plete the search goes up very rapidly with the number of rotatable bonds. Despite clever schemes to make the searches more efficient, the method described by Marshall et al.[1] is limited to a fairly small number of rotatable bonds (less than eight) or a fairly coarse (30° increment) search "grid". The computational time for the ensemble approach, which does not require a search, is independent of the number of rotatable bonds.

One important feature of distance geometry methods is that they are Monte Carlo methods. That is, conformational space within the distance constraints of the ensemble is randomly sampled. This has the drawback that we may have to take a large number of samples to be sure not to miss an interesting solution. Also we have to arrange the solutions into families for analysis. In contrast, a systematic search can in principle generate a complete set of distinct conformations (although, in practice, some interesting conformations may be missed if the grid is too coarse). For problems for which a complete set of solutions must be found, systematic search methods may be preferred. For problems in which a representative sample of solutions (or the fact that a solution does not exist) is sufficient, Monte Carlo methods may be applied to advantage.

Conclusion

We describe a new application of distance geometry in which two or more molecules are treated as an ensemble. With this approach we can find a common pharmacophore from a small set of biologically active molecules and generate coordinates for the set of molecules in their receptor-bound conformations such that their essential groups are superimposed. This approach has several advantages over previous methods of finding pharmacophore geometries, especially for molecules containing flexible rings.

We find only one pharmacophore geometry compatible with all four nicotinic agonists, assuming that the cationic center and a pair of atoms that form a dipole are important for activity. The pharmacophore triangle formed by the three atoms has sides 4.8, 4.0, and 1.2 Å. The pharmacophore geometry, compatible with previous models in the literature, can be reached by agonists and antagonists not in the original set used to derive the pharmacophore. We suggest that a specific arrangement of the pharmacophore triangle and the bulk of the volume of agonist molecules defines a "handedness" essential for agonist activity. By docking together the conformations of various agonists that achieve the pharmacophore geometry and that have the correct handedness, we can derive a volume occupied by agonists on the receptor.

Acknowledgment. This work was supported in part by the U.S. Army Medical Research and Development Command, Contract No. DAMD17-84-C-4111.

Registry No. (-)-Nicotine, 54-11-5; (-)-cytisine, 485-35-8; (-)-ferruginine methiodide, 85514-41-6; (-)-muscarone, 16980-76-0; *trans*-3,3′-bis O, 83800-31-1; strychnine, 57-24-9; trimethaphan, 7187-66-8; dihydro-β-erythroidine, 23255-54-1; (+)-muscarone, 4780-69-2; (+)-nicotine, 25162-00-9.

Structure–Taste Correlation of L-Aspartyl Dipeptides Using SIMCA Method

Yoshikatsu Miyashita, Yoshimasa Takahashi,† Chiyozo Takayama, Kazuo Sumi, Kazuya Nakatsuka, Takehiko Ohkubo, Hidetsugu Abe,† and Shin-ichi Sasaki*

School of Materials Science and Research Center for Chemometrics, Toyohashi University of Technology, Tempaku-cho Toyohashi 440, Japan. *Received November 26, 1984*

One of the pattern recognition techniques, the SIMCA method, has been applied to structure–taste studies on L-aspartyl dipeptides (L-Asp-NH-R). The sweet and bitter taste class models of the peptides were obtained by using five structural descriptors, such as molar refractivity, and four kinds of STERIMOL parameters. The classification rates were calculated to be 87% and 81% for sweet and bitter peptides, respectively. The SIMCA method has also suggested that two factors, shape and size, of the C-terminal amino acid moiety R in the dipeptides are extremely important to model their taste qualities.

Many attempts[1] have been carried out to develop new sweeteners. It is important to investigate the structure–taste relationships to obtain information of sweetener designing.

Recently, van der Heijden[2] investigated the quantitative structure–sweetness relationships of L-aspartyl dipeptide analogues. Iwamura has performed studies on the correlation between structure and taste potency of perillartine analogues[3] and the structure–sweetness relationships of L-aspartyl dipeptide analogues.[4]

Kier[5] examined a series of perillartines for their sweet or bitter taste by using molecular connectivity indices and discriminant analysis. We have been interested in applying the pattern recognition methods to several structure–activity problems.[6] With regard to the structure–taste relations, for example, the classification of perillartines into sweet and bitter classes was studied by the pattern recognition method.[7]

Ariyoshi[8] proposed that the structure–taste correlation of L-aspartyl dipeptide analogues depends on the common molecular features relating to the sweet peptide through the Fischer projection formula of dipeptides.

In this study an attempt has been made to classify the sweet and bitter class dipeptides by the SIMCA pattern

† Research Center for Chemometrics.

(1) Crosby, G. A.; Dubois, G. E.; Wingard, R. E. *Drug Design*; Ariens, E. J., Ed.; Academic: New York, 1979; Vol. 8, p 215.
(2) van der Heijden, A.; Brussel, L. B. P.; Peer, H. G. *Food Chem.* 1978, *3*, 207.
(3) Iwamura, H. *J. Med. Chem.* 1980, *23*, 308.
(4) Iwamura, H. *J. Med. Chem.* 1981, *24*, 572.
(5) Kier, L. B. *J. Pharm. Sci.* 1980, *69*, 416.
(6) Miyashita, Y.; Takahashi, Y.; Daiba, S.; Abe, H.; Sasaki, S. *Anal. Chim. Acta* 1982, *143*, 35.
(7) Takahashi, Y.; Miyashita, Y.; Tanaka, Y.; Abe, H.; Sasaki, S. *J. Med. Chem.* 1982, *25*, 1245. Takahashi, Y.; Miyashita, Y.; Tanaka, Y.; Hayasaka, H.; Abe, H.; Sasaki, S. *J. Pharm. Sci.* 1984, *73*, 737.
(8) Ariyoshi, Y.; Yasuda, N.; Yamatani, T. *Bull. Chem. Soc. Jpn.* 1974, *47*, 326.

