Monatin, Its Stereoisomers and Derivatives: Modeling the Sweet Taste Chemoreception Mechanism

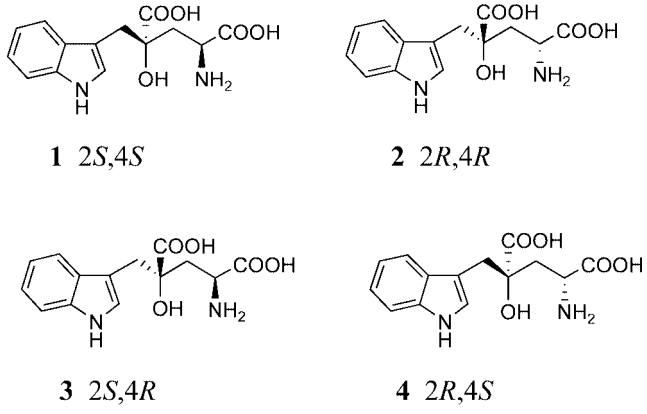
Angela Bassoli,*[a] Gigliola Borgonovo,[a] Gilberto Busnelli,[a] Gabriella Morini,[a] and Lucio Merlini[a]

Keywords: Sweet taste / Monatin / Molecular modeling / Sweet taste receptor / Pseudoreceptor modeling

The sweet natural compound monatin **1** has two stereogenic centers, and the 2*S*,4*S* absolute configuration has been attributed previously to the natural isomer. Among the four stereoisomers of monatin, three of them, particularly the 2*R*,4*R* isomer, tastes intensely sweet. The conformations of the four compounds have been studied by means of molecular modelling techniques. Both the diastereoisomeric forms show strong intramolecular hydrogen bonds which involve different functional groups and give rise to two different minimum energy conformations. The tertiary alcohol group in monatin seems to be indirectly involved in the generation of the taste, acting as an important contraint in generating the active conformation. The most important glucophores have been identified in the terminal $-NH_3^+$ and $-COO^-$ groups and in the indole ring by comparison with known topological models of sweet compounds and through the synthesis of appropriate derivatives in which some of these groups are lacking or modified. The relative affinity of each stereoisomer for its putative sweet taste receptor has been estimated semiquantitatively with the pseudoreceptor modelling technique. The predicted activity calculated with this technique is in good agreement with the experimental data and explains why the 2*R*,4*R* isomer (and not the natural 2*S*,4*S* isomer) is the sweetest of the series.

Introduction

Monatin **1** is a high-intensity sweet natural compound isolated from the roots of *Schlerochiton ilicifolius*, a spiny-leaved hardwood shrub growing in South Africa. Its structure [4-hydroxy-4-(indol-3-ylmethyl)glutamic acid] has been elucidated by Ackerman and coworkers.[1] The same authors also assigned the 2*S*,4*S* absolute configuration to the isolated compound. In a previous work,[2] we have described the synthesis of all the four stereoisomers of monatin and the results of the tasting trials. In our experiments, three of the four isomers tasted sweet, and unexpectedly, the 2*R*,4*R* isomer (compound **2**) was the sweetest of the series (Figure 1).



1 2*S*,4*S* **2** 2*R*,4*R*

3 2*S*,4*R* **4** 2*R*,4*S*

Figure 1. The four stereoisomers of monatin.

A similar finding has been reported recently in a patent:[3] all the four monatin isomers tested as sodium salts elicit a sweet taste, and the 2*R*,4*R* isomer is the sweetest. These results are of remarkable practical importance, since obtaining a mixture of isomers by synthesis is easier than the stereoselective preparation of single compounds. Interestingly, we also found a mixture of isomers during the analysis of an extraction sample of natural origin.[2] From a theoretical point of view, the low stereoselectivity observed in the binding of monatin isomers to the sweet taste receptor is intriguing, since often the stereoselectivity of taste chemoreception is very high. For instance, only the L, L isomer of the dipeptide aspartame is sweet, the others are tasteless or even bitter. It was particularly surprising that the sweetest monatin isomer is not the 2*S*,4*S* isomer, which was first isolated and identified as the natural sweet principle, but its enantiomer. Therefore, we started to study the structure–activity relationships for monatin and its stereoisomers in order to understand how each of them can bind to the putative sweet taste receptor with different strengths.

In developing structure–taste relationships of monatin and derivatives, we followed a three-step procedure: 1) se-

[a] Dipartimento di Scienze Molecolari Agroalimentari – DISMA, Sezione di Chimica, Università di Milano,
Via Celoria 2, 20133 Milano, Italia
Fax: +39-02-50316801
E-mail: angela.bassoli@unimi.it

arch for minimum energy conformation(s); 2) identification of important glucophores by use of topological models and through the synthesis of analogues; 3) modeling of pseudoreceptor.

Results and Discussion

The mechanism of action of sweet substances, as well as that of other flavors, has been under investigation for many years. From 2001, significant progress has been made in our understanding of the mechanism of sweet taste chemoreception, and almost contemporaneously, several independent groups of researchers[4–9] identified a new family of proteins of the GPCRs family (T1Rs), which form functional dimeric receptors that are able to bind several sweet compounds. Since that time, several papers contributed to unravel the biochemistry of the sweet taste receptor.[10,11] Very recently, the different functional roles of T1Rs subunits in the heteromeric sweet taste receptor has been proposed;[12] however, only the primary structure of the proteins is known, while the three-dimensional structure of the receptor and the binding mode of the active ligands to the receptor active site at a molecular level is not yet known.

In the last century, several different topological models were developed to describe the nature and the spatial arrangement of the glucophores of an ideal sweet compound and/or the recognition sites of the sweet taste receptor.[13–18]

These kinds of models are quite easy to use for qualitative interpretation of the data, but they do not give any quantitative information or any clue on the structure of the receptor active site. In the past, our research group has studied the (Q)SARs of many sweet tasting compounds by using topological and statistic models,[19–21] and recently, we developed a pseudoreceptor model[22] that is able to semiquantitatively predict the sweet taste of compounds belonging to many different chemical classes.

Conformational Studies

Monatin has a modified amino acid structure, in which it is possible to recognize a glutamic acid fragment and a tryptophan that has a hydroxyl group instead of an amino group; the two fragments are partially superimposed. It is possible to identify two polar blocks (corresponding to the terminal amino acid moiety and to the hydroxy acid at the C-4 position) and one hydrophobic area (corresponding to the indole ring). In contrast to the peptides, the principal skeleton of this molecule consists of a chain of C–C single bonds that elongates in space with no particular constraints beside the steric hindrance of the side substituents. Nevertheless, the presence of polar groups makes the formation of strong intramolecular hydrogen bonds possible, which characterize the calculated minimum energy conformations for the two diastereoisomers, *R***R** and *S***R** monatin (Figure 2).

