Articles

General Pseudoreceptor Model for Sweet Compounds: A Semiquantitative Prediction of Binding Affinity for Sweet-Tasting Molecules

Angela Bassoli,[†] Michael G. B. Drew,[‡] Lucio Merlini,[†] and Gabriella Morini*[,†]

Dipartimento di Scienze Molecolari Agroalimentari, Sezione di Chimica, Università di Milano, Via Celoria, 2-I 20133 Milano, Italy, and Department of Chemistry, University of Reading, Whiteknights, Reading RG6 6AD, United Kingdom

Received January 25, 2002

The chemical structures of sweet compounds are very different, ranging from sugars to amino acids and peptides or other compounds such as saccharin. The biological mechanism underlying the generation of sweet taste is still unknown, although in the past few years much research has provided evidence for the existence of a true chemoreception process, mediated by receptor proteins on the taste buds. In particular, the initial step of the process involves the reversible binding of the sweet compounds to their receptor(s). In this work, we have investigated this binding via a pseudoreceptor model, which has been developed using a training set of 24 compounds belonging to different families including sugars, peptides, and other intensive sweeteners. This model provided a correlation coefficient (r^2) of 0.985 between the calculated and the experimental free energies of binding, which are related to the molar relative sweetness, for the training set and is able to predict semiquantitatively free energies of ligand binding for an independent set of five test ligand molecules within 0.3−2.1 kcal mol^{-1} of the experimental values.

Introduction

The mechanism of action of sweet substances, as well as that of other tastants, has been under investigation for many years. In the last two years, significant progress has been made in our understanding of the mechanism of sweet taste chemoreception, as almost contemporaneously several independent groups of researchers[1−6] identified a new family of receptors, named T1R3, very similar to the orphan receptors T1R1 and T1R2, which had previously been identified.[7,8] These receptors could have a common function as demonstrated by Li et al.[9] who were able to obtain a functional dimeric receptor for sweet compounds after coexpression of T1R3 and T1R2. However, only the primary structure of the proteins is known, and the identity of the amino acids that make up the receptor site is yet to be established; consequentially, the three-dimensional arrangement of the amino acids of the binding site is not yet known. Therefore, alternative methodologies are still required to model the binding of sweet compounds to the receptor. Characteristics of the sweet taste receptorial system are the ability to recognize, with high specificity, molecules belonging to very different classes of compounds and the very low affinity toward the natural agonists. In the last century, several different models were developed to describe the nature and the topological arrangement of glucophores of an ideal sweet compound and/or the recognition sites of the sweet taste receptor.

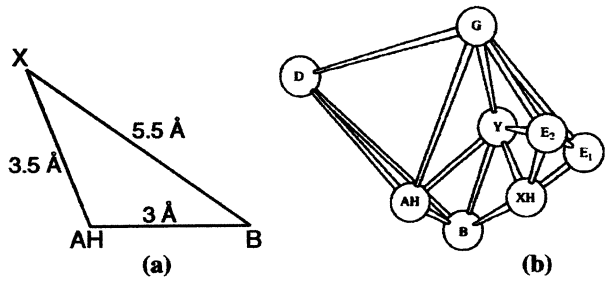

Figure 1. (a) Shallenberger−Acree−Kier model. The groups AH (H-bond donor) and B (H-bond acceptor) are engaged in two antiparallel hydrogen bonds with complementary sites on the receptor. Kier suggested a third site, X, of hydrophobic interaction. (b) The MPA Nofre−Tinti model describes an ideal sweet compound with eight glucophores, four high-affinity sites AH, B, G (corresponding to the AH, B, and X of panel a), and D, and four secondary sites Y, E_1, E_2, (hydrogen bond acceptor groups), and XH (hydrogen bond donor group).

The first useful model was proposed in 1967 by Shallenberger and Acree.[10] They recognized the existence in almost every sweet molecule of two functional groups (glucophores) corresponding to a hydrogen bond donor and a hydrogen bond acceptor, named AH and B, respectively. Their role in recognition involved the creation of two parallel hydrogen bonds at ca. 3.5 Å apart, with two complementary sites on the receptor protein. Kier[11] added a third interaction site (first called X) corresponding to a hydrophobic region of the molecule (Figure 1a).

* To whom correspondence should be addressed. Tel: +39.02.50316824. Fax: +39.02.50316801. E-mail: gabriella.morini@unimi.it.
[†] Università di Milano.
[‡] University of Reading.

10.1021/jm020833v CCC: $22.00
Published on Web 08/24/2002

This simple model found widespread acceptance, due to its ability to explain the sweetness of many structurally different compounds and also to interpret variations in sweetness due to geometrical or conformational differences, especially in sugars. It has also been of some heuristic value, contributing to the preparation of new sweet substances.

Substantial progress in the understanding of the molecular properties of organic compounds and in techniques of molecular modeling has provided tools from which more sophisticated models than the simple one of Shallenberger−Acree−Kier could be constructed. The first step in this direction was made by Temussi et al.[12,13] who, on the basis of an extensive conformational study of aspartame, which included NMR experiments and theoretical calculations, proposed a model for the sweet receptor. This model depicts the receptor as a hemihedral cavity with a definite shape and functionality, which includes the AH−B groups of Shallenberger and Acree. Moreover, it can explain the change from sweet to bitter taste of some enantiomeric compounds, such as some D- and L-amino acids. Modifications to this model were subsequently made by Goodman et al.[14] Conformational analysis and electrostatic potential calculations on some very sweet compounds have been used by Culberson and Walters[15] to derive a three-dimensional model of the sweet taste receptor, lately used to design ultrahigh potency sweeteners.[16] More recently, Walters et al.[17] suggested a pharmacophore model derived from extensive conformational analysis of some high potency sweeteners.

