Polish J. Chem., **75**, 1569–1582 (2001)

QSAR Computational Model of Nofre-Tinti Theory on Sweetness of Mono- and Disaccharides Composed by Pyranose Units

by W. Pietrzycki

Department of Chemistry, University of Agriculture, 31-120 Kraków, Mickiewicz Avenue 21, Poland

(Received February 13th, 2001; revised manuscript June 12th, 2001)

PM3 quantum chemical population analysis in the ground state was performed for about 32 anomeric ring-conformers of aldopyranoses and ketopyranoses, maltose and lactose, as well as lysine cation – sweet taste receptor moiety. Thermodynamic equation was formulated for RS_j relative sweetness, originated from LFER (linear free energy relationship) for sugar glycophore – sweet taste receptor interactions. From this equation, QSAR (Quantitative Structure – Activity Relationship) studies are carried out on a sweetness of 10 natural, most known and important mono- and disaccharides, composed by pyranose units. Using RS_j measured values, QSAR correlation equations were performed, which apply exclusively PM3 calculated quantum chemical parameters. The ln RS_j^{expl} demonstrates linear correlation against Q_j(O) oxygen as well as Q_j(H) hydrogen atomic electron net charges of 4-OH group in aldopyranoses and 2-OH group in ketopyranoses. These O and H atoms are B_1 and XH_1 subsites, respectively, which strongly predominate in {XH_2, XH_1, B_1, B_2, AH_1, AH_2} sweetener of *j*-th aldopyranose in Nofre-Tinti theory. On the other hand, L-sorbose represents B_1, AH_1, B_2 system, where B_2 subsite is *n*-electron donor situated on the ketopyranose O-1 oxygen atom. The β-D-fructopyranose sweetener provides additionally E_3 subsite (ketopyranose O-3 atoms). The NHOMO(pyr) → LUMO(recp) transition in this sugar reveals E_1(O-5) and E_2(O-4) subsites.

Key words: QSAR on sweetness of mono- and disaccharides, Nofre-Tinti theory of the pyranose sweetness, sweetness under thermodynamic control, XH_1, B_1 sweetener in the pyranoses

Sensory impressions became long ago a subject of medical investigations. Sensory physiology and "chemical senses" [1] opened a door to taste and olfactory chemistry. Sensory chemistry, especially the taste and smell chemistry, become a popular field of food chemistry. To-day the taste chemistry with its newest biochemical and molecular biology background is one of the most modern and attractive domain of bioorganic chemistry.

It is known that pairs of the functional groups, such as hydroxyl groups, aminogroups and ether oxygen are usually present in the sweet tasting compound. They were called "glycophores" [2,3,4]. The sweet taste-eliciting group for the sugars was a glycol (-CHOH-CHOH-) unit. According to Shallenberger, sweetener site pair of glycophore was marked by AH and B [2,3,4]. Birch and Lee have accepted [5] that AH is formed by H atom of 4-OH hydroxyl group, whereas B is O-3 atom of 3-OH group in aldopyranoses. On the other hand, ketopyranoses have another situated AH, B sweeteners [6,7]. They place AH on hydrogen of 2-OH group and B on O-1 atom.

The sweet taste receptors have analogous $\underline{AH}^+$, $\underline{B}^-$ dipoles. They interact with AH, B glycophore and form two hydrogen bonds. AH as well as B sweetener sites have their atomic net charges, so AH, B glycophore may be marked in $AH^{\delta(+)}$, $B^{\delta(-)}$ form. Hydrogen bonding, the form of glycophore---receptor intermolecular interaction is responsible for the sweetness of sugars. In other words, sweetness is not only caused by chemicals that possess a pair of complementary functional group. Sweetness is a concerted chemical reaction:

$$AH^{\delta(+)}, B^{\delta(-)} + \underline{AH}^+, \underline{B}^- \rightarrow \underline{AH}^+\text{---}B^{\delta(-)} + AH^{\delta(+)}\text{---}\underline{B}^- \quad (1)$$

Glycophore Receptor

This can be translated into thermodynamic dependences as well as QSAR correlations. Blanksma and Hoegen [8] have formulated the first QSAR correlation equation of the *RS* relative sweetness (in relation to sucrose) for nine 2-X- substituted 5-nitroanilines,

$$\log RS = 1.610\,\pi - 1.831\,\sigma + 1.729 \quad (2)$$

The σ is Hammett constant and π means the hydrophobic constant of the substituent. If one assumes that *K* is the equilibrium constant for the reaction of 2-X-5-nitroaniline---receptor complex formation, K_0 for 5-nitroaniline---receptor (X = H) complex formation and K_{suc} for sucrose---receptor formation, we can put in above equation $RS = K/K_{suc}$ and $1.729 = \log K_0/K_{suc}$. This equation takes an analogous form to the two-parameter Taft equation: $\log(K/K_0) = \rho_R\sigma_R + \rho_I\sigma_I$ [9]. Hence, a conception may arise that the *RS* should be described by K/K_{suc}.

A great progress of the sweet taste theory rised up after 1990. It was followed by the turning point in development of biochemistry of receptors [10,11] and molecular biology with cloning receptor DNAs techniques [10]. Rodbell and Gilman have won the Nobel Price in 1994 for the discovery of *G*-coupled proteins, which fulfil an important part in the signal transduction path. Biochemistry of the cyclic AMP-mediated transduction mechanism for sweet taste is presented in [12]. Structure of sweet taste receptor is considered as similar to the structure of other *G*-protein receptors. It shows a polipeptide chain, distinguished by seven transmembrane domain segments, TM I – TM VII helices, forming a pocket in which the sweet ligands are binded.

