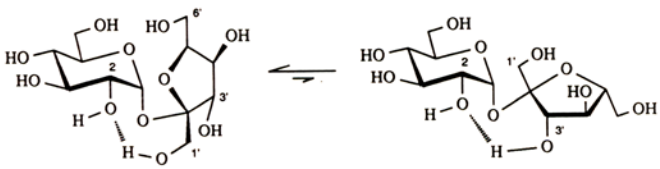


17



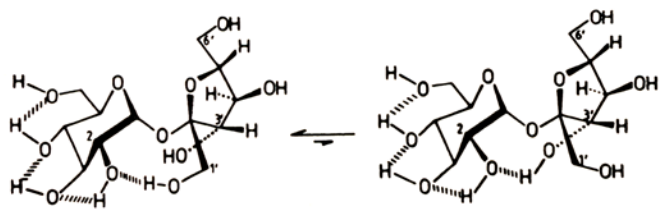
18

As inter- and intramolecular hydrogen bonding in the solid state is influenced by packing effects, it cannot be expected that the fixed overall shape as illustrated by formulation **15–18** is retained in solution, since the hydroxyl groups could satisfy their hydrogen bonding requirements by bonding with the solvent. Accordingly, it is surprising that through probing with several NMR criteria in conjunction with hard sphere *exo*-anomeric (HSEA) calculations, *Bock* and *Lemieux* [28] provided strong evidence that the overall conformation of sucrose is approximately the same in dimethyl sulfoxide and water solutions, and that this conformation is similar to that observed in the solid state. Only the longer, and, hence weaker 5-O^g ··· HO-1^f hydrogen-bond is disintegrated by solvation [28]. These conclusions are supported by detailed analysis of ^{13}C-NMR spin-lattice relaxation experiments of aqueous sucrose solutions, showing its conformation to be independent of temperature and concentration in the 0.1–1 molar range [29,30], and by recent SIMPLE ^{1}H-NMR isotope-shift measurements [31,32]. The latter revealed the presence of two intramolecular hydrogen-bonded conformations **19** and **20** for DMSO as well as for aqueous solutions of sucrose, in which the 2-oxygen of the glucose portion (i.e. O-2^g) acts as the acceptor for either the 1'-OH or 3'-OH of the fructose moiety:



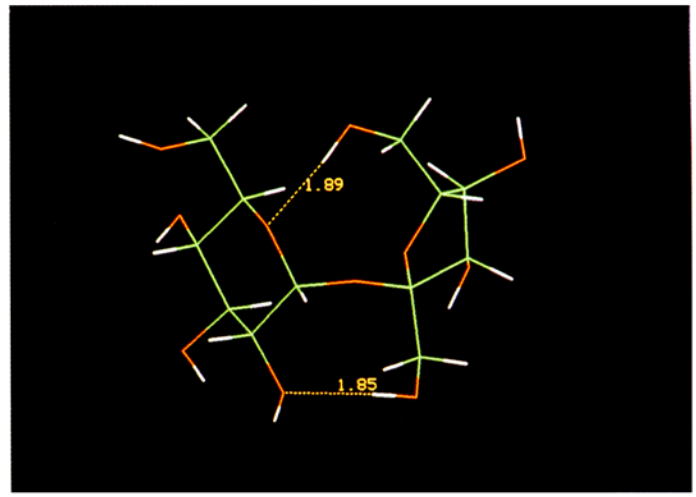
19 **20**

These two forms, which are graphic representations adapted from ref. 7, exist in a competitive equilibrium in which **19** is favoured over **20** in an approximate 2:1 ratio [31,32]. The ready interconversion of these two conformations becomes sterically understandable when using a more realistic sucrose representation, in which the clockwise arrangement of the hydrogen bond network [31] has been incorporated into the pyranoid ring:

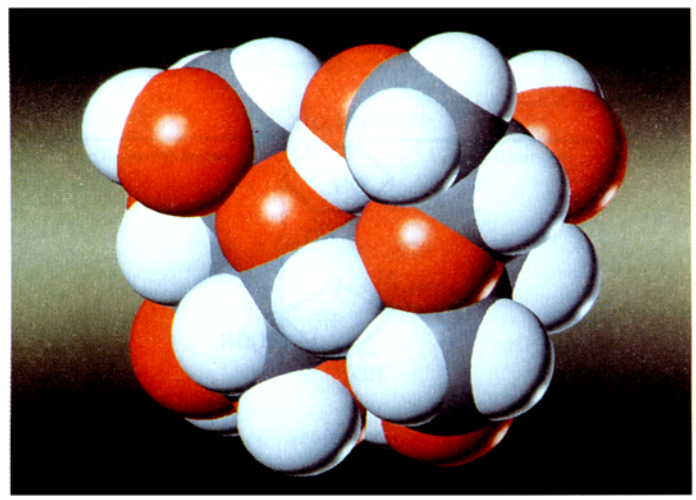


19 **20**

The proximity of 1'-OH and 3'-OH is clearly evident, the switch from one form to the other requiring a minor rotation within the intersaccharide linkage only, thus largely retaining the overall molecular geometry. Consistent with this picture of sucrose as a relatively rigid molecule in the crystal and in solution are recent molecular dynamics simulations in the vacuum [33], where the



a)



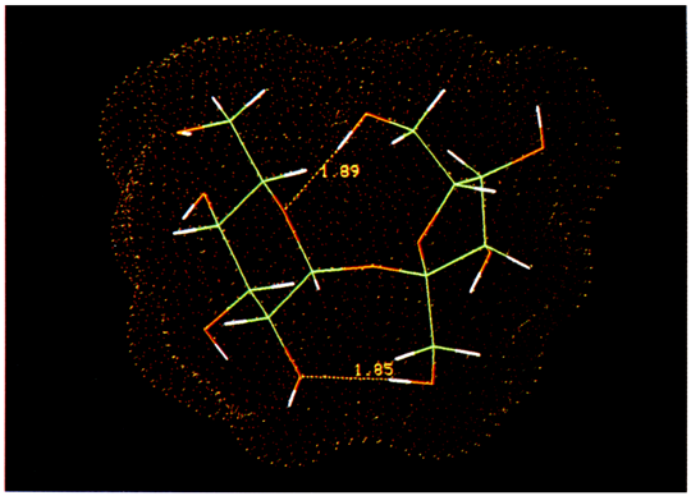
b)

Fig. 1. MOLCAD generated models of sucrose in the solid state, based on the neutron diffraction data of *Brown* and *Levy* [24]. Pictures are photographed from the computer screen. (a) *Dreiding* model representation including the two intramolecular hydrogen bonds (dotted lines) and their interatomic distances (in Å). (b) Space-filling CPK type form in the same orientation as above (white: hydrogen; grey: carbon; red: oxygen).

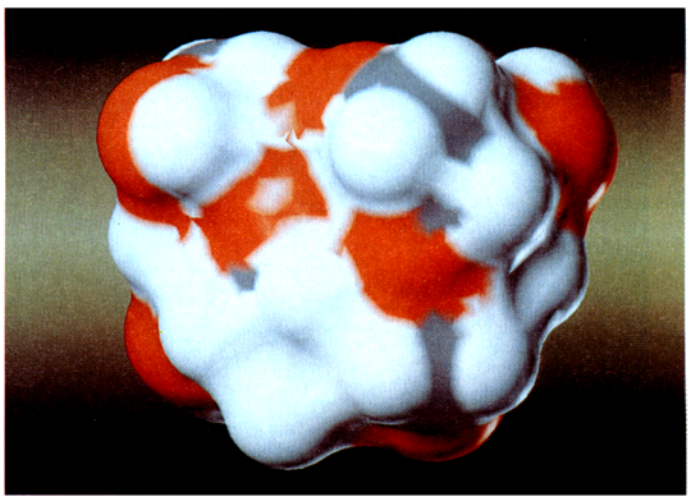
intramolecular hydrogen bond is retained in all but one of the five local minimum energy conformations.