As a result of a rational approach in the design of sweet molecules, Tinti and Nofre have been able to discover several series of hyperpotent sweeteners.[18] One of these series, the hyperpotent guanidinic sweeteners, include compounds with relative sweetness (RS) values over 200 000 times that of sucrose. Comparison of the molecular properties of these extremely sweet substances within qualitative structure−activity relationships (SAR) of various sweeteners led to the development of an eight site interaction model, based on the chemical nature and topology of an ideal sweet compound[19,20] (Figure 1b). The multipoint attachment (MPA) model characterizes an ideal sweet compound with eight glucophores, consisting of four high affinity sites AH, B, G (corresponding to AH, B, and X of Shallenberger−Acree−Kier), and D and four secondary sites Y, E_1, E_2, (hydrogen bond acceptor groups), and one hydrogen bond donor group, XH. At present, this is the most detailed model that is readily available and as such has been successfully used to explain the sweet taste of many compounds belonging to different classes and to design compounds with high RS.[21] In later work,[22] the same authors proposed an improved model in which they included eight specific amino acids of the sweet taste receptor, which were involved in 15 interactions with the glucophores. The MPA model has subsequently been reexamined, and the number of interacting amino acids has been increased to 10.[23] Unfortunately, the authors neither describe how they selected which amino acids to use in the model nor provide the three-dimensional arrangement of the amino acids. Moreover, this model—as well as the previous ones—does not provide any quantitative information about the affinity of the compounds toward the receptor model, which could be correlated with their RS values and hence be used for validation of their model.

The generally unsatisfactory nature of this type of qualitative modeling has led us to develop quantitative SAR (QSARs) for sweet-tasting compounds.[24] The development of QSARs is generally carried out with structurally similar compounds as, in general, a specific receptor is able to accommodate and recognize only very similar compounds. However, by contrast with many specific drug molecules, sweet-tasting molecules belong to very different classes of compounds, which could indicate either that several different taste receptors are present or that one taste receptor could accommodate all types of sweet-tasting molecules. The successful development of a QSAR for all types of sweet-tasting molecules would naturally favor the likelihood that there is only one sweet receptor that is suitable for different types of sweeteners. In this paper, we report the development of such a QSAR using the pseudoreceptor approach, previously applied by some of us to derive a three-dimensional binding site model for isovanillic sweet derivatives.[25] This model provides for the first time a general semiquantitative interpretation of sweetness potency.

Pseudoreceptor modeling[26,27] belongs to the so-called receptor mapping approaches, where a paucity of information concerning receptor structures has spawned techniques that project the properties of the bioactive ligands into three dimensions around their appropriately superimposed molecular framework. The resulting map provides steric, electrostatic, and lipophilic profiles used to identify the type and approximate position of receptor residues, or their functional groups, interacting with the ligand. This map can be used for subsequent molecular modeling and allows semiquantitative predictions of binding affinities for ligands. Although, in general, sequence and arrangement of the building blocks of a pseudoreceptor (e.g., amino acids residues) and its natural counterpart will only bear little resemblance, they should accommodate a series of ligands in a relatively similar binding pattern.

Experimental Section

Three-dimensional molecular models were built on a Silicon Graphics IRIS 4D-35GT, using the program InsightII/Discover, 97.0 (Biosym Technologies, San Diego, CA). The initial models were energy-refined by molecular mechanics techniques with conjugate gradients until a maximum energy derivative value of 0.008 kcal mol^{-1} $Å^{-1}$ was obtained using the CVFF force field. Conformational analysis was performed wherever necessary by molecular dynamics. For atomic partial charges of the ligand atoms, we used Mulliken charges calculated on the minimized structures using the MOPAC program[28] with the MNDO Hamiltonian.

Molar relative sweetness (MRS) values, obtained from the literature by converting the data usually given as a weight basis relative to a 2% sucrose solution, were converted to K_d values. For the mapping process, only relative K_d values were relevant (as the predicted values were obtained by regression, the absolute values were not required). Because the value of K_d for sucrose is unknown, it was arbitrarily set at 10^{-5} M in order to allow for a reasonable range of values for K_d of the considered compounds.

Experimental free energies of ligand binding were calculated from eq 1:

$$\Delta G^\circ_{exp} = RT \ln K_d \quad (1)$$

where $K_d = K_d$(sucrose)/MRS.

For pseudoreceptor modeling, the program PrGen 2.1 (SIAT Biographics Laboratory, Basel, CH) was used. Molecules in their lowest energy conformation were imported in PrGen and reminimized with the Yeti force field.[29]

In PrGen, free energies of ligand binding, $\Delta G°$, are estimated based on the approach of Blaney et al.[30] (eq 2)

$$E_{binding} \approx E_{ligand-receptor} - T\Delta S - \Delta G_{ligand\ solvation} + \Delta E_{internal,ligand} \quad (2)$$

The loss of entropy upon receptor binding was estimated following Searle and Williams.[31] Ligand solvation energies were calculated using the method of Still et al.[32] Algorithms to calculate these two quantities are included in PrGen 2.1. The fourth term corrects for the deviation of the ligand internal energy (while bound to the pseudoreceptor) from a strain-free reference conformation. To determine the ligand−receptor interaction energy, $E_{ligand-receptor}$, the program uses the force field Yeti. Free energies of ligand binding, ΔG°_{pred}, are then obtained by means of a linear regression (slope a, intercept b) between ΔG°_{exp} and $E_{binding}$ (eq 3)

$$\Delta G^\circ_{pred} = |a|E_{binding} + b \quad (3)$$

Results and Discussion

Pseudoreceptor Model for Guanidinic Compounds. As a starting point, we chose to build a pseudoreceptor for the class of the guanidinic hyperpotent sweeteners for several reasons. They provide a wide range of RS values as well as including the sweetest compounds known to date. Their structures contain several identifiable glucophores, they are large and quite rigid molecules, and the MPA model built using this family has been successfully used to explain, at least qualitatively, the sweet taste of many compounds that belong to different classes.[19] Figure 2 shows the structure and the MRS values (obtained from ref 20) of the 47 compounds used to derive the pseudoreceptor for the guanidinic compounds. The molecules were divided into a training set to create the model and a test set to validate the model. The training set consisted of 39 molecules. The test set consisted of the eight molecules **2**, **5**, **9**, **14**, **26**, **33**, **36**, and **42** (framed in Figure 2), which were selected to encompass a large range of sweetness values and include most distinct types of functional groups.

