(m, 3 H), 3.54–4.1 (m, 2 H), 3.89 (m, 1 H), 4.28 (pq, 2 H, J = 19.6 Hz) 4.5 (dd, 1 H, J = 12.8 and 8.8 Hz), 4.72 (m, 1 H), 4.78 (br d, 1 H, J = 7.6 Hz), 5.17 (m, 1 H), 6.83 (d, 1 H, J = 10.4 Hz), 7.27–7.63 (m, 7 H), 7.87 (dd, 1 H, J = 3.1 and 9.6 Hz); mass spectrum, m/e (relative intensity) 524 (M^+, not observed), 153 (4), 149 (2), 143 (2), 129 (1.5), 125 (2), 119 (2), 115 (2), 101 (8), 83 (9), 73 (42), 59 (45), 56 (31), 45 (83), 43 (100), 31 (87), 28 (85).

Amino Acid Analysis (control, L-amino acid oxidase treated) Found: Gly (1.97, 2.05), L-Phe (1.07, 0.03), D-Pro (1.0, 1.0), L-Dap (1.07, 0.59).

Acknowledgment. This work was supported in part by Biomedical Research Support Grant SO7RR-05456 and by CA 27985.

Structure-Sweetness Relationship of L-Aspartyl Dipeptide Analogues. A Receptor Site Topology

Hajime Iwamura

Department of Food Science and Technology, Faculty of Agriculture, Kyoto University, Kyoto 606, Japan.
Received September 2, 1980

The relationship between structure and the sweet potency of L-aspartyl dipeptide analogues was investigated by physicochemical parameters and regression analysis. The dipeptide analogues reported were divided into the following four classes: L-aspartic acid amides, L-aspartylaminoethyl esters, L-aspartylaminopropionates, and L-aspartylaminoacetates. The analysis carried out for each class indicated that the electron-withdrawing effect of the substituents directed to the peptide bond and the steric dimensions of the molecules are important in eliciting the sweet taste. The values of coefficients of the electronic σ^* terms in the correlations for L-aspartic acid amides, L-aspartylaminoethyl esters, and L-aspartylaminopropionates were ~0.7, indicating a common basic site on the receptor surface. The value for L-aspartylaminoacetates was ~1.5, and this value suggests, together with the factor of the participation of steric parameters, a closer or geometrically more proper fit to the receptor, explaining the generally higher potency of this class compared to the other three. The receptor model drawn based on these quantitative analyses appears to be consistent with other classes of sweeteners of apparently unrelated structures.

A wide variety of structurally unrelated compounds are known to elicit a common sweet taste. On the relationship between the structure and taste of these compounds, Shallenberger's A–H/B model[1] and Kier's dispersion site[2] are most commonly quoted, probably because of their diverse applicability to various kinds of sweet compounds. The limitations of these hypotheses are, however, obvious. Many nonsweet compounds which contain heteroatoms or double bonds possess an acidic proton and a basic site corresponding to the A–H and B sites, respectively, along with a dispersion site which is generally a methylene or methine group. In some cases, for example, in perillartines and 5-nitro- and 5-cyanoanilines, it is difficult to find the A–H site in a practical sense and in some cases, such as dihydrochalcones and phyllodulsins, multiple A–H/B units are a possibility.[3] Clearly, these oversimplified theories lack predictive value. One should, however, appreciate the attempts made by Shallenberger[1] and Kier[2] to seek common peculiarities among different classes of sweeteners. Efforts made since then to correlate sweetness or taste with chemical structure have been summarized in recent literature,[4] outlining the possibility of further developing structure–activity studies in this field. Although some workers have presented evidence in support of the existence of multiple sweet receptors and of differences in the sweet and bitter sites,[3] it is more useful in terms of predictive power to try to explain the possibly diverse classes of sweeteners in terms of a single, common receptor site.

Recently, the relationship between structure and taste of perillartine and nitro- and cyanoaniline derivatives has been quantitatively analyzed by physicochemical parameters and regression analysis,[5] indicating that the mode of interactions of these two classes of sweeteners is very similar. Furthermore, although only semiquatitatively, it has been shown that the sweet and bitter taste of the perillartine derivatives can be explained as a function of the steric dimensions of the molecules.[5] These ressults are suggestive of the usefulness or the predictive value of this approach. In this study, an attempt was made to correlate the relationship between the structure and the sweet potency of L-aspartyl dipeptide analogues in which activity had been estimated quantitatively and to compare the results with those of a previous paper.[5]

Methods

Activity data used for analyses were taken from literature reported by Mazur et al.,[6,7] Ariyoshi et al.,[8,9] Brussel et al.,[10] Fujino et al.,[11] and Miyoshi et al.[12]

The steric parameters used were calculated by the STERIMOL program developed by Verloop et al.[13] The L parameter expresses the length of a substituent along the bond axis which connects the substituent to the rest of the molecule. The W_r, W_l, W_u, and W_d parameters are the molecular width in directions perpendicular to the L axis and rectangular to each other. The W_r parameter is defmed as the width in the direction in which the longest chain of the substituent extends in the fully extended (staggered) conformation. W_l is the width in the direction opposite to W_r. The W_u and W_d parameters are the widths upward and downward, respectively, of the substituent when one views it from the con-

(1) R. S. Shallenberger and T. E. Acree, *Nature (London)*, **216**, 480 (1967).
(2) L. B. Kier, *J. Pharm. Sci.*, **61**, 1394 (1972).
(3) G. A. Crosby, G. E. DuBois, and R. E. Wingard, Jr., *Drug Des.*, **8**, 215 (1979).
(4) J. C. Boudreau, Ed., *ACS Symp. Ser.*, No. 115, 93, 149 (1979).
(5) H. Iwamura, *J. Med. Chem.*, **23**, 308 (1980).
(6) R. H. Mazur, A. H. Goldkamp, P. A. James, and J. M. Schlatter, *J. Med. Chem.*, **13**, 1217 (1970).
(7) R. H. Mazur, J. A. Reuter, K. A. Swiatek, and J. M. Schlater, *J. Med. Chem.*, **16**, 1284 (1973).
(8) Y. Ariyoshi, N. Yasuda, and T. Yamatani, *Bull. Chem. Soc. Jpn.*, **47**, 326 (1974).
(9) Y. Ariyoshi, *Agric. Biol. Chem.*, **40**, 983 (1976).
(10) L. B. P. Brussel, H. G. Peer, and A. van der Heijden, *Z. Lebensm. Unter-Forsch.*, **159**, 337 (1975).
(11) M. Fujino, M. Wakimasa, K. Tanaka, H. Aoki, and N. Nakajima, *Chem. Pharm. Bull.*, **24**, 2112 (1976).
(12) M. Miyoshi, K. Nunami, H. Sugano, and T. Fujii, *Bull. Chem. Soc. Jpn.*, **51**, 1433 (1978).
(13) A. Verloop, W. Hoogenstraaten, and J. Tipker, *Drug Des.*, **7**, 165 (1976).

0022-2623/81/1824-0572$01.25/0

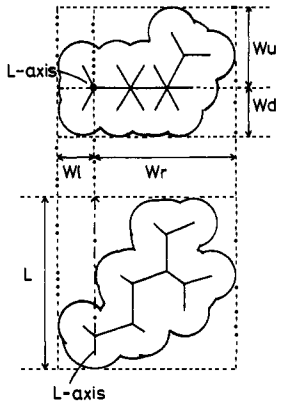


Figure 1. Schematic representation of the steric parameters, L, W_r, W_l, W_u, and W_d. The substituent used as the model is 4-methylamyl. In this and the following figures, the ends of the bars of the molecular models represent hydrogen atoms.

necting end along the bond axis L with W_r to the right. The relationship between the steric parameters is schematically shown in Figure 1. These parameters of a given substituent R_n are expressed by the subscript n, i.e., L_n, $(W_r)_n$, $(W_l)_n$, $(W_u)_n$, and $(W_d)_n$, unless otherwise noted.

When an aromatic ring is connected to an oxygen or sulfur atom, the dihedral angle between the ring and C–O or C–S bond of the C–O–ϕ or the C–S–ϕ is taken as 0°, taking the coplanarity due to the resonance effect into consideration. When the aromatic ring is connected to an alkyl chain, the angle is taken as 45°, minimizing steric constraint. Cyclobutyl and fencyl rings can not be built exactly from normal sp^3 carbons but can be approximated by them. Similarly, a cyclic methylene dioxy group is constructed approximately by normal sp^3 carbon and oxygen atoms. When an asymmetric carbon occurs in the substituent, i.e., in the case where the sweet potency is determined as an epimeric mixture, the steric parameter which concerns that center is expressed by the average value of the R and S configurations, unless otherwise mentioned.

The electronic parameter, σ^*, was estimated for the structure substituted on the common aspartylamino moiety, so that the electronic effect is directed to the peptide bond. Values were taken from or estimated according to the literature.[14,15] The value for $(CH_2)_n$-R was estimated by σ^* $(R) \times 0.34^n$, where 0.34 is the transmission factor.[16] The σ^* value of alkyl chains larger than n-butyl was approximated by that of the value for available n-butyl. When R is a higher alkyl than methyl in COOR, the σ^* value was approximated by that of COOMe. The value for branched structures was determined as the sum of the values for component structures e.g., σ^* $[CH(R_1)R_2] = \sigma^*$ $(CHR_1) + \sigma^*$ (CHR_2).

The hydrophobic parameter π was estimated also for the part of the molecule where the common aspartylamino moiety was eliminated, and the values were taken from or estimated according to the literature.[15,17] When the compound possesses an extra methyl at the amide nitrogen of the aspartylamino moiety, the value of 0.5 (π_{CH_3})[17] was added to that of the parent analogue. When appropriate π data were unavailable, it was estimated by $\pi(R) = \log P$ $(RH) - \log P$ (H_2),[18] where $\log P$ is the 1-octanol/water partition coefficient of the molecule.[15] The π values relative to the CH_2 of the ether and thioether linkage were estimated as $\log P$ $(EtOEt)$[15] – $\log P$ $(EtCH_2Et)$[15] and $\log P$ $(EtSEt)$[19] – $\log P$ $(EtCH_2Et)$[15] respectively.

(14) R. W. Taft, "Steric Effects in Organic Chemistry", M. S. Newman, Ed., Wiley, New York, 1965, p 556.
(15) C. Hansch and A. Leo, "Substituent Constants for Correlation Analysis in Chemistry", Wiley, New York, 1979.
(16) J. E. Leffler and E. Grunwald, "Rate and Equilibrium of Organic Reactions", Wiley, New York, 1963, p 171.
(17) A. Leo, C. Hansch, and D. Elkins, *Chem. Rev.*, 71, 525 (1971).
(18) C. Hansch, A. Vittoria, C. Silipo, and P. Y. C. Jow, *J. Med. Chem.*, 18, 546 (1975).
(19) J. Iwasa, T. Fujita, and C. Hansch, *J. Med. Chem.*, 8, 150 (1965).

