Servce (GM 29815). S.J.H. thanks the National Science Foundation for a Graduate Fellowship. We thank Dr. Kenneth Feingold for the gift of rat liver microsomes.

Registry No. 1, 73573-88-3; 5, 96575-19-8; 8, 91424-39-4; 9, 91424-40-7; 10, 91424-41-8; 11, 109720-91-4; 12, 109720-92-5; 13, 96555-56-5; 14, 109720-93-6; 15, 109720-94-7; 16, 109720-95-8; 17, 96555-51-0; 18, 96555-52-1; 19, 109720-96-9; 20, 109720-97-0; 21, 109720-98-1; 22, 109720-99-2; 23, 96575-20-1; 24, 96555-53-2; 25, 96555-54-3; 26, 96555-55-4; 27, 96555-58-7; 28, 109721-00-8; 29, 109721-01-9; (1*S*,2*S*,8*S*,8a*R*,5′*R*,2″*S*)-35 (R = *t*-BuMe$_2$Si, R^1 = Me), 109785-23-1; (1*R*,2*R*,8*R*,8a*S*,5′*R*,2″*S*)-35 (R = *t*-BuMe$_2$Si, R^1 = Me), 109785-24-2; (1*S*,2*S*,8*S*,8a*R*,5′*R*,2″*S*)-35 (R = H, R^1 = Me), 109721-02-0; 36, 109721-03-1; 37, 109744-47-0; (1*S*,2*S*,8*S*,8a*R*,5′*R*,2″*S*)-38, 96555-61-2; (1*R*,2*R*,8*R*,8a*R*,5′*R*,2″*S*)-38, 109721-04-2; (1*S*,2*S*,8*S*,8a*R*,5′*R*,2″*S*)-39, 79814-60-1; (1*R*,2*R*,8*R*,8a*S*,5′*R*,2″*S*)-39, 109785-25-3; (1*S*,2*S*,8*S*,8a*R*,5′*R*,2′′*S*)-40, 79896-20-1; (1*R*,2*R*,8*R*,8a*S*,5′*R*,2″*S*)-40, 109785-26-4; 41, 84173-31-9; 42, 109744-48-1; 43, 109721-05-3; 44, 109785-27-5; 45, 109721-06-4; 46, 79896-21-2; 47, 79896-19-8; 48, 84173-29-5; 49, 84173-30-8; 50, 109721-07-5; 51, 109785-28-6; 52, 109785-29-7; 53, 85540-02-9; 54, 85540-13-2; 55, 85540-03-0; 56, 85540-14-3; 57, 109744-49-2; 58, 109721-08-6; 59, 91424-35-0; 59 acid chloride, 109721-09-7; 60, 91424-36-1; 60 acid chloride, 109721-10-0; 61, 109721-11-1; 61 *t*-BuMe$_2$Si ether, 109721-12-2; 62, 109721-13-3; 62 *t*-BuMe$_2$Si ether, 109721-14-4; 63, 96555-62-3; 64, 96555-63-4; 65, 109721-15-5; 66, 109721-16-6; HMGR, 9028-35-7; *t*-BuMe$_2$SiOCH(CH$_2$CO$_2$Na)$_2$, 109721-17-7; (*R*)-(MeO)$_2$P(O)-CH$_2$C(O)CH$_2$CH(OH)CH$_2$CO$_2$H, 109721-18-8; diethyl 3-hydroxyglutarate, 32328-03-3; (*S*)-1-phenylethanol, 1445-91-6; (*R*)-1-phenylethanol, 1517-69-7; glutaric anhydride, 108-55-4.

Quantitative Structure–Activity Relationships of the Bitter Thresholds of Amino Acids, Peptides, and Their Derivatives

Masaaki Asao, Hajime Iwamura,* Miki Akamatsu, and Toshio Fujita

Department of Agricultural Chemistry, Faculty of Agriculture, Kyoto University, Kyoto 606, Japan.
Received December 9, 1986

Bitter thresholds of a total of 93 amino acids, peptides, and their derivatives were analyzed quantitatively by use of hydrophobicity parameters reported for amino acid side chains and those for the whole molecules estimated from partition coefficients obtained experimentally. We also explored the steric parameters that best explained the variation in the intensity of bitterness attributable to the molecular shape. The results showed that the total length along the zigzag peptide backbone chain of the molecule is an important factor. The bitterness of nonzwitterionic *N*-acyl and ester derivatives and that of neutral *N*-acyl ester derivatives were expressed by a single, common equation together with those of zwitterionic amino acids and peptides. Thus the interaction via the charge with the receptor site was probably not an indispensable factor for triggering of the bitter sensation. This study, together with earlier ones, may serve as a prototype of approaches toward unraveling structure–activity relationships of complex molecules like amino acids, peptides, and their derivatives that are of medicinal or agricultural importance.

Amino acids and peptides have long been studied by chemists because of their importance as flavoring constituents of foods as well as their significance in biological processes. The tastes of amino acids are various. Among them, bitter and sweet tastes have been extensively examined by a number of researchers. The results have been puzzling. The D enantiomers of some bitter L-amino acids such as leucine, phenylalanine, tryptophan, and tyrosine are sweet, but both enantiomers of some other amino acids, including alanine, serine, threonine, and ornithine, are sweet.[1] Many dipeptides and tripeptides are bitter. There is no simple correspondence for the tastes of component amino acids; for example, peptides D-Leu-Gly and D-Leu-D-Leu, which contain sweet amino acids, are bitter.[2] These complex features have made it difficult to obtain an overall view of their structure–activity relationships.

The state of structure–activity relationship studies of bitter compounds has been summarized by Belitz et al.[3] On the basis of data already reported, what we can say about the structural characteristics of bitter compounds is that there is always a polar function and a hydrophobic group within the molecule, the former probably affecting taste quality and the latter affecting taste intensity. Since the hydrophobic moieties are sterically various, the participation of steric factors has been suggested also. To obtain more information, a quantitative approach may be of use. For derivatives of amino acids and peptides, Gardner has investigated the relationship between the bitter thresholds and molecular connectivity, finding a significant correlation with the first-order-valence corrected index $^1\chi^v$.[4] This correlates with the partition coefficients of a wide range of compounds, so he suggested that the result is a reflection of the influence of hydrophobicity on bitterness. Despite the likelihood that the bitter intensity is also related to steric factors, the dimensional features on a whole-molecular basis of amino acids and peptides have not been parameterized and incorporated into quantitative regression analysis. In this study, we analyzed quantitatively the structure–bitterness relationships of these classes of compounds, by using the hydrophobic parameters derived from partition coefficients found experimentally[5,6] and exploring steric parameters that can explain the variation of the intensity of bitterness. This approach could be extended to derivatives of amino acids and peptides that have medicinal and agricultural importance, as well as to other classes of bitter compounds.