The lowest energy conformations established for the guanidinic compounds included an internal hydrogen bond between the Ar−NH and the carboxylate groups. It might be argued that this intramolecular hydrogen bond does not persist in solution or within the receptor site, but in this conformation, there is an optimal overlap between the glucophores and the corresponding AH, B, G, and D sites of the MPA model of Tinti and Nofre, and also, as these lowest energy conformations can be readily overlapped using the common guanidinium group, it seemed pertinent to use them in the calculations.

From the overlapped ligands, the program generates vectors for each functional group indicating steric, electrostatic, and lipophilic interactions. Individually chosen residues are then positioned at the tips of these vectors, and these residues taken together make up the pseudoreceptor. To choose the amino acids, structural

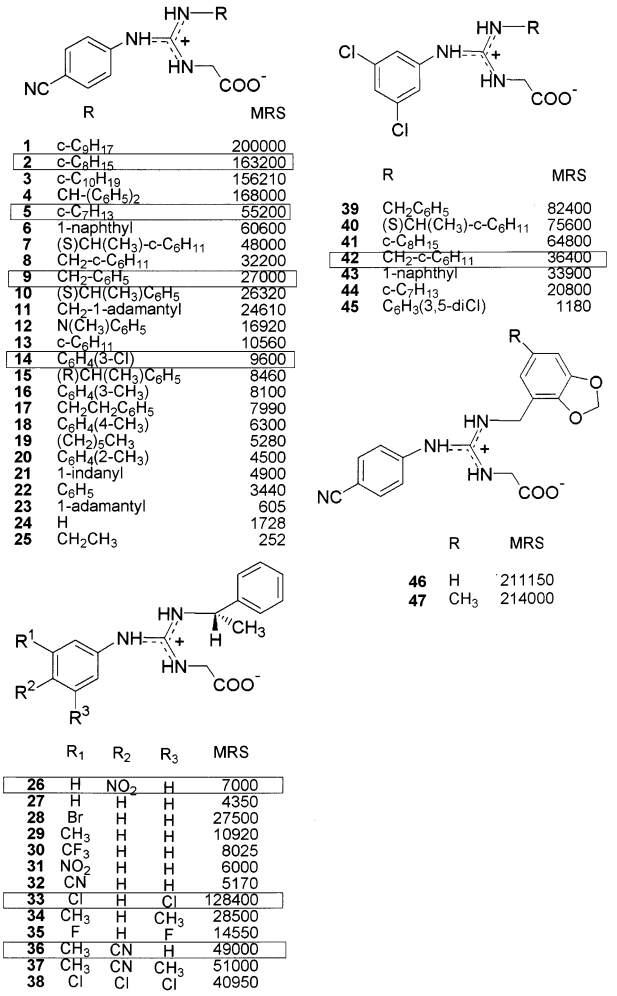

Figure 2. Molecular structure and MRS of the compounds used to derive. The compounds **2**, **5**, **9**, **14**, **26**, **33**, **36**, and **42** (framed) constitute the test set.

information from the monoclonal antibodies raised on a hyperpotent guanidinic derivative[33,34] was used. One of these antibodies includes within the receptor an Arg, which forms a salt bridge with the carboxylate moiety of the ligand, and a Glu that provides the negatively charged potential for interaction with the guanidinic group. We used Asp instead of Glu, because it is smaller, while maintaining the same binding feature. The two amino acids were appropriately positioned around the overlapped molecules, and then, additional residues were added. Each residue was chosen specifically to fit the type of interaction (steric, electrostatic, or lipophilic) of each vector as characterized by the adjacent structural features of the overlapped molecules. The structure of the resulting complex of training set molecules and 17 amino acid residues of the pseudoreceptor was then optimized using ligand equilibration, a protocol where correlation-coupled receptor optimization, which couples the rms deviation of observed and calculated $\Delta G°$ values together with the total energy of the complex with the ligands fixed, and free ligand relaxation within the fixed receptor are altered in an iterative fashion until the correlation between experimental and predicted free energies is maximized in the relaxed state. This procedure yielded a correlation coefficient (r^2) of 0.966 between experimental and predicted free energies of ligand binding, $\Delta G°$, and a rms deviation of 0.230 kcal

Table 1. Compounds Used to Derive the Pseudoreceptor Model for Guanidinic Compounds: MRS and Comparison of Experimental and Predicted Free Energies of Ligand Binding

compd	exp MRS (as $\Delta G°$, kcal mol^{-1})	pred MRS (as $\Delta G°$, kcal mol^{-1})	error in MRS (as $\Delta G°$, kcal mol^{-1})
	Training Set		
1	−13.810	−13.957	−0.147
3	−13.670	−13.614	0.056
4	−13.710	−13.570	0.140
6	−13.110	−12.933	0.177
7	−12.980	−13.012	−0.032
8	−12.750	−12.575	0.175
10	−12.630	−12.385	0.245
11	−12.590	−12.264	0.326
12	−12.370	−12.605	−0.235
13	−12.100	−12.017	0.083
15	−11.970	−11.618	0.352
16	−11.940	−11.989	−0.049
17	−11.940	−11.647	0.293
18	−11.800	−11.908	−0.108
19	−11.690	−11.715	−0.025
20	−11.600	−11.688	−0.088
21	−11.650	−12.142	−0.492
22	−11.440	−11.531	−0.091
23	−11.470	−11.441	0.029
24	−11.040	−10.860	0.180
25	−9.920	−10.292	−0.372
27	−11.580	−11.754	−0.174
28	−12.650	−12.695	−0.045
29	−12.120	−12.487	−0.367
30	−11.940	−11.638	0.302
31	−11.770	−11.527	0.243
32	−11.680	−12.183	−0.503
34	−12.680	−12.990	−0.310
35	−12.280	−11.973	0.307
37	−13.010	−12.806	0.204
38	−12.890	−13.242	−0.352
39	−13.290	−13.304	−0.014
40	−13.240	−13.247	−0.007
41	−13.150	−13.182	−0.032
43	−12.780	−12.816	−0.036
44	−12.490	−12.259	0.231
45	−10.820	−10.970	−0.150
46	−13.840	−13.583	0.257
47	−13.850	−13.820	−0.030
	Test Set		
2	−13.690	−13.805	−0.115
5	−13.060	−12.852	0.208
9	−12.640	−14.250	−1.610
14	−12.040	−13.402	−1.362
26	−11.860	−13.529	−1.669
33	−13.550	−12.841	0.709
36	−12.990	−13.383	−0.393
42	−12.820	−14.938	−2.118