In *R***R** monatin, a strong hydrogen bond (1.99 Å) involves the terminal $-NH_3^+$ group as the hydrogen donor and the oxygen atom of the C-4 hydroxyl group as the acceptor. This conformation is quite different from that obtained by Nakamura and coworkers[23] with the MM3 method, where no intramolecular hydrogen bonds were detected. The other diastereoisomer, *R**,*S** monatin, has a different minimum energy conformation with three intramolecular hydrogen bonds: one between the H atom of the hydroxyl group at C-4 and the carboxylate group at C-2, one between the H atom of COOH at C-4 and the O atom of the hydroxyl group at C-4; and another between the H atom of NH_3^+ at C-2 and the carboxylate group at C-2.

Identification of Glucophores

The identification of the glucophores (i.e. the functional groups involved in direct binding with the receptor) has often been made by measuring the distances between the functional groups and comparing them to those of known topological models. For sweet compounds, the models of Shallenberger and Acree, and Kier,[14,15] Temussi,[16] Goodman,[17] and the multipoint attachment model (MPA) of Tinti and Nofre[18] (which characterizes an ideal sweet compound with eight glucophores, which consists of four high affinity sites AH, B, G − corresponding to AH, B, X of Shallenberger–Acree–Kier − and D, and four secondary sites) are among the most utilized for this purpose. This simple method can be useful but suffers from many disadvantages: in fact, these distances are easily calculated when an atom in the structure is clearly identified as the "point" glucophore, but the calculation is very approximate when more undefined areas are considered. In monatin, the AH

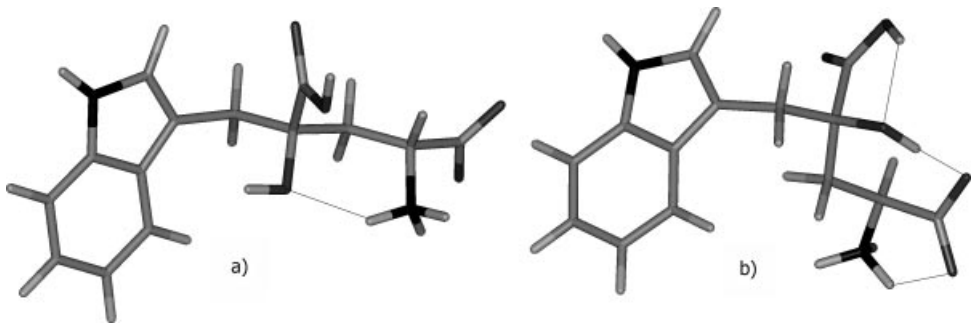


Figure 2. Minimum energy conformations: a) 2*R**,4*R** monatin; b) 2*R**,4*S** monatin.

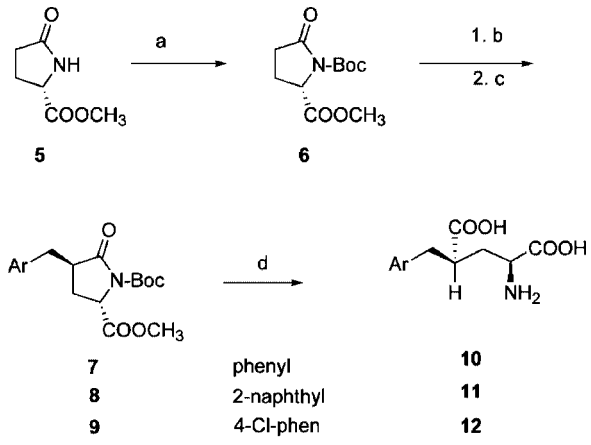
and B glucophores could likely correspond to the amino acid terminal group or to the quaternary hydroxy acid at C-4, whereas the indole ring is likely to correspond to the G hydrophobic glucophore. Owing to its magnitude, a great uncertainty occurs in the measure of the AH–G or B–G distance; depending on which atom is considered as the center of the G area, the distances could correspond to the central or terminal polar groups. Therefore, this method does not allow the unambiguous identification of the glucophores.

Another method is to identify glucophores by testing the importance of each functional group in the interaction with the receptor. Analogues of the target molecule that lack some functional groups or with modified ones can be synthesized and submitted to biological tests. If the activity is suppressed or reduced, it is possible to deduce an important role in activity. We applied this method to monatin and synthesized some analogues with modified functional groups and some that lacked functional groups.

Synthesis of Monatin Analogues

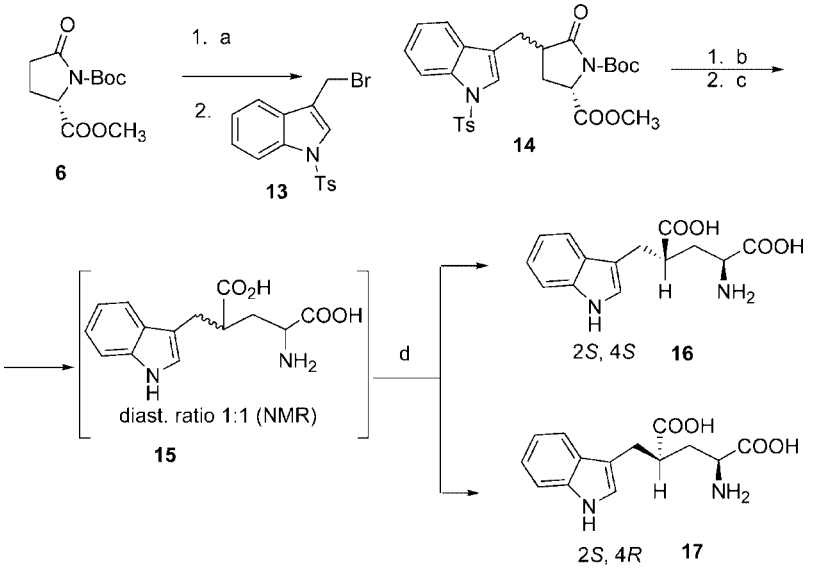
First we decided to plan the synthesis of some derivatives that lack the hydroxyl group. This group seems, in fact, to be important in the molecule, and the deoxy analogues are easier to access by synthesis. As a second criterion, in some cases, we substituted the indole ring (which is usually unstable under several reaction conditions and therefore needs special protective groups) with other hydrophobic aromatic rings which could, theoretically, give a similar interaction with the putative receptor. For this purpose, we used phenyl, naphthyl and 4-Cl-phenyl rings: the last one was chosen since it is known that 4-Cl-tryptophan is sweeter than tryptophan itself.