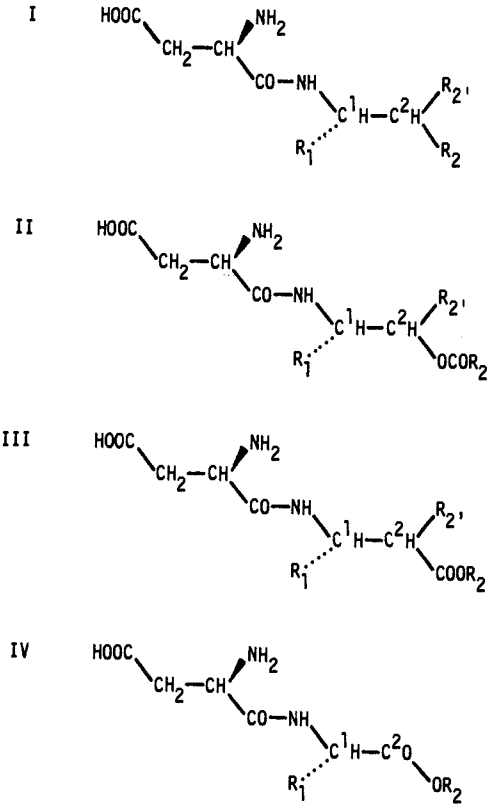


Figure 2. Structures of L-aspartyl dipeptide analogues which schematically represent the sweet, trans relationship of the R_1 to the aspartic amino group. The zigzag L-Asp-NHC^1C^2 backbone chain is fully extended and placed on the plane of the page with the L-aspartic amino group directed upward.

Results

Of the L-aspartyl dipeptide analogues reported by Mazur et al.,[6,7] Ariyoshi et al.,[8,9] Brussel et al.,[10] Fujino et al.,[11] and Miyoshi et al.,[12] the compounds whose sweet potency have been quantitatively estimated were analyzed and the activity, log SP, was expressed by the logarithm of the sweet potency relative to that of sucrose on a mole/mole basis. The compounds are divided by structure into the following four types: L-aspartic acid amides (I), L-Asp-$NHC^1H(R_1)C^2H(R_{2'})R_2$; L-aspartylaminoethyl esters (II), L-Asp-$NHC^1H(R_1)C^2H(R_{2'})OCOR_2$; L-aspartylaminopropionates (III), L-Asp-$NHC^1H(R_1)C^2H(R_{2'})COOR_2$; L-aspartylaminoacetates (IV), L-Asp-$NHC^1H(R_1)C^2OOR_2$. Here the carbons α and β to the amide nitrogen are expressed by the superscripts 1 and 2, respectively. In each type, the substituent R_1 is taken so that its length parameter L_1 is always smaller than that of the other C^1 substituent, i.e., $L_1 < L$ $[C^2H(R_{2'})R_2]$, $L_1 < L$ $[C^2H(R_{2'})OCOR_2]$, $L_1 < L$ $[C^2H(R_{2'})COOR_2]$, and $L_1 < L$ (C^2OOR_2). As has been noted by the authors above,[6-12] the absolute configuration at the C^1 carbon is critical for sweetness. By the above operation, the R_1 substituents of the sweet compounds become trans and those of the bitter, tasteless, or only slightly sweet compounds become cis to the aspartic acid amino group when the zigzag L-Asp-NHC^1C^2 backbone chain is placed on a plane as depicted by Figure 2. When the sweet potency had been determined as a mixture of these C^1 epimers, the value of 0.3 (log 0.5) was added to the activity data, assuming a 50:50 composition. Thus, in the case of the C^1 epimeric mixture, the steric parameters are those for the sweet configuration and are not the average of the R and S configuration. $R_{2'}$ is defined as a substituent whose length parameter $L_{2'}$ is always smaller than that of the other C^2 substituent. Alternative as-

sortment of the aspartyl dipeptide analogues are possible, but the one described above is most suitable for correlating structure vs. activity for this class of compounds.

Aspartic Acid Amides (I). Table I lists the activity data and the substituent parameters used. First, the compounds **1–24** reported by Mazur et al.[6,7] were analyzed and, of the various combinations of the steric and other substituent parameters as independent variables, eq 1 gave

$$\log SP = 0.73\sigma^* (\pm 0.46) + 2.23(W_r)_2 (\pm 1.82) - 0.25(W_r)_2^2 (\pm 0.21) - 3.25 \quad (1)$$

$$n = 24, r = 0.69, s = 0.35$$

the best correlation. In eq 1 and the following equations, n is the number of compounds included in the analysis, r is the multiple correlation coefficient, and s is the standard deviation. The figures in parentheses express the 95% confidence interval.

The positive coefficient of the electronic σ^* term indicates that the electron-withdrawing effect of the substituents enhances sweetness. This effect is mainly brought about by the COOMe group at C^1, because the σ^* values of the compounds with this function (compounds **19–24**) are 0.5–0.8, whereas those of the others are −0.2 to 0.0. The incorporation of the $(W_r)_2^2$ term shows that there is an optimum steric condition for sweetness in terms of the maximum width of the R_2 substituent in the direction of W_r. The colinearity between the terms, σ^* and $(W_r)_2$, is not significant, the squared simple correlation coefficient being 0.01. The hydrophobic parameter, π, whose squared simple correlation coefficients to σ^* and $(W_r)_2$ are 0.50 and 0.03, respectively, was found to not important.

The r value for eq 1 is not large enough, probably due to the narrow range of variations in the activity. The compounds reported by Ariyoshi et al.[8,9] (**51–58**), Brussel et al.[10] (**60–65**), and Dahlmans et al.[20] (**66**) include the same type of compounds as those analyzed above.[21] Thus, efforts were made to incorporate them into the analysis and improve the correlation.[22] Among the various combinations of parameters, eq 2 was the best for correlating the sweet potency of the whole set of compounds.

$$\log SP = 0.66\sigma^* (\pm 0.25) - 1.06\text{InA} (\pm 0.26) + 0.78(W_u)_1 (\pm 0.41) + 0.24(W_u)_2 (\pm 0.18) + 1.20(W_r)_2 (\pm 1.16) - 0.15(W_r)_2^2 (\pm 0.13) - 2.85 \quad (2)$$

$$n = 39, r = 0.92, s = 0.28$$

Use of the indicator variable, InA, for the data of Ariyoshi et al.[8,9] was necessary for correlation. It may correct the possible parallel difference in biological data between laboratories. The $(W_u)_1$ and $(W_u)_2$ terms were further incorporated into the correlation. The positive coefficients of these terms indicate that the thicker the R_1 and R_2 substituents in the W_u direction, the higher the activity. Table II shows the development of eq 2. The level of significance by Student's t test for the $(W_r)_2$ and $(W_r)_2^2$ terms finally incorporated into eq 2 is more than 95%, and that by the F test is 94.8%, leading us to adopt it as the one more significant than the others. Table III shows the degree of independence of the variables used.

The potency of compounds **25–50** reported by Mazur et al.[6,7] and of **59** reported by Ariyoshi et al.[8] was not definitely determined but qualitatively evaluated as slightly sweet, bitter, or tasteless. The predicted values calculated by eq 2 for these compounds are, however, listed in Table I. The values for the tasteless or slightly sweet compounds **25, 27–29, 31,** and **32** appear to be not in great conflict with the reported data, being less than 0.9. The activity of the DL compounds, **30, 37, 39, 42, 45,** and **46** is thought to be weak if the value of 0.3 is subtracted from the predicted values, because the D isomer (one of the C^1 epimers) of the mixture can be assumed to be only very weakly or a little active.[6] The bitter taste of the compounds **33, 34, 40,** and **41** suggests that the introduction of a methyl group to the bridged nitrogen atom makes the compounds bitter. The bulky methylenedioxy group on the benzene ring of compound **35** may deform the conformation of the receptor leading to the bitter taste, and the bulkiness due to this is not reflected in the steric parameters incorporated into eq 2. Similar causes are likely for the bitter compounds **38** and **39**. In the case of compound **38** with a polar $NHSO_2Me$ group at the molecular end, a hydrophilic or hydrogen-bonding interaction may be operative as well in the vicinity of that group, because the amino compound **59** is also bitter. Carbinol at C^1 appears to weaken the potency of the sweetness (compounds **47** and **48**). The steric dimensions of carbinol are not so different from those of the COOMe in terms of the STERIMOL parameters considered, and the corresponding COOMe compounds are sweet. Thus, their weak activity seems due to a position-specific hydrogen-bonding interaction with the receptor via OH. As mentioned above, the configuration at the C^1 carbon is critical for sweet potency and/or the taste quality. The configuration at other carbons appears to have a similar effect as in compounds **56** and **57**, although to a lesser extent. The L-erythro derivative **56** is almost twice as sweet as the L-threo compound **57**, although these two are both incorporated into the analysis because the deviations of the calculated values from the observed ones are permissible (0.28 and −0.14, respectively). The reportedly weak potency of compound **36** with OH and compound **49** with Me at C^2 may be due to a similar effect. The calculated values for compounds **43, 44,** and **50** suggests a significant sweet potency despite their respective reported bitterness, weak sweetness, and tastelessness. The reasons for these discrepancies remain uncertain, although a common feature of these compounds is a branch at the C^3 carbon which may be detrimental to sweetness. Further, more detailed studies are needed and also quantitative estimations of potency of the weakly active compounds, but, for the majority of the compounds, the activity data are rationalized by eq 2.

Aspartylaminoethyl Esters (II) and Aspartylaminopropionates (III). The substituents R_1, R_2, and $R_{2'}$ in Figure 2 are defined in Tables IV and VII. The steric parameters, L_2, $(W_r)_2$, $(W_l)_2$, $(W_u)_2$, and $(W_d)_2$, are, however, those of the $OCOR_2$ and $COOR_2$ moiety for convenience in comparing the results with those of the previous aspartic acid amides and the aspartylaminoacetates to follow.

(20) J. J. Dahlmans and W. H. Wilhelms, Netherlands Patent 7012897 (1972).

(21) Compound **66** has been also reported by Brussel et al.,[10] but its log SP value (0.90) estimated from their data always deviated from the predicted values during the development of eq 2. On the other hand, the sweetness of this compound, estimated as 2.00 in logarithmic terms according to a Netherland patent,[20] is in good accord with the predicted value. Thus, the activity data for this compound was taken from this patent, although the equation is essentially the same as that formulated excluding it.

(22) The compounds reported by Mazur et al.,[6,7] Ariyoshi et al.,[8,9] and Brussel et al.[10] overlap partially with each other. Thus, of the compounds reported by Ariyoshi et al.[8,9] and Brussel et al.,[10] those which were already incorporated in the derivation of eq 1, and of the compounds reported by Brussel et al.,[10] those which were also reported by Ariyoshi et al.,[8,9] were excluded from the analyses. The choice is arbitrary, but other choices caused little modification of eq 2.