mol^{-1}, which can be transformed into an uncertainty factor of 1.5 in the MRS value (Table 1, Figure 3).

To validate the pseudoreceptor model created using the 39 compounds in the training set, we then used the additional eight guanidinic compounds that constituted the test set. These molecules, which were not used to build the model, were added to the pseudoreceptor and subjected to free ligand relaxation while the pseudoreceptor was kept rigid. The predicted free energies of binding were then compared with the experimental values. The agreement between the data was good, with an rms deviation between experimental and predicted free energies of ligand binding, $\Delta G°$, of 1.247 kcal mol^{-1}. The pseudoreceptor model was further validated using a cross-validation method, leave-one-out procedure, in which each molecule was systematically removed in turn from the data set and a new model was derived using one less molecule. Then, each new model was used to predict the activity of the molecule that was not included in the new model set. Applying this method, we obtained a cross-validation correlation coefficient of 0.925 between predicted and experimental free energies of ligand binding. From these quantitative results, it can be concluded that the generation of a pseudoreceptor for the guanidinic compounds has been successful and that the model can be used to predict the free energies of ligand binding (or MRS) values of compounds in this family.

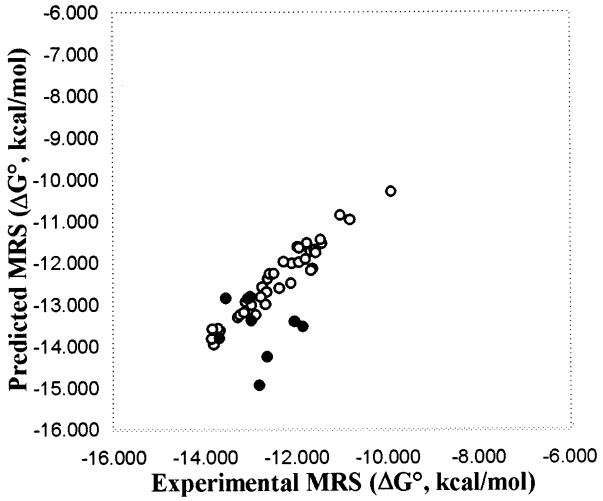

Figure 3. Pseudoreceptor for guanidinic compounds: plot of experimental MRS vs predicted MRS. Dots represent the training set; solid dots represent the test set.

Pseudoreceptor for Different Classes of Sweet Compounds. To transform this model for the guanidinic compounds into a general pseudoreceptor for sweet compounds, it was necessary to generate a more extensive set of molecules encompassing most of the sweeteners in use and many other sweet substances. The new set contained compounds **48**−**71** (Figure 4) plus the guanidinic derivatives **1**, **3**, **8**, **24**, and **33**. Five of these compounds, namely, **3**, **50**, **53**, **64**, and **71**, were kept aside as the test set with the remaining 24 molecules used as the training set. The lowest energy conformation of each molecule was generated as described above.

As all of these compounds belong to different chemical classes, it was not possible to superimpose them on the basis of the same structural feature as was done for the guanidinic compounds. However, these molecules share glucophoric functional groups that are described by the Shallenberger−Acree−Kier and in more detail by the Tinti−Nofre MPA model. Therefore, the coordinates of the Tinti and Nofre model were imported into PrGen, and the MPA model was oriented appropriately around the guanidinic compounds used to derive the pseudoreceptor for guanidinic compounds. At this point, the guanidinic compounds were removed from the model leaving the MPA model orientated within the pseudoreceptor. Next, the compounds in the new training set were inserted into the MPA model (and hence into the pseudoreceptor) in positions consistent with their postulated interaction sites within the MPA model. In this way, all of the AH, B, and G (and more if present) interaction sites of the different sweet compounds were overlapped. Figure 5 shows selected compounds in the