Under this biochemistry progress, Nofre and Tinti [13] have formulated the Multipoint Attachment Theory (MPA), which may explain the binding of sweet ligands with the receptor in a transmembrane pocket. Glycophore of the sweet compound is a sweetener, including **eight** interaction sites, described by B, AH, XH, G1, G2, G3, G4 and D. Every site, except D, owns two subsites. The full sweetener of sweet substance is shown in Fig. 1. Subsites B_1 and B_2 (such as COO^-, SO_3^- anionic group, oxygen nitrogen and halogen atoms) have negative atomic net charge, so they interact by Coulombic forces. On the other hand, they may be H-bond acceptors or

n-electron donors. Subsites AH_1, AH_2, XH_1 and XH_2 (such as hydrogen atoms in N–H or O–H groups) exhibit positive atomic net charges and they may interact by Coulombic forces, being also H-bond donors or acceptors of *n*-electron pair. The G-sites occur as two subsites, E_n and G_n (n = 1, 2, 3, 4). Subsites E_n are essentially *n*-electron donors (such as a nitrogen, oxygen or halogen atom). In the turn, G_n subsites represent often non-planar and weakly polar small group of atoms, such as CH_3, CH_2 and CH group or fluorine atom. These subsites are assumed to act with receptor recognition sites by atomic intermolecular steric interactions (dispersion, van der Waals forces). Finally, site D, which is often 4-cyano-phenyl group, is considered to act essentially with the receptor *via* its H-bond acceptor group (CN). Receptor TM-pocket is composed by eight aminoacid residues, which form receptor dipoles and recognition sites corresponding to sweetener sites. Lysine residue binds B site of sweetener, aspartic acid or glutaminic acid residues bind AH and XH sites. In the turn, threonine residue interacts with the E_n and G_n subsites as well as with D site of sweet-

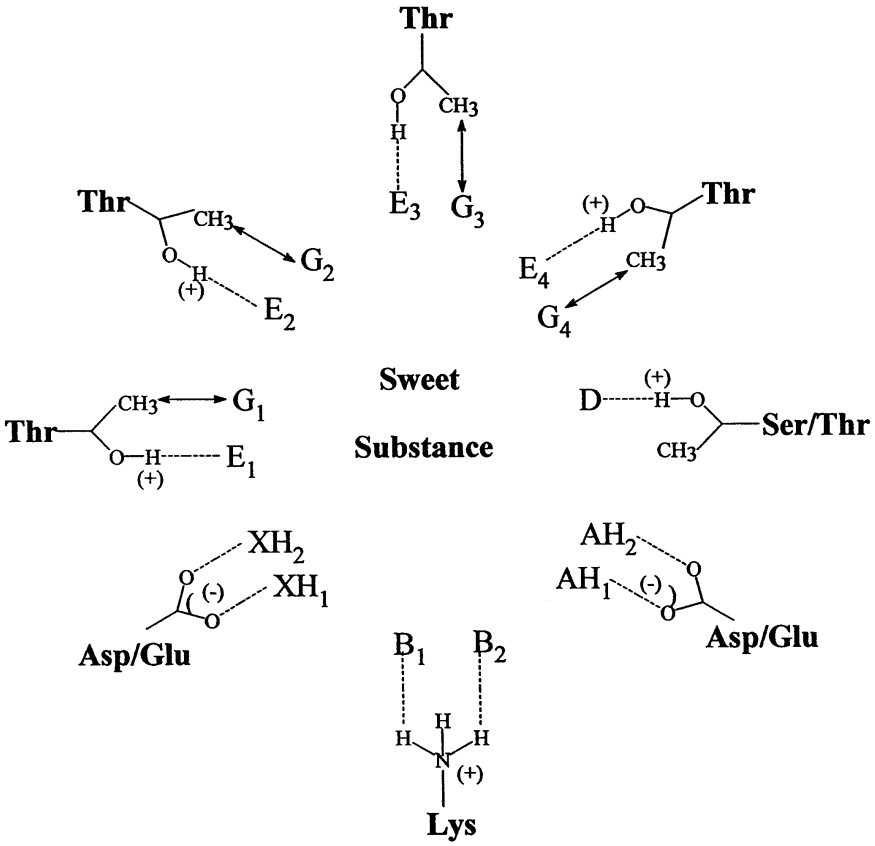

Figure 1. Activated state of sweet taste receptor pocket. Molecule of sweet substance, possessing all 15 sweetener subsites is binded by the corresponding 15 recognition points, aminoacid residues of sweet taste receptor.

ener. Probably seven from these eight aminoacid recognition sites are linked to different TM in receptor pocket. The schematic structure of sweet substance---receptor complex is presented in Fig. 1. From the mentioned 15 sweetener subsites, the **{XH_2, XH_1, B_1, B_2, AH_1, AH_2}** six subsite system is attributed to aldopyranoses [13]. On the other hand, Shallenberger theory [2–7] have shown that only two subsites, XH_1 and B_2 are active forming AH,B system, also in many non-sugar compounds [14–17]. In further theories, B, AH, X – three point sweetener theory of Kier [6,18] proves that only three subsites are active in the eight-site Nofre-Tinti sweetener. Both these theories seem to be particular cases of the general Nofre-Tinti theory. However, all three theories have a qualitative character and they are not suitable for the calculations of *RS* relative sweetness.

In this paper, the computational determination of the active subsites in {XH_2, XH_1, B_1, B_2, AH_1, AH_2} Nofre-Tinti sweeteners is carried out for the mono- and disaccharides, formed by pyranose units. Atypical QSAR correlation equations are formed, which apply theoretical, calculated quantum parameters in "Quantitative Structure" field, instead of experimental physicochemical quantities. Our theory of these equations originated from the far simplifications of intermolecular interaction theory and the thermodynamic suggestions mentioned. QSAR equations are suitable in the calculation of *RS* relative sweetness for an arbitrary pyranose and its some derivatives. They may be a useful tool for the sweetness technological forecasting of unknown sugars.