Table I. Sweet Potency and Physicochemical Parameters of Aspartic Acid Amide Derivatives (I)[a]

L-Asp-NHC^{1}H(R_1)C^{2}H($R_{2'}$)R_2

no.	R_1	R_2	$R_{2'}$	config at C^1	log SP obsd[b]	log SP calcd by eq 2	Δ log SP	σ*	$(W_u)_1$	$(W_r)_2$	$(W_u)_2$	InA	π
1	H	Ph-*p*-F	H		0.57	0.67	−0.12	0.07	1.00	3.11	1.70		2.82
2	H	c-hexyl	H		0.85	1.06	−0.21	−0.02	1.00	3.49	3.15		3.26
3	Me	Ph	H	L	1.56	1.33	0.23	−0.02	1.90	3.11	1.70		2.98
4	Me	Bzl	H	DL	0.89	0.91	−0.02	−0.08	1.90	6.02	1.90		3.48
5	Me	OPh	H	DL	1.19	1.55	−0.36	0.19	1.90	3.45	1.90		2.46
6	Me	Ph-*p*-F	H	DL	1.50	1.32	0.18	−0.03	1.90	3.11	1.70		3.12
7	Me	furfuryl	H	DL	1.16	1.29	−0.13	−0.03	1.90	3.09	1.60		2.19
8	Me	c-hexyl	H	L	1.57	1.70	−0.13	−0.12	1.90	3.49	3.15		3.56
9	Me	*n*-Pr	H	DL	1.60	1.33	0.27	−0.23	1.90	3.49	1.90		2.80
10	Me	*n*-Bu	H	DL	1.43	1.36	0.07	−0.23	1.90	4.42	1.90		3.30
11	Me	*i*-Bu	H	DL	1.54	1.67	−0.13	−0.23	1.90	4.21	3.16		3.10
12	Me	*n*-Bu	H	L	1.56	1.33	0.23	−0.23	1.90	3.49	1.90		3.30
13	Me	*i*-Bu	H	L	1.85	1.67	0.18	−0.23	1.90	4.21	3.16		3.10
14	Me	CH_2OEt	H	DL	1.14	1.49	−0.35	−0.04	1.90	4.33	1.90		0.80
15	Me	*n*-amyl	H	DL	1.17	1.26	−0.09	−0.23	1.90	4.95	1.90		3.80
16	Et	Ph	H	DL	0.89	1.32	−0.43	−0.04	1.90	3.11	1.70		3.48
17	Me_2	Ph	H		1.19	1.26	−0.08	−0.12	1.90	3.11	1.70		3.28
18	Me	*n*-amyl	Me	L	1.58	1.26	0.32	−0.23	1.90	4.95	1.90		4.10
19	COOMe	Ph	H	L	2.11	1.83	0.28	0.76	1.88	3.11	1.70		1.57
20	COOMe	c-hexyl	H	L	2.30	2.19	0.11	0.66	1.88	3.49	3.15		2.15
21	COOMe	*n*-Pr	H	L	1.49	1.82	−0.33	0.55	1.88	3.49	1.90		1.39
22	COOMe	*i*-Bu	H	DL	2.11	2.17	−0.06	0.55	1.88	4.21	3.16		1.69
23	COOMe	*n*-Bu	H	DL	1.90	1.85	0.05	0.55	1.88	4.42	1.90		1.89
24	COOMe	*n*-amyl	H	DL	2.07	1.76	0.31	0.55	1.88	4.95	1.90		2.39
25*	H	Ph	H		TL	0.69		0.08	1.00	3.11	1.70		2.68
26*	H	Ph	Me		B	0.62		−0.04	1.00	3.11	1.70		2.98
27*	H	OPh	H		S	0.91		0.29	1.00	3.45	1.70		2.16
28*	H	Ph-*p*-OH	H		S	0.67		0.04	1.00	3.11	1.70		2.01
29*	H	furfuryl	H		S	0.66		0.07	1.00	3.09	1.60		1.89
30*[c]	H	*i*-Bu	H		TL	1.04		−0.13	1.00	4.21	3.16		2.80
31*	H	*n*-Bu	H		S	0.73		−0.13	1.00	4.42	1.90		3.00
32*	H	*n*-amyl	H		S	0.63		−0.13	1.00	4.95	1.90		3.50
33*[d]	Me	Ph	H	L	B	1.33		−0.02	1.90	3.11	1.70		3.48
34*[d]	Me	Ph	H	D	B	1.33		−0.02	1.90	3.11	1.70		3.48
35*	Me	Ph-3,4-(OCH_2O)	H	DL	B	1.51		0.02	1.90	3.60	1.90		2.93
36*	Me	Ph	OH	DL	S	1.46		0.17	1.90	3.11	1.70		0.98
37*	Me	Ph-*p*-OH	H	DL	S	1.30		−0.06	1.90	3.11	1.70		2.31
38*	Me	Ph-$NHSO_2Me$	H	L	B	1.59		−0.03	1.90	3.42	2.52		1.80
39*	Me	2-indolyl	H	DL	B	1.21			1.90	5.43	1.70		2.99
40*[d]	Me	c-hexyl	H	L	B	1.70		−0.12	1.90	3.49	3.15		4.06
41*[d]	Me	c-hexyl	H	D	B	1.70		−0.12	1.90	3.49	3.15		4.06
42*	Me	Et	H	DL	TL	1.20		−0.23	1.90	2.97	1.90		2.30
43*	Me	*i*-Pr	H	L	B	1.50		−0.23	1.90	2.97	3.16		2.60
44*	Me	*s*-Bu	H	DL	S	1.63		−0.23	1.90	3.49	3.16		3.10
45*	Et	Et	H	DL	B	1.19		−0.25	1.90	2.97	1.90		2.80
46*	Et	*n*-Pr	H	DL	S	1.31		−0.25	1.90	3.49	1.90		3.30
47*	MeOH	Ph	H	L	S	1.52		0.27	1.90	3.11	1.70		1.18
48*	MeOH	Ph-*p*-OH	H	L	S	1.49		0.23	1.90	3.11	1.70		0.51
49*	COOMe	Et	Me	L	NS	1.69		0.55	1.88	2.97	1.90		1.19
50*	COOMe	*i*-Pr	H	L	NS	1.99		0.55	1.88	2.97	3.16		1.19
51	COOMe	Et	H	L	0.46	0.63	−0.17	0.55	1.88	2.97	1.90	1.00	0.89
52	COOMe	$(CH_2)_3NHAc$	H	L	0.27	0.29	−0.02	0.69	1.88	6.10	1.90	1.00	0.00
53	COOMe	$(CH_2)_2NHAc$	H	L	0.25	0.76	−0.51	0.70	1.88	5.06	1.90	1.00	−0.50
54	COOMe	Ph	Me	DL	0.95	0.70	0.25	0.66	1.88	3.11	1.70	1.00	1.87
55	COOMe	CH(Me)COOMe	H	DL	1.15	1.16	0.01	0.63	1.88	4.28	3.16	1.00	−0.22
56[e]	COOMe	*n*-Pr	OH	L	1.16	0.88	0.28	0.74	1.88	3.49	1.90	1.00	−0.61
57[f]	COOMe	*n*-Pr	OH	L	0.75	0.89	−0.14	0.74	1.88	3.49	1.90	1.00	−0.61
58	H	CH_2COOMe	H		0.13	−0.19	0.32	0.08	1.00	4.28	1.90	1.00	0.59
59*	COOMe	$(CH_2)_3NH_2$	H	L	B	0.88		0.69	1.88	4.42	1.90	1.00	0.15
60	COOMe	SEt	H	L	1.60	2.07	−0.47	0.86	1.88	3.97	1.90		−0.05
61	COOMe	*S*-*n*-Pr	H	L	2.11	2.05	0.06	0.86	1.88	4.51	1.90		0.45
62	COOMe	*S*-*i*-Pr	H	L	2.23	2.35	−0.12	0.86	1.88	4.49	3.16		0.25
63	COOMe	*S*-*t*-Bu	H	L	2.95	2.35	0.60	0.86	1.88	4.49	3.16		0.55
64	COOMe	*O*-*t*-Bu	H	L	2.15	2.37	−0.22	0.86	1.88	4.11	3.16		−0.51
65	COOMe	Ph-*p*-OH	H	L	2.15	1.80	0.35	0.72	1.88	3.11	1.70		0.90
66	COOMe	CH_2COOMe	H	L	2.00	2.00	0.00	0.76	1.88	4.28	1.90		−0.22

[a] Compounds not included in the regression analysis have asterisks. [b] Abbreviations used: TL, tasteless; B, bitter; S, sweet; NS, not sweet. [c] Compound derived from DL-aspartic acid. [d] Compound with an extra methyl at the amide nitrogen. [e] Compound derived from L-*erythro*-β-hydroxynorleucine. [f] Compound derived from L-*threo*-β-hydroxynorleucine.

Table II. Development of Equation 2

constant	σ^*	InA	$(W_u)_2$	$(W_u)_1$	$(W_r)_2$	$(W_r)_2^2$	r	s	$F_{1,x}$[a]
1.63		−0.99					0.63	0.51	24.23
1.42	0.85	−1.29					0.82	0.38	30.45
0.83	0.79	−1.22	0.27				0.86	0.34	8.13
−0.54	0.69	−1.13	0.30	0.73			0.90	0.30	12.05
−2.85	0.66	−1.06	0.24	0.78	1.20	−0.15	0.92	0.28	3.24[b]

[a] $F_{1,30;\alpha=0.05} = 4.17$. [b] The value of $F_{2,32}$. $F_{2,30;\alpha=0.05} = 3.32$; $F_{2,30;\alpha=0.10} = 2.49$. The level of significance is 94.8%.

Table III. Squared Correlation Matrix for Variables Used in the Derivation of Equation 2

	σ^*	$(W_r)_2$	$(W_u)_1$	$(W_u)_2$
$(W_1)_2$	0.01			
$(W_u)_1$	0.08	0.01		
$(W_u)_2$	0.01	0.03	0.00	
InA	0.12	0.01	0.01	0.01

Miyoshi et al.[12] have reported 55 L-aspartylaminoethyl ester type compounds (**67–121**). In addition to these, Ariyoshi et al.[8,9] have determined the sweet potency of four more (**122–125**) and Brussel et al.[10] two more compounds (**126** and **127**); the combined set of derivatives was analyzed. Equation 3 was derived from the data in Table IV for this set of compounds. Other combinations of variables gave inferior results.

$$\log SP = 0.67\sigma^* (\pm 0.48) + 3.36L_2 (\pm 1.03) - 0.29L_2^2 (\pm 0.08) + 4.18(W_u)_1 (\pm 0.88) - 0.85(W_u)_1^2 (\pm 0.18) - 0.53L_1 (\pm 0.18) - 11.33 \quad (3)$$

$$n = 51,\ r = 0.88,\ s = 0.27$$

As for the substituent R_1, the steric effect on sweet potency is represented by the width parameter $(W_u)_1$ and length parameter L_1 and, as for the $OCOR_2$ moiety, the length parameter L_2 is important. The activity varies parabolically with respect to the $(W_u)_1$ and L_2 parameters, the optimum value of $(W_u)_1$ being ~2.5 (0.25 nm) and that of L_2 ~5.8 (0.58 nm). The fact that the ρ value is positive indicates that the electron-withdrawing substituents favor activity. The magnitude of the value 0.67 (± 0.48) is suggestive of a similar electronic interaction with the receptor to the aspartic acid amide derivatives (I), although the σ^* term is least significant among those incorporated into eq 3, as shown by Table V. This seems to be due to the lesser variation of the σ^* values in this set of compounds. Table VI shows the degrees of independence of the variables used.