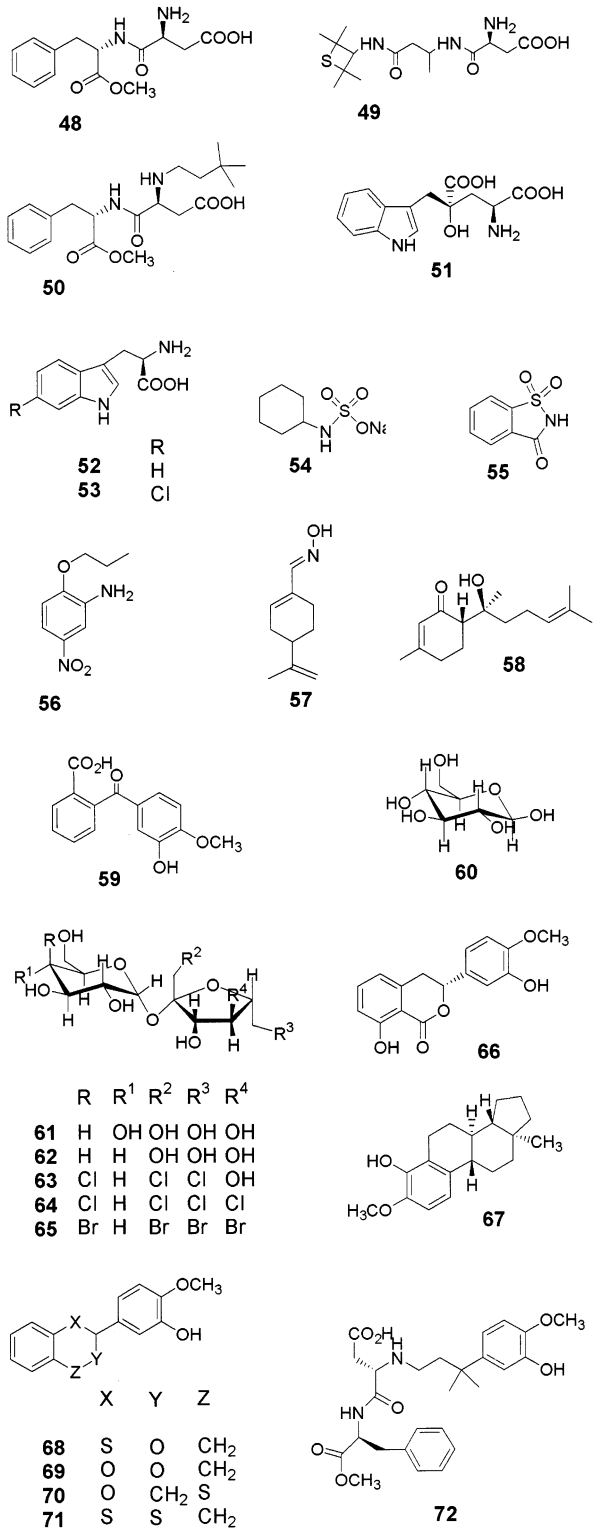

Figure 4. Molecular structure of compounds used to derive the general pseudoreceptor for sweet compounds (together with compounds **1**, **3**, **8**, **24**, and **33** of Figure 2). Compounds **3**, **50**, **53**, **64**, and **71** constitute the test set.

training set overlapped to the MPA model, abstracted from ref 19.

Sucrononic acid, **1**, is the compound presenting the most interaction sites: AH is NH(phenyl), B is COO^-, G is cyclononyl, and D is CN. For aspartame, **48**, AH is NH_3^+, B is COO^-, and G is the phenyl ring. For 6-Cl-D-tryptophan, **53**, AH is NH_3^+, B is COO^-, and G is 6-Cl-indolyl. For compound **68**, an isovanillic derivative, AH is OH, B is OCH_3, and G is the benzo-condensed phenyl ring.

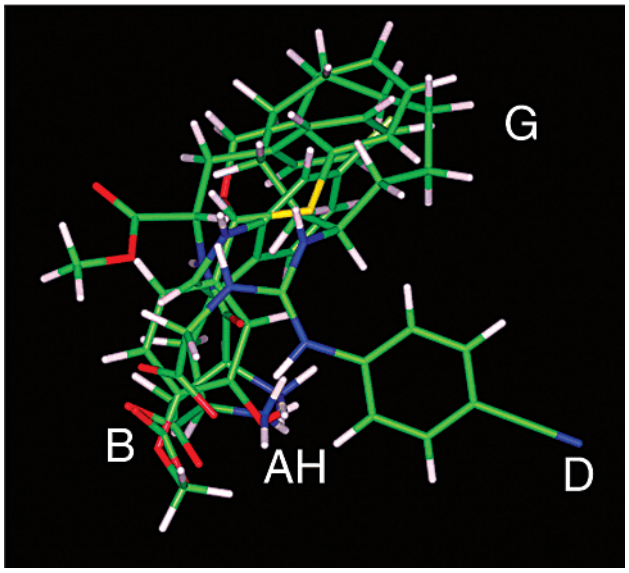

Figure 5. Compounds **1**, **48**, **53**, and **68** overlapped to the MPA model.

While this is clearly not an ideal procedure, as the fitting of the molecules to the MPA model can only be approximate and of course the accuracy of the MPA model is itself questionable, it does provide a starting orientation for the molecules within the pseudoreceptor. It should be noted that only an approximate positioning is required because one of the advantages of the pseudoreceptor method, over, for example, molecular field analysis, is that the positions (as well as the conformations) of all of the molecules are adjusted individually during the correlation-coupled optimization of the training set to provide the best fit with experimental data.

The pseudoreceptor model obtained for guanidinic compounds was used as the starting model for the creation of the general pseudoreceptor for most sweet-tasting compounds. An Asp residue was added in a position corresponding to the AH group of the MPA model, to account for an interaction of a hydrogen bond donor group, which in the previous guanidinic receptor was saturated by the interaction of the guanidinic NH group in an internal H-bond with the COO^- group of the guanidinic compounds. To make space for the added Asp, two hydrophobic amino acids included in the previous guanidinic pseudoreceptor were deleted. Solvation energies of the individual ligand molecules were not considered in the calculation (eq 2) because our ligand set contained both formally charged and uncharged species: this may cause problems as their difference in E_{sol} is typically in the range of 50 kcal mol^{-1}.[35] The complex of superimposed ligands of the training set and pseudoreceptor model was then optimized using the ligand equilibration protocol described above. This procedure yielded a correlation coefficient between experimental and predicted free energies of ligand binding, $\Delta G°$, of 0.985 and a rms deviation of 0.345 kcal mol^{-1} that can be transformed into an uncertainty factor of 1.8 in the MRS value (Table 2, Figure 6).