RESULTS AND DISCUSSION

Calculation of semiempirical quantum parameters: 32 anomeric ring-conformers of 10 most known natural sugars in bioorganic and food chemistry are taken into consideration. We use molecular editor and self-consistent procedures of HyperChem-5,0 program to estimate proper geometry of the molecules. Two geometric ring-conformers: *C*1 and 1*C* are considered. The *C*1 owns ring-oxygen atom, convex upwards, in the same direction as $-CH_2OH$ (see Fig. 2, glucoses), whereas the 1*C* ring-conformer, owns ring-oxygen atom as convex downwards (see Fig. 2, mannose).

Optimization of sugar geometry is carried out to find the lowest total molecular energy (or greatest heat of formation) as well as the greatest atomic absolute net charges of atoms in OH groups. The latter condition is important, because receptor dipoles join themselves with sugar glycophore by hydrogen bonds. Every anomeric ring-conformer possesses usually several OH conformers with different local minima of total molecular energy. Fortunately, the "*Add H & Model Build*" function under "*Build*" menu in HyperChem program excellently chooses the OH conformer with greatest atomic absolute net charges and low total energy for majority of aldopyranoses-*C*1. Then, MM+ molecular mechanics self-consistent procedure [19,20] with Fletcher-Rives convergence is used and PM3 quantum procedure applying Polak-Ribiere convergence in Hyperchem 5.0 standard [22]. If "*Add H & Model Build*" function does not work properly, turning molecule into another anomeric

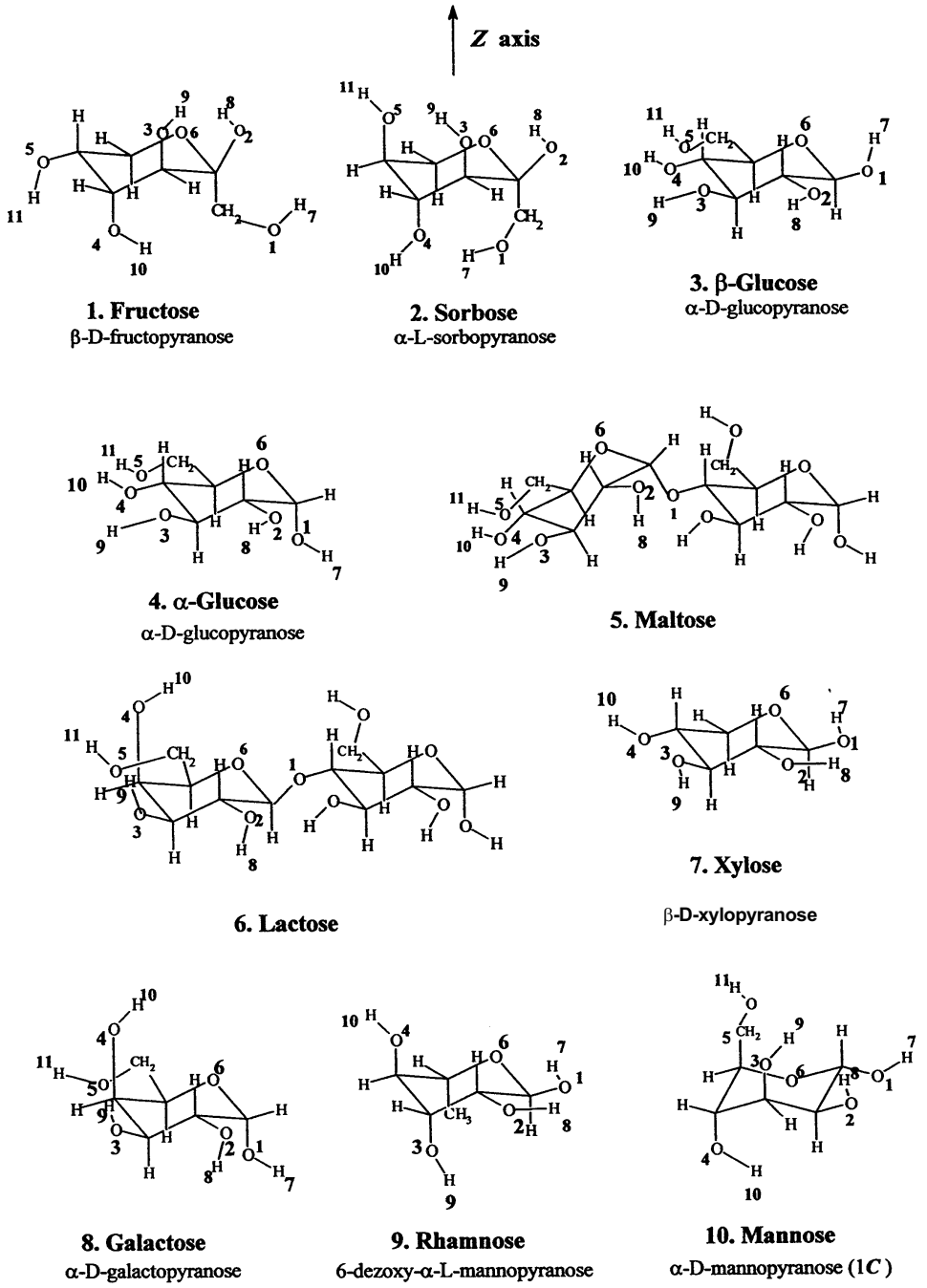

Figure 2. Structures of 10 basic sugars as a result of PM3 self-consistent geometry. Arbitrary continuous numeration of oxygen and hydroxylic pyranose atoms. Numeration of oxygen atoms converges with chemical numeration with exception for aldopyranoses: 6-th ring oxygen atom is 5-th in chemical numeration and 5-th CH_2OH oxygen atom is 6-th in chemical nomenclature.

ring-conformers (ketopyranoses -*C*1), one has to use first "*Add hydrogen*" function from "*Build*" menu, accompanied by twisting C-O-H handly around C—O bond and then MM+. 10 molecular structures are chosen from 32 anomeric ring-conformers using a criterion of the lowest total energy of molecule. They represent the mentioned 10 natural sugars and their structural formulas are shown in Fig. 2. PM3 population analysis was carried out for the self-consistent geometries of the all mentioned above molecules. It yields $Q_j(i)$ atomic net charges and $\varepsilon_k(j)$ molecular energies (in eV units) for every *j*-th sugar. Additionally, space diagrams of HOMO and NHOMO of every sugar are attentive analysed.