Bitter Thresholds. The threshold data of compounds **1–10**, **72–74**, **76**, **78**, **79**, **81–88**, and **90** were taken from literature reported by Wieser and Belitz in 1975,[1] and those of compounds **11–21**, **23–31**, **33–48**, **50**, **54–71**, and **91–99** were from literature reported by the same workers in 1976.[2] The values of 24 compounds (**1–5**, **7**, **12**, **14**, **16**, **18**, **19**, **27–29**, **34**, **50**, **55–58**, **60**, **61**, **73**, **76**, and **81**) in Table I were

(1) Wieser, H.; Belitz, H.-D. *Lebensm. Unters.-Forsch.* **1975**, *159*, 65.
(2) Wieser, H.; Belitz, H.-D. *Lebensm. Unters.-Forsch.* **1976**, *160*, 383.
(3) Belitz, H.-D.; Chen, W.; Jugel, H.; Stempfl, H.; Treleano, R.; Wieser, H. *Chem. Ind. (London)* **1983**, 23.
(4) Gardner, R. J. *J. Sci. Food Agric.* **1980**, *31*, 23.
(5) Fauchere, J.-L.; Pliska, V. *Eur. J. Med. Chem.—Chim. Ther.* **1983**, *18*, 369.
(6) Akamatsu, M.; Asao, M.; Iwamura, H.; Fujita, T., Kyoto University, Faculty of Agriculture, Kyoto, unpublished data.

0022-2623/87/1830-1873$01.50/0

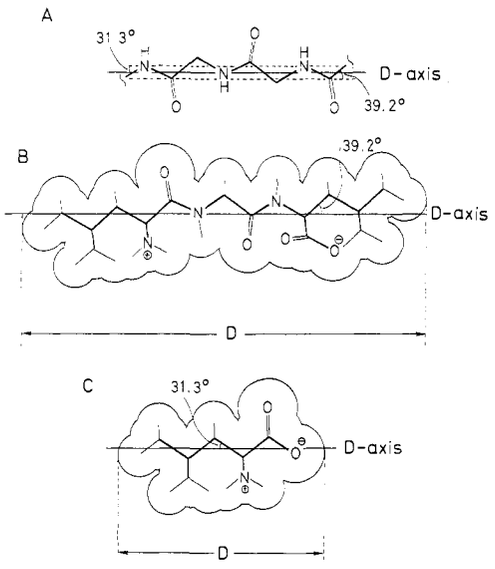

Figure 1. Definition of the length parameter D: A, schematic representation of the D axis that runs at the middle of the two dotted lines, one passing through the nitrogen atoms and the other passing through the carbonyl carbon atoms of the two peptide bonds that are one amino acid unit apart from each other; B and C, the D parameters of Leu-Gly-Leu and Leu, respectively.

also calculated by our own sensory panel. The collinearity between the previous and present values was examined by regression analysis to give eq 1. In this and following

$$\log (1/T) \text{ (Wieser et al.)} = \underset{(0.18)}{1.06} \log (1/T) \text{ (this study)} - \underset{(0.43)}{0.205} \quad (1)$$

$$n = 24,\ s = 0.20,\ r = 0.93$$

equations, n is the number of compounds, s is the standard deviation, r is the correlation coefficient, and the figures in parentheses are the 95% confidence intervals. The T(M) in the dependent variable is the center value of the reported range of the threshold molar concentrations, and that in the independent variable is the mean value of our determinations. The correspondence of the two terms was good, so our own data for compounds **22**, **32**, **49**, **51**, **52**, **53**, **75**, **77**, **80**, and **89**, the bitterness of which was not found by previous workers, were rectified by eq 1 and included in the analysis. The data are summarized in Table I.

Wieser and Belitz have also reported the taste properties of sets of natural and unnatural amino acids (**100–109** and **126–129**)[1] and peptide derivatives (**110–113** and **116–125**)[2] listed in Table IV. The taste of the rest (compounds **114** and **115**) was examined by us.

Steric Dimensions. To best express the steric features of the molecule, we first defined the length parameter D,[7] which corresponds to the maximum length of a molecule in the fully extended conformation in which the zigzag peptide backbone extends straight. It is measured as the length along the D axis that runs in the middle of the two straight lines; one passes through the two amide nitrogen atoms five atoms apart from each other in peptides with three or more amino acids, and the other passes through the corresponding amide carbon atoms (Figure 1A). The main chain was constructed so as to give the longest D. It follows that the side chains rather than the amino or hydroxycarbonyl groups of the terminal amino acids are incorporated into the main chain. When the C-terminal is glycine or alanine, however, the carboxylic acid moiety constitutes the main chain, since it is longer than the side groups H and Me. Similarly, when glycine is at the N-end, the main chain comes to include the amino group. Some of the N-acyl and ester derivatives give the longest D when acyl and ester groups are arranged along the D axis. By this definition, the angle between the D axis and the bond axis that links the terminal group at the C-end to the connecting atom becomes 39.2°, and that at the N-terminal becomes 31.3°. For compounds smaller or shorter than tripeptides, i.e., dipeptide and amino acid derivatives, the D axis was drawn according to this criterion. These situations are explained schematically in Figure 1. The values were calculated on the basis of the CPK model by use of a computer program made by Verloop et al.[8] for the estimation of the STERIMOL parameters and modified by us for this and similar purposes.[9] These calculations and the regression analyses to follow were done on a FACOM M-382 computer of the Data Processing Center of this university.