The new derivatives were obtained starting from *S*-pyroglutamic acid methylester **5**, which is a cheap chiral precursor that is able to undertake stereoselective alkylation reactions at the C-4 position.[24] The general procedure is shown in Scheme 1.

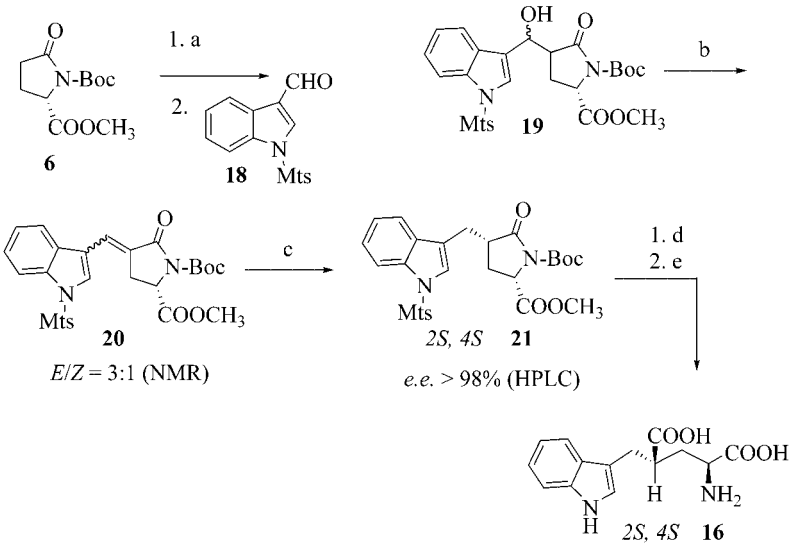


Scheme 1. Synthesis of deoxy monatin analogues by alkylation of pyroglutamate.

The alkylation reaction on *N*-Boc methylpyroglutamate **6** was diastereoselective, as expected from analogous alkylations described in the literature with benzyl bromide, 2-bromomethylnaphtalene, and 4-Cl-benzyl bromide, and gave compounds **7**, **8**, and **9**. The stereochemistry of these compounds was assigned as 2*S*,4*R* on the basis of comparison with literature data and by the NMR spectroscopic



Scheme 2. Synthesis of deoxy derivatives by aldol condensation.



a: LiHMDS, THF, –78 °C, BF_3, Et_2O; b: MsCl, DCM, Et_3N; c: H_2, PtO_2; d: HCl 1 N, reflux; e: MeOH, H_2O, K_2CO_3

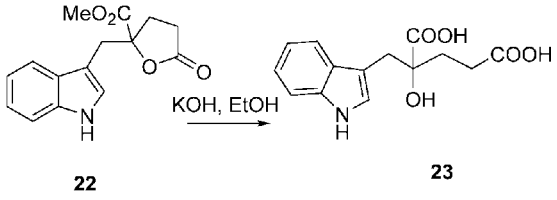
Scheme 3. Synthesis of compound **16** by aldol condensation.

spectra which showed a *trans* relationship between the substituents at the C-2 and C-4 positions. The acidic hydrolysis of the lactam ring gave derivatives **10**, **11**, and **12** in one step, respectively, without racemization of the chiral centers. These deoxy derivatives have the same stereochemistry of natural 2*S*,4*S* monatin **1** at both stereocenters (the absolute configuration in these analogues is 2*R*,4*S* as a result of the changes in the priority of the substituents).

A similar approach was used for the synthesis of the indole derivatives **16** and **17** (Scheme 2).

In this case the alkylation with bromide **13** was not selective; the resulting reaction mixture **14** was deprotected in two steps and chromatographed to give a mixture of diastereoisomers **15** in about a 1:1 ratio as shown by NMR spectroscopy. A satisfactory separation of the diastereoisomers **16** and **17** was achieved by RP-HPLC on an analytical scale only. Compound **16** was then obtained as a single enantiomer by a different route based on the aldol condensation of protected pyroglutamic acid **6** and aldehyde **18** (Scheme 3). The resulting mixture of aldols **19** was dehydrated to give the unsaturated product **20** as a 3:1 mixture of *E* and *Z* isomers (NMR spectroscopy), which was hydrogenated stereoselectively to give compound **21**. This latter compound was hydrolyzed and deprotected to give **16** as a single diastereoisomer. The attribution of the 2*S*,4*S* configuration to this product was made on the basis of comparison with analogous reactions described in the literature,[25] and confirmed by NMR mono- and bidimensional NOE experiments on compound **21**, that showed a *cis* relationship between the two hydrogen atoms in the positions C-2 and C-4.

A monatin analogue **23** that lacks the terminal amino group was prepared as a racemate by hydrolysis of lactone **22**[26] (Scheme 4).



Scheme 4. Synthesis of deamino derivative **23**.

Tasting Trials

All the new monatin analogues synthesized were submitted to preliminary tasting trials according standard procedures.[27] The related compound **24** (indole lactic acid) was also tested and is shown in Figure 3 for comparison.

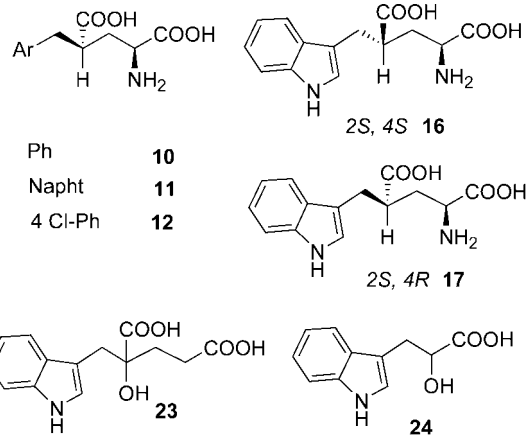


Figure 3. Tasteless monatin analogues.

None of the derivatives tasted sweet in the tasting trials. In particular, compound **17** was not sweet: this derivative has an identical structure and stereochemistry to that of

natural monatin, except for the absence of the hydroxyl group. This functional group seems therefore to be essential for the sweet taste of monatin, and this is confirmed by the observation that the biological activity disappeared in all the new deoxy derivatives. The minimum energy conformation of deoxymonatin **17** was also calculated and is shown in Figure 4.

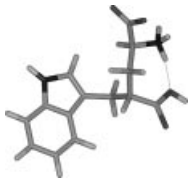


Figure 4. Minimum energy conformation of deoxyderivative **17**.