Basic sugars. Single regression analysis of sweetness: Measured relative sweetnesses of pyranoses, RS_j^{expl}, are taking into consideration. Pyranose chair-ring structure may be formed, taking into account D- and L- configurations, by 16 aldohexoses, 8 ketohexoses and 8 aldopentoses. Unfortunately, limited number of them provide valid values of experimental RS_j. In this paper, RS_j^{expl} data for 8 monosaccharides and 2 disaccharides are taken into consideration to calibrate *a*-coefficients in QSAR equations in the next sections. Sugar full names, abbreviations of these names and RS_j^{expl} relative sweetnesses are presented in Table 1. Structural forms of the above listed sugars are presented in Figure 2.

Table 1. Abbreviations of the full names of the saccharides submitted to QSAR correlation equations.

	Full name	Abbreviation	Mol. weight	RS_j^{expl} (%)
1.)	β-D-fructopyranose-*C*1	fructose	180.162	1.72
2.)	α-L-sorbopyranose-*C*1	sorbose	180.162	0.86
3.)	β-D-glucopyranose-*C*1	β-glucose	180.162	0.80
4.)	α-D-glucopyranose-*C*1	α-glucose	180.162	0.70
5.)	4-O-(α-D-glucopyranosyl)-α-D-glucopyranose (*C*1,*C*1)	maltose	342.308	0.33
6.)	4-O-(β-galactopyranosyl)-α-D-glucopyranose (*C*1,*C*1)	lactose	342.308	0.20
7.)	β-D-xylopyranose-*C*1	xylose	150.135	0.40
8.)	α-D-galactopyranose-*C*1	galactose	180.162	0.32
9.)	6-deoxy-α-L-mannopyranose-*C*1	rhamnose	164.162	0.33
10.)	α-D-mannopyranose-1*C*	mannose	180.162	0.30

The RS_j^{expl}(%) value usually expresses a sweetness of 10% solution of *j*-th sugar measured in respect to the sweetness of 10% sucrose solution. Such a method is very suitable when molecular structure and molecular weight of given sugar is unknown. On the other hand, the concentrations of solutions in the molecular theories are frequently expressed by mole/dcm^3 unit. Standard 10% solution of sucrose is about 0.3 molar. We assume that $RS_j(c)$ values are measured in the set of equimolar solutions. Recalculation of RS_j(%) values into $RS_j(c)$ is carried out in this paper according to:

$$RS_j(c) = \frac{d_0 M_j}{d_j M_0} RS_j(\%) \tag{3}$$

M_j is molecular weight of j-th sugar, whereas M_0 molecular weight of standard sugar = sucrose. Additionally, one accepts the simplification that solutions of different sugars under the same per cent concentrations have approximately the same densities ($d_j = d_0$). $RS_j^{\text{expl}}(\%)$ experimental values of fructose, maltose, xylose, galactose and rhamnose are taken as proportional to biochemical data of Filipowicz and Więckowski [21], in order to keep $RS_{sucrose}^{expl} = 1.00$ condition. RS^{expl} data for sorbose is taken from Tsuzuki and Yamazaki [22] using the similar procedure. Valuable data for β-glucose, α-glucose, mannose and lactose originate from Tomasik [23]. Electronic structure of molecule is an origin of its physicochemical properties. Observation of the atomic net charges of sugars, $Q_j(i)$, leads to interesting conclusions. We take into account only the atoms, which have significant absolute electron net charges, because they may produce electrostatic bonding interactions. These are oxygen and hydroxylic hydrogen atoms in aldo- and ketopyranoses. System of numeration of these atoms, presented in Fig. 2, was mentioned above. Analogous numeration in disaccharides is carried out only in nonreducing aldopyranose rings in maltose and lactose. According to Birch, Cowell and Eyton [24] only **one** pyranose moiety of disaccharide (nonreducing) reacts with the receptor. The $Q_j(i)$ atomic net charges of i-th oxygen or hydrogen atom for every j-th pyranose are collected in Table 1. At first, we take into account all aldopyranoses together with disaccharides formed by aldopyranose units, $j = 3, 4, \ldots 10$. The $\ln RS_j^{\text{expl}}(c)$ values demonstrate unexpectedly satisfactory correlation ***versus*** Q_j(O-4), and Q_j(H-10) atomic net charges in 4-OH group for these sugars. $\boldsymbol{R_{3\text{-}10}(i)}$ regression coefficients are high, 0.98121 and 0.96445, respectively. Remainder variances S^2 are very low. Corresponding regression equations for a set of $N = 8$ sugars are given below:

$$\ln RS_j(c) = -20.67255\, Q_j(\text{O-4}) - 8.08995 \qquad \text{for } j = 3, 4, 5, \ldots 10. \tag{4}$$

$$R_{3\text{-}10}(\text{O-4}) = 0.98121, \quad S^2_{3-10}(\text{O-4}) = 0.00739$$

$$\ln RS_j(c) = +23.79623\, Q_j(\text{H-10}) - 6.49491 \qquad \text{for } j = 3, 4, 5, \ldots 10. \tag{5}$$

$$R_{3\text{-}10}(\text{H-10}) = 0.96445, \quad S^2_{3-10}(\text{H-10}) = 0.01387$$

Fig. 3a presents the map of sweetener subsites in aldopyranose (D-glucose) according to Nofre and Tinti theory [13]. Both the latter equations show that O-4 and H-10 atoms of 4-OH group predominate as B_1 and XH_1 subsites in $\{XH_2, XH_1, B_1, B_2, AH_1, AH_2\}$ sweetener, (see Fig. 3b).