4 Possibilities for Interactive Graphics Display of Sucrose

Whilst crystallographers pioneered techniques to visualize, scrutinize, and manipulate three-dimensional molecular models – the ORTEP program [34] used in the sucrose representation **15** being a notable, early example – only their combination with computational chemistry led to molecular modeling at its present impressive level [35–38]. It offers the chemist an expanding arsenal of user-friendly tools with which molecules can be displayed in unexcelled, three-dimensional representations, relieving from the "drawing artistry" required to jot down on paper formulae of type **19** and **20**, for example. The additional possibility to impose vital chemical properties into the three-dimensional molecular representation, such as the electrostatic potential or –

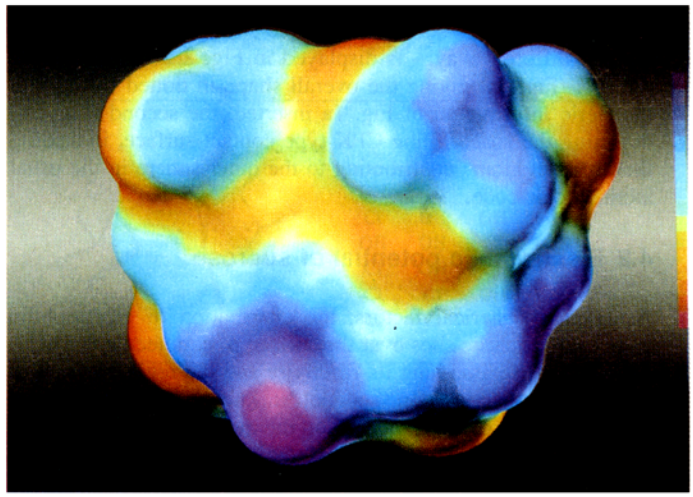


a)

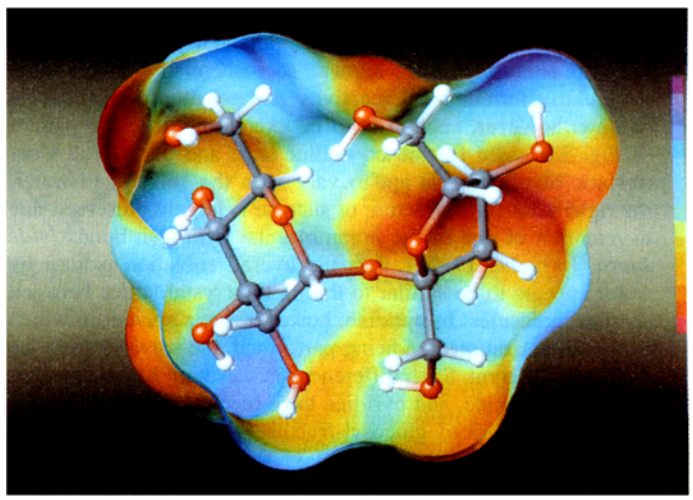


b)

Fig. 2. Contact surface (roughly equivalent to solvent-accessible surface) of crystalline sucrose in dotted form with *Dreiding* model insert (a), and in the space-filling CPK model representation (b), the individual atoms contributing to the outer sphere being indicated by the usual atomic colors (orientation of sucrose corresponds to that in Fig. 1).



a)



b)

Fig. 3. (a) Electrostatic potential molecular surface of sucrose, as obtained by MNDO-calculations, in color-coded form, red representing the negative, violet the positive maximum in charge density in relative terms. (b) For clarity and unequivocal orientation, a ball-stick model is inserted into the molecule in its half-opened form. As is clearly apparent the violet, i.e. most electropositive area is centered around the glucosyl-2-OH.

biologically more relevant – the hydrophobicity distribution over its surface, opens up an entirely new dimension of visualizing not only the structure of sucrose, but essentials of its chemical and physiological properties.

The various possibilities for structural display of the sucrose molecule are elaborated in the sequel utilizing the MOLCAD program* developed by *Brickmann* and coworkers [39] and the neutron diffraction-based structural data of *Brown* and *Levy* [24]. It is particularly simple to computer-generate pictures on a screen that correspond to the classical *Dreiding* model presentation (Fig. 1a). Therein, special attention is given to the intramolecular hydrogen bonds prevalent in crystalline sucrose, each being signified by dotted lines and the respective distances: 5-$O^g \cdots$HO-g^f (1,89 Å) and 2-$O^g \cdots$ HO-1^f (1.85 Å). Undoutedly, this representation is more exact and more lucid than the formulations **16**–**18**, mentioned above.

Fig. 1b displays the space-filling *Corey-Pauling-Koltun* (CPK)-type model of sucrose, that is as easily generated either in the orientation used in Fig. 1a, or in any other, producable by rotation on the screen. These CPK models, however – whether by computer-generation or by actual built using a model kit – display only the von der Waals surface of sucrose, which as such is not fully accessible to the solvent.

4.1 The contact surface of sucrose

A more "realistic" molecular surface is the contact- or *Connolly*-surface, that can readily be created by computers; it is defined [41–43] as the smooth molecule surface, which is accessible to a probe sphere with a radius of 1.4 Å, representing a solvent

* The program MOLCAD (for MOLecular Computer Aided Design) is an interactive, fast computer program for building and manipulating molecules and molecular systems. It is particularly suited to analyze and represent different physical molecular properties such as the electrostatic or hydrophobic potential on three-dimensional solid molecule surfaces, even of large molecules like proteins, zeolithes, and polymers. MOLCAD runs on Silicon-Graphics 4D-workstations and can be licenzed from the authors [40].

molecule like water in size. Such a contact surface – to be distinguished from the solvent-accessible surface [44] defined in a slightly different way – is displayed in Fig. 2 in two versions, the first (Fig. 2a) giving the overall shape in dotted form, with the steric information of Fig. 1a inserted, the second (Fig. 2b) depicting the space-filling CPK type contact-surface, indiçating the different atoms contributing to the outer sphere in the usual atomic color-code.

4.2 Electrostatic potential of sucrose

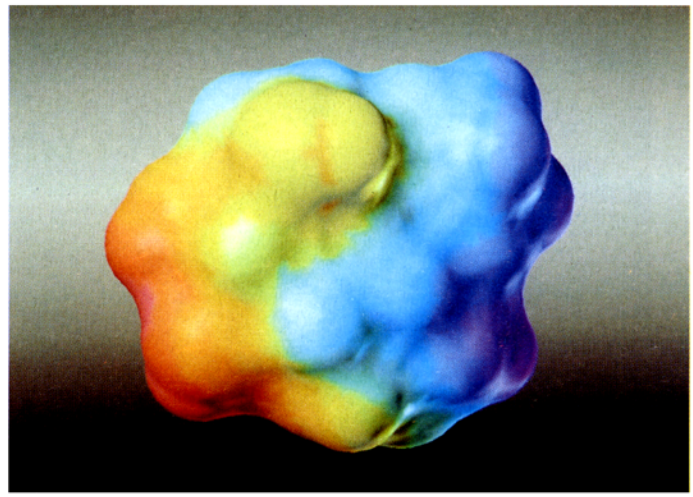
Aside purely sterical factors, the distribution of electrophilic and/or nucleophilic sites over the surface of a molecule is a decisive element in assessing its reactivity and the regioselectivities attainable in reactions. In this context, knowledge of the charge density or electrostatic potential on the molecular surface [45] is of vital importance. Calculation of this potential from the atomic charges obtained by semiempirical MNDO-methods using the MOPAC [38,46] program can indeed be accomplished, and, most conveniently, translated into a sixteen color-code, reaching from red (most negative) to violet (most positive). In the case of sucrose this is borne out in the pictures displayed in Fig. 3.