Results

In this study, amino acids, dipeptides, tripeptides, and their derivatives were analyzed. We first adopted, as the hydrophobic parameter, the π values for the side chains of amino acids calculated from the partition coefficients between 1-octanol and water at pH 7.0–7.2 of N-acetyl amino acid amides.[5]

First, we analyzed amino acids, dipeptides, and tripeptides separately and obtained the following equations. For amino acids:

$$\log (1/T) = \underset{(0.06)}{0.38\pi} + \underset{(0.40)}{0.51(D/10)} + \underset{(0.44)}{0.84} \quad (2)$$

$$n = 9,\ s = 0.08,\ r = 0.99$$

For dipeptides:

$$\log (1/T) = \underset{(0.10)}{0.46\textstyle\sum\pi} + \underset{(0.43)}{0.80(D/10)} - \underset{(0.53)}{0.19} \quad (3)$$

$$n = 52,\ s = 0.28,\ r = 0.89$$

For tripeptides:

$$\log (1/T) = \underset{(0.23)}{0.50\textstyle\sum\pi} + \underset{(0.73)}{0.56} \quad (4)$$

$$n = 9,\ s = 0.34,\ r = 0.89$$

The $\sum\pi$ in eq 3 and 4 is the sum of the π values for side chains of the component amino acids. The D value was scaled by 0.1 in the regressions to make the size comparable to that of the π or $\sum\pi$. The results indicate that the hydrophobicity and total length are among the governing factors. The insignificance of the D in eq 4 appears to arise from its poorer variation in the limited number of tripeptide compounds. Since the coefficient values of the π and $\sum\pi$ terms overlap in the series of compounds within the 95% confidence intervals and the coefficient values of the D overlap in eq 2 and 3, we combined the three sets of compounds to obtain eq 5. For an amino acid, $\sum\pi$

$$\log (1/T) = \underset{(0.07)}{0.44\textstyle\sum\pi} + \underset{(0.34)}{0.85(D/10)} - \underset{(0.23)}{0.65I_d} - \underset{(0.33)}{1.19I_t} + \underset{(0.37)}{0.44} \quad (5)$$

$$n = 70,\ s = 0.26,\ r = 0.91$$

(7) Nakayama, A.; Iwamura, H.; Fujita, T. *J. Med. Chem.* **1984**, *27*, 1493.

(8) Verloop, A.; Hoogenstraaten, W.; Tipker, J. *Drug Design*; Academic: New York, 1976; Vol. III, Chapter 4.

(9) Asao, M.; Iwamura, H., Kyoto University, Faculty of Agriculture, Kyoto, unpublished data, 1985.

stands for the π value of its side chain. I_d is an indicator variable that takes the value of unity for dipeptides and otherwise is 0, and I_t is the same kind of variable for tripeptides. The results were essentially the same as those for each series, except for the significance of the indicator variables.

N-Acyl and ester derivatives of some of the amino acids and peptides taste bitter, often very strongly. Thus, these classes of compounds were included in our sensory tests as well as in those of Wieser and Belitz.[1,2] Preliminary examinations of the structure–bitterness profile of these derivatives suggested that the same or similar factors that govern the bitterness of the parent amino acids and peptides also work for these types of compounds as well. We explored correlations for the whole set of compounds to obtain eq 6. As the hydrophobicity parameters of *N*-acetyl

$$\log (1/T) = \underset{(0.06)}{0.44}\sum\pi + \underset{(0.31)}{0.86}(D/10) - \underset{(0.15)}{0.50}I_d - \underset{(0.26)}{1.01}I_t + \underset{(0.15)}{0.41}I_{OR} + \underset{(0.19)}{0.40}I_{Pro} + \underset{(0.34)}{0.22} \quad (6)$$

$$n = 97,\ s = 0.25,\ r = 0.92$$

derivatives (**80–82**), methyl esters (**72–75**, **77**, **79**, **91–93**), and *N*-acetyl methyl esters (**86**, **88**, **97–99**), those of the corresponding amino acids or peptides were used as they stand. Possible hydrophobicity and other physicochemical differences were taken into consideration by use of indicator variables. The values for *N*-benzoyl derivatives (**83–85**) are those obtained by addition of the hydrophobicity difference between benzamide and acetamide, 1.91, to the values for the corresponding *N*-acetyl derivatives. The supplemental value was calculated by $f(NHCOC_6H_5) - f(NHCOCH_3)$, where f is the fragment constant for aliphatic compounds listed in the literature.[10] Similarly, for ethyl esters, the value of 0.54, obtained by $f(COOC_2H_5) - f(COOCH_3)$, was added to the π values of the corresponding methyl esters. The correlation was essentially the same as in eq 5 with respect to the π, D, I_d, and I_t terms. I_{OR} is the indicator variable for the compounds with an ester function and takes 1 for these and 0 for the others. Unexpectedly, the indicator variable for *N*-acyl compounds was not significant. The possible difference in the hydrophobicity and other physicochemical differences may cancel out each other between the *N*-acyl and corresponding N-unsubstituted compounds.

The log $(1/T)$ values of proline and its derivatives always deviated from the calculated values throughout the correlations. Thus the indicator variable I_{Pro}, which takes 1 for these compounds and otherwise 0, was introduced in eq 6 and probably represented steric factors inherent to the cyclic structure. The value of *N*-acetyl-Phe-Leu methyl ester (**99**) also deviated from the predicted one. This was again probably because its molecular shape was segregative from all others; it has bulky branches at the middle of the molecule when constructed according to the definition, and this, as shown by Figure 2, makes the molecule very thick and wide. Steric factors other than D are thus quite likely to operate for this compound as well, probably making its bitterness pronounced. Since the compounds that are congeneric to **99** in terms of steric shape are not now available, we excluded the compound from the analysis. The use of an indicator variable term or steric thickness and/or width parameters makes no sense for the single compound. It would, however, be an interesting, future problem to explore the structure–bitterness correlations

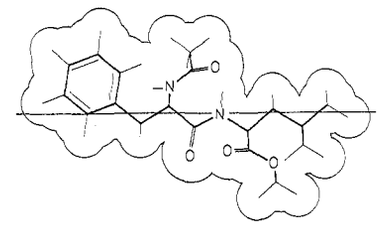

Figure 2. Shape of *N*-acetyl-Phe-Leu methyl ester (**99**).

with a number of congeners. Tyrosine (**6**) was not included since it is sparingly soluble in water at room temperature and thus a reliable threshold value could not be obtained.