Ketopyranose map of subsites in $\{XH_2, XH_1, B_1, B_2, AH_1, AH_2\}$ sweetener of fructopyranose is unknown. Solely a map of $\{E_1, E_2, E_3, E_4\}$ fructofuranose subsweetener in sucrose was presented in [13]. However, the comparative analysis procedure of atomic net charge distribution in ketopyranoses and aldopyranoses leads in this paper to a such ketopyranose map (see Fig. 3c). It converges excellently with the AH,B

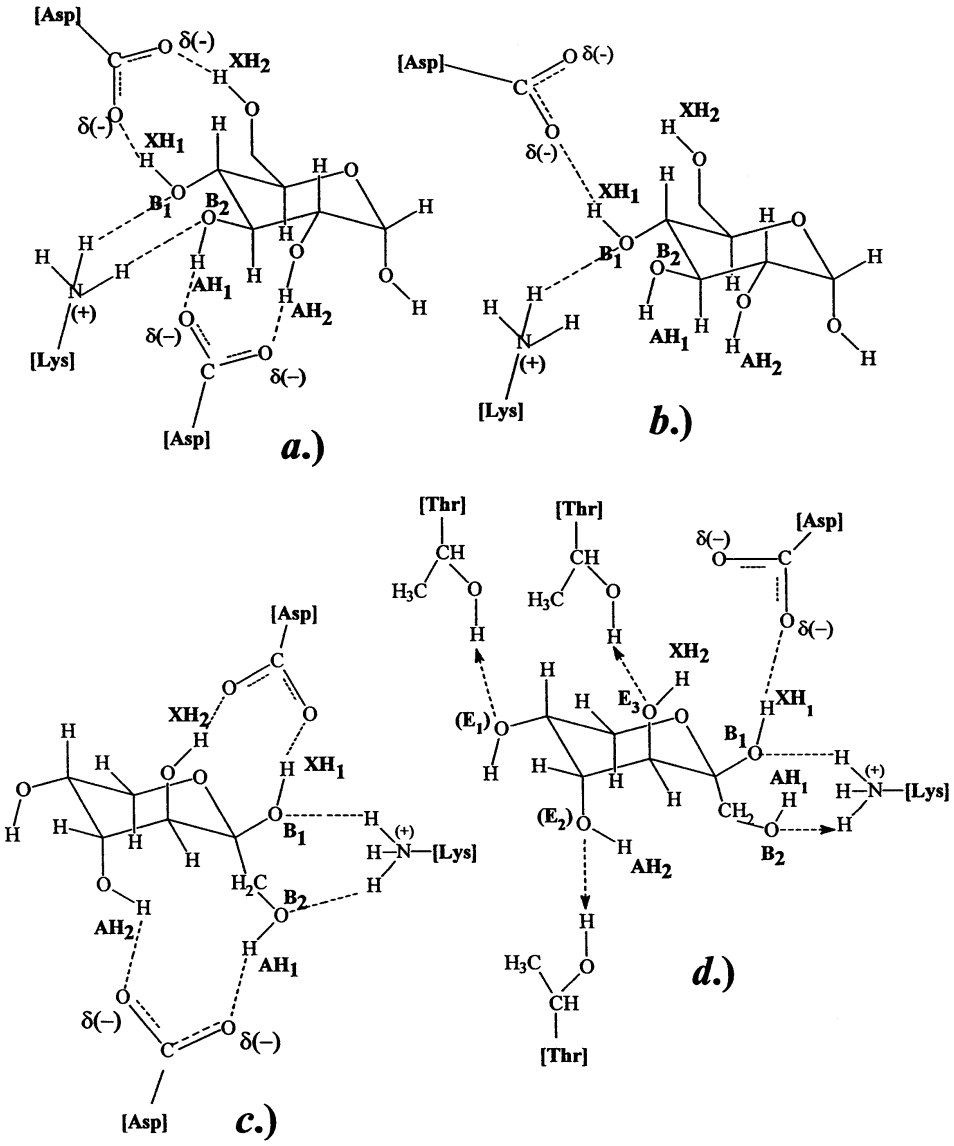

Figure 3. Structure of sugar---receptor complexes: (***a***) – Binding of {XH_2, XH_1, B_1, B_2, AH_1, AH_2} six subsite sweetener of α-D-glucopyraranose by receptor aminoacid residues according to Nofre-Tinti theory [13]. (***b***) – Calculated structure of α-D-glucopyranose---receptor complex with active sweetener subsites. **(c)** – Theoretical map of {XH_2, XH_1, B_1, B_2, AH_1, AH_2} six subsites in β-D-fructopyranose. (***d***) – Calculated structure of β-D-fructopyranose---receptor complex with active sweetener subsites.

Shallenberger glycophore for fructose [17]. The H-10 and O-3 atoms in aldopyranoses correspond to H-8 and O-1 atoms in ketopyranoses and they form XH_1,B_2 = AH,B glycophore. In Table 2 the ketopyranose atomic numbers are placed in the parentheses.

If we include sorbose and fructose atoms to (4) and (5), we receive $\boldsymbol{R}_{1\text{-}10} = 0.87763$ and 0.75854 correlation coefficients, respectively.

Table 2. Correlation of the net charges of oxygen and hydroxylic hydrogen atoms against experimental $RS_j^{expl}(c)$ relative sweetnesses for 0.3 molar solutions, sucrose $RS^{expl} = 1.0$. The $R(i)$ are correlation coefficients. Number in the parentheses are attributed to fructose and sorbose atoms. XH_1, XH_2, AH_1, AH_2, B_1 and B_2 are the names of sweetener subsites in Tinti-Nofre notation.