The correlation of the electrostatic potential patterns thus unveiled in sucrose with the known "sucrochemistry" is tempting, particularly ferreting out reactivity and selectivity profiles. As of now, however, caution is to be exercised, since the pattern of Fig. 3 is valid for the crystalline state and the overall electrostatic distribution is apt to vary with the solvent. In water, the rigid solid-state structure of sucrose is largely preserved due to retention of one of the intramolecular hydrogen bonds (*vide supra*, formulae **19–20**), thus the overall solvation pattern by water appears to be similar to that in the crystal lattice between sucrose molecules. In pyridine, however, i.e. a solvent in which many derivatizations are being performed, the solvation may be distinctly different. Such differences in solvation between water, dimethyl sulfoxide, and pyridine have recently been established for the five α-linked glucosyl-fructoses isomeric with sucrose [47] and are likely to be found in sucrose as well. In addition, the electrostatic distribution over the surface of sucrose may vary with the reagent utilized for ensuing reactions. Nevertheless, electrostatic potential surfaces show promise of having substantial predictive value in further exploiting the chemistry of sucrose.

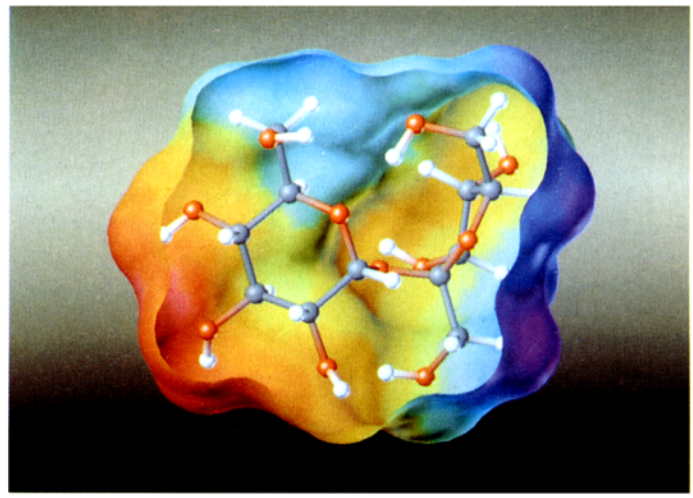
4.3 The hydrophobicity potential of sucrose

The interaction of a biologically active molecule with its receptor – as, for example, of sucrose with the taste bud initiating the sweet sensation – is not only determined by the spatial complementarity of their three-dimensional molecular geometry and the electrical conformity of their charges on the surface, but also by the hydrophobic interactions between their contact surfaces [48,49]. Since there is now sufficient evidence to assume that the sweet-taste receptor is proteinaceous and that the main driving-force for the initial binding is hydrophobic in nature [50–52], the elicitation of the sweet-taste responce is likely to involve an interaction between a hydrophobic cleft in the surface of the protein and a hydrophobic portion of sucrose [28,52]. Consequently, for a new insight into the sweetness of sucrose, the *van der Waal* surface and the electrical potential of the effector molecule is not adequate but must be complemented by an appropriate estimation and representation of the hydrophobicity around the molecule.

The hydrophobic effect [53] of a molecule describes its tendency to orient in non-homogenous ways, like water to exclude hydrophobic substances and *vice versa*. Usually, the hydrophobicity of a molecule is only seen as an over-all quality and



a)



b)

Fig. 4. MOLCAD-generated hydrophobicity potential of sucrose in the crystalline state, red representing most hydrophilic, blue the most hydrophobic portion of the molecule (in relative values). For identification of the atoms contributing to hydrophobicity the ball-stick model is inserted into an opened form. The orientation of sucrose used here is different from that in Figs. 1–3; it was chosen to illustrate the opposite hydrophilic and hydrophobic sites of the molecule.

expressed as the partititon coefficient in the water/*n*-octanol system. First attempts to calculate this "one-dimensional" property led to an incremental system, assigning a hydrophobic potential to molecular fragments or atoms [54]. Taking into account that the hydrophobicity potential of each fragment or atom decreases exponentially with increasing distance, it it possible to calculate it at every point of the molecule's surface [48,49]. It must be stressed though, that the hydrophobic potential is empirically defined and calculated; it is not proportional to the electrostatic potential, since, for example, $-NH_3^+$ and $-COO^-$ have opposite electrical charges, but similar contribution to hydrophobicity.

Application of this computational methodoly [48,49] to sucrose and expressing it in the sixteen color-code on its surface, yields the hydrophobicity (or molecular lipophilicity) potential represented in Fig. 4, the respective legend explaining the meaning. Especially informative is the half-opened model with the stick and ball model insert, which allows unequivocal correlation of the most hydrophobic (i.e. violet) section of the molecule to the fructose portion, involving the region characterized by the

CH-protons H-1^f, H-1'^{f} ··· H-3^f ··· H-5^f ··· H-6^f, and H-6'^{f}. The hydrophilic (red) section is distinctly separated therefrom, quasi at the "other side" of the molecule (Fig. 4), with its major intensity centered around the 3-oxygen of glucose.

5 Preliminary Assessment of the Hydrophobicity Potential Profile of Sucrose

The topography of the region in sucrose amenable to engagement in hydrophobic bonding has been assessed [28] on the basis of CPK models to include parts of the surface described by H-3^f, O-5^f, O-5^g, and the surface involving H-1^f, H-1'^{f}, O-1^f ··· O-2^g, H-1^g, and H-2^g. Comparison with the computational data presented in Fig. 4a and b, however, reveals them to be distinctly different, particularly with respect to the non-contribution of the glucose portion to the hydrophobicity of the molecule. In view of the fact that – at least in the hands of the authors of this article – it is essentially impossible to delineate with any certainty hydrophobic or hydrophilic regions from a CPK model of sucrose, the computationally generated hydrophobicity profile of Fig. 4 is deemed the more realistic one. By consequence, the hydrophobic cleft of the taste protein may conceivably be portrayed to correspond to the hydrophobic region in the fructose part of sucrose (cf. Fig. 5):

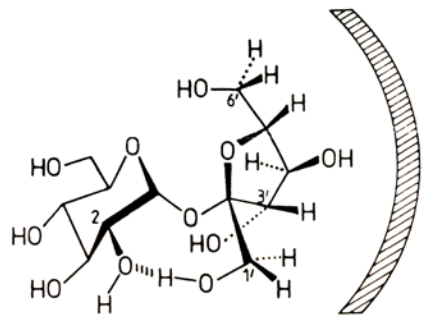
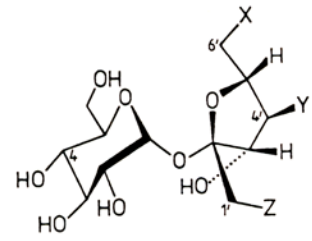


Fig. 5. Conceivable hydrophobic cleft of the taste bud protein corresponding to the hydrophobic fructose portion of sucrose.

In keeping with this notion, it would have to be surmised that an increase of hydrophobicity in the fructose portion – i.e. along the violet region in Fig. 4 – favours binding to the taste bud receptor protein and, hence, enhances sweetness. The data collected in Table 1 appear to support this rationalization, since replacement of the fructose hydroxyl groups at the 1'-, 4'-, and/or 6'-position by chlorine uniformly leads to compounds sweeter than sucrose [55]. Thereby, it is probably not without significance that, with a factor of 3500, the sweetness enhancement is largest in the 4',6'-dichloro-sucrose, i.e. in the compound in which hydrophobicity is substantially increased with maintenance of the fructose-1'-OH, involved in hydrogen bonding to the glucose-2-OH in the crystal and in aqueous solution (cf. above).
In this context, the glucose-4-OH group, close to the hydrophilic region of sucrose and contributing to it (cf. Fig. 4) appears to play a weighty role, because its replacement (with inversion) by chlorine substantially extends the fructose-located hydrophobic region into the glucose portion.

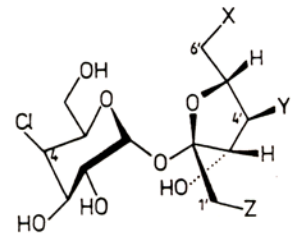
The sweetness enhancements obtained thereby are amazing (cf. Table 2), most remarkable being the steady increase of sweetness with the degree of substitution, particularly evident when going from 4,1',6'-trichloro-*galacto*-sucrose ("sucralose®", 650 × sweeter than sucrose) to its 4'-fluoro- (1000 ×), 4'-chloro- (2000 ×), 4'-bromo (3000 ×) and 4'-iodo (7500 ×) derivatives.