0022-2623/86/1829-0906$01.50/0

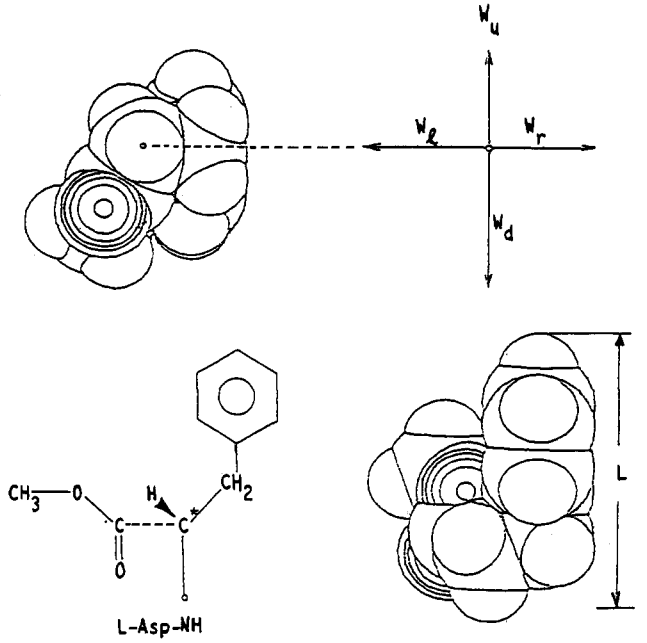


Figure 1. STERIMOL parameter.

recognition method using physicochemical descriptors.

Data Set

The taste of L-aspartyl dipeptide (L-Asp-NH-R) varies from sweet to bitter or tasteless in accordance with the difference of chemical structure of their C-terminal amino acid moieties R.

The data set used in this study consists of 70 sweet (class 1), 21 bitter (class 2), and 17 tasteless L-aspartyl dipeptides (class 0), taken from references. Table I lists the chemical structures, configurations, taste qualities, descriptor values of the dipeptides, and the corresponding reference numbers.

It is believed that the taste response might be closely related to the size, shape, and functionality of a molecule.[1] According to this postulation, these three factors can be described by molar refractivity (MR), the STERIMOL constants (L, W_r, W_l, W_u, W_d), and the Taft's σ^* constant, respectively. These descriptors stand for physicochemical constants for the remaining substructure R by eliminating the common L-aspartyl amino moiety from the dipeptide. These descriptor values were taken from the books edited by Leffler[9] and Hansch.[10]

The STERIMOL W constants were calculated by the STERIMOL program prepared by Iwamura.[3]

A space-filling representation of aspartame (L-Asp-L-Phe-OMe) is illustrated in Figure 1 to give the definition of these W constants. At first the bulkiest group is set on the right hand side, and N, C*, and C atoms are put on a paper. L is the length of the substructure R along the N–C* axis. W_r constant is defined as the right-hand width of R and W_l is the opposite one. W_u and W_d mean the width upward and downward, respectively.

The conformation of substructure R is fixed in the form of the fully extended (staggered) conformation. For aspartame we have adopted the $F_I D_{II}$ conformation, which was proposed to be an active conformer by using the NMR data and the partitioned energy model of Lelj et al.[11]

Method

We have employed the SIMCA method developed by Wold.[12] Minimum introduction of this method will be described below. The basic idea of this method is to model the data of each class separately. The class models are described by principal component models. $x_{ik}^{(q)}$ is the kth

$$x_{ik}^{(q)} = \alpha_k^{(q)} + \sum_{\beta=1}^{r_q} z_{i\beta}^{(q)} t_{k\beta}^{(q)} + \epsilon_{ik}^{(q)} \quad (1)$$

descriptor value for a compound i in class q. The parameters $\alpha_k^{(q)}$, $z_{i\beta}^{(q)}$ and $t_{k\beta}^{(q)}$ ($\beta = 1, 2, ..., r_q$) are determined so as to minimize the variance of the residuals $\epsilon_{ik}^{(q)}$. α_k is the mean of descriptor k, $z_{i\beta}^{(q)}$ is the βth principal component, $t_{k\beta}^{(q)}$ is the kth component of βth eigenvector, and r_q is the number of components in class q. r_q is estimated by a cross-validation technique.

The tolerance interval of class q is given by $RSD^{(q)}$, which is defined by

$$RSD^{(q)} = [\sum_{i}^{n_q} \sum_{k}^{d} (\epsilon_{ik}^{(q)})^2 / (n_q - r_q - 1)(d - r_q)]^{1/2} \quad (2)$$

where n_q is the number of samples in class q, and d is the number of descriptors. Classification of a sample is made by the comparison of its class fit distance with other samples distances when fitted to their respective models. Distance is calculated by

$$D_i^{(q)} = \sum_{k=1}^{d} [(\epsilon_{ik}^{(q)})^2 / (d - r_q)]^{1/2} \quad (3)$$

The SIMCA method also provides information for reducing the number of descriptors. The discrimination power of descriptor k to discriminate class p from class q is defined by

$$\phi_k^{p,q} = \left[\frac{(S_{k,p}^{(q)})^2 + (S_{k,q}^{(p)})^2}{(S_{k,p}^{(p)})^2 + (S_{k,q}^{(q)})^2} \right]^{1/2} - 1 \quad (4)$$

where $S_{k,p}^{(q)}$ is the residual standard deviation of descriptor k when the samples in class p fit to the class model q. The closer to zero $\phi_k^{p,q}$ is, the lower discriminating power of