Saccharide	Electronic densities on oxygen atoms						$RS_j^{expl}(c)$
j	*i* = 1 (5)	2 (4)	3 (1) B_2	4 (2) B_1	5 (3)	6	(0.3 molar)
1. fructose	–0.29501	–0.31847	–0.30466	–0.34226	–0.30557	–0.28858	0.905
2. sorbose	–0.30436	–0.31526	–0.33722	–0.30981	–0.28114	–0.29255	0.453
3. β-glucose	–0.28809	–0.29963	–0.32084	–0.34304	–0.32462	–0.28540	0.421
4. α-glucose	–0.31256	–0.30047	–0.32325	–0.34333	–0.32269	–0.29263	0.368
5. maltose	–0.28537	–0.28249	–0.29969	–0.34170	–0.32126	–0.27804	0.330
6. lactose	–0.24201	–0.32404	–0.30393	–0.31812	–0.30154	–0.26509	0.200
7. xylose	–0.33371	–0.31249	–0.31787	–0.30594	---	–0.29854	0.175
8. galactose	–0.33399	–0.28616	–0.28676	–0.30438	–0.30241	–0.29085	0.168
9. rhamnose	–0.33553	–0.31434	–0.31755	–0.30507	---	–0.29516	0.158
10. mannose	–0.28802	–0.28412	–0.33062	–0.29832	–0.32730	–0.28128	0.158
$R_{1\text{-}10}(i)$	0.22618	0.24502	0.08768	0.75854	0.18050	0.04856	
$R_{3\text{-}10}(i)$	0.28617	0.18175	0.10979	0.98121	0.48264	0.12136	

Saccharide	Electronic densities on hydroxylic hydrogen atoms					$RS_j^{expl}(c)$
j	*i* = 7 (11)	8 (10) AH_2	9 (7) AH_1	10 (8) XH_1	11 (9) XH_2	(0.3 molar)
1. fructose	0.19597	0.20235	0.19661	0.23321	0.20781	0.905
2. sorbose	0.19717	0.20331	0.21511	0.21892	0.19686	0.453
3. β-glucose	0.19503	0.20535	0.21240	0.23087	0.20255	0.421
4. α-glucose	0.19757	0.20653	0.21190	0.22910	0.20067	0.368
5. maltose	0.05553	0.19140	0.20249	0.22885	0.20126	0.330
6. lactose	0.00601	0.21809	0.20506	0.21109	0.18970	0.200
7. xylose	0.22445	0.20717	0.20932	0.20607	---	0.175
8. galactose	0.20942	0.20146	0.19515	0.19458	0.18796	0.168
9. rhamnose	0.22268	0.20874	0.20912	0.19108	---	0.158
10. mannose	0.19601	0.19237	0.21545	0.19600	0.20393	0.158
$R_{1\text{-}10}(i)$	0.01734	0.09902	0.15437	0.87763	0.62003	
$R_{3\text{-}10}(i)$	0.19621	0.07892	0.17876	0.96445	0.50111	

Theory. Multiple regression analysis of sweetness: We assume that K_j is equilibrium constants of reaction (1) for j-th sugar, whereas K_0 is the equilibrium constant of a such reaction for standard sugar (sucrose) under same conditions: the same receptor and molar concentration of the solution, the same temperature, T. Taking into account the quasi-Taft equation (2) and our first observations of the experimental data (4,5), we can suggest a definition of $RS_j(c)$ relative sweetness of j-th sugar as a physicochemical quantity:

$$RS_j(c) = \frac{K_j}{K_0} \tag{6}$$

Now, we call attention to some properties of the reaction system, which very simplify the theoretical treatment.

(i) All the sugars of the set react with the same taste receptor. Geometry of the receptor, especially $NH_3^{(+)}$ and $COO^{(-)}$ dipoles are common for all sugars and unchanged. Since $NH_3^{(+)}$ cation of lysine is electron acceptor center, we calculate ε_{LU}(Recp) energy of LUMO of this cation. ε_{LU}(Recp) equals –4.8538 eV.

(ii) All the sugars have a similar geometry of the AX_1 and B_1 sweetener subsites. They are situated at pyranose chair-ring unit as common for all sugars considered.

(iii) The sugar receptor-glycophore complexes, Fig. 3b, have a similar geometry for every sugar of the set.

Under above simplifications and taking into account (6), LFER thermodynamic equation [25] presented by Shorter is reduced in quantum chemistry by Gołębiewski [25] to the equation:

$$\ln(RS_j(c)) \cong -\frac{E_{int}(j)}{RT} + \frac{E_{int}(0)}{RT} \tag{7}$$

$E_{int}(j)$ is formation energy of receptor---glycophore of j–th sugar complex. Analogous $E_{int}(0)$ energy for receptor---sucrose complex is constant for all considered sugars. Hence, (7) for sweetness takes the form:

$$\ln(RS_j(c)) = -\frac{E_{int}(j)}{RT} + a_0 \tag{8}$$

The thermodynamic conception of sweetness (8) was confirmed by Höltje and Kier [26]. They examinated the RS_j of 1-X-2-amino-4-nitrobenzenes with several X substituents as glycophores. The 3-methylindole was taken in the character of receptor. The $E_{int}(j)$ intermolecular interaction energy was calculated for exactly assumed geometry of every j-th 1-X-2-amino-4-nitrobenzene---receptor complex. They have found a linear correlation of $log\ RS_j^{expl}$ with $E_{int}(j)$ calculated energies.

In this paper we build $E_{\mathrm{int}}(j)$ term by far approximations of intermolecular interaction theory [27]. **Coulombic model** (atomic net charge model) is the first step of approximation. This model is suitable for description of properties of Coulombic aldopyranose sweetener, {XH_1, XH_2, B_1, B_2, AH_1, AH_2}, with strong domination of B_1 and XH_1 subsites. Monopol-monopol Coulombic interactions of $Q_j(B_1)$ and $Q_j(XH_1)$ net charges with $Q_j(k)$ and $Q_j(l)$ atomic net charges of taste receptor are considered. This interaction may be described by

$$E_{int}(j) = Coul = RT\left[Q_j(B_1)\sum_k \frac{Q(k)}{r(B_1,k)} + Q_j(XH_1)\sum_l \frac{Q(l)}{r(XH_1,l)} \right] \tag{9}$$