The L-aspartylaminopropionate type of compounds (III) has been reported by Miyoshi et al.[12] (compounds **128–152**), Ariyoshi et al.[8,9] (compounds **153–155**), Brussel et al.[10] (compounds **156** and **157**), and Fujino et al.[11] (compound **158**) (Table VII). The calculations for these compounds were not completely successful because the electronic parameter σ^* and the hydrophobic parameter π were interchangeable without causing a significant difference in the statistical parameters. This appears to be because of the smaller variations in structure due to the fewer number of derivatives. As indicated by eq 3 above, by eq 2 in the preceding section, and by the equations in the following section, however, the electronic variable σ^* is always important in correlating the activity. Thus, the equation with the best correlation, eq 4, was selected from the equations including the σ^* term. The squared simple correlation coefficient between the σ^* and $(W_r)_2$ terms is 0.09. The level of significance by the F test of the equation with only the steric parameters $(W_r)_2$ and $(W_r)_2^2$ is 90.3%, as shown in Table VIII and the development of eq 4, and this level also is due to the lesser variation in structure in this set of compounds. Since the structures of aspartylaminoethyl esters and aspartylaminopropionates are closely related, except for the alkoxy carbonyl moiety, and since the coefficients of the σ^* term in eq 3 and 4 overlap within the 95% confidence interval, these two sets of the compounds were combined. In eq 5 thus derived, the

$$\log SP = 0.73\sigma^* (\pm 0.36) + 2.75L_2\,[OCOR] (\pm 0.86) - 0.24L_2^2\,[OCOR] (\pm 0.07) + 4.13(W_u)_1\,[OCOR] (\pm 0.86) - 0.83(W_u)_1^2\,[OCOR] (\pm 0.18) - 0.55L_1\,[OCOR] (\pm 0.16) + 4.25(W_r)_2\,[COOR] (\pm 1.25) - 0.45(W_r)_2^2\,[COOR] (\pm 0.15) - 9.34 \quad (5)$$

$$n = 72,\ r = 0.89,\ s = 0.27$$

parameters for L-aspartylaminoethyl esters (II) are marked by OCOR and those for L-aspartylaminopropionates (III) by COOR in brackets.

In this combined form, the hydrophobic parameter π was not significant at all, the squared simple correlation coefficient to σ^* being 0.53. As indicated by the σ^* term, both types of compounds interact electrostatically with a basic site on the receptor surface at either the amide hydrogen or carbon atom. The difference in the significance of the steric parameters is thus though to reflect differences in molecular shape or conformation which would result from the exchange of the carbonyl function and oxygen atom and would bring about a somewhat different mode of interaction. Previously, Miyoshi et al.[12] noted that a CPK model of a L-aspartylaminoethyl ester [OCOR] differs in shape from that of a corresponding L-aspartylaminopropionate [COOR] and attributed the weaker activity of the latter to the difference in the carboalkoxy moiety. The fact that more steric parameters were necessary to correlate the activity of the ethyl esters [OCOR] suggests that this type of compound is located closer to the spatial walls of the receptor cavity, strengthening binding and enhancing activity. On the other hand, the propionate type of compounds [COOR] is believed to interact sterically only at the region where the R_2 substituent is located.

The L-aspartylaminoethyl esters, **113–121** and **127**, and the L-aspartylaminopropionates, **144–152** and **158**, have been reported as not being sweet and so were excluded from the analysis. Their sweet potencies estimated according to eq. 5 are, however, listed in Table IV and VII. Subtracting the epimeric factor 0.3 (log 0.5) from the calculated values of L-Asp-NH-DL-compounds, the data for compounds **115–121** and **127** (Table IV) and compounds **144**, **145**, and **147–150** (Table VII) are not incompatible with the reported weak potency, the values being ~0.7. The dimethylcyclohexyl substituent of compound **113** is so large that it may not be accomodated well at the receptor site. A bulky substituent like benzyl at the C^1

$$\log SP = 0.64\sigma^* (\pm 0.59) + 2.36(W_r)_2 (\pm 2.27) - 0.24(W_r)_2^2 (\pm 0.26) - 5.16 \quad (4)$$

$$n = 21,\ r = 0.64,\ s = 0.27$$

Table IV. Sweet Potency and Physicochemical Parameters of L-Aspartylaminoethyl Esters (II)[a]

L-Asp-NHC^{1}H(R_1)C^{2}H($R_{2'}$)OCOR_2

no.	R_1	R_2	$R_{2'}$	confign at C^1	log SP obsd[b]	log SP calcd by eq 5	Δ log SP	σ^*	L_1	$(W_u)_1$	L_2	π
67	H	Me	H		0.28	0.74	−0.46	0.31	2.06	1.00	4.74	0.09
68	H	Et	H		0.61	0.98	−0.37	0.31	2.06	1.00	5.99	0.59
69	H	*n*-Pr	H		0.56	0.74	−0.18	0.31	2.06	1.00	6.81	1.09
70	H	*i*-Pr	H		1.16	0.98	0.18	0.31	2.06	1.00	5.99	0.89
71	H	c-Pr	H		1.03	0.98	0.05	0.31	2.06	1.00	6.05	0.62
72	H	c-Bu	H		0.99	0.99	0.00	0.31	2.06	1.00	5.73	1.03
73	H	*i*-Bu	H		1.36	0.74	0.62	0.31	2.06	1.00	6.81	1.19
74	Me	Me	H	D	1.61	1.77	−0.16	0.21	2.87	1.90	4.74	0.39
75	Me	Et	H	D	1.86	2.02	−0.16	0.21	2.87	1.90	5.99	0.89
76	Me	vinyl	H	D	1.85	1.99	−0.14	0.21	2.87	1.90	6.17	0.59
77	Me	*n*-Pr	H	DL	1.58	1.78	−0.20	0.21	2.87	1.90	6.81	1.39
78	Me	*i*-Pr	H	D	1.98	2.02	−0.04	0.21	2.87	1.90	5.99	1.19
79	Me	C(Me)=CH$_2$	H	D	1.99	1.99	0.00	0.21	2.87	1.90	6.17	0.89
80	Me	c-Pr	H	D	2.18	2.01	0.17	0.21	2.87	1.90	6.05	0.92
81	Me	*n*-Bu	H	D	0.60	0.80	−0.20	0.21	2.87	1.90	8.06	1.89
82	Me	*s*-Bu	H	D	2.08	1.80	0.28	0.21	2.87	1.90	6.77	1.69
83	Me	*i*-Bu	H	D	2.08	1.78	0.30	0.21	2.87	1.90	6.81	1.69
84	Me	*t*-Bu	H	D	2.28	2.02	0.26	0.21	2.87	1.90	5.99	1.49
85	Me	2-Me-c-Pr	H	D	1.68	1.89	−0.21	0.21	2.87	1.90	6.53	1.22
86	Me	c-Bu	H	D	2.24	2.03	0.22	0.21	2.87	1.90	5.73	1.33
87	Me	CH(Et)$_2$	H	D	1.53	1.78	−0.25	0.21	2.87	1.90	6.81	2.19
88	Me	2-Me-c-Bu	H	D	2.13	1.99	0.14	0.21	2.87	1.90	6.17	1.63
89	Me	c-amyl	H	D	1.70	1.76	−0.06	0.21	2.87	1.90	6.84	1.74
90	CH$_2$OH	Me	H	D	1.23	1.37	−0.14	0.50	3.97	1.90	4.74	−1.41
91	CH$_2$OH	Et	H	D	1.99	1.62	0.37	0.50	3.97	1.90	5.99	−0.91
92	CH$_2$OH	c-Pr	H	D	1.96	1.62	0.34	0.50	3.97	1.90	6.05	−0.88
93	CH$_2$OH	*i*-Pr	H	D	1.48	1.62	−0.14	0.50	3.97	1.90	5.99	−0.61
94	CH$_2$OH	*n*-Pr	H	D	0.58	0.40	0.18	0.50	3.97	1.90	8.06	−0.41
95	CH$_2$OH	c-Bu	H	D	1.71	1.63	0.08	0.50	3.97	1.90	5.73	−0.47
96	CH$_2$OH	*i*-Bu	H	D	1.30	1.38	−0.08	0.50	3.97	1.90	6.81	−0.31
97	CH$_2$OH	*s*-Bu	H	D	1.21	1.38	−0.17	0.50	3.97	1.90	6.81	−0.11
98	CH$_2$OH	*c*-amyl	H	D	0.92	1.37	−0.45	0.50	3.97	1.90	6.84	−0.08
99	Et	Me	H	DL	1.16	1.07	0.09	0.19	4.11	1.90	4.74	0.89
100	Et	Et	H	DL	1.48	1.32	0.16	0.19	4.11	1.90	5.99	1.39
101	Et	*i*-Pr	H	DL	1.38	1.32	0.06	0.19	4.11	1.90	5.99	1.69
102	Et	c-Pr	H	DL	1.38	1.31	0.07	0.19	4.11	1.90	6.05	1.42
103	Et	c-Bu	H	DL	1.40	1.33	0.07	0.19	4.11	1.90	5.73	1.83
104	CHOMe	Me	H	DL	0.48	0.92	−0.44	0.49	4.78	1.90	4.74	−1.11
105	CHOMe	Et	H	DL	1.21	1.17	0.04	0.49	4.78	1.90	5.99	−0.61
106	*i*-Pr	Me	H	D	1.58	0.99	0.59	0.19	4.11	3.16	4.74	1.19
107	*i*-Pr	Et	H	D	1.60	1.31	0.29	0.29	4.11	3.16	5.99	1.69
108	*i*-Pr	*i*-Pr	H	D	1.10	1.24	−0.14	0.19	4.11	3.16	5.99	1.99
109	*i*-Pr	c-Pr	H	D	1.22	1.23	−0.01	0.19	4.11	3.16	6.05	1.72
110	*i*-Pr	c-Bu	H	D	1.06	1.24	−0.18	0.19	4.11	3.16	5.73	2.13
111	*i*-Pr	*n*-Pr	H	D	0.40	1.00	−0.60	0.19	4.11	3.16	6.81	2.19
112	*n*-Pr	Me	H	DL	0.66	0.62	0.04	0.18	4.92	1.90	4.74	1.39
113*	Me	3,3-Me$_2$-c-hexyl	H	D	NS	1.81		0.21	2.87	1.90	6.75	1.93
114*	Me	c-hexyl	H	D	NS	0.82		0.21	2.87	1.90	8.04	2.15
115*	Me	Ph	H	D	NS	0.68		0.18	2.87	1.90	8.15	2.50
116*	Et	*n*-Bu	H	DL	NS	0.10		0.19	4.11	1.90	8.06	2.39
117*	Et	c-hexyl	H	DL	NS	0.12		0.19	4.11	1.90	8.04	2.65
118*	*n*-Pr	Et	H	DL	NS	0.87		0.18	4.92	1.90	5.99	1.89
119*	*i*-Bu	Me	H	DL	NS	0.53		0.18	4.92	3.16	4.74	1.69
120*	*s*-Bu	Me	H	DL	NS	0.90		0.18	4.92	2.53	4.74	1.69
121*	Bzl	Me	H	DL	NS	0.25		0.39	5.86	1.99	4.74	2.07
122	COOMe	Me	H	L	0.91	1.29	−0.38	0.99	4.73	1.88	4.74	−1.02
123	COOMe	Et	H	L	1.53	1.54	−0.01	0.99	4.73	1.88	5.99	−0.52
124	COOMe	*n*-Pr	H	L	1.43	1.30	0.13	0.99	4.73	1.88	6.81	−0.02
125	COOMe	*i*-Pr	H	L	1.65	1.54	0.11	0.99	4.73	1.88	5.99	−0.22
126	COOMe	*t*-Bu	H	L	1.81	1.54	0.27	0.99	4.73	1.88	5.99	0.08
127*	COOMe	Ph	H	L	NS	0.20		0.96	4.73	1.88	8.15	1.09

[a] Compounds not included in the regression analysis have asterisks. [b] For abbreviations, see footnote *b* of Table I.

carbon in compound **146** appears to be detrimental to activity, as is the bulky COOMe at the C^2 carbon in compounds **151** and **152**. Compound **158** is the only propionate ester among the compounds reported by Fujino et al.,[11] and the reason for the discrepancy between the predicted and reported data is uncertain. The calculated value for compound **114** shows moderate sweetness, and why it does not taste sweet is also unclear. Some of the reportedly unsweet compounds may be bitter.