(9) Leffler, J. E.; Grunwald, E. *Rates and Equilibria of Organic Reaction*; Wiley: New York, 1963.
(10) Hansch, C.; Leo, A. *Substituent Constants for Correlation Analysis in Chemistry and Biology*; Wiley: New York, 1979.
(11) Lelj, F.; Tancredi, T.; Temussi, P. A.; Toniolo, C. *J. Am. Chem. Soc.* **1976**, *98*, 6669.
(12) Wold, S.; Sjostrom, M. *Chemometrics; Theory and Application*; Kowalski, B. R., Ed.; American Chemical Society: Washington, DC, 1977.
(13) Mazur, R. H.; Goldkamp, A. H.; James, P. A.; Schlatter, J. M. *J. Med. Chem.* **1970**, *13*, 1217.
(14) Ariyoshi, Y. *Agric. Biol. Chem.* **1976**, *40*, 983.
(15) Ariyoshi, Y. *Kagaku Sosetsu*, No. 14 (Chemistry of Taste and Smell), The Chemical Society of Japan, Japan Scientific Society Press: Tokyo, 1976; p 85.
(16) Miyoshi, M.; Nunami, K.; Sugano, H.; Fujii, T. *Bull. Chem. Soc. Jpn.* **1978**, *51*, 1433.
(17) Ariyoshi, Y.; Yasuda, N.; Yamatani, T. *Bull. Chem. Soc. Jpn.* **1974**, *47*, 326.
(18) Inglett, G. E. *Symposium: Sweeteners*; Avi: Westport, CT, 1974.
(19) Ney, K. H. *Z. Lebensm. Unters. Forsch.* **1971**, *147*, 64.
(20) Mazur, R. H.; Schlatter, J. M.; Goldkamp, A. H. *J. Am. Chem. Soc.* **1969**, *91*, 2684.
(21) Brussel, L. B. P.; Peer, H. G.; van der Heijden, A. *Z. Lebensm. Unters. Forsch.* **1975**, *159*, 337.
(22) Gardner, R. J. *J. Sci. Food Agric.* **1980**, *31*, 23.
(23) Ariyoshi, Y. *Food Taste Chemistry*; Boudreau, J. C., Ed.; American Chemical Society: Washington, DC, 1979.
(24) Kawai, M.; Chorev, M.; Marin-Rose, J.; Goodman, M. *J. Med. Chem.* **1980**, *23*, 420.

Table I. Chemical Structure, Configuration, Taste Quality, and Descripter Value[a]

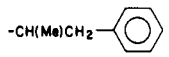
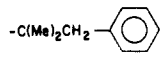
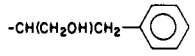
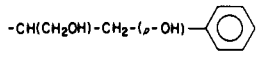
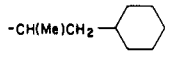
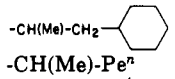
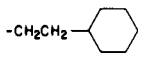
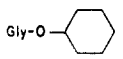
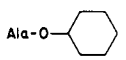
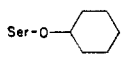
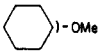
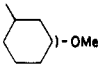
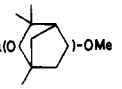
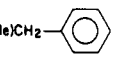
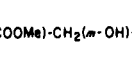
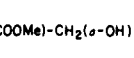
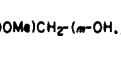
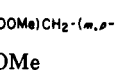
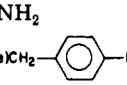
no.	NH–R	config	taste class	MR	L	W_r	W_l	W_u	W_d	σ^*	ref
1		L	sweet	39.3	8.33	3.15	2.76	3.11	3.16	0.03	13
2			sweet	44.0	8.33	3.15	2.76	3.16	3.16	−0.12	13
3		L	sweet	40.8	8.33	3.15	2.76	3.11	3.16	0.29	13
4		L	sweet	42.6	9.01	3.15	2.76	3.11	3.16	0.26	13
5		L	sweet	40.6	8.22	4.94	2.76	3.16	3.16	−0.26	13
6		D	sweet	40.6	8.22	4.94	2.76	3.16	3.16	−0.26	13
7	-CH(Me)-Pe^n	L	sweet	33.6	8.22	5.87	2.76	1.90	3.16	−0.27	13
8	-CH(Me)-Pe^i	L	sweet	33.6	6.78	5.18	2.76	3.16	3.16	−0.26	13
9	-CH(Me)-Pe^i	D	sweet	33.6	6.78	5.18	2.76	3.16	3.16	−0.26	13
10	-CH(Me)-He^n	L	sweet	38.0	9.03	6.39	2.76	1.90	3.16	−0.27	13