The absence of the hydroxyl group changes the minimum energy conformation; in this case, an intramolecular hydrogen bond (1.85 Å) is formed between the terminal amino group and the oxygen atom of the carboxylic acid at the C-5 position. As a result, the relative topology of potential glucophores is completely different from that observed for *S*,*S* monatin. A similar conformation is also observed for all the other deoxy derivatives (**10–12** and **16**). It is therefore possible that the hydroxyl group plays an important role in the taste of monatin, not as a glucophore itself, but indirectly, by introducing an important constraint which generates an active conformation. Other evidence that supports this hypothesis is the fact that indole lactic acid **24** is itself not sweet, as we would expect if the main AH and B glucophores corresponded to the hydroxy acid moiety. On the other hand, the terminal amino group is also important for the taste, since the corresponding derivative **23**, which lacks this group, is completely inactive (tasteless).

Pseudoreceptor Modeling

The results seen above seem to indicate that the main glucophores in monatin likely correspond to the indole (G) and the terminal amino acid fragment (AH and B). We started with this assumption and compared the obtained minimum energy conformation of the monatin stereoisomers with the Tinti and Nofre multipoint attachment model. The four compounds **1**–**4** have been compared to the MPA model by overlapping the AH, B, and G sites with the amino, C-2 carboxylate and indole groups, respectively, to find the best possible superimposition (Figure 5).

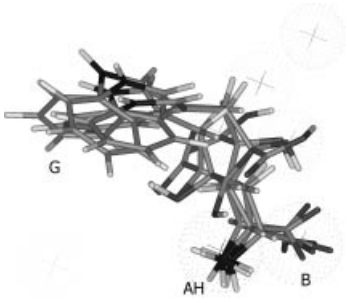


Figure 5. The four stereoisomers of monatin superimposed on the MPA model. Glucophores are indicated by grey-dot spheres; only AH, B and G are labeled for clarity.

Qualitatively, all the compounds appear to fit the model. In particular, when the main glucophores $-NH_3^+$ and $-COO^-$ are set as AH and B sites, the indole ring is always lying in the large hydrophobic area corresponding to the G site, so that it is hard to predict any difference in activity between the compounds. In fact, only a qualitative comparison is possible with this methodology as well as with similar methods, e.g. with the Goodman model,[17] whereas more refined information is required in order to give a theoretical explanation of the observed biological activity.

Therefore, we have applied the so-called pseudoreceptor modeling technique to obtain some semiquantitative information on the binding energy between the ligands and the putative receptor. For this aim, we used the pseudoreceptor model developed by our group to be able to explain and predict the affinity of several sweet-tasting compounds belonging to different classes.[22] The four monatin stereoisomers in their minimum energy conformation have been inserted in the pseudoreceptor in positions consistent with the MPA model (utilized as a tool also during the development of the general pseudoreceptor model to overlap compounds belonging to different classes for which it was not possible to superimpose on the basis of the same structural feature) and then subjected to free ligand relaxation, while the pseudoreceptor was kept rigid. The predicted free energies of binding were then compared with the experimental values. It is important to underline that the individual conformations of the four isomers after free ligand relaxation are not significantly different from the starting ones. Table 1 shows the comparison between the predicted free energies of ligand binding, $\Delta G°$, for isomers **1**–**4** and the experimental activity obtained by sensory evaluation.

Table 1. Taste activity of the four stereoisomers by sensory analysis and molecular modeling.

	Reference	Monatin stereoisomers			
Compound		**1**	**2**	**3**	**4**
Absolute configuration		2*S*,4*S*	2*R*,4*R*	2*S*,4*R*	2*R*,4*S*
Relative Sweetness (RS)	this paper	350	1000	tasteless	250
	[3][a]	50	2700	300	1300
	[1]	1200	–	–	–
$\Delta G°_{binding}$ (pseudoreceptor)		–10.15	–13.02	–11.46	–11.32

[a] Tested as Na or Na/NH_4^+ salts.

Among the four stereoisomers of monatin the 2*R*,4*R* isomer has the highest affinity for the pseudoreceptor. Therefore, the taste activity predicted by the pseudoreceptor method is in good agreement with the experimental finding that this isomer, (and not its natural enantiomer 2*S*,4*S*) is the sweetest of the series. Interestingly, the highest calculated $\Delta G°$ (–10.15) is that of 2*S*,4*S* monatin, and this stereoisomer is therefore predicted to be the less active of the series. This result is in contrast with the sensory evaluation experimental data, where we found compound **3** to be the least active – it was tasteless at the concentration used. Nevertheless, the result is in agreement with the data indicated by the more recent sensory analysis experiments made by the Ajinomoto Co. Inc.[3]

The pseudoreceptor approach also gives some insights into the binding mode of the active compounds to the pseudoreceptor. Figure 6 shows compound **2** inside the pseudoreceptor and the interactions between the ligand and the surroundings amino acid residues.

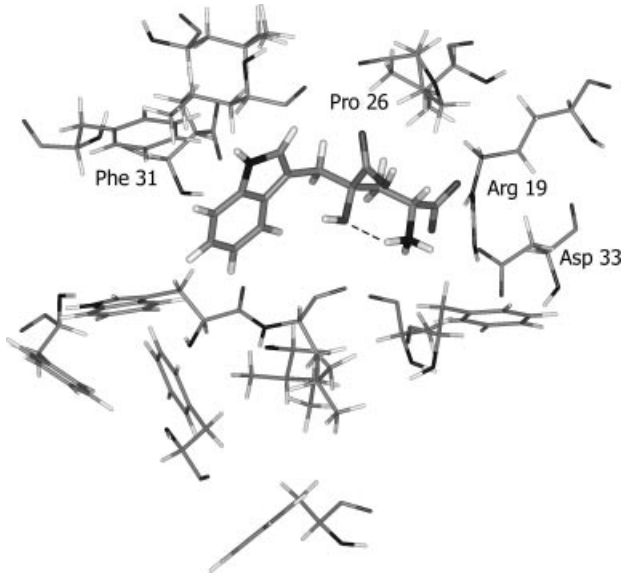


Figure 6. 2*R*,4*R* monatin (compound **2**) in the pseudoreceptor cavity.

The most important interactions in determining the binding affinity of the compounds are (in decreasing order of calculated energy) those involving the charged terminal amino group (Asp, 33) and C-2 carboxylate (Arg, 19); followed by the C-4 carboxylic acid (Pro, 26) and its hydrophobic interaction with indole (Phe, 31). In this model, the hydroxyl group of monatin lies in the center of the cavity and does not seem to be involved in strong interactions with any amino acid residue.

The intramolecular hydrogen bond between the –OH and the $-^{+}NH_3$ groups is retained inside the pseudoreceptor. This situation confirms that the most important interactions should involve the terminal charged amino acid groups and indole as glucophores, while the alcohol group is of key importance in keeping the active conformation.