Table 2. Compounds Used to Derive the General Pseudoreceptor: Common Name (When Available), Comparison of Experimental and Predicted Free Energies of Ligand Binding, and Experimental and Predicted MRS

compd	common name	exp MRS (as $\Delta G°$, kcal mol^{-1})	pred MRS (as $\Delta G°$, kcal mol^{-1})	error in MRS (as $\Delta G°$, kcal mol^{-1})	exp MRS	pred MRS
			Training Set			
1	sucrononic acid	−13.810	−13.219	0.591	200 000	72 500
8		−12.750	−12.861	−0.111	32 200	39 200
24		−11.040	−11.006	0.034	1728	1620
33		−13.550	−13.756	−0.206	128 400	182 250
48	aspartame	−9.700	−9.664	0.036	172	160
49	alitame	−11.110	−11.534	−0.424	1937	4000
51	monatin	−10.740	−10.585	0.155	1025	790
52	D-tryptophan	−8.480	−8.356	0.124	21	20
54	cyclamate	−8.600	−8.849	−0.249	26	40
55	saccharin	−9.660	−9.705	−0.045	161	170
56	P4000	−11.210	−11.337	−0.127	2293	2860
57	perillartine	−10.710	−10.373	0.337	966	550
58	hernandulcin	−10.510	−10.159	0.351	691	380
59		−9.780	−9.558	0.222	198	135
60	D-glucose	−5.920	−5.992	−0.072	0.26	0.29
61	sucrose	−6.703	−6.476	0.227	1	0.68
62		−6.670	−6.659	0.011	0.95	0.93
63	sucralose	−10.560	−10.279	0.281	755	470
65		−12.220	−11.508	0.712	13 012	3850
66	phyllodulcin	−10.320	−10.740	−0.420	502	1025
67		−8.880	−9.431	−0.551	42	110
68		−12.280	−12.213	0.067	14 426	19 200
69		−11.600	−11.693	−0.093	4527	5250
70		−10.190	−11.040	−0.850	401	1720
			Test Set			
3		−13.670	−15.786	−2.116	156 210	5 953 900
50	neotame	−12.120	−11.862	0.258	11 057	7050
53	6-Cl-D-tryptophan	−10.670	−10.208	0.462	906	420
64		−11.300	−10.981	0.319	2674	1550
71		−12.370	−14.142	−1.772	16 968	353 600

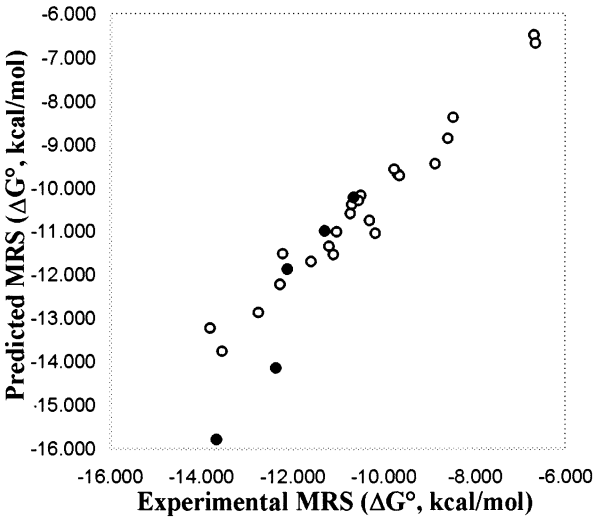

Figure 6. General pseudoreceptor for sweet compounds: plot of experimental MRS vs predicted MRS. Dots represent the training set; solid dots represent the test set.

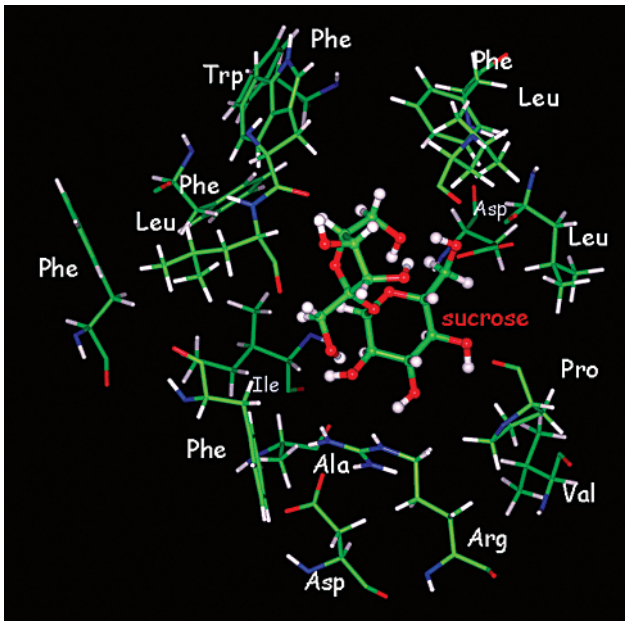

Figure 7. Three-dimensional structure of the general pseudoreceptor model for sweet compounds.

The general pseudoreceptor for the sweet compounds consists of 16 amino acids (Figure 7). Most of the residues used are hydrophobic amino acids (five Phe, three Leu, and one each of Ala, Trp, Pro, Ile, and Val), which form two distinct hydrophobic binding pockets in the pseudoreceptor corresponding to the G and D sites of the Tinti and Nofre model. However, the strongest interactions between the sweet-tasting molecules and the pseudoreceptor are found with the three polar residues, one Arg and two Asp. The Arg shows a polar interaction with hydrogen bond acceptor groups of the sweet compounds (the B site of the Tinti and Nofre model), while the two Asp residues show a polar interaction with the hydrogen bond donor groups of the sweet compounds (the AH and XH glucophores of the Tinti and Nofre model).

The test set (compounds **3**, **50**, **53**, **64**, and **71**) was initially added to the pseudoreceptor in positions consistent with the MPA model and then subjected to free ligand relaxation. This procedure involved minimizing the interaction energy of the ligand with the fixed pseudoreceptor and did not involve minimizing $\Delta\Delta G°$ values. For the test set, the rms deviation between

experimental and predicted free energies of ligand binding, $\Delta G°$, was 1.268 kcal mol^{-1}. The cross-validation correlation coefficient using the leave-one-out procedure was calculated as 0.966.