Table 1.
Relative Sweetness of Sucrose in which the Fructose Hydroxyl Groups at 1', 4', and 6' are Replaced by Halogen [55].



X	Y	Z	Relative Sweetness
OH	OH	OH	1 (sucrose)
Cl	OH	OH	20
OH	OH	Cl	20
Cl	OH	Cl	76
Cl	Cl	OH	3500
Cl	Cl	Cl	100
Br	Br	Br	30

Table 2.
Relative Sweetness of 4-Chloro-4-deoxy-*galacto*-sucroses in Relation to the Halogen Substitution Pattern in the Fructose Portion [55].



X	Y	Z	Relative Sweetness (sucrose = 1)
OH	OH	OH	5
Cl	OH	OH	120
Cl	Cl	OH	220
Cl	OH	Cl	650 (sucralose)
Cl	F	Cl	1000
Cl	Cl	Cl	2000
Cl	Br	Cl	3000
Cl	I	Cl	7500

The taste properties of deoxy-sucroses and of a variety of its methyl ethers [52] allow similar rationalizations as to their relationships between sweetness and substitution (i.e. hydrophobicity) pattern. An assessment is not given since sweetness effects are substantially smaller and require considerably more detailed modellings to arrive at meaningful conclusions.

In this context, it its tempting to scrutinize *Shallenberger*'s AH-B concept [56] or the more advanced AH-B-X theory of *Kier* [57] in terms of the electrostatic and the hydrophobicity potential profile emerging from the calculations (Figures 3 and 4), particularly in view of X being a hydrophobic factor which acts in harmony with the AH-B unit to guide or lock the sweet compound into the proteinaceous receptor site. What clearly evolves from the distinctly hydrophobic (i.e. violet) region of sucrose in Fig. 4, is that the X-part of the AH-B-X triangle is to be placed into the fructose portion, conceivably within the 1^f-CH_2-3^f-H area; this X-site, obviously, can vary within a certain range and reach over, via the 6^f-CH_2, to the glucose portion (the axial side of C-4^g, in particular), to account for the dramatic increase of sweetness on increase of hydrophobicity in the 4^f-, 6^f-, and 4^g-portions, e.g. by chlorine substitution (cf. Tables 1 and 2).

Correspondingly, the AH-B unit of the glucophore – AH being the proton donor, B the proton acceptor in simultaneous hydrogen-bonding to a complementary AH-B system of the receptor protein – is most likely to be located at the diol grouping made up by the 2-OH and 3-OH of the glucose portion which turns out to be the most hydrophilic (red) region of the molecule (cf. Fig 4). The pronounced tendency of the equatorial glucose 2-OH to engage in intramolecular hydrogen bonding is not only prevalent in sucrose (cf. Fig. 1a), but also in other disaccharides containing α-glucosyl residues (cf. Table 3).

Table 3.
Intramolecular Hydrogen Bonding of the Glucosyl-2-OH (2-O^g) in the X-ray Structures of α-Glucosyl-disaccharides.

Compound	Intramolecular hydrogen bond	Distance (Å)	ref. (X-ray data)
Sucrose	2-O^g ···HO-1^f	1.85	24
Isomaltulose	2-O^gH··· O-2^f	2.28	58
Turanose	2-O^g ···HO-4^f	2.11	59
β-Maltose	2-O^gH··· O-$2^{g'}$	1.84	60
Methyl β-maltoside	2-O^gH··· O-$3^{g'}$	1.86	61
α-maltose	2-O^g ···HO-$3^{g'}$	2.06	62
Phenyl α-maltoside	2-O^g ···HO-$3^{g'}$	1.74, 1.90[a]	63
Sucralose[b]	2-O^gH··· O-3^f	1.86	64

[a] Corresponding to two different molecular conformations realized in the crystal lattice.
[b] The direction of the intramolecular hydrogen bond is reversed to 2-O^g···HO-3^f on dissolution in dimethyl sulfoxide [65].

Apparently, any sterically available OH-group in the second sugar moiety can participate in this hydrogen bonding interaction. It is interesting to note that the glucosyl-2-OH (2-O^g in Table 3) can take part as the proton acceptor, as in sucrose, turanose, and α-maltose, or as the proton donor (isomaltulose, β-maltose, sucralose) with comparable intensity, whereas there exists some indication that the 2-OH group of glucose reverts to the acceptor in solution (cf. sucralose) [64,65].

The glucosyl-2-OH thus being capable of functioning – in the solid state – as the acceptor or donor part in intramolecular hydrogen bonding, both, the AH or the B part may, conceivably, be assigned to it. Support in favour of the glucosyl-2-OH being the AH portion of the AH-B-X tripartite system can be derived from the distribution of the electrostatic potential over the surface of sucrose as displayed in Fig. 3. There, unequivocally, the intensely violet glucosyl-2-OH area is the center of the highest positive charge density, and, accordingly, should exhibit a pronounced tendency to participate in hydrogen bonding as the donor, i.e. AH part. These rationalizations, by consequence, entail the B portion of the AH-B-X tripartite system to be allotted to the glucosyl-3-OH. This leads – taking into account the two forms of sucrose prevalent in aqueous solution [31,32] – to the assignments made in Fig. 6.

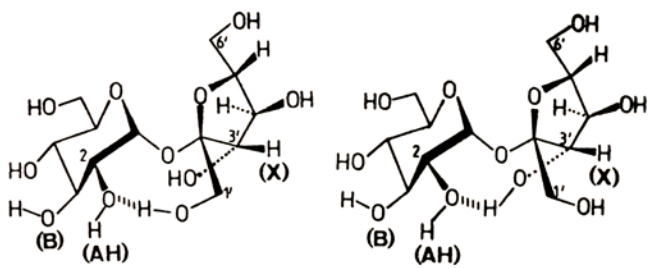


Fig. 6. Location of the AH-B-X glucophore ("sweetness triangle") in sucrose emerging from the electrostatic potential distribution (Fig. 3) and from the hydrophobicity potential profile (Fig. 4).

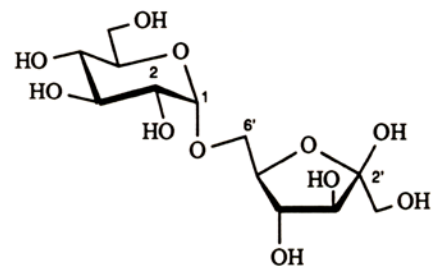
In these assignments, the notion is fostered that the intramolecular hydrogen bond, as drawn in (a) and (b), plays an integral role in eliciting the sweetness response such that it strongly enhances the H-donor (AH) capacity of the glucosyl-2-OH in its hydrogen-bonding interaction with a complementary acceptor group within the taste bud receptor. Indeed, all intensely sweet halo-sucroses listed in Tables 1 and 2 are capable of establishing the 2-O^g ··· HO-1^f (a) or the 2-O^g ··· HO-3^f (b) arrangement, which for some of these compounds (inclusively sucralose) has even been proved to prevail in dimethyl sulfoxide solution [65].

Additional support may be found, for example, in the observation that 3-ketosucrose is similar in sweetness to sucrose (retention of the hydrogen bond acceptor B-site) [66], whilst on steric intervention at the glucosyl-3-OH (e.g. by inversion to *allo*-sucrose) sweetness is lost altogether [66]. Continuing this line of thought, it would have to be surmised, that 3-O^g-methyl sucrose should be sweet, whilst the 2-O^g-methyl derivative, like the 2-deoxy- and the 2-keto-sucroses should be not – predictions that remain to be proved.