Examination of eq 5 and 6 showed that the coefficient of the indicator variable I_t for tripeptides is nearly twice as large as that of I_d for dipeptides. It is likely, therefore, that the terms mostly represent the hydrophobicity difference arising from the difference between the backbone structures, in other words, the difference in the number of peptide bonds. This finding prompted us to measure and compare the partition coefficients between 1-octanol and water at pH 7.0 of amino acids, dipeptides, tripeptides, and *N*-acetyl amino acids and dipeptides. The details of the study and the data will be reported separately as a different item,[6] but only the final result is shown here as eq 7, by which we estimated the hydrophobicity of the whole molecule. The I_p is the indicator variable that takes

$$\log P = \underset{(0.062)}{0.985}\sum\pi - \underset{(0.110)}{0.720}I_p + \underset{(0.141)}{0.532}E_s'^c(N) + \underset{(0.105)}{0.292}\sum E_s'^c - \underset{(0.173)}{0.675}I_{Ac} - \underset{(0.110)}{3.251} \quad (7)$$

$$n = 42,\ s = 0.149,\ r = 0.987$$

1 for dipeptides, 2 for tripeptides, and 0 for others, indicating that a peptide bond contributes to hydrophobicity to about the same extent in dipeptides and tripeptides. The $E_s'^c$ is a steric parameter defined by $E_s' - 0.306(3 - n)$, where E_s' is that reported by MacPhee et al.[11] and n is the number of α-hydrogen atoms of aliphatic substituents. The $E_s'^c$ parameter represents not only the effect of steric bulk but also that of branching of the alkyl chain.[12] The $E_s'^c(N)$ expresses the $E_s'^c$ of the side chain of an N-terminal amino acid, and $\sum E_s'^c$ is the sum of the values of the remainders. I_{Ac} is the indicator variable that takes 1 for *N*-acetyl derivatives and 0 for other derivatives, its coefficient indicating that N-acetylation lowers the hydrophobicity.

Reinvestigation of eq 6 with the use of log P gave eq 8. Proline derivatives **15**, **40**, **49**, and **52** were excluded from the analysis because of the lack of the $E_s'^c$ data mentioned above, but compounds **4**, **50**, **51**, **53**, and **56** were included since their log P values have been determined experimentally.[6] In this final equation, the indicator variables

$$\log (1/T) = \underset{(0.06)}{0.47}\log P + \underset{(0.21)}{1.02}(D/10) + \underset{(0.14)}{0.37}I_{Ac} + \underset{(0.14)}{0.34}I_{OR} + \underset{(0.25)}{0.52}I_{Pro} + \underset{(0.36)}{1.58} \quad (8)$$

$$n = 93,\ s = 0.24,\ r = 0.92$$

I_d and I_t were not at all significant. Instead, the indicator term I_{Ac} for *N*-acyl derivatives became significant. Its positive coefficient suggests that the N-acylation makes

(10) Hansch, C.; Leo, A. *Substituent Constants for Correlation Analysis in Chemistry and Biology*; Wiley: New York, 1979.

(11) MacPhee, J. A.; Panaye, A.; Dubois, J.-E. *Tetrahedron* **1978**, *34*, 3553.

(12) Takayama, C.; Akamatsu, M.; Fujita, T. *Quant. Struct.–Act. Relat.* **1985**, *4*, 149.

Table I. Bitter Thresholds and Physicochemical Properties of Amino Acids, Peptides, and Their Derivatives