Conclusions

Generally, stereochemistry is a strong discriminating factor in the chemoreception mechanism of taste recognition. The sweet natural compound 2*S*,4*S* monatin and its three stereoisomers are a curious exception to this rule, since at least three of them (or all four, following other authors) tasted sweet in sensory analysis experiments. In these molecules, the absolute configuration of the two stereogenic centers is not so important to the binding with the receptor. In particular, the configuration at C-4 is not crucial; this could depend on the fact that the substituents on this carbon atom are not directly involved in the interaction with the receptor. In fact, in the minimum energy conformation, the molecule is quite extended and the central groups are sterically hindered; moreover, the tertiary hydroxyl group is engaged in a strong intramolecular hydrogen bond that prevents other interactions. The low stereoselectivity in chemorecognition of monatin isomers was not easily foreseen, since these molecules are structurally related to tryptophan and this compound, as many amino acids, is sweet in its *R*-D form but bitter in the *S*-L form and shows high stereochemical discrimination in taste chemoreception.

These results therefore suggest that all factors have to be considered when predicting the chemosensory properties of new compounds on the basis of the structure and stereochemistry of known leads. The pseudoreceptor modeling calculations are useful in the prediction of the relative potency of sweet compounds and in the interpretation of experimental data from sensory analysis, since they give semiquantitative information on the binding energy and some insights into the possible interaction of active ligands with the putative receptor protein until the detailed three-dimensional structure of the receptor active site(s) is clarified.

Experimental Section

Molecular Modeling

Three-dimensional molecular models were built on a Silicon Graphics O2, using the programs InsightII and Discover, 97.0 (Accelrys, 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 by molecular dynamics. Monatin and other amino acid derivatives were minimized in vacuo as zwitterions as they are supposed to be in solution. For atomic partial charges of the ligand atoms, Mulliken charges calculated on the minimized structures using the MOPAC program[28] with the MNDO Hamiltonian were used.

For pseudoreceptor modeling, the program PrGen 2.1[29] (SIAT Biographics Laboratory, Basel, CH) was used. Molecules in their lowest-energy conformation were imported in PrGen and reminimized with the Yeti force field.[30]

The equation used in PrGen to correlate free energies of ligand binding $\Delta G°$ with relative sweetness is: $\Delta G°_{exp.} = RT \cdot \ln K_d$, where $K_d = K_d(\text{sucrose})/\text{RS}$. Details on PrGen and on the general pseudoreceptor model for sweet compounds developed by our group can be found in ref.[22]

Synthesis of Monatin Derivatives

Reagents were of commercial grade purity; solvents were dried with standard procedures; m.p. are uncorrected. ^{1}H NMR spectra were recorded with Bruker AMX-300 and AMX-600 instruments, by using TMS as internal standard; *J* values are given in Hz. Mass spectra were recorded with a Finnigan 4021 spectrometer; HPLC experiments were made with a Varian PROSTAR instrument; HPLC-MS analysis were made with an Agilent SL 1100 series instrument with ESI and ion trap system. HRMS were made with a Bruker Daltonics APEX II ICR-FTMS instrument, by using the ESI ionization mode. Optical rotation data were recorded with a Jasco J 810 spectropolarimeter.

1-*tert*-Butyl 2-Methyl (2*S*)-5-Oxopyrrolidine-1,2-dicarboxylate (6): White solid (7.19 g, 65%). M.p. 67–68. ^{1}H NMR ($CDCl_3$): δ = 1.48 (s, 9 H, Boc), 1.98–2.10 (m, 1 H), 2.25–2.65 (m, 3 H), 3.78 (s, 3 H, OCH_3), 4.63 (dd, ^1J = 3.7, ^2J = 11.2 Hz, 1 H, 5-H) ppm. IR (nujol): 2990, 1800, 1750, 1700 cm^{-1}.$[\alpha]_D^{25}$ = –39.61 (*c* = 1.02, EtOH). *m*/*z* (%): 244 [M$^+$] (5), 188 (40), 84 (80), 59 (100).

General Procedure for the Alkylation (Scheme 1): A 1 M solution of LiHMDS (13.57 mL, 13.57 mmol) was added to **6** (3.0 g, 12.34 mmol) in dry THF under nitrogen at –78 °C. After 1 h, the bromide was added dropwise (10% molar excess) and stirred for 2.5 h. The reaction mixture was quenched with saturated NH_4Cl solution and extracted with diethyl ether. The organic phase was dried on Na_2SO_4, filtered, and the solvents evaporated to dryness. The product was purified by flash chromatography.

1-*tert*-Butyl 2-Methyl (2*S*,4*R*)-4-Benzyl-5-oxopyrrolidine-1,2-dicarboxylate (7): White solid (53%, *ee* > 98%). M.p. 98–99 °C. ^{1}H NMR ($CDCl_3$): δ = 1.50 (s, 9 H, Boc), 2.08 (m, 2 H), 2.58 (m, 1 H), 2.85 (m, 1 H), 3.30 (m, 1 H), 3.70 (s, 3 H, OCH_3), 4.40 (dd, ^1J = 3.4, ^2J = 8.7 Hz, 1 H, 2-H), 7.15–7.30 (m, 6 H, ar) ppm. *m*/*z* (%): 334 [M$^+$], 303 (5), 278 (20), 233 (100), 174 (80), 129 (10), 91 (30), 57 (70). $[\alpha]_D^{25}$ = –38,7 (*c* = 0.78, $CHCl_3$).

1-*tert*-Butyl 2-Methyl (2*S*,4*R*)-4-(2-Naphthylmethyl)-5-oxopyrrolidine-1,2-dicarboxylate (8): White solid (21%, *ee* > 98%). M.p. 143–146 °C. ^{1}H NMR ($CDCl_3$): δ = 1.48 (s, 9 H, Boc), 2.08 (m, 2 H), 2.88 (m, 1 H), 3.02 (m, 1 H), 3.48 (m, 1 H), 3.75 (s, 3 H, OCH_3), 4.50 (dd, ^1J = 3.7, ^2J = 7.6 Hz, 1 H, 2-H), 7.30–7.87 (m, 7 H, ar) ppm. *m*/*z* (%): 383 [M$^+$] (15), 327 (10), 283 (60), 224 (20), 167 (20), 141 (100). $[\alpha]_D^{25}$ = –9.7 (*c* = 0.89, $CHCl_3$).