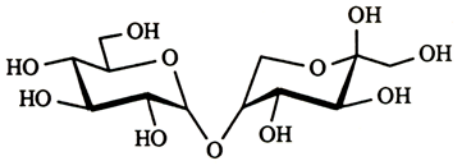
Whilst these rationalizations appear to be reasonable, especially the extension of the tripartite AH-B-X theory to comprise an intramolecular hydrogen bonding for enhancement of the H-donor capacity of the AH part – at least for sucrose derivatives – it is to be noted that the computation-based assignments of Fig. 6 differ substantially from those made on the basis of a huge number of structure-sweetness considerations [8,52,55]. At the present state of affairs, there can be no doubt, that further investigations in general, and along this computer modelling vein in particular, are needed – not only to probe into the validity of assignments in Fig. 6 but to prove the validity of the AH-B-X concept altogether, its major shortcoming presently being the absence of any predictive value. Nevertheless, the approach outlined here towards unravelling structure-sweetness relationships has high potential for providing new insights into the subtleties of sweetness and, hopefully, for leading to concepts of higher predictive value than those presently available.

6 3D-Representation and Hydrophobicity Potentials of Isomaltulose, Leucrose, Glucosyl-α(1→1)-mannitol, and Glucosyl-α(1→6)-sorbitol

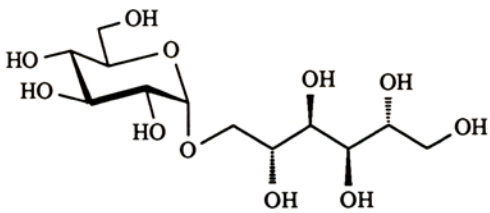
In view of the far-reaching implications inherent in the hydrophobicity potential profile of sucrose it was deemed appropriate to probe with this approach into some sucrose-related sugar substitutes, such as isomaltose (**21**), leucrose (**22**), and the terminally α-glucosylated mannitol (GPM, **23**) and sorbitol (GPS, **24**).



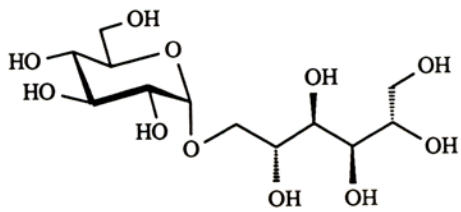
Isomaltulose (21)
D-Glucopyranosyl-α(1→6)-D-fructofuranose



Leucrose (22)
D-Glucopyranosyl -
α (1→5) -
D-fructopyranose



GPM (23)
D-Glucopyranosyl –
α(1→1) -
D-mannitol

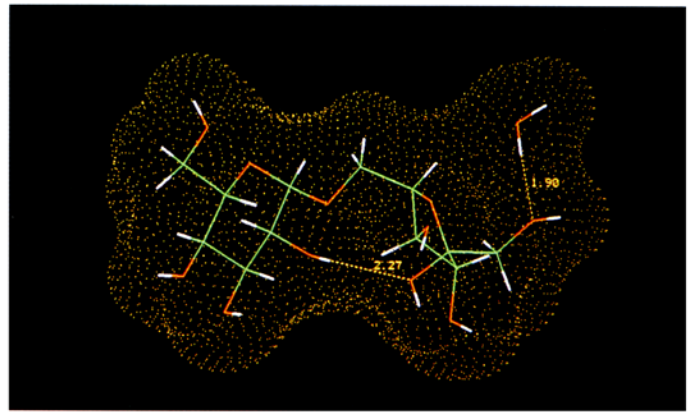


GPS (24)
D-Glucopyranosyl –
α(1→6) -
D-sorbitol

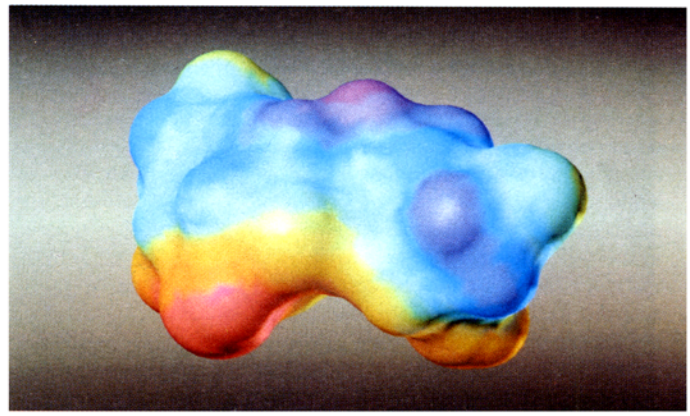
Due to their sweetening properties, all four compounds are accessible on an industrial scale [67,68], the latter two already being approved for use as a low-caloric, low-cariogenic sweetener. Since determinations of the crystal structures are available for each of the compounds [58,69–71], their 3D-geometrical characteristics in the representations used for sucrose, as well as their hydrophobicity profiles can be computer-generated with the X-ray coordinates as input.

In the case of isomaltulose (**21**), which crystallizes as a monohydrate, the X-ray structure [58] reveals (Fig. 7a) the glucose moiety in the $^4C_1$ conformation and the fructofuranose portion in a $^4T_3$ twist form – as is the case of sucrose. The glucosyl-2-OH engages in a comparatively weak (2.28 Å, cf. Table 3) intramolecular hydrogen bonding to the fructosyl-2-O, the glucosyl-2-OH functioning as the donor.

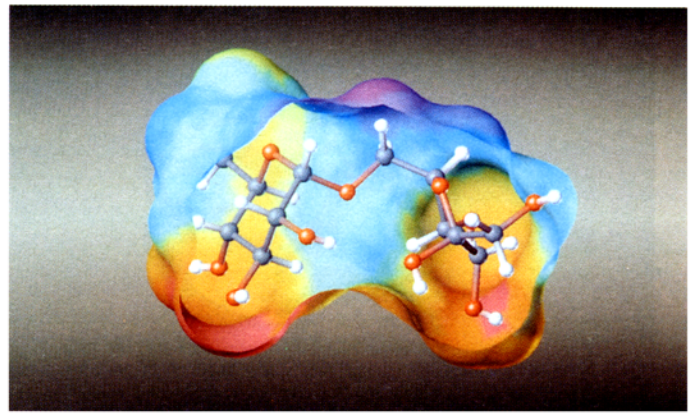
The molecular hydrophobicity potential of isomaltulose is presented in Figures 7b and 7c, wherein the water of crystallization has been removed to better account for the overall shape of the molecule before dissolution, i.e. before solvation with water molecules. Another reason for omitting of the crystal water in the hydrophobicity potential surface representations is to get a more feasible visualization of the hydrophobic interactions with a protein, this not only being pertinent for isomaltulose, but for GPM as well. The separation of hydrophobic and hydrophilic regions on the contact surface is clearly evident, yet their correlation



a)



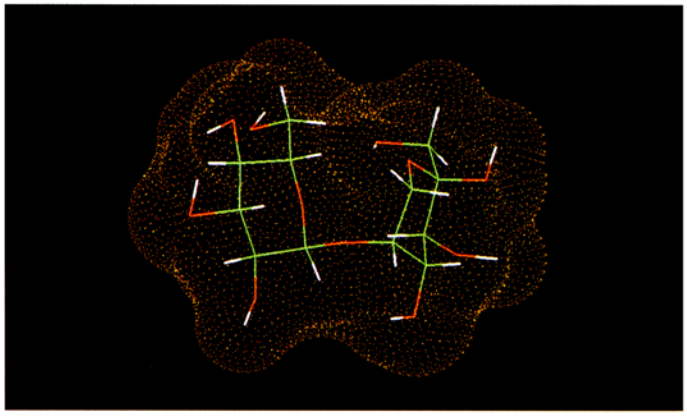
b)



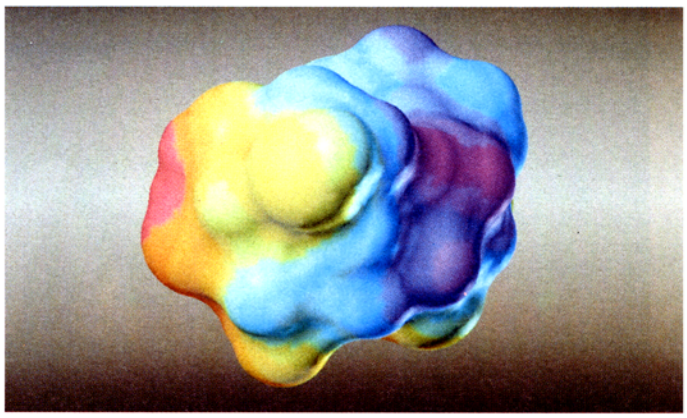
c)

Fig. 7. (a) Dotted contact surface of β-D-isomaltulose monohydrate (**21**·H_2O) with *Dreiding* type model as insert, based on X-ray structural data [58]. (b) Color-coded molecular lipophilicity potential profile (red for hydrophilic, blue for hydrophobic regions). (c) Hydrophobicity profile representation in open form with a ball-stick model inserted. In (b) and (c), the water of crystallization has been left off to simulate the molecule's shape before solvation.