The r (B_1,k) means the distance between B_1 sweetener and k receptor atoms. The r(XH_1,l) distance is defined analogously. Summations over k and l define unknown a_1 and a_2 values respectively, as common for a whole set of the sugars, since a glycophore—receptor complexes exibit a similar geometry for all sugars. Thus, QSAR correlation equation may be formulated for the **Coulombic model** in the form:

$$\ln(RS_j(c)) = a_0 + a_1Q_j(B_1) + a_2Q_j(XH_1) \tag{10}$$

On the other hand, ln $RS_j^{\mathrm{expl}}(c)$ demonstrates a poor correlation with the $Q_j(B_2)$ charge, see Table 1. B_2 subsite occurs only as n-electron donor. Therefore, *CT* orbital term should be added to right side of the latter equation [28–30]. We receive the **Coulombic + *CT* frontier orbital** model described by QSAR equation:

$$\ln(RS_j(c)) = a_0 + a_1Q_j(B_1) + a_2Q_j(XH_1) + a_3\frac{1}{\varepsilon_{LU}^{\mathrm{Re}cp} - \varepsilon_{HO}(j)} + a_4\frac{1}{\varepsilon_{LU}^{\mathrm{Re}cp} - \varepsilon_{NHO}(j)} \tag{11}$$

By the far approximations, two latter terms represent a semiquantitative $-E_{\mathrm{CT}}^{(2)}$ energy of charge-transfer interactions, which arises from the transfer of certain part of n-electron pair of some sweetener atoms to LUMO of Lysine-NH_3^+ receptor moiety. The n-electron pairs of B_2 and B_1 subsite atoms dominate in HOMO and NHOMO of pyranoses. Additionally, fructose NHOMO reveals contributions of O-5, O-4 and O-3 n-electron pairs.

The a_0, a_1, a_2, a_3 and a_4 parameters in QSAR equations (10–11) are common for all sugars from the set. They are estimated by least-square procedure, R coefficient of multiple correlation and remainder variance, S^2, are calculated according to Czermiński formulas [31]. One uses RS_j(%) values in the practice. The RS_j(c) values calculated in (10,11) may be recalculated to RS_j(%) values from (3) under $d_j = d_0$ condition. The above theory is suitable for analysis of B_1, B_2 and XH_1, XH_2, AH_1, AH_2 Coulombic net charge subsites as well as B_1, B_2 and E_1, E_2, E_3, E_3 n-electron donors. On the other hand, (11) does not possess suitable terms for investigation of XH_1, XH_2, AH_1, AH_2 n-electron acceptors and G_1, G_2, G_3, G_4 steric (dispersion) subsites.

Atomic net charges of O-4 and H-10 atoms in 4-OH group of aldopyranoses show an excellent correlation with ln $RS_j^{expl}(c)$, as in equations (4) and (5). So, these atoms may be considered as sweet taste centers in aldopyranoses. They are identified as B_1 and XH_1 subsites of Nofre-Tinti sweetener. The XH_1 subsite is equivalent to AH Shallenberger sweetener. The B_2 subsite (B-Shallenberger sweetener) does not undergo net charge correlations. Thus, (10) transforms under a least squares procedure into QSAR correlation equation:

$$\ln(RS_j(c)) = -7.7722 - 15.7423 \times Q_j(B_1) + 5.9721 \times Q_j(XH_1) \qquad (12)$$

with $R = 0.98316$, $S^2 = 0.00795$ for $j = 3, 4, \ldots 10$.

The RS_j relative sweetnesses, calculated by the above equation, are presented in Table 3. The latter equation points for pure Coulombic interactions in the receptor—sweetener complex, (Fig. 3b), which are called by "ionic interaction" in Nofre-Tinti criterion. Energies of such interactions, calculated from $-RT \ln RS_j(c)$, are equal to –3.974 kcal/mole and –3.439 kcal/mole for β-glucose and mannose, respectively. Under the high value of R multiple correlation coefficient, missing of $a_3Q_j(B_2)$ term in (12) indicates that B (=B_2) sweetener is inactive in AH,B system of aldopyranoses. Structure of aldopyranose---receptor complex is presented in Fig. 3b. (12) is very suitable in the $RS_j^{exp}(c)$ calculations for aldopyranoses. If one use $RS_j(\%)$ values, $RS_j^{exp}(c)$ should be recalculated to $RS_j(\%)$ using (3).

The (12) reveals a worse correlation since ketopyranoses are included into QSAR equation, yielding especially poor results for sorbose and fructose. Therefore, (11) should be take into consideration. Using a least squares procedure to (11), the following QSAR equation is received:

$$\ln(RS_j(c)) = -6.9450 - 3.8383 \times Q_j(B_1) + 30.8712 \times Q_j(XH_1) + \frac{32.2882}{\varepsilon_{LU}^{Recp} - \varepsilon_{HO}(j)} -$$

$$- \frac{48.0007}{\varepsilon_{LU}^{Recp} - \varepsilon_{NHO}(j)}; \quad \text{with } R = 0.97215, \ S^2 = 0.03433 \ \text{ for } j = 1,2,3,\ldots 10. \qquad (13)$$

The results of calculation based on the latter equation are collected in Table 3. They are excellent for sorbose and satisfactory for fructose. For aldopyranoses the results are worse than these received from (12). Therefore, (13) is suitable only in RS_j calculation for ketopyranoses. The *n*-electron pair of O-1(B_2) atom dominates in contribution to the sorbose HOMO and NHOMO. Thus, sorbose owns B_1, XH_1, B_2 sweetener, in which B_2 is pure *n*-electron donor subsite (Fig. 3d). The consideration of HOMO(pyranose) → LUMO(receptor) transition is sufficiently for sorbose. On the other hand, additionally consideration of the second transition, NHOMO(pyranose) → LUMO(receptor), is necessary for fructose. Space diagram of fructose NHOMO orbital reveals a great contribution of O-3 oxygen *n*-electrons (E type subsite) as well as O-5 and O-4 oxygen *n*-electron pairs. The O5, O4 and O3 *n*-electron donors converge with fructofuranose E_1, E_2 and E_3 subsites [13], respectively. The second *CT* or-

bital term in (13) emulates, in case of fructose, the interactions of these subsites with Thr receptor recognition subsites (Fig. 3d). One concludes that fructose reveals B_1, XH_1, B_2, E_1, E_2 E_3 sweetener.