Aspartylaminoacetates (IV). The definition for the steric parameters of R_1 is the same as the one in the preceding sections. The parameters with the subscript 2 are those for the OR_2 moiety, for convenience in comparing the results with those of the preceding sections. First,

Table V. Development of Equation 3

constant	$(W_u)_1$	$(W_u)_1^2$	L_1	L_2	L_2^2	σ^*	r	s	$F_{1,x}$[a]
-0.82	2.16	-0.49					0.48	0.47	7.10
-1.18	3.48	-0.73	-0.35				0.65	0.41	16.04
-11.72	3.76	-0.79	-0.35	3.53	-0.30		0.85	0.29	23.95[b]
-11.33	4.18	-0.85	-0.53	3.36	-0.29	0.67	0.88	0.27	8.12

[a] $F_{1,40;\alpha=0.05} = 4.09$. [b] The value of $F_{2,45}$. $F_{2,40;\alpha=0.05} = 3.23$

Table VI. Squared Correlation Matrix for Variables Used in the Derivation of Equation 3

	σ^*	$(W_u)_1$	L_2
$(W_u)_1$	0.02		
L_2	0.00	0.04	
π	0.51	0.08	0.07

compounds **159–185** in Table IX reported by Ariyoshi et al.[8,9] were analyzed, giving eq 6 as the best correlation.

$$\log SP = 1.77\sigma^* (\pm 1.05) + 2.21L_2 (\pm 1.38) - 0.19L_2^2 (\pm 0.13) + 0.32L_1 (\pm 0.14) - 0.64\text{In} (\pm 0.22) - 7.03 \quad (6)$$

$$n = 27,\ r = 0.90,\ s = 0.24$$

Examination of the data and the preliminary calculations indicated that there exists a parallel difference in the sweet potency between the compounds derived from D-threonine **167–173** and the allo counterparts **174–180**. Thus, the indicator variable In was introduced for the compounds of the latter type. The negative coefficient of this term indicates that the allo configuration within the substituent R_1 is sterically unfavorable for closer binding to the receptor. Of the steric parameters, the substituent lengths L_1 and L_2 are important for correlating activity. There is an optimum steric condition for activity in terms of L_2, the value of which was estimated as 5.8 (0.58 nm). The positive coefficient of the L_1 term indicates that the longer the R_1 substituent, the higher the activity. The large positive coefficient, 1.77 (±1.05), of the σ^* term means that electronic interaction with the receptor is much more significant in this series of compounds than for those mentioned above. Addition of the hydrophobic parameter π did not improve the correlation. Table X shows the development of eq 6 and Table XI the degree of independence of the variables used in eq 6 and π.

In addition to the compounds analyzed above, Mazur et al.[7] have determined the taste of compounds **186–196**, and Fujino et al.[11] have determined the taste of an additional 21 compounds (**197–217**) of this type. Thus, these were further incorporated into the analyses, giving eq 7 as the best correlation. The indicator variable, InM, was

$$\log SP = 1.54\sigma^* (\pm 0.51) + 2.46L_2 (\pm 1.04) - 0.21L_2^2 (\pm 0.09) + 0.37\text{InM} (\pm 0.29) + 1.49L_1 (\pm 0.61) - 0.19L_1^2 (\pm 0.07) - 0.74\text{In} (\pm 0.29) + 1.05(W_1)_2 (\pm 0.25) - 10.87 \quad (7)$$

$$n = 56,\ r = 0.95,\ s = 0.34$$

Table VII. Sweet Potency and Physicochemical Parameters of L-Aspartylaminopropionates (III)[a]

L-Asp-NHC^{1}H(R_1)C^{2}H($R_{2'}$)COOR$_2$

no.	R_1	R_2	$R_{2'}$	config at C^1	log SP obsd[b]	log SP calcd by eq 5	Δ log SP	σ^*	$(W_r)_2$	π
128	Me	Et	H	D	1.16	0.74	0.42	0.13	4.29	0.89
129	Me	*i*-Pr	H	D	0.84	0.74	0.10	0.13	4.29	1.19
130	Me	*s*-Bu	H	D	0.60	0.82	-0.22	0.13	4.83	1.69
131	Me	*t*-Bu	H	D	0.68	0.74	-0.06	0.13	4.29	1.49
132	Me	c-hexyl	H	D	0.64	0.82	-0.18	0.13	4.83	2.15
133	Me	Et	H	DL	0.88	0.72	0.16	0.11	4.29	1.39
134	H	*s*-Bu	H		0.73	0.90	-0.17	0.23	4.83	1.39
135	H	*t*-Bu	H		0.36	0.81	-0.45	0.23	4.29	1.19
136	H	$CH(Et)_2$	H		0.60	0.90	-0.30	0.23	4.83	1.89
137	H	CH(Me)-*n*-Pr	H		0.68	0.44	0.24	0.23	5.76	1.89
138	H	c-amyl	H		1.16	0.90	0.26	0.23	4.83	1.44
139	H	c-hexyl	H		1.22	0.90	0.32	0.23	4.83	1.85
140	H	*i*-Pr	Me		0.18	0.72	-0.54	0.11	4.29	1.19
141	H	*t*-Bu	Me		0.51	0.72	-0.21	0.11	4.29	1.49
142	H	c-amyl	Me		0.88	0.81	0.07	0.11	4.83	1.74
143	H	c-hexyl	Me		0.94	0.81	0.13	0.11	4.83	2.35
144*	Et	*i*-Pr	H	DL	NS	0.72		0.11	4.29	1.69
145*	Et	c-hexyl	H	DL	NS	0.80		0.11	4.83	2.65
146*	Bzl	Et	H	L	NS	0.87		0.31	4.29	2.57
147*	H	*n*-Bu	H		NS	0.44		0.23	5.76	1.59
148*	H	2-Me-c-hexyl	H		NS	0.62		0.23	5.54	2.15
149*	H	Me	Bzl		NS	0.07		0.26	3.36	2.07
150*	H	Me	COOMe		NS	0.21		0.46	3.36	-1.02
151*	H	c-amyl	COOMe		NS	1.07		0.46	4.83	0.33
152*	H	c-hexyl	COOMe		NS	1.07		0.46	4.83	0.74
153	Me	Me	H	DL	0.43	-0.03	0.46	0.13	3.36	0.39
154	H	Me	H		0.11	0.04	0.07	0.23	3.36	0.09
155	H	*i*-Pr	H		0.70	0.81	-0.11	0.23	4.29	0.89
156*	COOMe	Me	H	L	0.60	0.54	0.06	0.91	3.36	-1.02
157	COOMe	*t*-Bu	H	L	1.30	1.31	-0.01	0.91	4.29	0.08
158*	COOMe	c-hexyl	H	L	NS	1.39		0.91	4.83	0.74

[a] Compounds not included in the regression analysis have asterisks. [b] For abbreviations, see footnote *b* of Table I.

Table VIII. Development of Equation 4

constant	$(W_r)_2$	$(W_r)_2^2$	σ^*	r	s	$F_{1,x}$
−3.63	1.80	−0.18		0.48	0.30	2.66[a]
−5.16	2.36	−0.24	0.64	0.64	0.27	5.19[b]

[a] $F_{2,18;\alpha=0.05} = 3.56$; $F_{2,18;\alpha=0.10} = 2.62$. The level of significance is 90.3%. [b] $F_{1,17;\alpha=0.05} = 4.45$.

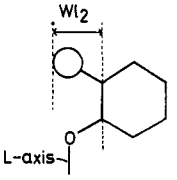


Figure 3. Schematic representation of the $(W_l)_2$ parameter used for the derivation of eq 7 for L-aspartylaminoacetates (IV). The figure depicts the OR_2 moiety of the cyclohexyl derivative.

used for the data of Mazur et al.,[7] but it was not necessary for the data of Fujino et al.[11] The highly active compounds, **202**, **203**, **206**, **209**, and **215**, reported by Fujino et al.[11] commonly possess a methyl substituent at the 2 position of the cyclohexane ring. This structural characteristic is manifested by the $(W_l)_2$ parameter when it is measured from the C_1–C_2 bond axis of the cyclohexane ring as depicted by Figure 3 rather than the L axis as defined under Methods. As seen from Table IX, the values of these compounds are 2.8 (0.28 nm) to 3.0 (0.3 nm) and those for the other compounds are constant at ~1.5 (0.15 nm). Thus, the exceptionally high potency of these compounds is attributable to this fact, when one looks for the cause in steric dimensions in addition to the effect of their large σ^* values, 1.36.

Compound **216** is not sweet. Because the corresponding **204** is a sweet compound, the big projection, isopropyl, at the 6 position of the cyclohexane moiety appears to be detrimental to sweetness and its bulkiness is not reflected by the steric parameters used. The predicted values of the unsweet compound **196** and the weakly sweet **217** indicate rather potent sweetness. The bulky, branched R_1 substituent, isobutyl, of the former and the *tert*-butyl at the molecular end of the latter may be unfavorable for binding to the receptor cavity. These observations and the rather large deviations of the observed potencies of both *tert*-butyl compound **197** and the 4-methylcyclohexyl derivative **205** from the calculated values are suggestive of the participation of other steric factors than those considered here. Correlation will be improved by further subdivision of the steric parameters and with larger numbers of these kinds of compounds. Tables XII and XIII are the development of the equation and the squared correlation matrix for variables used in the derivation of eq 7.

Discussion

The results of the present study indicate that the variations in the sweet potency of L-aspartyl dipeptide analogues are governed mainly by variations in the steric and electronic properties of the compounds.