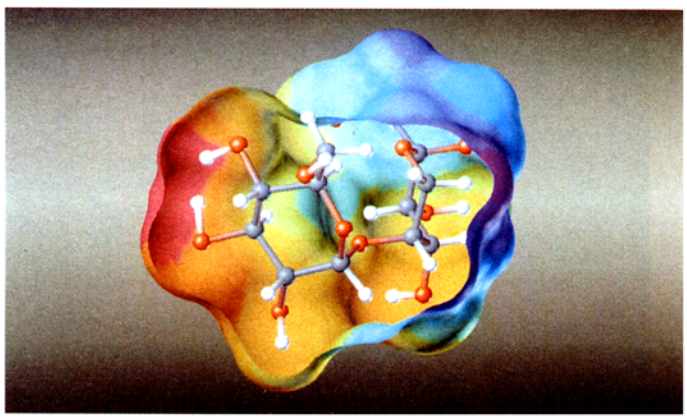
with the sweet response in the taste bud receptor – isomaltulose has about 40% of the sweetness of sucrose – must clearly await detailed investigations of the molecular shape in aqueous solution, i.e. whether the intramolecular hydrogen bond (cf. Fig. 7a) that gives the molecule a rigid overall conformation (cf. Fig. 7) is retained or not.



a)



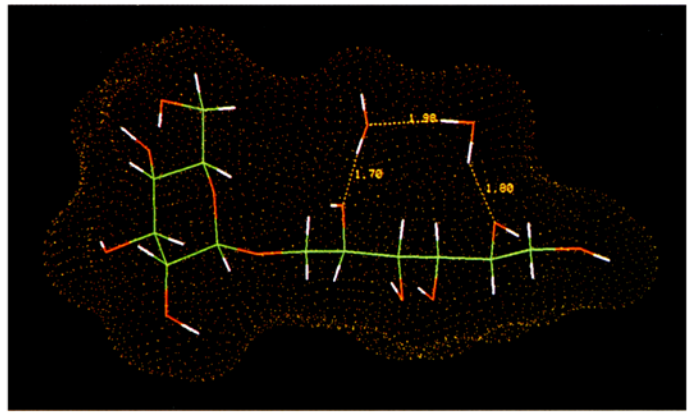
b)



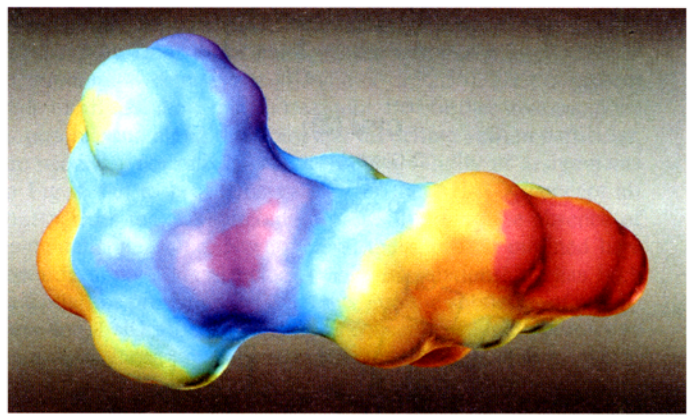
c)

Fig. 8. Leucrose (**22**): X-ray data-derived [69] solid state contact surface with stick model insert (a), and hydrophobicity potential profile (b, c).

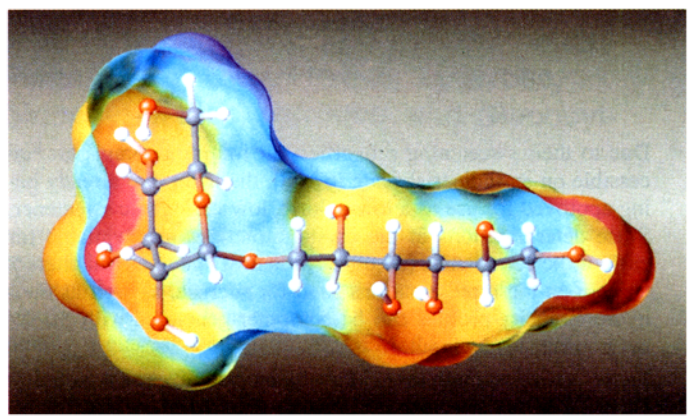
The α(1→5)-intersaccharidic linkage in leucrose (**22**) entails pyranoid forms for both monosaccharide units as evidenced in its X-ray structure [69] by the $^4C_1$ conformation of glucose and a $^2C_5$ chair for the fructose portion (Fig. 8a). Unlike isomaltulose, leucrose, in its crystalline β-anomeric form, does not develop an intramolecular hydrogen bond. In the hydrophobicity potential profile (Fig. 8b and c) a distinct distribution of hydrophobic and hydrophilic regions is observed, whose significance again remains to be interpreted.



a)



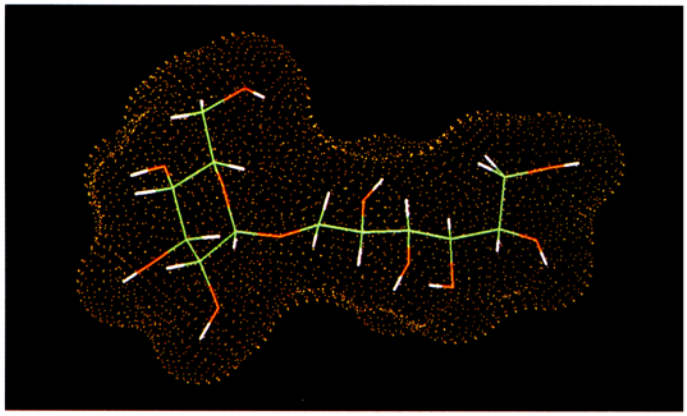
b)



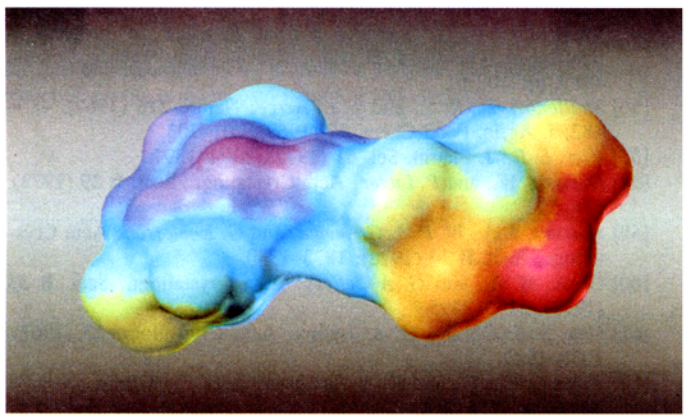
c)

Fig. 9. Glucopyranosyl-α(1→1)-mannitol (GPM) dihydrate (**23**·2H_2O): (a) dotted contact surface for solid state with stick model insert as developed from X-ray structural data [70]; the double water bridge spanning O-2 and O-5 of the mannitol portion is accentuated. (b) and (c) molecular lipophilicity potential profile in color-code (violet for most hydrophobic region), leaving off the water of crystallization to extricate the basic molecular geometry.