11	-CH(Me)-Bu^n	L	sweet	28.8	6.97	4.94	2.76	1.90	3.16	−0.26	13
12			sweet	35.9	8.22	4.94	1.52	3.16	3.16	−0.16	13
13	Gly-OMe		sweet	16.5	5.98	4.28	1.52	1.90	1.90	0.82	14
14	Gly-OPr^n		sweet	25.8	8.04	5.71	1.52	1.90	1.90	0.82	14
15	Gly-OPr^i		sweet	25.8	6.59	5.03	1.52	3.16	3.16	0.82	15
16	Gly-OBu^n		sweet	30.4	8.86	6.21	1.52	1.90	1.90	0.82	16
17			sweet	37.5	7.97	6.71	1.52	3.16	3.16	0.82	14
18	β-Ala-OMe		sweet	21.1	6.78	4.81	1.52	1.90	1.90	0.28	17
19	β-Ala-OPr^i		sweet	30.4	8.01	5.54	1.52	3.16	3.16	0.28	17
20	(α-Me)Ala-OPr^i		sweet	36.0	6.59	5.03	2.76	3.16	3.16	0.62	18
21	Ala-OMe	D	sweet	21.1	5.98	4.28	2.76	1.90	3.16	0.72	19
22	Ala-OEt	D	sweet	25.7	6.79	4.78	2.76	1.90	3.16	0.72	19
23	Ala-OPr^n	D	sweet	30.3	8.04	5.71	2.76	1.90	3.16	0.72	19
24	Ala-OPr^i	D	sweet	30.3	6.59	5.03	2.76	3.16	3.16	0.72	19
25	Ala-OBu^n	D	sweet	34.9	8.86	6.21	2.76	1.90	3.16	0.72	19
26	Ala-Pe^n	D	sweet	39.5	10.1	7.15	2.76	1.90	3.16	0.72	19
27		D	sweet	42.1	7.97	6.71	2.76	3.16	3.16	0.72	16
28	γ-Abu-OMe		sweet	25.7	8.03	5.73	1.52	1.90	1.90	0.10	17
29	Abu-OMe	D	sweet	25.7	5.98	4.28	2.76	1.90	3.16	0.71	17
30	(α-Et)Abu-OPr^i		sweet	44.3	6.59	5.03	2.76	3.16	3.16	0.59	18
31	Abu-OPr^n	D	sweet	35.0	8.04	5.71	2.76	1.90	3.16	0.71	17
32	Abu-OPr^i	D	sweet	35.0	6.59	5.03	2.76	3.16	3.16	0.71	19
33	NVa-OMe	L	sweet	30.4	6.17	4.42	3.42	1.90	4.29	0.69	17
34	Val-OPr^n	D	sweet	39.7	8.04	5.71	3.63	1.90	3.99	0.70	15
35	Val-OPr^i	D	sweet	39.7	6.59	5.03	3.63	3.16	3.99	0.70	19
36	(α-Me)Val-OPr^i	D	sweet	44.3	6.59	5.03	3.63	3.16	3.99	0.60	18
37	NVa-OPr^n	D	sweet	39.7	8.04	5.71	3.49	1.90	4.41	0.69	14
38	NLe-OMe	L	sweet	35.0	6.97	4.94	3.42	1.90	4.29	0.66	19
39	NLe-OEt	L	sweet	40.7	6.97	4.94	3.42	1.90	4.29	0.66	19
40	ILe-OPr^i	D	sweet	36.0	6.59	5.03	3.49	3.16	4.41	0.70	19
41	HyNle-OMe (erythro)	L	sweet	37.5	6.97	4.94	3.42	2.52	4.29	0.98	17
42	HyNle-OMe (threo)	L	sweet	37.5	6.97	4.94	3.42	1.90	4.29	0.98	17
43	Cap-OMe	L	sweet	44.3	9.03	6.39	3.42	1.90	4.29	−0.10	15
44	Phe-OMe	L	sweet	46.5	8.33	3.15	3.42	3.11	4.29	0.90	20
45	HPhe-OMe	L	sweet	46.8	8.22	4.94	3.42	3.16	4.29	0.66	13
46	(p-NH_2)Phe-OMe	L	sweet	49.8	9.05	3.15	3.42	3.11	4.29	0.87	15
47	Tyr-OMe	L	sweet	54.0	9.01	3.15	3.42	3.11	4.29	0.87	20
48	Ser(But^i)-OMe	L	sweet	41.2	8.04	5.51	3.42	3.16	4.29	1.06	17
49	Ser(But)-OMe	L	sweet	32.3	6.85	5.56	3.42	3.16	4.29	1.06	15
50	Ser-OMe	D	sweet	22.6	5.98	4.28	2.76	1.90	3.16	1.03	17
51	Ser-OEt	D	sweet	27.2	6.79	4.78	2.76	1.90	3.16	1.03	17
52	Ser-OPr^n	D	sweet	31.8	8.04	5.71	2.76	1.90	3.16	1.03	17
53	Ser-OPr^i	D	sweet	31.8	6.59	5.03	2.76	3.16	3.16	1.03	17
54	Ser-OBu^n	D	sweet	36.5	8.86	6.21	2.76	1.90	3.16	1.03	17
55	Ser-OBu^i	D	sweet	36.5	8.04	5.51	2.76	3.16	3.16	1.03	17
56		D	sweet	43.7	7.97	6.71	2.76	3.16	3.16	1.03	17
57	Met-OMe	L	sweet	36.0	6.36	4.69	3.42	1.90	4.29	1.10	20
58	Thr-OMe	D	sweet	36.6	5.98	4.28	2.76	1.90	3.77	1.04	17
59	Thr-OPr^n	D	sweet	37.6	8.04	5.71	2.76	1.90	3.77	0.98	17
60	Thr-OPr^i	D	sweet	37.6	6.59	5.03	2.76	3.16	3.77	0.98	17
61	*a*Thr-OMe	D	sweet	28.3	5.98	4.28	3.55	1.90	3.16	0.98	17
62	*a*Thr-OPr^n	D	sweet	37.6	8.04	5.71	3.55	1.90	3.16	0.98	17