The overlap of the positive coefficient of the σ^* terms in eq 2 and 5 within a 95% confidence interval suggests that aspartic acid amides (I), aspartylaminoethyl esters (II), and aspartylaminopropionates (III) interact in a similar fashion with the basic site of the receptor surface, probably via hydrogen bonding at the amide nitrogen atom. The difference in the importance of the steric parameters to activity is thought to reflect a somewhat different mode of interactions due to the differences in overall structure. The positive coefficients of the $(W_u)_1$ and $(W_u)_2$ terms in eq 2 for the aspartic acid amides (I) indicate that the width effects of the R_1 and R_2 substituents are not yet supraoptimum within the compounds analyzed. On the other hand, the parabolic relationship between the $(W_r)_2$ term and sweet potency suggests that compounds with $(W_r)_2$ values greater than ~4.0 (0.4 nm), the optimum value, invade the spatial wall of the receptor cavity which is located in the $(W_r)_2$ direction. The L_2 [OCOR] term in eq 5 shows that there exists a spatial wall in the L_2 direction of the aspartylaminoethyl esters (II). Similarly, the $(W_r)_2$ [COOR] term indicates that the aspartylaminopropionates (III) interact with the receptor so that their R_2 substituents face to the spatial wall in the $(W_r)_2$ direction. This variation in the mode of interaction between these two classes of compounds appears to be caused by the difference in backbone structure, suggesting a possible additional interaction site, steric or electronic, on a receptor surface where the carboalkoxy group is located.

The coefficient of the σ^* term in eq 7 for the aspartylaminoacetates (IV) is more than twice those of eq 2 and 5. This means that this series of compounds binds tighter with the basic site of the receptor than do the other three

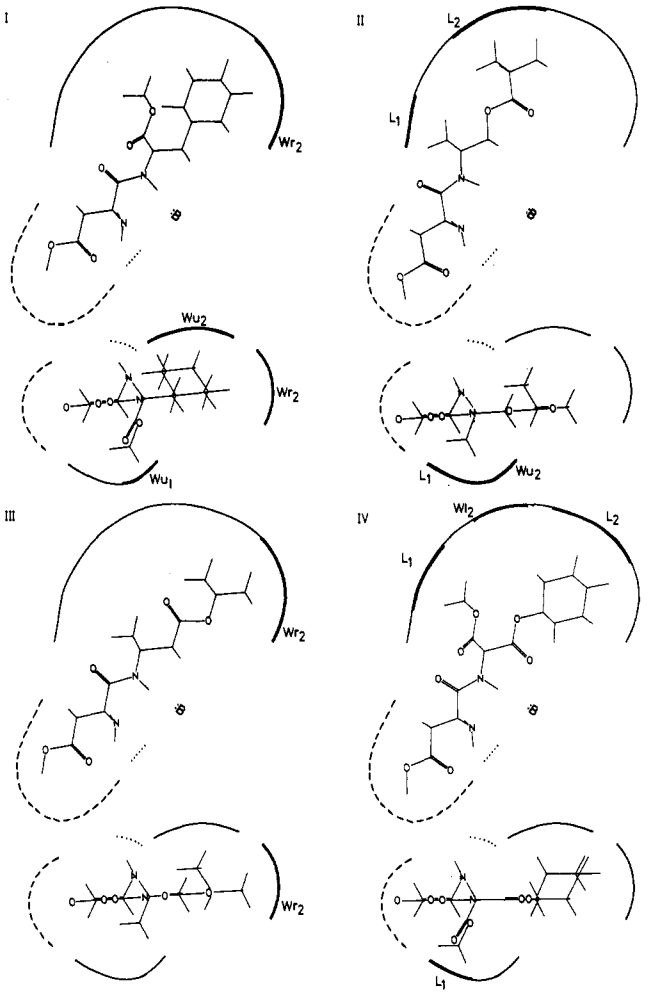


Figure 4. L-Aspartyl dipeptide–receptor binding model. The solid lines represent walls identified by quantitative analyses and the broken and dotted lines are those drawn by qualitative interpretations. :B is the basic site of the receptor. The compounds used as models are as follows: I, L-Asp-NHCH(COOMe)CH_2-c-hexyl (**20**); II, L-Asp-NH(Me)CH_2OCO-*i*-Pr (**78**); III, L-Asp-NHCH(Me)CH_2COO-*i*-Pr (**129**); IV, L-Asp-NHCH(COOMe)-c-hexyl (**201**). These are chosen as the representatives of L-aspartic acid amides (I), L-aspartylaminoethyl esters (II), L-aspartylaminopropionates (III), and L-aspartylaminoacetates (IV), respectively.

Table IX. Sweet Potency and Physicochemical Parameters of L-Aspartylaminoacetates (IV)[a]

L-Asp-NHC^{1}H(R_1)COOR_2

no.	R_1	R_2	config at C^1	log SP obsd[b]	log SP calcd by eq 7	Δ log SP	σ*	L_1	L_2	$(W_1)_2$	InM	In	π
159	Et	Me	D	1.04	1.00	0.04	0.56	4.11	3.98	1.53			0.39
160	CH_2OH	Me	D	1.49	1.48	0.01	0.87	3.97	3.98	1.53			−1.91
161	CH_2OH	Et	D	1.92	1.96	−0.04	0.87	3.97	4.80	1.53			−1.51
162	CH_2OH	*n*-Pr	D	2.39	2.15	0.24	0.87	3.97	6.08	1.53			−1.01
163	CH_2OH	*i*-Pr	D	1.96	1.96	0.00	0.87	3.97	4.80	1.53			−1.21
164	CH_2OH	*n*-Bu	D	1.75	1.91	−0.16	0.87	3.97	6.86	1.53			−0.51
165	CH_2OH	*i*-Bu	D	2.21	2.15	0.06	0.87	3.97	6.05	1.53			−0.71
166	CH_2OH	c-hexyl	D	1.72	2.15	−0.43	0.87	3.97	6.05	1.53			−0.15
167[c]	CH(OH)Me	Me	D	1.26	1.29	−0.03	0.75	4.11	3.98	1.53			−1.61
168[c]	CH(OH)Me	Et	D	1.93	1.78	0.15	0.75	4.11	4.80	1.53			−1.11
169[c]	CH(OH)Me	*n*-Pr	D	2.08	1.96	0.12	0.75	4.11	6.05	1.53			−0.61
170[c]	CH(OH)Me	*i*-Pr	D	1.93	1.78	0.15	0.75	4.11	4.80	1.53			−0.81
171[c]	CH(OH)Me	*n*-Bu	D	1.41	1.73	−0.32	0.75	4.11	6.86	1.53			−0.11
172[c]	CH(OH)Me	*i*-Bu	D	1.97	1.96	0.01	0.75	4.11	6.05	1.53			−0.31
173[c]	CH(OH)Me	c-hexyl	D	1.44	1.96	−0.52	0.75	4.11	6.05	1.53			0.15
174[d]	CH(OH)Me	Me	D	0.71	0.55	0.16	0.75	4.11	3.98	1.53		1.00	−1.61
175[d]	CH(OH)Me	Et	D	0.66	1.04	−0.38	0.75	4.11	4.80	1.53		1.00	−1.11
176[d]	CH(OH)Me	*n*-Pr	D	1.51	1.22	0.29	0.75	4.11	6.05	1.53		1.00	−0.61
177[d]	CH(OH)Me	*i*-Pr	D	0.91	1.04	−0.13	0.75	4.11	4.80	1.53			−0.81
178[d]	CH(OH)Me	*n*-Bu	D	1.23	0.99	0.24	0.75	4.11	6.86	1.53			−0.11
179[d]	CH(OH)Me	*i*-Bu	D	1.36	1.22	0.14	0.75	4.11	6.05	1.53			−0.31
180[d]	CH(OH)Me	c-hexyl	D	1.31	1.22	0.09	0.75	4.11	6.05	1.53			0.15
181	Et	*n*-Pr	D	1.86	1.67	0.19	0.56	4.11	6.05	1.53			1.39
182	H	Me		0.68	0.48	0.20	0.68	2.06	3.98	1.53			−0.41
183	H	*n*-Pr		0.98	1.15	−0.17	0.68	2.06	6.05	1.53			0.59
184	H	c-hexyl		1.02	1.15	−0.13	0.68	2.06	6.05	1.53			1.35
185	*n*-Pr	*n*-Pr	D	1.56	1.50	0.06	0.55	4.92	6.05	1.53			1.89
186	Me	Me	D	1.20	1.16	0.04	0.58	2.87	3.98	1.53	1.00		−0.11
187	Me	Et	D	1.75	1.64	0.11	0.58	2.87	4.80	1.53	1.00		0.39
188	Me	*n*-Pr	D	2.10	1.82	0.28	0.58	2.87	6.05	1.53	1.00		0.89
189	Me	*i*-Pr	D	1.95	1.64	0.31	0.58	2.87	4.80	1.53	1.00		0.69
190	Me	*n*-Bu	D	0.90	1.59	−0.69	0.58	2.87	6.86	1.53	1.00		1.39
191	Me	*n*-amyl	D	0.68	0.68	0.00	0.58	2.87	8.11	1.53	1.00		1.89
192	Et	*i*-Pr	D	2.15	1.85	0.30	0.56	4.11	4.80	1.53	1.00		1.19
193	*i*-Pr	*i*-Pr	D	2.15	1.84	0.32	0.55	4.11	4.80	1.53	1.00		1.49
194	*n*-Pr	*i*-Pr	DL	1.43	1.69	−0.26	0.55	4.92	4.80	1.53	1.00	1.00	1.69
195	*s*-Bu	*i*-Pr	D	0.53	0.95	−0.42	0.55	4.92	4.80	1.53	1.00		1.99
196*	*i*-Bu	*i*-Pr	D	NS	1.69		0.55	4.92	4.80	1.53	1.00		1.99
197	COOMe	*t*-Bu	DL	2.03	2.62	−0.59	1.36	4.73	4.80	1.53			−0.42
198	COOMe	Et_2-carbinyl	DL	2.79	2.81	−0.02	1.36	4.73	6.05	1.53			0.28
199	COOMe	Et_2-carbinyl	DL	3.06	2.81	0.25	1.36	4.73	6.05	1.53			0.88
200	COOMe	c-amyl	DL	2.92	2.63	0.29	1.36	4.73	5.56	1.35			−0.17
201	COOMe	c-hexyl	DL	3.14	2.81	0.33	1.36	4.73	6.05	1.53			0.24
202	COOMe	*cis*-2-Me-c-hexyl	DL	3.92	4.12	−0.20	1.36	4.73	6.05	2.78			0.54
203	COOMe	*trans*-2-Me-c-hexyl	DL	4.11	4.12	−0.01	1.36	4.73	6.05	2.78			0.54
204	COOMe	3-Me-c-hexyl	DL	3.02	2.57	0.45	1.36	4.73	6.86	1.53			0.54
205	COOMe	4-Me-c-hexyl	DL	1.87	2.81	−0.94	1.36	4.73	6.05	1.53			0.54
206	COOMe	2,6-Me_2-c-hexyl	DL	4.02	4.12	−0.10	1.36	4.73	6.05	2.78			0.84
207	COOMe	3,3,5-Me_3-c-hexyl	DL	2.76	2.57	0.19	1.36	4.73	6.86	1.53			1.14
208	COOMe	*l*-bornyl	DL	3.16	3.35	−0.09	1.36	4.73	6.05	2.05			1.15
209	COOMe	fencyl	DL	4.79	4.33	0.46	1.36	4.73	6.05	2.98			1.15
210	COOEt	Et		1.35	2.02	−0.67	1.36	5.95	4.80	1.53			−0.52
211	COOEt	*t*-Bu	DL	2.32	2.21	0.11	1.36	5.95	6.05	1.53			0.78
212	COOEt	c-amyl	DL	2.44	2.03	0.41	1.36	5.95	5.59	1.35			0.33
213	COOEt	c-hexyl	DL	2.68	2.21	0.47	1.36	5.95	6.05	1.53			0.74
214	COOEt	*trans*-2-Me-c-hexyl	DL	3.09	3.52	−0.43	1.36	5.95	6.05	2.78			1.04
215	COOEt	fencyl	DL	4.05	3.73	0.32	1.36	5.95	6.05	2.98			1.65
216*	COOMe	3-Me-6-*i*-Pr-c-hexyl	DL	NS	2.57		1.36	5.95	6.86	1.53			1.64
217*	COOEt	*t*-Bu	DL	S	2.02		1.36	5.95	4.80	1.53			0.08