no.	compd	log (1/T) obsd	log (1/T) calcd[a]	Δ	$\Sigma\pi^b$	log P^c	D
1	Val	1.68	1.44	0.23	1.22	−2.08[d]	8.14
2	Leu	1.92	1.79	0.13	1.70	−1.61[d]	9.37
3	Ile	1.96	1.74	0.22	1.80	−1.72[d]	9.37
4	Pro	1.59	1.21	0.38	0.72	−2.50[d]	7.76
5	Phe	2.19	1.98	0.21	1.79	−1.46[d]	10.52
6	Tyr[e]	2.30	(1.66)	0.65	0.96	−2.31	11.23
7	Trp	2.30	2.35	−0.05	2.25	−1.06[d]	12.32
8	Lys	1.07	0.81	0.26	−0.99	−4.23	11.76
9	Arg	1.13	0.93	0.19	−1.01	−4.25	13.02
10	His	1.32	1.18	0.14	0.13	−3.12	10.31
11	Gly-Val	1.13	1.32	−0.20	1.22	−2.98	11.04
12	Gly-Leu	1.68	1.65	0.03	1.70	−2.55	12.32
13	Gly-D-Leu	1.67	1.65	0.02	1.70	−2.55	12.32
14	Gly-Ile	1.70	1.65	0.05	1.80	−2.56	12.32
15	Gly-Pro[f]	1.35	1.26[g]	0.09	0.72		9.48
16	Gly-Phe	1.80	1.88	−0.09	1.79	−2.31	13.47
17	Gly-D-Phe	1.80	1.88	−0.09	1.79	−2.31	13.47
18	Gly-Trp	1.89	2.28	−0.39	2.25	−1.84	15.18
19	Gly-Tyr	1.77	1.60	0.17	0.96	−3.12	14.47
20	Ala-Val	1.16	1.43	−0.27	1.53	−2.78	11.15
21	Ala-Leu	1.70	1.76	−0.05	2.01	−2.35	12.43
22	Ala-Phe	1.72	1.99	−0.27	2.10	−2.11	13.58
23	Val-Gly	1.19	1.25	−0.06	1.22	−3.29	11.76
24	Val-Ala	1.16	1.36	−0.21	1.53	−3.04	11.76
25	Val-Val	1.71	1.54	0.17	2.44	−2.82[d]	12.43
26	Val-Leu	2.00	2.02	−0.02	2.92	−2.07[d]	13.71
27	Leu-Gly	1.72	1.56	0.16	1.70	−2.90	12.99
28	Leu-Ala	1.72	1.67	0.05	2.01	−2.65	12.99
29	Leu-Leu	2.35	2.43	−0.08	3.40	−1.46[d]	14.94
30	Leu-D-Leu	2.26	2.43	−0.17	3.40	−1.46	14.94
31	D-Leu-D-Leu	2.26	2.43	−0.17	3.40	−1.46	14.94
32	Leu-Phe	2.75	2.69	0.06	3.49	−1.15[d]	16.09
33	Leu-Trp	3.40	2.97	0.43	3.95	−0.93	17.80
34	Leu-Tyr	2.46	2.30	0.16	2.66	−2.22	17.09
35	Ile-Gly	1.68	1.51	0.17	1.80	−3.00	12.99
36	Ile-Ala	1.68	1.63	0.05	2.11	−2.75	12.99
37	Ile-Val	2.05	1.97	0.08	3.02	−2.17	13.66
38	Ile-Leu	2.26	2.30	−0.04	3.50	−1.74	14.94
39	Ile-Ile	2.26	2.26	0.00	3.60	−1.82	14.94
40	Ile-Pro[f]	2.40	2.29[g]	0.11	2.52		12.11
41	Ile-Trp	3.05	2.92	0.13	4.05	−1.03	17.80
42	Ile-Asn	1.49	1.29	0.20	1.20	−3.87	14.83
43	Ile-Asp	1.37	1.15	0.22	1.03	−4.04	14.23
44	Ile-Gln	1.49	1.61	−0.13	1.58	−3.45	16.05
45	Ile-Glu	1.37	1.36	0.01	1.16	−3.87	15.50
46	Ile-Lys	1.65	1.39	0.26	0.81	−4.21	17.34
47	Ile-Ser	1.49	1.48	0.01	1.76	−3.18	13.49
48	Ile-Thr	1.49	1.60	−0.11	2.06	−2.95	13.66
49	Pro-Ala[f]	1.32	1.56[g]	−0.24	1.03		11.38
50	Pro-Leu	2.22	2.34	−0.12	2.42	−2.41	13.34
51	Pro-Ile	2.33	2.27	0.06	2.52	−2.56	13.34
52	Pro-Tyr[f]	1.80	2.21[g]	−0.40	1.68		15.49
53	Pro-Phe	2.80	2.62	0.19	2.51	−2.07	14.48
54	Phe-Gly	1.77	1.85	−0.08	1.79	−2.52	14.14
55	Phe-Leu	2.87	2.68	0.19	3.49	−1.17[d]	16.09
56	Phe-Pro	2.70	2.82	−0.14	2.51	−1.36	13.26
57	Phe-Phe	3.10	2.95	0.15	3.58	−0.85[d]	17.24
58	Phe-Tyr	3.13	2.59	0.54	2.75	−1.84	18.24
59	Trp-Glu	1.56	2.11	−0.55	1.61	−2.92	18.45
60	Trp-Trp	3.60	3.58	0.03	4.50	−0.27[d]	20.75
61	Tyr-Leu	2.40	2.33	0.07	2.66	−2.09	16.81
62	Ser-Leu	1.49	1.64	−0.15	1.66	−2.85	13.54
63	Gly-Gly-Leu	1.13	1.71	−0.59	1.70	−3.21	15.94
64	Gly-DL-Leu-Gly	1.26	1.51	−0.25	1.70	−3.21	13.99
65	Gly-Leu-Tyr	2.52	2.25	0.27	2.66	−2.53	18.09
66	Leu-Gly-Gly	1.13	1.62	−0.49	1.70	−3.56	16.61
67	Leu-Gly-Leu	2.26	2.40	−0.14	3.40	−2.31	18.56
68	Leu-Val-Leu	2.70	2.75	−0.05	4.62	−1.57[d]	18.56
69	Leu-Leu-Leu	2.87	3.04	−0.17	5.10	−0.94[d]	18.56
70	Leu-Gln-Leu	2.52	2.19	0.33	3.18	−2.76	18.56
71	Leu-Glu-Leu	2.00	2.00	0.00	2.76	−3.18	18.56
72	Leu-OEt	2.16	2.69	−0.54	2.24	−1.04	12.31
73	Phe-OMe	2.40	2.48	−0.08	1.79	−1.49	12.23
74	D-Phe-OMe	2.46	2.47	−0.02	1.79	−1.49	12.23
75	Tyr-OMe	1.75	2.17	−0.42	0.96	−2.31	12.94

Table I (Continued)

no.	compd	log (1/T) obsd	log (1/T) calcd[a]	Δ	Σπ[b]	log P[c]	D
76	Tyr-OEt	2.35	2.54	−0.19	1.50	−1.77	14.18
77	Trp-OMe	2.76	2.87	−0.11	2.25	−1.06	14.03
78	Trp-OEt	2.84	3.24	−0.42	2.79	−0.50	15.26
79	His-OMe	1.56	1.69	−0.13	0.13	−3.12	12.02
80	*N*-Ac-Leu	1.65	1.88	−0.23	1.70	−2.61	11.20
81	*N*-Ac-Phe	1.96	2.11	−0.15	1.79	−2.37	12.36
82	*N*-Ac-Trp	1.96	2.45	−0.49	2.25	−2.01[d]	14.06
83	*N*-Bz-Gly	2.30	2.21	0.09	1.91	−1.99	11.63
84	*N*-Bz-Ala	2.30	2.33	−0.03	2.22	−1.74	11.63
85	*N*-Bz-D-Ala	2.30	2.33	−0.03	2.22	−1.74	11.63
86	*N*-Ac-Leu-OMe	2.46	2.21	0.24	1.70	−2.61	11.20
87	*N*-Ac-Phe-OEt	2.76	2.81	−0.05	2.33	−1.83	13.46
88	*N*-Ac-D-Phe-OMe	2.70	2.45	0.25	1.79	−2.37	12.36
89	*N*-Ac-Trp-OEt	3.40	3.21	0.19	2.79	−1.37	15.26
90	*N*-Ac-Tyr-OEt	2.56	2.50	0.06	1.50	−2.65	14.18
91	Gly-Leu-OMe	2.16	1.99	0.17	1.70	−2.55	12.32
92	Leu-Gly-OMe	2.05	2.07	−0.02	1.70	−2.90	14.70
93	Phe-Leu-OMe	3.82	2.97	0.85	3.49	−1.27	16.09
94	*N*-Ac-Gly-Leu	1.70	1.94	−0.24	1.70	−3.28	14.83
95	*N*-Ac-Leu-Gly	1.68	1.75	−0.07	1.70	−3.28	12.99
96	*N*-Ac-Phe-Leu	3.00	2.77	0.23	3.49	−1.77	16.09
97	*N*-Ac-Gly-Leu-OMe	2.40	2.28	0.12	1.70	−3.28	14.83
98	*N*-Ac-Leu-Gly-OMe	2.35	2.26	0.09	1.70	−3.28	14.70
99	*N*-Ac-Phe-Leu-OMe[e]	4.52	(3.10)	1.42	3.49	−1.77	16.09

[a] Unless otherwise noted, the values were calculated by eq 8. [b] The values of *N*-acetyl and methyl ester derivatives are those of the corresponding unprotected compounds. The values of *N*-benzoyl derivatives are those obtained by addition of the hydrophobicity difference between benzamide and acetamide. The values of ethyl ester derivatives are those obtained by addition of the difference between aliphatic ethyl and methyl esters. [c] Unless otherwise noted, the values were estimated by eq 7. The values of methyl esters are those of the corresponding free acids, and the values of ethyl esters are those obtained as described in footnote *a*. [d] Found experimentally in ref 6. [e] Excluded from regression analysis, but the calculated value by eq 8 is shown in parentheses. [f] Excluded from the analysis of eq 8. [g] The value was calculated by eq 6.