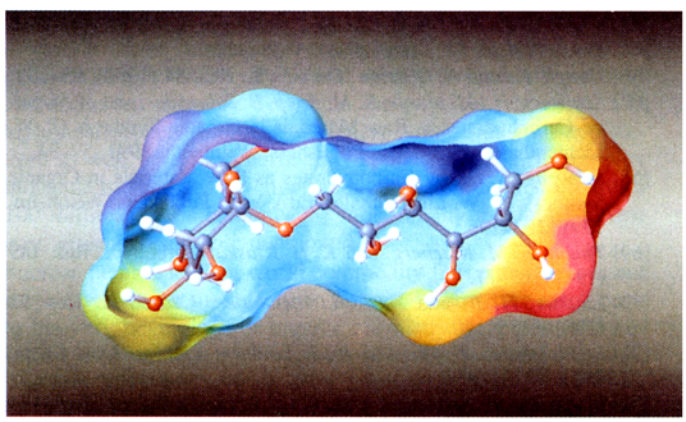
The molecular geometry of GPM dihydrate (**23**·2 H_2O) as revealed by X-ray structural analysis [70], is depicted in Fig. 9a. This representation nicely reveals the mannitol chain in a nearly planar zigzag conformation and the two molecules of crystal water to be fixed via hydrogen bonds establishing a rather unusual "double water bridge" between two oxygen atoms that



a)



b)



c)

Fig. 10. Glucopyranosyl-α(1→6)-sorbitol (GPS, **24**): (a) X-ray structure [71] derived contact surface and molecular geometry for the solid state, the stick model insert clearly showing the kink of the terminal hydroxymethyl group in the sorbitol chain. (b) and (c) color-coded hydrophobicity profile in an orientation different from that in (a) to account for full visualization of hydrophobic (violet) and hydrophilic (red) regions.

are five bonds apart. The corresponding 2-epimeric sorbitol analog, GPS (**24**), crystallizes anhydrous and exhibits in its X-ray structure [71] (Fig. 10a) a linear extended-chain backbone running from C-2 of glucose to the penultimate O-2, whereby the terminal sorbitol-CH_2OH is bent off to avoid unfavourable 1,3-interactions between O-2 and O-4 that would be operative in the extended-chain rotameric form, realized in **23**.

The hydrophobicity profiles for GPM and GPS in Figures 9 and 10 appear to be surprisingly similar if in the former the crystal water is left off for comparability, both then representing the molecular shape of the molecules before solvation (i.e. before dissolution in water). That, in fact, the side chain conformation in aqueous solution is very similar to that found in 5the solid state, is evidenced by the close correlation of the coupling constants along the alditol chains [72] with the respective X-ray-derived [70,71] dihedral angles. Nevertheless, an assessment of the lipophilicity potentials in terms of their biological significance has to await further studies.

7 Epilog

In the preceeding pages of this account, we have attempted to demonstrate that fairly advanced modelling modes for the molecular geometry of sucrose, for its contact surface, its electrostatic potential, and, most notably, for its hydrophobicity potential profile, provide unusually powerfurl tools for *building*, *visualizing*, and *analyzing* models of carbohydrates that are particularly suited to bring structure-activity (sweetness) relationships to an entirely new level. Thereby, the individual 3D-representations given are not limited to expose new insights only, but to stimulate new lines of thought *and* new experiments to prove them. Irrespective of whether these ideas and experiments are particularly ingenious, of aesthetic attraction, and/or practicable (or not), they are apt to be a stimulation for the progress in the chemistry and biochemistry of sucrose, not the least in its use as an organic raw material.

Acknowledgments

This work was supported by the Bundesministerium für Forschung und Technik. We are particularly grateful to Dipl.-Ing. *M. Waldherr-Teschner* and Dipl.-Inform. *M. Knoblauch* for their kind assistance in securing the graphic displays, to Prof. Dr. *J. Brickmann* for using his MOLCAD program and the Silicon Graphics workstation, and to Prof. *H. J. Lindner* for lucid discussions on the subject.

Bibliography

[1] This account is a modified version of an article contained in the monograph "Carbohydrates as Organic Raw Materials", VCH Publishers, Weinheim, 1991, developed from a Workshop Conference held at the Technische Hochschule Darmstadt, April 11–12, 1990.
[2] *Prout, W.*: Phil. Trans. **1** (1827), 355–388.
[3] *Liebig, J.*: Poggendorffs Ann. Phys. Chem. **31** (1834), 321–360.
[4] *Peligot, E.*: J. Prakt. Chem. **15** (1838), 65–113.
[5] *Berzelius, J.*: Poggendorffs Ann. Phys. Chem. **47** (1839), 289–322.
[6] *Dubrunfaut, A.-P.*: Compt. Rend. Acad. Sci. **29** (1849), 51–55; **42** (1856), 901–905.
[7] *Levi, I.*, and *C. B. Purves*: Adv. Carbohydr. Chem. **4** (1949), 1–35.
[8] *James, C. E., L. Hough*, and *R. Khan*: Prog. Chem. Org. Nat. Prod. **55** (1989), 117–184.
[9] *Tollens, B.*: Ber. Dtsch. Chem. Ges. **16** (1883), 921–924; formula: p. 923.
[10] *Fischer, E.*: Ber. Dtsch. Chem. Ges. **26** (1893), 2400–2412; formula: p. 2405.
[11] *Tollens, B.*: Kurzes Handbuch der Kohlenhydrate. Verlag J. A. Barth, Leipzig 1914, p. 363.
[12] *Haworth, W. N.*, and *J. Law*: J. Chem. Soc. **109** (1916), 1314–1325; formula p. 1319.