Table I (Continued)

no.	NH–R	config	taste class	MR	L	W_r	W_l	W_u	W_d	σ^*	ref
63	*a*Thr-OPr^i	D	sweet	37.6	6.59	5.03	3.55	3.16	3.16	0.98	17
64	Lys(Ac)-OMe	L	sweet	48.9	10.1	7.22	3.42	1.90	4.29	0.83	17
65	Orn(Ac)-OMe	L	sweet	44.3	8.88	6.61	3.42	1.90	4.29	0.84	17
66	Glu-$(OMe)_2$	L	sweet	36.6	8.03	5.73	3.42	1.90	4.29	0.91	21
67	Ama(O )-OMe		sweet	43.7	7.50	6.03	3.42	4.05	4.29	1.64	22
68	Ama(O )-OMe		sweet	48.4	7.97	6.71	3.42	3.16	4.29	1.64	22
69	Ama(O )-OMe		sweet	53.0	7.97	6.71	3.42	3.16	4.29	1.64	22
70	Ama(O )-OMe		sweet	64.8	8.22	6.49	3.42	4.41	4.29	1.64	22
71	-CH(Me)CH_2- 	D	tasteless	39.3	8.33	3.15	2.76	3.16	3.11	0.03	13
72	-CH_2CH_2- 		tasteless	34.7	8.33	3.15	1.52	3.11	3.11	0.08	13
73	-CH(Me)-Pr^n	D	tasteless	24.3	6.17	4.42	2.76	3.16	1.90	−0.21	13
74	-CH(Me)-Pr^n	L	tasteless	24.3	6.17	4.42	2.76	1.90	3.16	−0.26	13
75	-CH(Me)-Pe^n	D	tasteless	33.4	8.22	5.87	2.76	3.16	1.90	−0.27	13
76	Abu-OMe	L	tasteless	26.4	5.98	4.28	2.76	3.16	1.90	0.71	17
77	Phe-NH_2	L	tasteless	43.4	8.33	3.15	2.54	3.11	3.41	0.73	20
78	-CH(COOMe)-CH_2(*m*-OH)- 	D	tasteless	48.3	8.33	5.20	3.42	4.29	1.90	0.91	24
79	-CH(COOMe)-CH_2(*o*-OH)- 	D	tasteless	47.3	8.33	5.20	3.42	4.29	1.90	0.87	24
80	-CH(COOMe)CH_2-(*m*-OH, *p*-OMe)- 	L	tasteless	55.1	10.3	5.20	3.42	2.87	4.29	0.99	24
81	-CH(COOMe)CH_2-(*m*,*p*-OMe)- 	L	tasteless	60.2	10.3	6.23	3.42	2.87	4.29	0.99	24
82	Ser-OMe	L	tasteless	22.6	5.98	4.28	2.76	3.16	1.90	1.03	15
83	Met-OMe	D	tasteless	38.6	7.42	4.98	3.42	4.29	1.90	0.07	20
84	Cys(Me(O_2))-OMe	L	tasteless	33.6	6.17	4.42	3.42	2.52	4.29	1.31	20
85	Tyr-OMe	D	tasteless	48.3	9.01	3.15	3.42	4.29	3.11	0.87	24
86	Thr-OMe	L	tasteless	28.3	5.98	4.28	3.42	3.16	1.90	0.98	20
87	Tyr-NH_2	L	tasteless	45.2	9.01	3.15	2.54	3.11	3.41	0.70	20
88	-CH(Me)CH_2-  -$NHSO_2CH_3$	L	bitter	52.8	11.1	4.56	2.76	3.42	3.16	0.08	13
89	-CH(Me)-Bu^i	D	bitter	28.9	6.17	4.21	2.76	3.16	3.16	−0.26	13
90	-CH(Me)-Bu^i	L	bitter	28.9	6.17	4.21	2.76	3.16	3.16	−0.26	13
91	Ala-OMe	L	bitter	21.7	5.98	4.28	2.76	3.16	1.90	0.72	20
92	Val-OMe	L	bitter	31.0	5.98	4.28	3.63	3.99	1.90	0.70	20
93	NVal-OEt	L	bitter	36.0	6.79	4.78	3.49	4.41	1.90	0.69	14
94	Leu-OMe	L	bitter	35.1	6.17	4.21	3.42	3.16	4.29	0.66	20
95	Leu-OPr^i	D	bitter	45.3	6.59	5.03	3.70	3.16	4.41	0.66	15, 19
96	ILeu-OMe	L	bitter	35.0	6.17	4.42	3.42	3.16	4.29	0.66	20
97	Cap-OEt	L	bitter	49.8	9.03	6.39	3.42	1.90	4.29	0.66	23
98	Phe	L	bitter	40.9	8.33	3.15	2.56	3.11	3.45	1.13	20
99	Phe-NHMe	L	bitter	48.2	8.33	3.15	3.49	3.11	4.42	0.84	20
100	Phe-NMe_2	L	bitter	52.6	8.33	3.15	3.44	3.11	4.32	0.84	20
101	Phe-OMe	D	bitter	46.5	8.33	3.15	3.42	4.29	3.11	0.90	20
102	Thr(COPr^i)-OMe	L	bitter	47.4	8.04	5.51	3.42	3.16	4.29	1.12	17
103	Lys-OMe	L	bitter	40.4	8.13	5.78	3.42	1.90	4.29	0.83	17
104	Trp-OMe	L	bitter	56.6	8.67	4.58	3.97	4.29	1.90	0.82	20
105	Trp-OMe	D	bitter	56.6	8.67	4.58	3.97	1.90	4.29	0.82	20
106	Tyr	L	bitter	42.7	9.00	3.15	2.56	3.11	3.45	1.10	20
107	Tyr-NHMe	L	bitter	50.0	9.01	3.15	3.49	3.11	4.42	0.81	20
108	Tyr-NMe_2	L	bitter	56.5	9.01	3.15	3.44	3.11	4.32	0.81	20

[a] Other abbreviations used:[15] Pr^n, *n*-propyl; Pr^i, *i*-propyl; Bu^n, *n*-butyl; But^i, *i*-butyroyl; Pe^n, *n*-pentyl; Pe^i, *i*-pentyl; He^n, *n*-hexyl; Bu^t, *tert*-butyl; Cap, capryline = α-aminooctanoic acid; *a*Thr, allothreonine; HyNle, β-hydroxynorleucine; HPhe, β-cyclohexyl-α-alanine.

descriptor k is, on the other hand, the more above one the higher power.

The modeling power of a descriptor k over all classes is defined by

$$\psi_k = 1 - (S_k/S_{k,x}) \quad (5)$$

where S_k is the residual standard deviation of descriptor k over all the data in the training set and $S_{k,x}$ is the standard deviation of the training set data. A value of ψ_k close to one indicates higher modeling power.

Results and Discussion

The analyses were done with the ARTHUR81 package. Additional programs were prepared in our laboratory. In

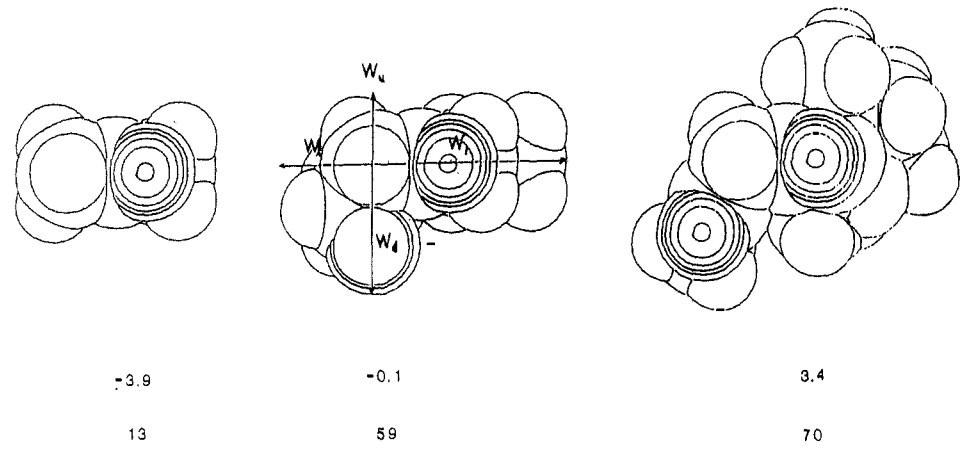


Figure 2. Substructure of sweet L-aspartyl dipeptides.