Table II. Development of Eq 8

const	log P	D	I_{AC}	I_{OR}	I_{Pro}	r	s	$F_{X,Y}$[a]
3.34	0.54					0.77	0.39	$F_{1,91}$ = 133.33
2.04	0.50	0.86				0.84	0.33	$F_{2,90}$ = 112.13
1.84	0.50	0.95	0.44			0.89	0.29	$F_{3,89}$ = 108.30
1.68	0.48	0.99	0.35	0.31		0.91	0.26	$F_{4,88}$ = 100.77
1.58	0.47	1.02	0.37	0.34	0.52	0.92	0.24	$F_{5,87}$ = 98.62

[a] *F* statistic for significance of the addition of each variable.

Table III. Squared Correlation Matrix for Variables Used in Eq 8

	log P	D	I_{Ac}	I_{OR}
D	0.16			
I_{AC}	0.03	0.14		
I_{OR}	0.16	0.09	0.28	
I_{Pro}	0.06	0.04	0.10	0.10

amino acids and peptides about 2 times more bitter, due to changes in the steric and/or electronic properties of the molecules but not to changes in the hydrophobicity. As the log *P* values of methyl esters, those of the corresponding amino acids or peptides were used. For ethyl esters, the hydrophobicity difference between aliphatic ethyl and methyl esters (0.54) was added to the value of the corresponding methyl esters. Unfortunately, we could not measure the log *P* value of esters, because of their lability. Otherwise, we would have studied whether the doubled bitterness of the esters compared to the parents is due to increased hydrophobicity only or due to other factors as well. Table II shows the development of the final eq 8, and Table III shows the degree of independence of the variables we considered.

Discussion

The total length *D* of the molecule in the extended conformation was found to be an important factor that governs the bitterness of amino acids, peptides, and their derivatives. The conformation at the site of action, or the active conformation, may, however, not necessarily be the extended one for compounds with flexible skeletal structures like those studied here. When the coefficient of *D* is positive, the molecules with a longer total length may be able to take on a conformation closer to the one that can fit the receptor best (the optimal shape). The compound best in terms of *D* appears to be longer in the extended form than any of the compounds studied here. The case where the optimal *D* is apparent has been previously documented in the analysis of insect juvenile hormone mimics, compounds with a long zigzag aliphatic chain.[7]

The side-chain groups on the main chain project out in various directions depending on their position and configuration. Thus we considered also the width and thickness of the molecule perpendicular to the *D* axis (data not shown), but they were not significant at all. In this respect, little difference was found between the bitter intensities of enantiomers of amino acid methyl esters **73** and **74** and *N*-benzoyl amino acids **84** and **85** and enantiomers and diastereomers of dipeptides **12** and **13**, **16** and **17**, and **29**, **30**, and **31**. These were predicted well by eq 8, suggesting that the projection of chiral substituents into either the *R* or *S* direction has little effect on the taste property. However, while *N*-acetyl L-amino acids (**80** and **81**) are bitter, their D enantiomers are not (they are reported to be sour/neutral).[1] The D counterparts of sweet L-amino acids are bitter. Why the configurational effect is specific

Table IV. Prediction of Bitter Thresholds[a]

no.	compd	log (1/T) obsd	log (1/T) calcd[b]	Δ	Σπ	log P[c]	D
100	Gly	NB	0.77		0.00	−3.21[d]	6.74
101	Ala	NB	0.94		0.31	−2.89[d]	6.86
102	Asn	NB	0.94		−0.60	−3.41	9.25
103	Asp	NB	0.60		−0.77	−4.01	8.69
104	Gln	NB	1.19		−0.22	−3.15	10.54
105	Glu	NB	0.79		−0.64	−3.88	9.94
106	Ser	NB	0.86		−0.04	−3.30[d]	7.96
107	Thr	NB	1.06		0.26	−2.91[d]	8.14
108	Cys	S	1.64		1.54	−1.73	8.51
109	Met	S	1.85		1.23	−1.84[d]	11.02
110	Gly-Gly	NB	0.87		0.00	−3.81	10.37
111	Gly-DL-Ala	NB	0.98		0.31	−3.56	10.37
112	DL-Ala-Gly	NB	0.97		0.31	−3.61	10.48
113	DL-Ala-DL-Ala	NB	1.09		0.62	−3.36	10.48
114	Pro-Gly	NI[e]	1.42[f]		0.72		11.38
115	Pro-Val	NI[e]	2.02[f]		1.94		12.05
116	Gly-Gly-Gly	NB	0.93		0.00	−4.47	13.99
117	D-Leu-Gly	B	1.56		1.70	−2.90	12.99
118	Ala-Ile-Ala	B	1.73		2.42	−2.78	14.11
119	Ala-Ala-Leu	B	1.93		2.32	−2.77	16.05
120	Gly-Ala-Leu	B	1.84		2.01	−2.97	16.05
121	Phe-Gly-Gly-Phe	2.90	3.38	−0.48	3.58	−1.51[d]	24.49
122	Phe-Gly-Phe-Gly	2.90	2.96	−0.06	3.58	−1.55	21.38
123	Leu-Pro-Phe-Asp-Gln-Leu	3.82			5.09		29.43
124	Leu-Pro-Phe-Ser-Gln-Leu	3.82			5.65		29.43
125	Ala-Ala-Ala-Ala-Ala-Ala	NB			1.86		24.97
126	norleucine	1.70	1.94	−0.24	1.70	−1.58	10.65
127	norvaline	1.32	1.65	−0.33	1.37	−1.90	9.37
128	ornithine	NB	0.63		−1.11	−4.34	10.53
129	2-aminobutyric acid	1.01	1.27	−0.26	0.82	−2.44	8.10

[a] Abbreviations used: B, bitter; NB, not bitter; S, sulfurous; NI, not identifiable. [b] Unless otherwise noted, the values were calculated by eq 8. [c] Unless otherwise noted, the values were estimated by eq 7 or in ref 6. [d] Determined experimentally in ref 6. [e] Citrous or fruity taste. [f] Calculated by eq 6.

to these compounds is an unanswered question.