[13] *Hudson, C. S.*: J. Am. Chem. Soc. **38** (1916), 1566–1575; formula p. 1567.
[14] *Fischer, E.*: Ber. Dt. Chem. Ges. **24** (1891), 2683–2687.
[15] a) *Charlton, W., W. N. Haworth,* and *S. Peat*: J. Chem. Soc. **1926**, 89–101; formula p. 99. b) *Haworth, W. N.,* and *E. L. Hirst*: J. Chem Soc. **1926**, 1858–1868; formula p. 1864.
[16] *Haworth, W. N.*: The Constitution of Sugars. Arnold and Co., London 1929, pp. 70–71.
[17] *Pigman, W. W.*: Chemistry of the Carbohydrates. Acad. Press, New York 1948, p. 446.
[18] Ref. [6], p. 25.
[19] *Morrison, R. T.,* and *R. N. Boyd*: Organic Chemistry. Allyn and Bacon, Inc., Boston 1959, p. 789.
[20] *Hough, L.*: Presentation at the 4th European Carbohydr. Symp., Darmstadt 1987; cf. Nachr. Chem. Tech. Lab. **38** (1990), 860.
[21] *Collins, P. M.*: Carbohydrates. Chapman and Hall, London 1987, frontispiece and formula p.466.
[22] Chemist's Personal Software Series. ChemText 1.3. Molecular Design Ltd., San Leandro, Cal., USA; Labo **12** (1988), 32–36.
[23] *Brown, G. M.,* and *H. A. Levy*: Science **141** (1963), 921–923.
[24] *Brown, G. M.,* and *H. A. Levy*: Acta Crystallogr., Sect. **B 29** (1973), 790–797.
[25] *Hanson, J. C., L. C. Sieker,* and *L. H. Jensen*: Acta Crystallogr., Sect. **B 29** (1973), 797–808.
[26] *Lichtenthaler, F. W.*: Dtsch. Zahnärztl. Zeitschr. **37** (1982), S46–S49; formula p. S48.
[27] *Hough, L.*: Chem. Soc. Rev. **14** (1985), 357–374.
[28] *Bock, K.,* and *R. U. Lemieux*: Carbohydr. Res. **100** (1982), 63–74.
[29] *McCain, D. C.,* and *J. L. Markley*: Carbohydr. Res. **152** (1986), 73–80.
[30] *McCain, D. C.,* and *J. L. Markley*: J. Am. Chem. Soc. **108** (1986), 4259–4264.
[31] *Christofides, C.,* and *D. B. Davies*: J. Chem. Soc., Chem. Commun. **1985**, 1533–1534.
[32] *Davies, D. B.,* and *J. C. Christofides*: Carbohydr. Res. **163** (1987), 269–274.
[33] a) *Tran, V. H.,* and *W. J. Brady*: Biopolymers **29** (1990), 961–976. b) *Tran, V. H.,* and *W. J. Brady*: Biopolymers **29** (1990), 977–997. c) *Tran, V. H.,* and *W. J. Brady*: in: Computer Modeling of Carbohydrate Molecules (A. D. French, J. W. Brady, Eds.). ACS Symposium Series # 40, Am. Chem. Soc., Washington DC, 1990. pp. 213–226.
[34] *Johnson, C. K.*: ORTEP II: A Fortran Thermal-Ellipsoid Plot Program for Crystal Structure Illustration. Oak Ridge National Laboratory, Report ORNL-3794, Oak Ridge, Tenn., 1965.
[35] For a Pertinent Review, see: *Cohen, N. C., J. M. Blaney, C. Humblet, P. Gund,* and *D. M. Barry*: J. Med. Chem. **33** (1990), 883–894.
[36] *Richards, W. G.*: Computer-Aided Molecular Design. IBC Technical Servies Ltd., London 1989.
[37] *French, A. D.,* and *J. W. Brady* (Eds.): Computer Modeling of Carbohydrate Molecules. ACS Symposium Series # 430, Am. Chem. Soc., Washington DC, 1990.
[38] *Lipkowitz, K. B.,* and *D. B. Boyd*: Reviews in Computational Chemistry. VCH Verlagsgesellschaft, Weinheim 1990.
[39] *Brickmann, J.,* and *M. Waldherr-Teschner*: Labo **10** (1989), 7–14.
[40] Brickmann, J.: Institut für Physikalische Chemie, Technische Hochschule Darmstadt, Petersenstr. 20, D-6100 Darmstadt, Germany.
[41] *Richards, F. M.*: Ann. Rev. Biophys. Bioeng. **6** (1977), 151–176.
[42] *Connolly, M. L.*: J. Appl. Cryst. **16** (1983), 548–558.
[43] *Connolly, M. L.*: Science **221** (1983), 709–713.
[44] *Lee, B.,* and *F. M. Richards*: J. Mol. Biol. **55** (1971), 379–400.
[45] *Weiner, P. K., R. Langridge, J. M. Blaney, R. Schaefer,* and *P. A. Kollman*: Proc. Natl. Acad. Sci. USA **79** (1982), 3754–3758.
[46] *Stewart, J.-J. P.*: MOPAC: A Semiempirical Molecular Orbital Program. Quantum Chem. Prog. Exch., Program No. 455 (1983).
[47] *Lichtenthaler, F. W.,* and *S. Rönninger*: J. Chem. Soc., Perkin Trans. 2, **1990**, 1489–1497.
[48] *Furet, P., A. Sele,* and *N. C. Cohen*: J. Mol. Graphics **6** (1988), 182–189.
[49] *Fauchère, J.-L., P. Quarendon,* and *L. Kaetterer*: J. Mol. Graphics **6** (1988), 202–206.
[50] *Deutsch, E. W.* and *C. Hansch*: Nature **211** (1966), 75.
[51] *Lemieux, R. U.*: Frontiers Chem., Plenary Keynote Lect. 28th IUPAC Congr., 1981 (Laidler, K. J., Ed.), Pergamon, Oxford, U.K., 1982, pp. 3–24.
[52] *Lee, C.-K.*: Adv. Carbohydr. Chem. Biochem. **45** (1987), 199–351; pp. 223 ff., in particular.
[53] *Tanford, C.*: The Hydrophobic Effect. Wiley, New York 1973.
[54] a) *Ghose, A. K.,* and *G. M. Grippen*: J. Comput. Chem. **7** (1986), 565–577. b) *Ghose, A. K., A. Pritchett,* and *G. M. Grippen*: J. Comput. Chem. **9** (1988), 80–90.
[55] a) *Hough, L.,* and *R. Khan*: Trends Biol. Sci. **3** (1978), 61–63. b) *Jackson, G., M. R. Jenner, R. A. Khan, C. K. Lee, K. S. Mufti, G. D. Patel,* and *E. B. Rathbone* (Tate & Lyle PLC): Brit. Pat. 2,088,855 (1982); Eur. Pat. Appl. EP 73,093 (1983); Chem. Abstr. **99** (1983) 54127j. c) *Hough, L.,* and *R. Khan*: in: Progress in Sweeteners (Grenby, T. H., Ed.), Elsevier Appl. Science, London 1989, pp. 102 ff.
[56] *Shallenberger, R. S.,* and *T. E. Acree*: Nature **216** (1967), 480–482; J. Agric. Food Chem. **17** (1969), 701–703.
[57] *Kier, L. B.*: J. Pharm. Sci. **61** (1972), 1394–1397.
[58] *Dreissig, W.,* and *P. Luger*: Acta Crystallogr., Sect. **B 29** (1973), 514–521.
[59] *Kanters, J. A., W. P. J. Gaykema,* and *G. Roelofson*: Acta Crystallogr., Sect. **B 34** (1978), 1873–1880.
[60] *Gress, M. E.,* and *G. A. Jeffrey*: Acta Crystallogr., Sect. **B 33** (1977), 2490–2495.
[61] *Chu, S. S. C.,* and *G. A. Jeffrey*: Acta Crystallogr. **23** (1967), 1038–1049.
[62] *Takusagawa, F.,* and *R. A. Jacobsen*: Acta Crystallogr., Sect. **B 34** (1978), 213–218.
[63] *Tanaka, I., N. Tanaka, T. Ashida,* and *M. Kakudo*: Acta Crystallogr., Sect. **B 32** (1976), 155–160.
[64] *Kanters, J. A., R. L. Scherrenberg, B. R. Leeflang, J. Kroon,* and *M. Mathlouthi*: Carbohydr. Res. **180** (1988), 175–182.
[65] *Christofides, J. C., D. B. Davies, J. A. Martin,* and *E. B. Rathbone*: J. Am. Chem. Soc. **108** (1986), 5738–5743.
[66] *Hough, L.,* and *E. O'Brien*: Carbohydr. Res. **84** (1980), 95–102.
[67] *Schiweck, H., M. Munir, K. M. Rapp, B. Schneider,* and *M. Vogel*: in: Carbohydrates as Raw Materials in Organic Synthesis (F. W. Lichtenthaler, Ed.). VCH Publishers, Weinheim, 1991, 57–94.
[68] *Schwengers, D.*: in: Carbohydrates as Raw Materials in Organic Synthesis (F. W. Lichtenthaler, Ed.). VCH Publishers, Weinheim, 1991, 183–196.
[69] *Thiem, J., M. Kleeberg,* and *K. H. Klaska*: Carbohydr. Res. **189** (1989), 65–77.
[70] *Lindner, H. J.,* and *F. W. Lichtenthaler*: Carbohydr. Res. **93** (1981), 135–140.
[71] *Lichtenthaler, F. W.,* and *H. J. Lindner*: Liebigs Ann. Chem. **1981**, 2372–2383.
[72] *Munir, M., B. Schneider,* and *H. Schiweck*: Carbohydr. Res. **164** (1987), 477–485.

Address of authors: Prof. Dr. *Frieder W. Lichtenthaler*, Dipl.-Ing. *Stefan Immel*, and Dipl.-Ing. *Uwe Kreis*, Institut für Organische Chemie, Technische Hochschule Darmstadt, D-6100 Darmstadt (Germany), Petersenstraße 22, Tel. 0 61 51/16 23 76.

(Received: December 18, 1990).