Table II. Discrimination Power and Modeling Power of Seven Descriptors for $r_1 = 1$ and $r_2 = 2$

descriptor	MR	L	W_r	W_l	W_u	W_d	σ^*
discrimination power	0.992	1.141	1.738	3.181	1.592	3.361	1.289
modeling power	0.518	0.202	-0.005	0.267	0.225	0.491	0.049

Table III. Recognition Rate (Percent)

		computed class	
		1	2
true class	1	42.9	57.1
	2	14.3	85.7
	av	64.3%	

this study, all of the sweet and bitter compounds were included in a training set while tasteless compounds were included in a test set. It is believed that sweet and bitter compounds act at two different receptor sites.[1] Tasteless compounds will not interact or will only slightly interact with these two receptors.

In the initial stages of the analysis, the descriptors were autoscaled to give equivalent variance of one and means of zero.

The number of components for each class model is determined by a double cross-validation method. These numbers are one and two for class 1 and 2, respectively; $r_1 = 1, r_2 = 2$. Modeling power and discrimination power for these class models are shown in Table II. To choose intrinsic features, the discrimination power and modeling power were used. W_r and σ^* are eliminated because of low modeling power. Therefore, the pattern vector is five-dimensional. For the training set, principal component models were calculated. Samples are classified according to the class fit distance D. The classification results are shown in Table III. The average recognition rate is 66.9%. Because of the poor results, dimensionality for the classes was reconsidered and reset to $r_2 = 1$. Modeling power and discrimination power for these two class models are shown in Table IV. L and σ^* are eliminated because of the lower modeling power. The modeling powers of W_r and W_u showed lower values too, but their discrimination powers are rather high values. Thus they are retained and the data become five-dimensional vectors. The dimensionality of each class model was redetermined by the cross-validation method; $r_1 = 1, r_2 = 1$. The classification results are shown in Table V. The recognition rates for class 1 and 2 are 87.1% and 81.0%, respectively. The average recognition rate is elevated to 84.0%. Class models for sweet and bitter dipeptides are given by eq 6 and 7.

$$\begin{aligned} \mathrm{MR} &= -0.155 + 0.517z^{(1)} + \epsilon_{\mathrm{MR}} \\ W_r &= 0.211 + 0.025z^{(1)} + \epsilon_{W_r} \\ W_l &= -0.159 + 0.599z^{(1)} + \epsilon_{W_l} \\ W_u &= -0.196 + 0.265z^{(1)} + \epsilon_{W_u} \\ W_d &= -0.021 + 0.551z^{(1)} + \epsilon_{W_d} \end{aligned} \tag{6}$$

$$\begin{aligned} \mathrm{MR} &= 0.516 + 0.464z^{(2)} + \epsilon_{\mathrm{MR}} \\ W_r &= -0.704 + 0.076z^{(2)} + \epsilon_{W_r} \\ W_l &= 0.530 + 0.150z^{(2)} + \epsilon_{W_l} \\ W_u &= 0.654 - 0.437z^{(2)} + \epsilon_{W_u} \\ W_d &= 0.069 + 0.752z^{(2)} + \epsilon_{W_d} \end{aligned} \tag{7}$$

For the class 1, RSD equals 0.754 and for the class 2, RSD equals 0.820. Each descriptor is related to only one component.

The descriptor values for each class can be estimated from the model equations. The average values of the descriptors are given, as follows.

$z^{(1)} = 0$, the nonscaled standard pattern vector for sweet compound is expressed by $\mathbf{x} = (36.9, 5.21, 2.88, 2.57, 3.49)$.

$z^{(2)} = 0$, the nonscaled standard pattern vector for bitter compound is expressed by $\mathbf{x} = (43.0, 4.23, 3.30, 3.19, 3.56)$.

$z^{(1)}$ varies from -3.9 to 3.4. Three typical sweet L-aspartyl dipeptides, which are deduced from the above standard vector, are shown in Figure 2.

The statistical data for the analysis are given in Table VI. In this table, $z^{(1)}$ and $z^{(2)}$ show the principal components of sweet and bitter classes, respectively. The distances of sample from each principal component model are also shown.

Table IV. Discrimination Power and Modeling Power of Seven Descriptors for $r_1 = 1$ and $r_2 = 1$

descriptor	MR	L	W_r	W_l	W_u	W_d	σ^*
discrimination power	1.767	0.942	2.090	3.196	2.572	1.660	1.203
modeling power	0.445	0.107	0.004	0.320	0.074	0.243	0.087

Table V. Recognition Rate (Percent)

		computed class	
		1	2
true	1	87.1	12.9
class	2	19.0	81.0
	av	84.0%	

In order to visualize the five-dimensional descriptor space, the space was transformed to two dimensionality by the Karhunen–Loève method. The resulting plot was shown in Figure 3. This plot includes tasteless compounds. The cumulative percent variance is 69.2.

In conclusion, models for structural variation in sweet and bitter dipeptides were obtained. These models indicate structural requirements for sweet dipeptides. The plausible substructures for sweet dipeptides can be sug-

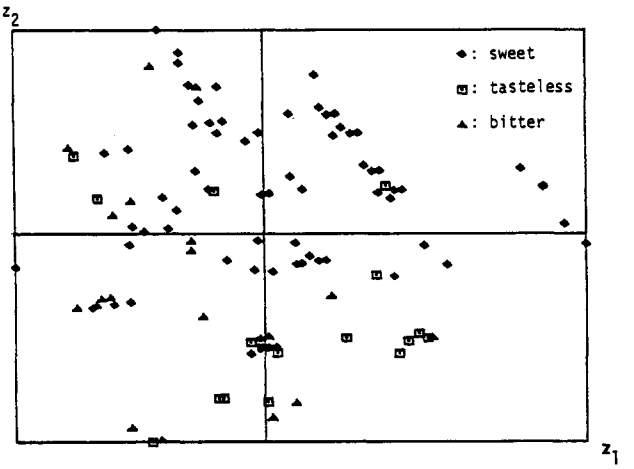


Figure 3. K–L plot of aspartyl dipeptides.