[a] Compounds not included in the regression analysis have asterisks. [b] For abbreviations, see footnote *b* of Table I. [c] Compounds derived from D-threonine. [d] Compounds derived from D-allothreonine.

classes of compounds. If the interaction is a hydrogen-bonding one, it is very sensitive to the geometry of binding. The steric dimensions of this series of compounds are most suitable for the formation of a substrate–receptor complex with the proper geometry for hydrogen bonds. The parabolic relationship between the activity and the L_1 and L_2 terms shows that both the R_1 and R_2 moieties are not bound so tightly to the receptor; therefore, the sweetness increases first with the increase in substituent length and then decreases beyond the optimum, ~3.7 (0.37 nm) and ~5.5 (0.55 nm) for L_1 and L_2, respectively. This suggests proper room for a hydrogen-bonding interaction. The sweet potencies of compounds **197–215** are outstanding. According to eq 7, this is, first of all, attributable to their

Table X. Development of Equation 6

constant	σ^*	In	L_1	L_2	L_2^2	r	s	$F_{1,x}$[a]
−0.12	2.14					0.41	0.45	5.15
0.04	2.11	−0.53				0.64	0.34	9.48
−1.06	1.86	−0.63	0.34			0.79	0.32	12.89
−7.03	1.77	−0.64	0.32	2.21	−0.19	0.90	0.23	9.95[b]

[a] $F_{1,23;\alpha=0.05} = 4.28$. [b] The value of $F_{2,21}$. $F_{2,21;\alpha=0.05} = 3.47$.

Table XI. Squared Correlation Matrix for Variables Used in the Derivation of Equation 6

	σ^*	L_1	L_2	In
L_1	0.01			
L_2	0.00	0.01		
In	0.00	0.04	0.00	
π	0.52	0.04	0.27	0.02

large σ^* values, which are brought about by the two ester carbonyls bound β to the bridged amide nitrogen atom. Compounds **203** and **210** are among the sweetest compounds known, and the $(W_l)_2$ term with the positive coefficient in eq 7 explains this in addition to the factors described above. The shape of these compounds is thought to be the best fit for the receptor cavity at the nearby site where the methylcyclohexyl and fencyl moieties are located.

The trifluoroacetamidosuccinanilic acid derivatives, $CF_3CONHCH(CH_2COOH)CONHC_6H_4$-X, reported by Lapidus et al.,[23] are structurally very similar to the compounds studied here. Their activity data were not analyzed quantitatively because of an insufficient number of data points. It is worthy to note here, however, that the sweet potency of some of these derivatives is very high (that of the most active one is 3000 times that of sucrose[23]) despite their being *N*-(trifluoroacetyl) derivatives of aspartyl dipeptides, in which the aspartic acid amino group has to be unsubstituted for sweetness.[24] Their mode of steric interaction with the receptor may thus differ from those of the compounds analyzed here. However, the high sweet potency appears attributable at least partly to their large σ^* values due to the trifluoroacetyl group directly bound to the amide nitrogen and the two other carbonyl functions.

Figure 4 was drawn to show schematically the above considerations on the binding of the L-aspartyl dipeptide analogues to the sweet taste receptor. The upper map is the view from above when the backbone chain of the molecule is placed on the plane of the page[25] and the lower one is the view from the H atom along the OH axis of the aspartic acid end, placing the plane of the upper map perpendicular to the page. The solid lines express the spatial barriers of the receptor site drawn according to the

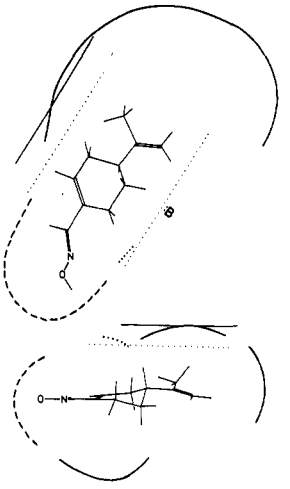


Figure 5. Arrangement of the perillartine molecule in the sweet receptor. The straight solid lines express the spatial walls and the straight dotted lines the bitter barriers drawn according to the previous perillartine–receptor binding model.[5] The molecular model was constructed by the STERIMOL program.[13]

equations developed in this study. The thickened lines in each panel are the spatial walls, which show the steric parameters incorporated into the equation for each series of compounds. Accordingly, it is understood that $(W_r)_2$ in Figure 4 (I and III) corresponds to L_2 in Figure 4 (IV) for the aspartylaminoacetates. As seen from the results of Mazur et al.,[6,24] the epimers of the sweet compounds either at the aspartic acid moiety or at the C_1 carbon are tasteless, bitter, or only slightly sweet. Thus, there may exist additional spatial barriers at the nearby position of the concerned atoms facing the epimeric substituents, and these are expressed by dotted lines. It has also been reported that the taste becomes bitter when the L-aspartic acid moiety of compound **3** is replaced by L-glutamic acid,[6] which is one methylene unit longer than the original sweet compound. Furthermore, the compounds where the aspartic acid moiety of compound **19**, L-aspartyl-L-phenylalanine, is replaced by various amino acids are bitter,[24] i.e., in the case of alanyl-, glycyl-, histidyl-, isoleucyl-, leucyl-, lysyl-, norleucyl-, norvalyl-, phenylalanyl-, prolyl-, sarcosyl-, threonyl-, tryptophyl-, tyrosyl-, and valylphenylalanines. Some of these amino acids are smaller and some others larger than the parent aspartic acid, and some of them lack the carboxylic acid end. These observations appear to suggest the existence of an interaction site or sites closer to the aspartic acid end which seems to influence taste quality. Probably a subtle conformational change in the receptor caused by the interaction of these dipeptides at this region causes the bitter sensation. These considerations are depicted by the broken line in Figure 4.

Previous analyses of the taste potency of perillartine and 5-nitro- and 5-cyanoaniline derivatives have indicated that the modes of interaction of these two classes of compounds with the receptor are very similar and that the consecutive structural change transforms the sweet taste into a bitter one.[5] Thus, it is meaningful to test the validity of the sweet receptor model drawn in Figure 4 on other classes of

(23) M. Lapidus and M. Sweeney, *J. Med. Chem.*, 16, 163 (1973).
(24) R. H. Mazur, J. M. Schlatter, and A. H. Goldkamp, *J. Am. Chem. Soc.*, 91, 2684 (1969).
(25) The backbone conformation is the fully extended one as described under Methods; therefore, each dihedral angle of the chain is 180°. This conformation is slightly different from, but in most important respects the same as, that proposed by Lelj et al.[26] for the most populated conformer in aqueous solution. Within the error of the experiments and calculations made by Lelj et al.,[26] their model is essentially the same as that constructed by the conventional extended conformation. At any rate, the principles of the drawing itself are little influenced by conformation and, so far as these two conformations are concerned, the resulting schemata of the receptor site is much the same.
(26) F. Lelj, T. Tancredi, P. A. Temussi, and C. Toniolo, *J. Am. Chem. Soc.*, 98, 6669 (1976).

Table XII. Development of Equation 7

constant	σ^*	$(W_1)_2$	In	L_1	L_1^2	L_2	L_2^2	InM	r	s	$F_{1,x}$[a]
−0.06	2.23								0.75	0.66	68.42
−1.31	1.58	1.11							0.86	0.51	35.37
−1.04	1.41	1.09	−0.63						0.89	0.46	11.79
−3.97	1.43	1.12	−0.79	1.46	−0.18				0.92	0.41	8.13[b]
−9.97	1.26	1.08	−0.81	1.41	−0.17	2.28	−0.20		0.94	0.36	9.30[b]
−10.87	1.54	1.05	−0.74	1.49	−0.19	2.46	−0.21	0.37	0.95	0.34	6.38

[a] $F_{1,40;\alpha=0.05} = 4.09$. [b] The value of $F_{2,x}$. $F_{2,40;\alpha=0.05} = 3.23$.

Table XIII. Squared Correlation Matrix for Variables Used in the Derivation of Equation 7

	σ^*	L_1	L_2	$(W_1)_2$	In	InM
L_1	0.47					
L_2	0.07	0.01				
$(W_1)_2$	0.22	0.11	0.03			
In	0.06	0.02	0.01	0.02		
InM	0.26	0.13	0.02	0.03	0.00	
π	0.00	0.02	0.19	0.08	0.05	0.20

sweeteners. Figure 5 shows that the perillartine molecule can be easily accomodated to it without causing significant discrepancies with the previously proposed perillartine–receptor binding model.[5] The spatial walls and bitter barriers of the previous model are depicted in Figure 5 by the straight solid and the dotted lines, respectively, and these are not in conflict with the considerations above on the aspartyl dipeptide analogues and their receptor model. Sucrose is also found to fit the receptor model, so far as overall molecular shape is concerned (Figure 6). The sweet potency of sucrose and other sugar derivatives is very weak compared to those of perillartines, L-aspartyl dipeptides, and most other synthetic sweeteners. This seems attributable in part to the rather ragged, puckered molecular shape of these sugar derivatives. In the quantitative analysis of the perillartine derivatives,[5] the hydrophobic parameter, log P (the logarithm of the 1-octanol/water partition coefficient), was important in correlating activity, together with steric parameters similar to those used in this study. Thus, the low potency of sugars may also be attributable to the polyol structure. Further examination of the applicability of this model to other diverse classes of sweeteners, as well as further refinement of the model itself, will be the concern of future studies.