The pseudoreceptor model was also validated by randomization tests in which the values of MRS were reassigned randomly to different molecules of the training set, and a new correlation-coupled minimization was performed. The correlation coefficient between observed and calculated $\Delta G°$ values was then calculated and compared to the correlation coefficient obtained with the correct data. The procedure was repeated 10 times, and all r^2 values ($r^2 < 0.1$ one time; $0.1 < r^2 < 0.6$ six times; $0.6 < r^2 < 0.7$ three times) were much less than the 0.985 obtained with the correct data, thus validating the original calculation.

Thus, three methods of validation, one external using the test set and two internal using cross-correlation and randomization, have demonstrated that this pseudoreceptor model created for a variety of different sweet tastants has good predictive power. Further confirmation of the predictive power of this model consists of the good prediction of the sweetness value of a compound recently synthesized (**72** in Figure 4).[36,37] The experimental MRS value is 70 000 ($\Delta G°$ −13.20 kcal mol^{-1}) while our model predicted a $\Delta G°$ of −12.53 kcal mol^{-1}, corresponding to a MRS of 22 200, which is a good result taking into account the size of the error inherent in the sweetness measurements.

Conclusions

Our model is able to explain and predict the sweet taste of compounds belonging to different families. Its structural features are in agreement with the preexisting models suggested for the sweet taste receptor, but it also provides for the first time a a semiquantitative evaluation of sweetness activity. The recent discovery of a putative sweet taste receptor gene strengthens the hypothesis for the existence of a GPCR-mediated chemoreception mechanism for sweet tastants. The pseudoreceptor model provides an explanation for the taste of known compounds and can predict with some accuracy the taste of new derivatives; it could therefore be used as a valid tool to model ligand−receptor interactions and to provide information relevant to the structure of the binding set once the identity of the residues making up the binding site within the protein are identified.

The fact that it has proved possible to fit successfully a large number of sweet-tasting molecules from different families into the model not only confirms the usefulness of this pseudoreceptor technique but also adds some support toward the hypothesis that there is just one sweet taste receptor for these families.

Acknowledgment. Financial support by the EC (contract FAIR CT-98-4040) and by the University of Milano is gratefully acknowledged.

References

(1) Max, M.; Shanker, Y. G.; Huang, L.; Rong, M.; Liu, Z.; Campagne, F.; Weinstein, H.; Damak, S.; Margolskee, R. F. Tas1r3, encoding a new candidate taste receptor, is allelic to the sweetresponsiveness locus Sac. *Nat. Genet.* **2001**, *28*, 58−63.

(2) Montmayeur, J. P.; Liberles, S. D.; Matsunami, H.; Buck, L. B. A candidate taste receptor gene near a sweet taste locus. *Nat. Neurosci.* **2001**, *4*, 492−498.

(3) Kitagawa, M.; Kusakabe, Y.; Miura, H.; Ninomiya, Y.; Hino, A. Molecular genetic identification of a candidate receptor gene for sweet taste. *Biochem. Biophys. Res. Commun.* **2001**, *283*, 236−242.

(4) Nelson, G.; Hoon, M. A.; Chandrashekar, J.; Ryba, N. J. P.; Zuker, C. S. Mammalian sweet taste receptor. *Cell* **2001**, *106*, 381−390.

(5) Li, X.; Inoue, M.; Reed, D. R.; Hunque, T.; Puchalski, R. B.; Tordoff, M. G.; Ninomiiya, Y.; Beauchamp, G. K.; Bachmanov, A. A. High-resolution genetic mapping of the saccharin preference locus (Sac) and the putative sweet taste receptor (T1R1) gene (Gpr70) to mouse distal Chromosome 4. *Mamm. Genome* **2001**, *12*, 13−16.

(6) Sainz, E.; Korley, J. N.; Battey, J. F.; Sullivan, S. L. Identification of a novel member of the T1R family of putative taste receptors. *J. Neurochem.* **2001**, *77*, 896−903.

(7) Hoon, M. A.; Adler, E.; Lindemeier, J.; Battey, J. F.; Ryba, N. J. P.; Zuker, C. S. Putative mammalian taste receptors: a class of taste specific GPCRs with distinct topography selectivity. *Cell* **1999**, *96*, 541−551.

(8) Chandrashekar, J.; Mueller, K. L.; Hoon, M. A.; Adler, E.; Feng, L.; Guo, W.; Zuker, C. S.; Ryba, N. J. P. T2Rs functions as bitter taste receptors. *Cell* **2000**, *100*, 703−711.

(9) Li, X.; Staszewski, L.; Durick, K.; Zoller, M.; Adler, E. Human receptors for sweet and umami taste. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 4692−4696.

(10) Shallenberger, R. S.; Acree, T. E. Molecular theory of sweet taste. *Nature* **1967**, *216*, 480−482.

(11) Kier, L. B. A molecular theory of sweet taste. *J. Pharm. Sci.* **1972**, *61*, 1394−1397.

(12) Castiglione Morelli, M. A.; Lelj, F.; Naider, F.; Tallon, M.; Tancredi, T.; Temussi, P. A. Conformation-activity relationship of sweet molecules. Comparison of aspartame and naphthimidazolesulfonic acids. *J. Med. Chem.* **1990**, *33*, 514−520.

(13) Kamphuis, J.; Lelj, F.; Tancredi, T.; Toniolo, C.; Temussi, P. A. SAR of sweet molecules: conformational analysis of two hypersweet and two conformationally restricted aspartame analogues. *Quant. Struct.-Act. Relat.* **1992**, *11*, 486−491.

(14) Yamazaki, T.; Benedetti, E.; Kent, D.; Goodman, M. Conformational requirements for sweet-tasting peptides and peptidomimetics. *Angew. Chem., Int. Ed. Engl.* **1994**, *33*, 1437−1451.