Table VI. Principal Component and Distances of Sample from Principal Component Model

				distance						distance	
no.	class	$z^{(1)}$	$z^{(2)}$	class 1	class 2	no.	class	$z^{(1)}$	$z^{(2)}$	class 1	class 2
1	1	−0.075		1.067	0.651						
2	1	0.209		1.136	0.672	47	1	2.238		1.083	0.672
3	1	0.010		1.081	0.652	48	1	1.588		0.305	0.679
4	1	0.112		1.104	0.663	49	1	1.085		0.665	0.973
5	1	0.059		0.531	0.554	50	1	−1.434		0.743	1.461
6	1	0.059		0.531	0.554	51	1	−1.162		0.501	1.337
7	1	−0.774		0.492	1.374	52	1	−0.880		0.451	1.372
8	1	−0.332		0.479	0.716	53	1	−0.437		0.512	0.719
9	1	−0.332		0.479	0.716	54	1	−0.602		0.626	1.421
10	1	−0.513		0.711	1.458	55	1	−0.160		0.488	0.779
11	1	−1.068		0.440	1.308	56	1	0.276		0.912	1.257
12	1	−1.429		0.989	1.468	57	1	0.816		0.765	0.922
13	1	−3.928		0.408	2.112	58	1	−0.193		0.658	1.023
14	1	−3.368		0.418	2.098	59	1	−0.103		0.553	1.200
15	1	−1.998		0.920	1.602	60	1	0.340		0.437	0.663
16	1	−3.096		0.731	2.152	61	1	−0.333		0.970	1.230
17	1	−1.297		1.240	1.851	62	1	0.227		0.776	1.197
18	1	−3.655		0.176	2.052	63	1	0.669		0.644	0.516
19	1	−1.726		0.935	1.610	64	1	1.605		1.164	1.422
20	1	−0.200		0.474	0.622	65	1	1.330		0.914	1.215
21	1	−1.519		0.791	1.514	66	1	0.874		0.746	1.096
22	1	−1.247		0.541	1.384	67	1	2.065		0.858	1.124
23	1	−0.965		0.466	1.407	68	1	2.023		0.714	1.137
24	1	−0.522		0.542	0.764	69	1	2.284		0.782	1.161
25	1	−0.693		0.612	1.446	70	1	3.401		1.373	1.751
26	1	−0.410		1.029	1.707	88	2		−0.153	1.040	0.775
27	1	0.185		0.886	1.252	89	2		−1.235	0.734	0.695
28	1	−3.373		0.421	2.103	90	2		−1.235	0.734	0.695
29	1	−1.259		0.654	1.357	91	2		−2.861	1.108	0.858
30	1	0.270		0.616	0.610	92	2		−2.672	1.654	0.501
31	1	−0.699		0.457	1.308	93	2		−2.669	1.738	0.622
32	1	−0.256		0.476	0.640	94	2		0.377	0.713	0.633
33	1	0.492		0.944	1.141	95	2		1.143	0.194	0.510
34	1	1.035		0.713	1.033	96	2		0.387	0.655	0.641
35	1	1.478		0.378	0.533	97	2		2.032	0.907	1.069
36	1	1.738		0.274	0.441	98	2		−0.429	1.088	0.775
37	1	1.206		0.742	0.992	99	2		1.145	1.005	0.605
38	1	0.765		0.755	0.989	100	2		1.256	1.055	0.641
39	1	1.087		0.674	0.782	101	2		−0.981	1.546	0.851
40	1	1.439		0.535	0.733	102	2		1.093	0.245	0.602
41	1	1.132		0.424	0.662	103	2		1.512	0.707	0.997
42	1	0.906		0.703	0.893	104	2		−1.447	1.979	1.470
43	1	1.325		0.847	1.128	105	2		2.386	0.981	0.691
44	1	1.814		1.001	0.578	106	2		−0.338	1.110	0.778
45	1	1.891		0.257	0.374	107	2		1.237	1.016	0.614
46	1	2.001		1.023	0.594	108	2		1.454	1.128	0.740

gested by the parameter values based on the class models. And according to the suggestion, the peptide with the following structure is expected to have a sweet taste: L-Asp-NH-CH_2OH.

Acknowledgment. We thank the Computer Center, Institute for Molecular Science, for affording facilities for computation and thank Hoogenstraaten for sending the STERIMOL program.

Registry No. 1, 25352-43-6; 2, 25352-57-2; 3, 30239-25-9; 4, 58889-59-1; 5, 25352-56-1; 6, 25352-46-9; 7, 25353-70-2; 8, 25353-69-9; 9, 25353-72-4; 10, 101145-65-7; 11, 101145-66-8; 12, 25352-48-1; 13, 22839-89-0; 14, 59917-64-5; 15, 101145-67-9; 16, 74216-15-2; 17, 59917-63-4; 18, 51871-24-0; 19, 51871-18-2; 20, 39614-06-7; 21, 39613-86-0; 22, 39613-88-2; 23, 39613-91-7; 24, 39613-94-0; 25, 50833-41-5; 26, 50833-42-6; 27, 101145-68-0; 28, 51871-23-9; 29, 52945-91-2; 30, 101145-69-1; 31, 52946-15-3; 32, 39613-97-3; 33, 52945-89-8; 34, 101145-70-4; 35, 39614-00-1; 36, 101145-71-5; 37, 59917-61-2; 38, 49558-29-4; 39, 101145-72-6; 40, 77096-46-9; 41, 52946-00-6; 42, 52993-06-3; 43, 101223-13-6; 44, 22839-47-0; 45, 26270-66-6; 46, 101145-73-7; 47, 22839-51-6; 48, 101145-74-8; 49, 37622-39-2; 50, 52946-01-7; 51, 52946-02-8; 52, 52946-03-9; 53, 52946-04-0; 54, 52946-05-1; 55, 52946-06-2; 56, 52946-07-3; 57, 22839-92-5; 58, 52946-08-4; 59, 52946-10-8; 60, 52946-11-9; 61, 53022-01-8; 62, 53022-03-0; 63, 53022-04-1; 64, 52945-94-5; 65, 52945-95-6; 66, 37465-92-2; 67, 58406-52-3; 68, 58406-53-4; 69, 43188-74-5; 70, 43188-75-6; 71, 25352-44-7; 72, 25352-45-8; 73, 101145-75-9; 74, 101145-76-0; 75, 25353-71-3; 76, 52945-90-1; 77, 5241-71-4; 78, 72683-65-9; 79, 72683-70-6; 80, 72683-74-0; 81, 72683-76-2; 82, 101145-77-1; 83, 22839-93-6; 84, 22839-88-9; 85, 72683-62-6; 86, 22839-48-1; 87, 22840-07-9; 88, 101145-78-2; 89, 30239-44-2; 90, 30239-43-1; 91, 22838-86-4; 92, 22839-52-7; 93, 59917-62-3; 94, 22839-91-4; 95, 50833-44-8; 96, 22839-90-3; 97, 101145-79-3; 98, 13433-09-5; 99, 15368-70-4; 100, 22849-03-2; 101, 22839-65-2; 102, 101145-80-6; 103, 52945-93-4; 104, 22839-49-2; 105, 22839-50-5; 106, 22840-03-5; 107, 22840-08-0; 108, 22840-09-1.