1-*tert*-Butyl 2-Methyl (2*S*,4*R*)-4-(4-Chlorobenzyl)-5-oxopyrrolidine-1,2-dicarboxylate (9): White solid (32%, *ee* > 98%). M.p. 128 °C. ^{1}H NMR ($CDCl_3$): δ = 1.50 (s, 9 H, Boc), 1.94–2.12 (m, 2 H), 2.70 (m, 1 H), 2.92 (m, 1 H), 3.25 (m, 1 H), 3.72 (s, 3 H, OCH_3), 4.50 (dd, ^1J = 2.0, ^2J = 9.1 Hz, 1 H, 2-H), 7.10 (d, 2 H), 7.30 (d, 2 H) ppm. *m*/*z* (%): 267 (15), 216 (5), 210 (25), 208 (50), 125 (20), 116 (30). $[\alpha]_D^{25}$ = –31.7 (*c* = 0.5, $CHCl_3$).

General Procedure for Hydrolysis: A solution of **7**, **8**, or **9** was refluxed in HCl (6 N) for 4 h. The reaction was monitored with TLC and the spots detected with ninhydrine. The solvent was evaporated at reduced pressure.

(2*S*,4*R*)-2-Amino-4-benzylpentanedioic Acid (10): The crude oil was purified through an ion exchange resin (Amberlist IRA-68) eluting with aqueous NH_3 (0.05 N) to give a white solid (0.264 g, 75%). ^{1}H NMR (DMSO): δ = 1.75 and 2.10 (two m, 2 H), 2.78–3.00 (m, 3 H), 3.80 (m, 1 H, 2-H), 7.15–7.30 (m, 5 H, ar), 8.30 (br. s, 2 H, -NH_2) ppm. *m*/*z* (%): 219 (60), 201 (20), 174 (50), 146 (15), 128 (20), 117 (20), 91 (100). HRMS (ESI, pos): 238.10755 (M + 1); calcd. 237.10011. $[\alpha]_D^{25}$ = –14 (*c* = 0.45, MeOH).

(2*S*,4*R*)-2-Amino-4-(2-naphthylmethyl)pentanedioic Acid (11): White solid (0.138 g, 87%). M.p. (aq. ethanol) 158–160 °C. ^{1}H NMR (DMSO): δ = 1.78 (m, 1 H), 2.10 (m, 1 H), 2.78–3.00 (m, 3 H), 3.78–3.85 (m, 1 H, 2-H), 7.30–7.81 (m, 7 H, ar), 8.20 (broad s, 2 H, NH_2) ppm. *m*/*z* (%): 269 [M$^+$ – H_2O] (40), 224 (10), 181 (10), 152 (20), 141 (100), 115 (50), 84 (40). $[\alpha]_D^{25}$ = +16.4 (*c* = 0.53, MeOH). HRMS (ESI, pos): 288.12256 (M + 1); calcd. 287.11576.

(2*S*,4*R*)-2-Amino-4-(4′-chlorobenzyl)pentanedioic Acid (12): The crude product was purified with a cation exchange resin (Amberlist IR-120) eluting with NH_3 solution (1 N). Recrystallization from ethanol gave the pure product (0.30 g, 89%). ^{1}H NMR (600 MHz,$CDCl_3$): δ = 2.14 (m, 1 H, 3a-H), 2.29 (m, 1 H, 3b-H), 2.68–2.75 (m, 2 H, 4-H and 5a-H), 3.11 (m, 1 H, 5b-H), 3.98 (dd, 1 H, 2-H), 7.00–7.25 (m, 4 H, ar) ppm. *m*/*z* (%): 208 (60), 143 (20), 127 (85), 125 (100). $[\alpha]_D^{25}$ = –4.0 (*c* = 0.2, MeOH).

3-(Bromomethyl)-1-[(4-methylphenyl)sulfonyl]-1*H*-indole (13): Compound **13** was obtained from indole-3-carboxyaldehyde in three steps by protection of the nitrogen atom with *p*-toluenesulphonyl chloride, reduction to alcohol with $NaBH_4$, and reaction with bromine and PPh_3 in CCl_4. White solid (1.30 g, 64%). M.p. 132 °C (dec.). ^{1}H NMR ($CDCl_3$): δ = 2.35 (s, 3 H, CH_3), 4.60 (s, 2 H, CH_2Br), 7.20–7.40 (m, 5 H, ar), 7.65–7.90 (m, 4 H, ar) ppm. *m*/*z* (%): 365 [M$^+$] (2), 284 (100), 155 (20), 129 (25), 102 (20), 91 (50), 65 (20).

(2*S*,4*S*)- and (2*S*,4*R*)-4-[Hydroxy(1-tosylsulfonil-1*H*-indol-3-yl)-methyl]-5-oxopyrrolidine-1,2-dicarboxylic Acid (14, Diastereoisomeric Mixture): White solid (23%). M.p. 81–82 °C.^{1}H NMR ($CDCl_3$): δ = 1.51 (s, 18 H, 2 Boc), 1.68 (m, 1 H, 4-H), 1.95 (m, 3 H), 2.34 (s, 6 H, 2 CH_3), 2.75 (m, 2 H), 2.94 (m, 2 H), 3.33 (m, 2 H), 3.74 (s, 3 H, OCH_3), 3.77 (s, 3 H, OCH_3), 4.49 (m, 2 H, two 2-H), 4.1–8.0 (m, 18 H, ar and indole) ppm. *m*/*z* (%): 526 [M$^+$] (22), 426 (80), 284 (100), 272 (90), 211 (30), 168 (22), 155 (60), 130 (45), 91 (90), 65 (18), 57 (30). HRMS (ESI, pos): 549.16416 [M + Na]; calcd. 526.17737.

(2*S*,4*R*)-2-Amino-4-(1*H*-indol-3-ylmethyl)pentanedioic Acid (16): White solid. M.p. 128–130 °C. ^{1}H NMR (aceton-d6): δ = 1.89 (ddd, ^1J = 12.7, ^2J = 8.0, ^3J = 8.9 Hz, 1 H, 3a-H), 2.64 (ddd, ^1J = 12.7, ^2J = 8.0, ^3J = 8.6, 1 H, 3b-H), 2.87 (ddd, ^1J = 14.4, ^2J = 9.6, ^3J = 1.0, 1 H, 5a-H), 2.92 (m, 1 H, 4-H), 3.29 (ddd, ^1J = 14.4 Hz, ^2J = 3.7, ^3J = 1.3, 1 H, 5b-H) 4.29 (t, ^1J = 8.0 Hz, 1 H, 2-H), 6.9–6.5 (m, 6 H, indole) ppm. *m*/*z* (%): (LCMS – ESI): 276 [M$^+$] (35), 258 (55), 130 (100). $[\alpha]_D^{25}$ = +41.5 (*c* = 0.27, MeOH).