The bitterness of the nonzwitterionic and neutral species, i.e., *N*-acyl, ester, and *N*-acyl ester derivatives, was expressed by a single eq 6 or 8, together with that of the zwitterionic amino acids and peptides. This result suggests that the interaction with the receptor site via the charge is not an indispensable requisite for triggering of the bitter sensation.

The positive sign of the $\sum\pi$ term in eq 6 and the log P term in eq 8 may reflect the partitioning from a polar aqueous medium, saliva, onto the hydrophobic receptor cavity of the tongue, rather than the transport process to the site. A similar observation and discussion have been made previously in the analysis of a class of sweeteners, perillartines, the coefficient value (0.63) of the hydrophobicity term of which overlaps with the present one within the 95% confidence intervals.[13]

Table IV lists the activity calculated by eq 8 or 6 of the compounds the taste of which is flat or not identifiable. Since the lowest detection limit for bitterness is probably about 1 in terms of log $(1/T)$, as seen from the data of Table I, the nonbitterness or low bitterness of most of these compounds is predicted well by eq 8, the calculated values being about 1 or less. Similarly, the intensities of the bitter peptides D-Leu-Gly (**117**), Ala-Ile-Ala (**118**), Ala-Ala-Leu (**119**), and Gly-Ala-Leu (**120**), the threshold values of which have not been calculated exactly, were predicted by eq 8 and are listed in Table IV. We think that the values are reliable. Our panel could not perceive the possible bitterness of cysteine (**108**) and methionine (**109**), because of their strong sulfurous odor and taste. The possible bitterness of Pro-Val (**114**) was not recognized by our panel; the taste was somewhat citrous. Similarly, Pro-Gly (**115**) tasted fruity. Although these and some other exceptions may exist, eq 8 and/or eq 6 may be of value in food manufacturing in prediction of the bitterness of peptides that have not been tasted so far in the pure state.

In this study, we confined our analysis to compounds smaller than tetrapeptides, since there are few bitter compounds the size of which is equal to or larger than tetrapeptides. However, we examined the predictability or extensibility of eq 8 to the tetrapeptide **121** and **122**, for which hydrophobicity data are available.[6] The correspondence between the observed and calculated log $(1/T)$ values was good, suggesting that eq 8 is applicable also to tetrapeptides and probably to larger polypeptides, if the log P values are available. Of the three hexapeptides in Table IV, **123** and **124** are very bitter whereas **125** is not. On the basis of the results of eq 8 and the above discussion, we think that the taste is attributable to their hydrophobicity and molecular lengths. Although their log P values are not known, the hydrophobicities of **123** and **124** seem to be large enough for bitterness, but that of **125** seems not to be, judging from their $\sum\pi$ values. Moreover, the D value is larger in the first two.

Fauchère has estimated the π values of several unnatural amino acids not directly by the measurement of partition coefficients but by chromatography,[14] and the bitterness of some, norleucine (**126**), norvaline (**127**), ornithine (**128**), and 2-aminobutyric acid (**129**), has been examined by Wieser and Belitz.[1,2] We calculated their log $(1/T)$ values by eq 8. The predictability was very good, as shown in Table IV, suggesting that eq 8 is extensible also to these unnatural compounds.

(13) Iwamura, H. *J. Med. Chem.* **1980**, *23*, 308.

(14) Fauchere, J.-L. *QSAR in Design of Bioactive Compounds. Proceedings of the 1st Telesymposium on Medicinal Chemistry*; J. R. Prous: Barcelona, 1984; p 135.

Amino acid and peptide derivatives have a variety of medicinally and agriculturally important activities. QSAR approaches to these classes of compounds have been scarce, probably because of insufficient hydrophobicity data and the lack of a way to express their steric features. The work of Fauchere and Pliska[5] provided us with the basis for the present study with respect to hydrophobicity. They analyzed some medicinal peptides by using steric bulk parameters,[14] but their results can perhaps be further elaborated. Directional steric parameters based on the CPK model are useful in the study of structure–activity relationships of dipeptide sweeteners.[15] The dimensional parameter D has been developed for analysis of long-chain terpenoid molecules with insect juvenile hormone activity.[13] The results led to the development of new classes of active compounds.[16] The present study and the related ones cited above may serve as prototypes of how to unravel structure–activity puzzles of such complex molecules as amino acids, peptides, and their derivatives.

Experimental Section

Test Compounds. Compounds **1**–**8**, **10**, and **75** were kindly provided by the Tanabe Seiyaku Co., Ltd., and others were purchased from the Sigma Chemical Co., Wako Pure Chemical Industries, Ltd., and Tokyo Kasei Kogyo Co. Ltd.

Bitter Thresholds. The preparation of the test solution and the methods of evaluation were essentially the same as those reported by Wieser and Belitz.[1,2] Briefly, a series of solutions of increasing concentration in which one solution was twice as strong as the preceding one was prepared, and a 2–3-mL portion of each was tested by each person. The panel consisted of six persons. When the compound was a hydrochloride or an N-acyl derivative, the solution was neutralized with 0.1 N NaOH. The standard error of the determination was within ±15%.