Structure-Activity Relationships of Dopaminergic 5-Hydroxy-2-aminotetralin Derivatives with Functionalized *N*-Alkyl Substituents

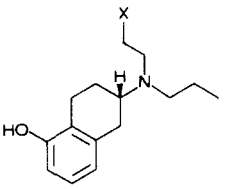
Max P. Seiler,* André P. Stoll, Annemarie Closse, Willy Frick, Annelise Jaton, and Jean-Marie Vigouret

Sandoz Ltd., Preclinical Research, CH-4002 Basel, Switzerland. *Received September 4, 1985*

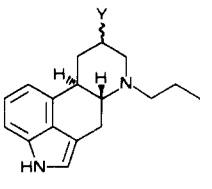
5-Hydroxy-2-aminotetralin derivatives in which one *N*-alkyl substituent carries a functional group have been prepared and their dopaminergic activities compared with those of 5-hydroxy-2-(di-*n*-propylamino)tetralin (5-OH-DPAT) and known ergolines. Several members of the series demonstrated high affinities in dopamine (DA) receptor binding and DA agonist properties in the rotational behavior model in the range of known potent ergolines. The results suggest that the accessory binding site for the larger *N*-alkyl substituent of the 5-hydroxy-2-aminotetralins can accommodate various neutral and bulky functionalities and is probably identical with the site(s) to which the 8-substituents of the ergolines bind.

Dopaminergic activity has been identified in a variety of structural types. The tetracyclic ergoline derivatives with three asymmetric centers figure certainly among the most complex ones, whereas the phenolic 2-aminotetralins seem to contain the minimal structural requirements for longer acting, metabolically stabilized dopamine (DA) analogues.[1] We have chosen the 2-aminotetralin skeleton as the starting point for the design of new clinically useful dopaminergic drugs to investigate whether these relatively simple structures could compete in in vivo tests with the prominent dopaminergic activities of the ergolines.

By comparing phenolic 2-aminotetralins with either a primary or a tertiary amino group in vitro, we have observed that on DA_1 and DA_2 receptor subtypes *N,N*-di-*n*-propylation has no influence upon the activity of 7-hydroxyaminotetralin, the most active primary amine, but leads to an increase in activity of the corresponding 5-hydroxy derivative, rendering 5-hydroxy-2-(di-*n*-propylamino)tetralin (5-OH-DPAT) the most potent member of the series. This increase in activity (and affinity in DA receptor binding) upon N,N-dialkylation has been interpreted in the sense that the *N*-propyl substituents of the 5-hydroxylated aminotetralins can reach accessory binding sites.[2,3] Earlier investigations have revealed that the DA receptor can accommodate N,N-disubstituted DA analogues with one *N*-alkyl substituent not larger than *n*-propyl, whereas the structural requirements for the second N-substituent are less stringent.[4-10] Similarly, *n*-butyl substitution at the basic nitrogen of the ergolines results in compounds with strongly reduced dopaminergic activities,[11] suggesting that this N-substituent corresponds to the DA N-substituent with reduced steric freedom. It is tempting to speculate that the larger N-substituent of the N,N-dialkylated 5-hydroxyaminotetralins I (X = alkyl) could reach at the DA receptor the site(s) of the important 8-substituent Y of the ergolines II. We have therefore



I



II

(1) Cannon, J. G. *Ann. Rev. Pharmacol. Toxicol.* **1983**, *23*, 103.
(2) Seiler, M. P.; Markstein, R. *Mol. Pharmacol.* **1982**, *22*, 281.
(3) Seiler, M. P.; Markstein, R. *Mol. Pharmacol.* **1984**, *26*, 452.
(4) McDermed, J. D.; McKenzie, G. M.; Phillips, A. P. *J. Med. Chem.* **1975**, *18*, 362.
(5) Cannon, J. G.; Lee, T.; Goldman, H. D. *J. Med. Chem.* **1977**, *20*, 1111.
(6) Ginos, J. Z.; Cotzias, G. C.; Doroski, D. *J. Med. Chem.* **1978**, *21*, 160.
(7) Cannon, J. G.; Hsu, F. L.; Long, J. P.; Flynn, J. R.; Costall, B.; Naylor, R. J. *J. Med. Chem.* **1978**, *21*, 248.
(8) Hacksell, U.; Svensson, U.; Nilsson, J. L. G.; Hjorth, S.; Carlsson, A.; Wikstroem, H.; Lindberg, P.; Sanchez, D. *J. Med. Chem.* **1979**, *22*, 1469.
(9) Wikstroem, H.; Sanchez, D.; Lindberg, P.; Arvidsson, L. E.; Hacksell, U.; Johansson, A.; Nilsson, J. L. G.; Hjorth, S.; Carlsson, A. *J. Med. Chem.* **1982**, *25*, 925.
(10) Wikstroem, H.; Andersson, B.; Sanchez, D.; Lindberg, P.; Arvidsson, L. E.; Johansson, A. M.; Nilsson, J. L. G.; Svensson, K.; Hjorth, S.; Carlsson, A. *J. Med. Chem.* **1985**, *28*, 215.
(11) Closse, A.; Frick, W.; Jaton, A.; Vigouret, J. M., personal communication.

0022-2623/86/1829-0912$01.50/0