(15) Muller, G. W.; Madigan, D. L.; Culberson, J. C.; Walters, D. E.; Carter, J. S.; Klade, C. A.; DuBois, G. E.; Kellogg, M. S. High-potency sweeteners derived from β-amino acids. In *Sweeteners Discovery, Molecular Design and Chemoreception*; Walters, D. E., Orthoefer, F. T., DuBois, G. E., Eds.; American Chemical Society: Washington, DC, 1991; pp 113−125.

(16) DuBois, G. E.; Walters, D. E.; Kellogg M. S. The rational design of ultrahigh-potency sweeteners. In *Sweet Taste Chemoreception*; Mathlouthi, M., Kanters, J. A., Birch, G. G., Eds.; Elsevier: London, 1993; pp 237−267.

(17) Walters, D. E.; Prakash, I.; Desai, N. Active conformation of neotame and other high-potency sweeteners. *J. Med. Chem.* **2000**, *43*, 1242−1245.

(18) Tinti, J. M.; Nofre, C. Design of sweeteners. In *Sweeteners Discovery, Molecular Design and Chemoreception*; Walters, D. E., Orthoefer, F. T., DuBois, G. E., Eds.; American Chemical Society: Washington, DC, 1991; pp 88−99.

(19) Tinti, J. M.; Nofre, C. Why does a sweetener taste sweet? A new model. In *Sweeteners Discovery, Molecular Design and Chemoreception*; Walters, D. E., Orthoefer, F. T., DuBois, G. E., Eds.; American Chemical Society: Washington, DC, 1991; pp 206−213.

(20) Nofre, C.; Tinti, J. M. In quest of hyperpotent sweeteners. In *Sweet Taste Chemoreception*; Mathlouthi, M., Kanters, J. A., Birch, G. G., Eds.; Elsevier: London, 1993; pp 205−236.

(21) Nofre, C.; Tinti, J. M. Neotame: discovery, properties, utility. *Food Chem.* **2000**, *69*, 245−257.

(22) Nofre, C.; Tinti, J. M.; Glaser, D. Evolution of the sweetness receptor in primates. II. Gustatory responses of nonhuman primates to nine compounds known to be sweet in man. *Chem. Senses* **1996**, *21*, 747−762.

(23) Glaser, D.; Wanner, M.; Tinti, J. M.; Nofre, C. Gustatory responses of pigs to various natural and artificial compounds known to be sweet in man. *Food Chem.* **2000**, *68*, 375−385.

(24) Bassoli, A.; Drew, M. G. B.; Hattotuwagama, C. K.; Merlini, L.; Morini, G.; Wilden, G. R. H. Quantitative Structure−Activity Relationships of Sweet Isovanillyl Derivatives. *QSAR* **2001**, *20*, 3−16.

(25) Bassoli, A.; Merlini, L.; Morini, G.; Vedani, A. A three-dimensional receptor model for isovanillic sweet derivatives. *J. Chem. Soc., Perkin Trans. 2* **1998**, 1449−1454.

(26) Vedani, A.; Zbinden, P.; Snyder, J. P.; Greenidge, P. A. Pseudoreceptor modelling: the construction of three-dimensional receptor surrogates. *J. Am. Chem. Soc.* **1995**, *117*, 4987−4994.

(27) Gurrath, M.; Müller, G.; Höltje, H. D. Pseudoreceptor modelling in drug design. Applications of Yak and PrGen. *Perspect. Drug Discovery Des.* **1998**, *12−14*, 135−157.

(28) Stewart, J. J. P. A semiempirical molecular orbital program. *J. Comput.-Aided Mol. Des*. **1990**, *4*, 1−105.
(29) Vedani, A.; Hutha, D. W. A new force field for modeling metalloproteins. *J. Am. Chem. Soc*. **1990**, *112*, 4759−4767.
(30) Blaney, J. M.; Weiner, P. K.; Dearing, A.; Kollman, P. A.; Jorgensen, E. C.; Oatley, S. J.; Burridge, J. M.; Blake, J. F. Molecular mechanics simulation of protein−ligand interactions: binding of thyroid hormone analogues to prealbumin. *J. Am. Chem. Soc*. **1982**, *104*, 6424−6434.
(31) Searle, M. S.; Williams, D. H. The cost of conformational order: entropy changes in molecular associations. *J. Am. Chem. Soc*. **1992**, *114*, 10690−10697.
(32) Still, W. C.; Tempczyk, R. C.; Hawley, R. C.; Hendrickson, T. Semianalytical treatment of solvation for molecular mechanics and dynamics. *J. Am. Chem. Soc*. **1990**, *112*, 6127−6129.
(33) Guttad, L. W.; Shan, L.; Anchin, J. M.; Linthicum, D. S.; Edmuson, A. B. Local and transmitted conformational changes on complexation of an anti-sweetener fab. *J. Mol. Biol*. **1994**, *236*, 247−274.
(34) Guttad, L. W.; Shan, L.; Broomell, C.; Ramsland, P. A.; Fan, Z.; Anchin, J. M.; Linthicum, D. S.; Edmuson, A. B. The three-dimensional structure of a complex of a murine fab (NC10.14) with a potent sweetener (NC174): an illustration of structural diversity in antigen recognition by immunoglobulins. *J. Mol. Biol*. **2000**, *302*, 853−872.
(35) Vedani, A.; Zbinden, P. *PrGen 2.1, Tutorial*; Biographics Laboratory 3R: 1998; p 16.
(36) Nakamura, R.; Takahashi, M.; Amino, Y.; Sakamoto, K. Pharmacophore model of high-potency sweeteners. Presented at the 2nd IUPAC International Symposium on Sweeteners, Hiroshima, November 13−17, 2001; Poster P-02.
(37) Goodman, M.; Del Valle, J. R.; Amino, Y. The molecular basis of taste of peptide and peptidomimetics sweeteners. Presented at the 2nd IUPAC International Symposium on Sweeteners, Hiroshima, November 13−17, 2001; Poster PL-02.

JM020833V