(2*S*, 4*S*)-2-Amino-4-(1*H*-indol-3-ylmethyl)pentanedioic Acid (17): ^{1}H NMR (aceton-d6): δ = 1.92 (m, 1 H, 3a-H), 2.27 (m, 1 H, 3b-H), 2.91 (dd, ^1J = 5.4, ^2J = 14.2 Hz, 1 H, 5a-H), 2.99 (m, 1 H, 4-H), 3.14 (m, 1 H, 5b-H) 4.27 (t, ^1J = 8.8, 1 H, 2-H), 7.1–8.1 (m, 6 H, indole) ppm. *m*/*z* (%): (LCMS-ESI): 276 [M$^+$] (5), 258 (100), 130 (50), 102 (10). $[\alpha]_D^{25}$ = +35.5 (*c* = 0.38, MeOH).

Compounds **16** and **17** were separated by RP-HPLC eluting under gradient conditions from eluent A (H_2O + 1% AcOH) to eluent B (MeOH) in 10 minutes, flow 1 mL min^{-1}. Retention times (min): 12.92 (compound **16**) and 12.26 (compound **17**).

Compound **16** was also obtained as follows: Compound **21** (0.520 g, 0.9 mmol) [see below], a mixture of MeOH/H_2O 3:1 (15 mL), and K_2CO_3 were refluxed for 4 h. After removal of the solvent, the crude product was purified by flash chromatography (CH_2Cl_2/MeOH/NH_3/H_2O, 8:3:0.2:0.2). White solid (0.12 g, 38%). Spectral data identical to those reported above.

1-(Mesitylsulfonyl)-1*H*-indole-3-carbaldehyde (18): Obtained from 1H-indole-3-carbaldehyde by reaction with mesityl chloride and NaH in dry DMF at 0 °C. White solid (44%). M.p. 129–131 °C. ^{1}H NMR ($CDCl_3$): δ = 2.23 (s, 3 H, CH_3), 2.49 (s, 6 H, two CH_3),

6.94–7.30 (m, 7 H, indole and ar), 10.50 (s, 1 H, CHO) ppm. *m/z* (%): 327 [M^+] (100), 183 (10), 119 (50), 86 (30), 84 (45).

(2*S*)-4-[Hydroxy(1-mesitylsulfonyl-1*H*-indol-3-yl)methyl]-5-oxopyrrolidine-1,2-dicarboxylic Acid (19): A 1 M solution of LiHMDS in THF (7.3 mmol) was added to a solution of **6** (1.48 g, 6.09 mmol) in dry THF whilst stirring at –78 °C under nitrogen. The reaction mixture was stirred for 1 h at –78 °C prior to addition of a solution of **18** (2.2 g, 6.75 mmol) in THF and $BF_3{\cdot}Et_2O$ (2.16 mL, 6.75 mmol). The mixture was stirred for 2.5 h at –78 °C, quenched with saturated NH_4Cl, and extracted with diethyl ether. The combined organic phases were dried with Na_2SO_4, filtered, and the solvents evaporated in vacuo. Flash chromatography (hexane/ethyl acetate, 45:55) gave mixture of **19** (0.89 g, 26%). ^1H NMR ($CDCl_3$): δ = 1.51 (s, 18 H, 2Boc), 1.85 (m, 1 H, 4a-H), 1.99 (m, 1 H, 4b-H), 2.31 (s, 3 H, CH_3), 2.53 (s, 6 H, two CH_3), 3.29 (m, 1 H, 3-H), 3.75 (s, 3 H, OCH_3), 4.58 (dd, $^{1,2}J$ = 9.2, 1 H, 5-H), 4.8 (s, 1 H, OH), 5.05 (d, ^1J = 9.2, 1 H, 6-H), 6.95 (s, 2 H, ar), 7.2–7.8 (m, 5 H, indole) ppm. *m/z* (%): 571 [M + 1] (5), 470 (3), 465 (10), 328 (15), 209 (10), 183 (18), 143 (40), 119 (100), 91 (100).

(2*S*)-4-[Hydroxy(1-mesitylsulfonyl-1*H*-indol-3-yl)methylene]-5-oxopyrrolidine-1,2-dicarboxylic Acid (20): The aldol mixture **19** (0.70 g, 1.23 mmol) was dissolved in dry CH_2Cl_2 and treated with methanesulphonyl chloride (0.1 mL, 1.35 mmol) and TEA (1.46 mL, 13.5 mmol). After 2 d at room temperature, the reaction mixture was quenched with water and extracted with CH_2Cl_2. The organic layer was dried with Na_2SO_4, and the solvent evaporated under reduced pressure. After purification of the crude product by flash chromatography (hexane/ethyl acetate, 6:4), a white solid product (0.39 g, 57%) was obtained as a mixture of *E*/*Z* isomers. M.p. 85–88 °C. ^1H NMR ($CDCl_3$): δ = 1.55 (s, 18 H, Boc *E* and *Z*), 2.27 (s, 3 H, *p*CH_3 *Z*), 2.3 (s, 3 H, *p*CH_3 *E*), 2.52 (s, 6 H, OCH_3 of *E*), 2.58 (s, 6 H, OCH_3 *Z*), 2.85–2.9 (m, 2 H, 4a-H *E* and *Z*), 3.25–3.32 (m, 2 H, 4b-H *E* and *Z*), 3.78 (s, 3 H, OCH_3 *Z*), 3.8 (s, 3 H, OCH_3 *E*), 4.68 (dd, ^1J = 4.5, ^2J = 10.0 Hz, 1 H, 5-H *Z*), 4.78 (dd, ^1J = 3.5, ^2J = 10,0 Hz, 1 H, 5-H *E*), 6.95 (s, 2 H, ar of *Z*), 6.97 (s, 2 H, ar of *E*), 7.1 (br. t, 1 H, CH=C *Z*), 7.19–7.31 (m, 8 H, indole *Z* and *E*), 7.77 (s, 1 H, 8-H *E*), 7.78 (s, 1 H, 8-H *Z*), 7.81 (t, ^1J = 3.0, 1 H, CH=C, *E*) ppm. *m/z* (%): 552 [M^+] (30), 452 (100), 393 (10), 327 (10), 269 (40), 209 (30), 119 (30), 71 (20). $[\alpha]_D^{25}$ = +23 (*c* = 1.0, $CHCl_3$).