Table 3. The comparison of calculated RS_j^{calc} (8 sugar and 10 sugar sets) and measured RS_j^{expl} relative sweetnesses and energies of the HOMO and NHOMO frontier obitals (eV) – for the most important mono- and disaccharides.

Saccharide	E_{HOMO}	E_{NHOMO}	0.3 molar solutions			10% per cent solutions		
			Eq.(12), 8-th	Eq.(13), 10-th	RS_j^{expl} (*c*)	Eq.(12), 8-th	Eq.(13), 10-th	RS_j^{expl} (%)
Sucrose					1.000			1.000
1. fructose	–10.5305	–11.2778	---	0.806	0.905	---	1.531	1.720
2. sorbose	–10.4198	–11.1229	---	0.426	0.453	---	0.809	0.860
3. β-glucose	–10.8646	–11.1030	0.370	0.445	0.421	0.704	0.845	0.800
4. α-glucose	–10.8478	–10.9731	0.368	0.364	0.368	0.700	0.691	0.700
5. maltose	–10.7133	–10.8524	0.358	0.346	0.330	0.358	0.346	0.330
6. lactose	–10.7842	–10.9875	0.222	0.204	0.200	0.222	0.204	0.200
7. xylose	–10.8024	–11.2521	0.178	0.227	0.175	0.406	0.517	0.400
8. galactose	–10.9168	–11.1869	0.162	0.132	0.168	0.308	0.251	0.320
9. rhamnose	–10.4651	–11.1738	0.161	0.180	0.158	0.335	0.375	0.330
10. mannose	–10.6060	–11.0141	0.149	0.145	0.158	0.283	0.276	0.300
R coeff. of multiple correlation:			0.98316	0.97214				
S^2 remainder variances:			0.00795	0.03433				

REFERENCES

1. Acree T.E., *Handbook of Sensory Physiology*, vol. **4**, *Chemical Senses*, Part 2, ed. Beidler L.M., Springer-Verlag, Berlin Heidelberg and NY, 1971.
2. Shallenberger R.S., *Food Technol*., **52**, 72 (1998).
3. Shallenberger R.S., *J. Food Sci.*, **28**, 584 (1963).
4. Shallenberger R.S., *Agr. Sci. Rev.*, **2**, 11 (1964).
5. Birch G.G. and Lee C.K., *The Theory of Sweetness* in *Developments in Sweetness 1*. Applied Science Publishers Ltd., London, 1979, pp. 165–186.
6. Kier L.B., *J .Pharm. Sci.*, **61**, 1394 (1972).
7. Shallenberger R.S. and Acree T.E., *Nature*, **216**, 480 (1967).
8. Blanksma J.J. and Hoegen D., *Rec. trav. chim.*, **65**, 333 (1946).
9. Pietrzycki W. and Tomasik P., *J. Chim. Phys.*, **90**, 2055 (1993).
10. Ruffolo R.R., Jr., Nichols A.J., Stadel J.M. and Hieble J.P., *Pharmacol. Rev.*, **43**, 475 (1991).
11. Nowak Z. and Zawilska B., *Receptors. Structure*, *Characteristic*, *Function*, (in Polish), PWN, Warszawa, 1977.
12. Brand J.G. and Feigin A.M., *Food Chem*., **56**, 199 (1996).
13. Nofre C. and Tinti J.M., *Food Chem*., **56**, 263 (1996).
14. Shallenberger R.S., *New Scientist*, **407**, 569 (1964).
15. Shallenberger R.S., Acree T.E. and Guild W.E., *J. Food Sci.*, **30**, 560 (1965).
16. Shallenberger R.S. and Acree T.E., *Nature*, **216**, 480 (1967).
17. Shallenberger R.S. and Acree T.E., *J. Agr. Food Chem.*, **17**, 701 (1969).

18. Höltje H.D. and Kier L.B., *Pharm. Sci.*, **211**, 75 (1977).
19. Hypercube, Inc., 419 Philip Street, Waterloo, Ontario N2L 3X2, Canada.
20. Allinger N.L., *J. Am. Chem. Soc.*, **99**, 8127 (1977).
21. Filipowicz B. and Więckowski W., *Biochemistry*, Vol. I, (in Polish), PWN Łódź-Warszawa, 1965, p. 223.
22. Tsuzuki T. and Yamazaki J., *Biochem. Z.*, **323**, 525 (1953).
23. Tomasik P., in *Food Chemistry,* (in Polish), Ed. Sci.-Tech., Warszawa, 2000, p. 166.
24. Birch G.G., Cowell N.D. and Eyton D., *Food Technol*., **5**, 227 (1970).
25. Shorter J., *Correlation Analysis in Organic Chemistry*, (in Polish), PWN, Warszawa, 1980; Gołębiewski A., *Quantum Chemistry of Organic Compounds*, (in Polish), PWN, Warszawa, 1973, pp. 379–380.
26. Höltje H.D. and Kier L.B., *J. Pharm. Sci.*, **63**, 1722 (1974).
27. Pietrzycki W., *Acta Phys. Polon.*, **A76**, 421 (1989).
28. Pietrzycki W., *Polish J. Chem.*, **74**, 89 (2000).
29. Moskal J., Moskal A. and Pietrzycki W., *J. Chem. Soc*., *Perkin II*, 1893 (1977).
30. Salem L., *J. Am. Chem. Soc.*, **90**, 543 (1968).
31. Czermiński J.B., Iwasiewicz A., Paszek Z. and Sikorski A., *Statistical Methods for Chemists*, (in Polish), PWN, Warszawa, 1992, p. 259–261.