Registry No. **1**, 72-18-4; **2**, 61-90-5; **3**, 73-32-5; **4**, 147-85-3; **5**, 63-91-2; **6**, 60-18-4; **7**, 73-22-3; **8**, 56-87-1; **9**, 74-79-3; **10**, 71-00-1; **11**, 1963-21-9; **12**, 869-19-2; **13**, 688-13-1; **14**, 19461-38-2; **15**, 704-15-4; **16**, 3321-03-7; **17**, 34258-14-5; **18**, 2390-74-1; **19**, 658-79-7; **20**, 3303-45-5; **21**, 3303-34-2; **22**, 3061-90-3; **23**, 686-43-1; **24**, 27493-61-4; **25**, 3918-94-3; **26**, 3989-97-7; **27**, 686-50-0; **28**, 7298-84-2; **29**, 3303-31-9; **30**, 17665-02-0; **31**, 38689-30-4; **32**, 3063-05-6; **33**, 5156-22-9; **34**, 968-21-8; **35**, 868-28-0; **36**, 24787-73-3; **37**, 41017-96-3; **35**, 26462-22-6; **39**, 42537-99-5; **40**, 37462-92-3; **41**, 13589-06-5; **42**, 59652-59-4; **43**, 54532-76-2; **44**, 59652-60-7; **45**, 42516-53-0; **46**, 22677-60-7; **47**, 6403-14-1; **48**, 59652-61-8; **49**, 6422-36-2; **50**, 52899-07-7; **51**, 51926-51-3; **52**, 19786-36-8; **53**, 13589-02-1; **54**, 721-90-4; **55**, 3303-55-7; **56**, 7669-65-0; **57**, 2577-40-4; **58**, 17355-18-9; **59**, 36099-95-3; **60**, 20696-60-0; **61**, 17355-10-1; **62**, 6665-16-3; **63**, 14857-82-0; **64**, 59685-29-9; **65**, 4306-24-5; **66**, 1187-50-4; **67**, 19408-48-1; **68**, 58337-01-2; **69**, 10329-75-6; **70**, 56243-96-0; **71**, 58337-00-1; **72**, 2743-60-4; **73**, 2577-90-4; **74**, 21685-51-8; **75**, 1080-06-4; **76**, 949-67-7; **77**, 4299-70-1; **78**, 7479-05-2; **79**, 1499-46-3; **80**, 1188-21-2; **81**, 2018-61-3; **82**, 1218-34-4; **83**, 495-69-2; **84**, 2198-64-3; **85**, 17966-60-8; **86**, 1492-11-1; **87**, 2361-96-8; **88**, 21156-62-7; **89**, 2382-80-1; **90**, 840-97-1; **91**, 54793-73-6; **92**, 27560-15-2; **93**, 38155-19-0; **94**, 29852-55-9; **95**, 4033-42-5; **96**, 38155-16-7; **97**, 59652-62-9; **98**, 17554-10-8; **99**, 38155-11-2; **100**, 56-40-6; **101**, 56-41-7; **102**, 70-47-3; **103**, 56-84-8; **104**, 56-85-9; **105**, 56-86-0; **106**, 56-45-1; **107**, 72-19-5; **108**, 52-90-4; **109**, 63-68-3; **110**, 556-50-3; **111**, 926-77-2; **112**, 1188-01-8; **113**, 2867-20-1; **114**, 2578-57-6; **115**, 52899-09-9; **116**, 556-33-2; **117**, 997-05-7; **118**, 58336-99-5; **119**, 54865-20-2; **120**, 22849-49-6; **121**, 40204-87-3; **122**, 59005-83-3; **123**, 109552-74-1; **124**, 58337-03-4; **125**, 10576-91-7; **126**, 327-57-1; **127**, 6600-40-4; **128**, 70-26-8; **129**, 80-60-4.

(15) Iwamura, H. *J. Med. Chem.* **1981**, *24*, 572.
(16) Nakayama, A.; Iwamura, H.; Niwa, A.; Nakagawa, Y.; Fujita, T. *J. Agric. Food Chem.* **1985**, *33*, 1034.

Synthesis of Some [N-(2-Haloalkyl)amino]tetralin Derivatives as Potential Irreversible Labels for Bovine Anterior Pituitary D_2 Dopamine Receptors

A. W. Hall,[†] R. J. K. Taylor,*[†] S. H. Simmonds,[‡] and P. G. Strange*[‡]

Department of Biochemistry, The Medical School, Queen's Medical Centre, Nottingham NG7 2UH, U.K., and School of Chemical Sciences, University of East Anglia, Norwich, NR4 7TJ, U.K. *Received June 13, 1986*

A series of hydroxylated 2-aminotetralins were prepared in the search for irreversible labels for D_2 dopamine receptors. N-2-Haloacetyl and N-2-haloalkyl substituents were chosen as potential receptor alkylating groups. Titrimetric studies were carried out on [N-(chloroethyl)-N-methylamino]tetralins **10**, **10a**, **24**, and **26** to demonstrate that aziridinium ions were formed as reactive intermediates from these compounds. This observation was confirmed by ^{1}H NMR studies on compound **10**. The majority of the aminotetralins prepared showed reasonably high affinity binding to anterior pituitary D_2 dopamine receptors and exhibited agonist properties. Structure–activity results are presented together with preliminary studies designed to identify irreversible receptor binding agents. [N-(2-Chloroethyl)-N-propylamino]-6,7-dihydroxytetralin hydrobromide (**18**) proved most promising in these studies.

Many of the important physiological actions of dopamine are mediated via its binding to D_2 dopamine receptors, and these receptors are also key sites of action of antipsychotic and anti-Parkinsonian drugs. The receptors have been extensively characterized recently by using the ligand binding technique,[1] the mechanism of action of the receptor has been probed,[2,3] and progress has been made toward isolation of the receptor protein.[4]

A selective irreversible label for the receptor would be an extremely useful tool for studying the receptor, and recently several photolabile-modified dopamine antagonists have been prepared, e.g., azidoclebopride,[5,6] azidosulpiride,[7] and N-(p-azido-m-iodophenethyl)spiperone.[8] An agonist-derived irreversible label would also be useful for studying the receptor, and N-(chloroethyl)norapomorphine was prepared[9] in this regard. Although early studies in-

[†]University of East Anglia.
[‡]Queen's Medical Centre.

(1) Seeman, P. *Pharmacol. Rev.* **1980**, *32*, 229.
(2) Enjalbert, A.; Bockaert, J. *Mol. Pharmacol.* **1983**, *23*, 576.
(3) Simmonds, S. H.; Strange, P. G. *Neurosci. Lett.* **1985**, *60*, 267.
(4) Strange, P. G. *Trends Pharmacol. Sci.* **1983**, *4*, 188.
(5) Wouters, W.; Laduron, P. *Biochem. Soc. Trans.* **1985**, *13*, 1103.
(6) Niznik, H. B.; Guan, J. H.; Neumeyer, J. L.; Seeman, P. *Mol. Pharmacol.* **1985**, *27*, 193.
(7) Redouane, K.; Sokoloff, P.; Schwartz, J. C.; Hamdi, P.; Mann, A.; Wermuth, C. C.; Roy, J.; Morgat, J. L. *Biochem. Biophys. Res. Commun.* **1985**, *130*, 1086.
(8) Amlaiky, N.; Caron, M. G. *J. Biol. Chem.* **1985**, *260*, 1983.

0022-2623/87/1830-1879$01.50/0