(2*S*,4*S*)-4-[Hydroxy(1-mesitylsulfonyl-1*H*-indol-3-yl)methyl]-5-oxopyrrolidine-1,2-dicarboxylic Acid (21): M.p. 58–60 °C. ^1H NMR ($CDCl_3$): δ = 1.50 (s, 9 H, Boc), 1.72 (m, 1 H, 3a-H),2.30 (s, 3 H, CH_3), 2.52 (s, 6 H, two CH_3), 2.43 (m, 1 H, 3b-H), 2.84 (m, 1 H, CH_2), 2.95 (m, 1 H, 4-H), 3.40 (dd, 1 H, CH_2), 4.50 (dd, ^1J = 9.6, ^2J = 11.2, 1 H), 6.97 (s, 2 H), 7.20 (m, 2 H), 7.31 (m, 1 H), 7.4 (s, 1 H), 7.5 (m, 1 H) ppm. *m/z* (%): 554 [M^+] (20), 522 (10), 396 (10), 318 (10), 266 (10), 242 (100) 183 (10). $[\alpha]_D^{25}$ = +18.2 (*c* = 1.0, $CHCl_3$). The absolute configuration was assessed by NMR mono- and bi-dimensional NOE experiments (Supplementary Information available).

2-Hydroxy-2-(1*H*-indol-3-ylmethyl)pentanedioic Acid (23): Lactone **22**[26] (71 mg, 0.26 mmol) was stirred in aqueous EtOH and KOH (0.5 mmol) at room temperature for 2 d. Acidic workup followed by extraction with ethyl acetate gave compound **23**. White solid (25 mg, 30%). ^1H NMR ([D_6]acetone): δ = 2.40 (m, 4 H, CH_2), 3.45 (m, 2 H, CH_2-indole), 7.10–7.70 (m, 5 H, indole), 10.20 (br. s, 1 H, NH) ppm. *m/z* (%): 287 [M^+] (10), 259 (5), 229 (5), 200 (5), 155 (20), 149 (30).

Acknowledgments

We thank MIUR (Italian Ministry for Research and the University) and University of Milano for financial support (programs COFIN 2002 and FIRST 2003); Fondazione Fratelli Confalonieri (Milano) for a PhD fellowship to G. Busnelli; Damien Kelly, National University of Ireland, Galway, for contribution to experimental work.

[1] R. Vleggaar, L. G. J. Ackerman, P. S. Steyn, *J. Chem. Soc., Perkin Trans. 1* **1992**, 3095–3098.
[2] A. Bassoli, G. Borgonovo, G. Busnelli, M. G. B. Drew, G. Morini, *Eur. J. Org. Chem.* **2005**, 1652–1658.
[3] WO Pat. 2003045914, **2003**.
[4] M. Max, Y. G. Shanker, L. Huang, M. Rong, Z. Liu, F. Campagne, H. Weinstein, S. Damak, R. F. Margolskee, *Nature Genetics* **2001**, *28*, 58–63.
[5] J. P. Montmayeur, S. D. Liberles, H. Matsunami, L. B. Buck, *Nature Neuroscience* **2001**, *4*, 492–498.
[6] M. Kitagawa, Y. Kusakabe, H. Miura, Y. Ninomiya, A. Hino, *Biochem. Biophys. Res. Commun.* **2001**, *283*, 236–242.
[7] G. Nelson, M. A. Hoon, J. Chandrashekar, N. J. P. Ryba, C. S. Zuker, *Cell* **2001**, *106*, 381–390.
[8] X. Li, M. Inoue, D. R. Reed, T. Hunque, R. B. Puchalski, M. G. Tordoff, Y. Ninomiiya, G. K. Beauchamp, A. A. Bachmanov, *Mamm. Genome* **2001**, *12*, 13–16.
[9] E. Sainz, J. N. Korley, J. F. Battey, S. L. Sullivan, *J. Neurochem.* **2001**, *77*, 896–903.
[10] S. Damak, M. Rong, K. Yasumatsu, Z. Kokrashvili, V. Varadarajan, S. Zou, P. Jiang, Y. Ninomiya, R. F. Margolskee, *Science* **2003**, *301*, 850–853.
[11] G. Q. Zhao, Y. Zhang, M. A. Hoon, J. Chandrashekar, I. Erlenbach, N. J. Ryba, C. S. Zuker, *Cell* **2003**, *115*, 255–266.
[12] H. Xu, L. Staszewski, H. Tang, E. Adler, M. Zoller, X. Li, *PNAS* **2004**, *101*, 14258–14263.
[13] For a rewiev: D. E. Walters, *J. Chem. Ed.* **1995**, *72*, 680–683.
[14] R. S. Shallenberger, T. E. Acree, *Nature* **1967**, *216*, 480–482.
[15] L. B. Kier, *J. Pharm. Sci.* **1972**, *61*, 1394–1397.
[16] M. A. Castiglione Morelli, F. Lelj, F. Naider, M. Tallon, T. Tancredi, P. A. Temussi, *J. Med. Chem.* **1990**, *33*, 514–520.
[17] T. Yamazaki, E. Benedetti, D. Kent, M. Goodman, *Angew. Chem. Int. Ed. Engl.* **1994**, *33*, 1437–1451.
[18] J. M. Tinti, C. Nofre, D. Glaser, *Chem. Senses* **1996**, *21*, 747–762.
[19] A. Bassoli, M. G. B. Drew, C. K. Hattotuwagama, L. Merlini, G. Morini, G. R. H. Wilden, *QSAR* **2001**, *20*, 3–16.
[20] A. Bassoli, L. Merlini, G. Morini, A. Vedani, *J. of Chem. Soc., Perkin Trans. 2* **1998**, 1449–1454.
[21] A. Arnoldi, A. Bassoli, G. Borgonovo, M. G. B. Drew, L. Merlini, G. Morini, *Journal of Agricultural and Food Chemistry* **1998**, *46*, 4002–4010.
[22] A. Bassoli, M. G. B. Drew, L. Merlini, G. Morini, *J. Med. Chem.* **2002**, *45*, 4402 –4409.
[23] K. Nakamura, K. Kogiso, T. Nakajima, H. Kayahara, *Peptide Sci.* **2004**, *40*, 61–64.
[24] J. Ezquerra, C. Pedregal, A. Rubio, Y. Yruretagoyena, A. Escribano, S. Sanchez-Ferrando, *Tetrahedron* **1993**, *49*, 8665–8678.
[25] J. Ezquerra, C. Pedregal, A. Rubio, Y. Yruretagoyena, *J. Org. Chem.* **1995**, *60*, 2925–2930.
[26] US Patent 5994559, **1999**.
[27] A. Arnoldi, A. Bassoli, G. Borgonovo, L. Merlini, G. Morini., *J. Agric. Food Chem.* **1997**, *45*, 2047–2054.
[28] J. J.P Stewart, *J. Comput.-Aided Mol. Des.* **1990**, *4*, 1–105.
[29] A. Vedani, P. Zbinden, J. P. Snyder, P. A. Greenidge, *J. Am. Chem. Soc.* **1995**, *117*, 4987–4994.
[30] A. Vedani, D. W Hutha, *J. Am. Chem. Soc.* **1990**, *112*, 4759–4767.

Received: December 28, 2004