As for hydrophobicity, and in contrast to the perillartine derivatives, the parameter π was not significant in correlating the activity of the aspartyl dipeptide analogues. The highly hydrophilic aspartic acid moiety may cover a variation in the hydrophobicity of the rest of the molecule in the process of partioning from saliva onto the receptor site of the tongue. The hydrophobic parameter may, however, be incorporated into the equations for the set of compounds with more diverse variations in hydrophobicity. Throughout the development of the equations, except for eq 7, the squared simple correlation coefficients between π and σ^* are rather larger than those between π and other steric parameters. This is perhaps because the electron-withdrawing groups in the present compounds are at the same time somewhat polar. Thus, an examination of compounds with electron-withdrawing but without polar substituents should be carried out both to refine the analyses and to develop compounds of higher potency.

While this work was being carried out, van der Heijden et al.[27] reported a similar QSAR study, but they restricted their analyses to the aspartyl amino acid methyl esters

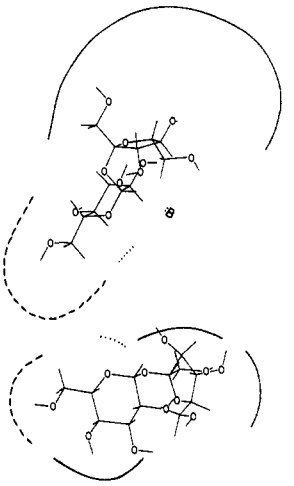


Figure 6. Arrangement of the sucrose molecule in the sweet receptor. The molecular model was constructed by the STERIMOL program.[13] The conformation of the glucose moiety is taken as chair with the substituents at carbon-2 to carbon-5 equatorial and that of the fructose moiety approximated by the flat one.

[L-Asp-NHCH(COOMe)R] only. The analyses were performed using the STERIMOL (length and width) parameters of the R moiety, as well as the parachor parameter P,[28] which is nearly equivalent to the molecular volume, and the hydrophobic parameter f described by Rekker.[29] Within this limited number of compounds, they suggest the importance of a strong hydrophobic character, a high molecular volume, and a narrow width for the substituent R in sweetness, according to the equation $\log S = 0.193f + (1.491 \times 10^{-2})P - 3.11B_5 - 0.172$ ($n = 31$, $r = 0.91$, $s = 0.40$), where S is the sweet potency times sucrose on a molar basis and the STERIMOL parameter B_5 is the maximum width of R. The discrepancy from our results appears to be due to their disregard of the electronic factor. Their analysis using the σ^* values as the electronic parameter resulted in the σ^* term entering into the above equation, the ρ value being 1.30 (±0.95), improving the correlation, and the r and s values being 0.93 and 0.35, respectively. They excluded other aspartyl dipeptide analogues from the analysis because of the expected complexity of the calculations due to variations in the two moieties, i.e., the substituents R_1 and R_2 in the present notation, and because of their relatively weak sweet potency, less than 300 times that of sucrose. With these compounds, if they had considered the electronic factor they would have reached a similar conclusion to ours, probably accompanied by a subdivision of the steric and/or parachor parameters. Worthy to note here is, however, the fact that the hydrophobic parameter is important for the compounds chosen by Heijden and his co-workers.[27] The possible significance of the hydrophobicity of the com-

(27) A. van der Heijden, L. B. P. Brussel, and H. G. Peer, *Chem. Senses Flavour*, 4, 141 (1979).

(28) O. R. Quayle, *Chem. Rev.*, **53**, 439 (1953).

(29) R. F. Rekker, "The Hydrophobic Fragmental Constant", Elsevier Scientific, Amsterdam, 1977.

pounds was discussed above, but these workers's results appear to suggest that the hydrophobicity depends on the selection of compounds. This is a reemphasis of the necessity to examine compounds with more diverse structural variations.

The results of the present and the previous[5] quantitative studies strongly suggest a close relationship between sweet receptors for various kinds of compounds, as well as a close relationship between sweet and bitter receptors. The receptor model drawn in this study is different from that proposed recently by Temussi et al.,[30] which was based on qualitative interpretations of the many kinds of strongly sweet compounds, endorsing Shallenberger's A-H/B theory.[1] These acidic and basic sites are conventionally assigned in the aspartyl dipeptide analogues to the aspartic amino and carboxylic acid moieties, respectively.[3,27,30] The present results are not at all helpful in determining the basic interaction site, but they suggest an acidic site at either the amide hydrogen or carbon atom. The aspartic amino moiety is far apart from the electronic effect, being shielded by a methine group.

Acknowledgment. The author expresses his sincere thanks to Professor Toshio Fujito for his stimulating discussions and advice and to Professor Koichi Koshimizu for his support of this work. The regression analyses were carried out on a FACOM 230/75 computer of the Data Processing Center of this university.

(30) P. A. Temussi, F. Lelj, and T. Tancredi, *J. Med. Chem.*, 21, 1154 (1978).

Aromatic Retinoic Acid Analogues. Synthesis and Pharmacological Activity

Marcia I. Dawson,* Peter D. Hobbs, Rebecca L. Chan, Wan-Ru Chao, and Victor A. Fung

Bio-Organic Chemistry Laboratory, SRI International, Menlo Park, California 94025. *Received October 9, 1980*

Aromatic analogues of (*all-E*)- and 13(*Z*)-retinoic acids have been synthesized as potential chemopreventive agents for the treatment of epithelial cancer. In the *E* series, (1*E*,3*E*)-1-(4-carboxyphenyl)-2-methyl-4-(2,6,6-trimethyl-1-cyclohexen-1-yl)-1,3-butadiene (**7a**), its ethyl ester **5a**, and the epoxy ethyl ester **14** displayed excellent activity in the assay for the inhibition of tumor promotor-induced mouse epidermal ornithine decarboxylase, while (1*E*,3*E*)-1-(4-carboethoxy-3-methylphenyl)-2-methyl-4-(2,6,6-trimethyl-1-cyclohexen-1-yl)-1,3-butadiene (**5b**) was inactive. The 13(*Z*) analogues, (*E*)-1-(2-carboxyphenyl)-4-methyl-6-(2,6,6-trimethyl-1-cyclohexen-1-yl)-1,3,5-hexatriene (**19**) and (*E*)-1-(2-hydroxyphenyl)-4-methyl-6-(2,6,6-trimethyl-1-cyclohexen-1-yl)-1,3,5-hexatriene (**27**), had minimal activity.

Evidence that synthetic retinoids are capable of suppressing or reversing the transformation of premalignant epithelial cells to the malignant state[1] has prompted the search for new structurally modified retinoids that may possses enhanced prophylactic and therapeutic activity and reduced systemic toxicity (hypervitaminosis A).[2] The recent report[3] on the synthesis and favorable biological activity of a series of aromatic analogues of retinoic acid has prompted us to present our own results on a similar series of aromatic analogues.[4]

Our compounds were designed to probe what structural constraints on retinoid conformation are necessary for biological activity. The *p*-carboxyphenyl trienes **7a** and **7b** could be considered as analogues of (*all-E*)-retinoic acid. Carbons 1 to 4 of the aromatic ring would correspond to carbons 11 to 14 of the retinoid chain, in which the (*E*)-11,12 and (*E*)-13,14 double bonds are held in an *s*-cis or cisoid conformation.[5] The *o*-methyl substituent on the aromatic ring of **7b** would correspond to the C-20 methyl group of retinoic acid.

The *o*-carboxyphenyl tetraene **19** and the *o*-hydroxyphenyl tetraene **27** could be envisioned as analogues of 13(*Z*)-retinoic acid, which has been shown by Sporn et al.[6] to prevent nitrosamine-induced bladder lesions in the rat and by Hixson et al.[7] to be less toxic than the (*all-E*)-acid in the mouse. Carbons 1 and 2 of the aromatic ring of **19** and **27** would correspond to carbons 13 and 14 of the retinoid chain. In contrast to 13(*Z*)-retinoic acid, isomeri-

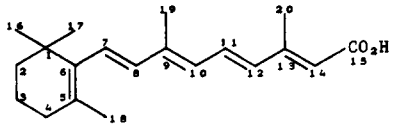
(1) (a) Sporn, M. B.; Newton, D. L.; Smith, J. M.; Acton, N.; Jacobson, A. E.; Brossi, A. In "Carcinogens: Identification and Mechanism of Action"; Griffin, A. C.; Shaw, C. R., Eds., Raven Press: New York, 1979; pp 441-453. (b) Sporn, M. B.; Dunlop, N. M.; Newton, D. L.; Smith, J. M. *Fed. Proc., Fed. Am. Soc. Exp. Biol.* **1976**, *35*, 1332. (c) Sporn, M. B.; Dunlop, N. M.; Newton, D. L.; Henderson, W. R. *Nature (London)* **1976**, *263*, 110. (d) Sporn, M. B.; Newton, D. L. *Fed. Proc., Fed. Am. Soc. Exp. Biol.* **1979**, *38*, 2528. (e) Todaro, G. J.; DeLarco, J. E.; Sporn, M. B. *Nature (London)*, **1978**, *276*, 272.

(2) (a) Acton, N.; Brossi, A.; Newton, D. L.; Sporn, M. B. *J. Med. Chem.* **1980**, *23*, 805. (b) Dawson, M. I.; Hobbs, P. D.; Kuhlmann, K.; Fung, V. A.; Helmes, C. T.; Chao, W.-R. *Ibid.* **1980**, *23*, 1013. (c) Dawson, M. I.; Hobbs, P. D. *Carbohydr. Res.* **1980**, *85*, 121. (d) Davalian, D.; Heathcock, C. H. *J. Org. Chem.* **1979**, *26*, 4988. (e) Davalian, D.; Heathcock, C. H. *Ibid.* **1979**, *26*, 4458. (f) Pawson, B. A.; Chan, K.-K.; DeNoble, J.; Han, R. L.; Piermattie, V.; Specian, A. C.; Srisethnil, S. *J. Med. Chem.* **1979**, *22*, 1059. (g) Pawson, B. A.; Cheung, H.-C.; Han, R.-J.; Trown, P. W.; Buck, M.; Hansen, R.; Bollag, W.; Ineichen, U.; Pleil, H.; Rüegg, R.; Dunlop, N. M.; Newton, D. L.; Sporn, M. B. *Ibid.* **1977**, *20*, 918. (h) Welch, S. C.; Gruber, S. *Ibid.* **1979**, *22*, 1532.

(3) Loeliger, P.; Bollag, W.; Mayer, H. *Eur. J. Med. Chem.* **1980**, *15*, 9.

(4) A preliminary account of this work was presented at the Second Chemical Congress of the North American Continent, Las Vegas, Nev., Aug, 1980.

(5) For structural comparisons standard retinoid numbering has been used:



Similar proton and carbon atoms in the aromatic analogues have been denoted by the subscript R. The aryl carbon atoms of those retinoids have been denoted as 1′ to 6′. The position bearing the polyene substituent is numbered C-1′ and the remaining positions are numbered in the direction of lowest numerical assignment to the other substituents.

(6) Sporn, M. B.; Squire, R. A.; Brown, C. C.; Smith, J. M.; Wenk, M. L. *Science* **1977**, *195*, 487.

(7) Hixson, E. J.; Denine, E. P. *Toxicol. Appl. Pharmacol.* **1978**, *44*, 29.

0022-2623/81/1824-0583$01.25/0