Chapter 4

Sweet Peptides and Proteins

Synthetic Studies

Yasuo Ariyoshi, Masanori Kohmura, Yasunobu Hasegawa, Masafumi Ota, and Noriki Nio

Central Research Laboratories, Ajinomoto Company, Inc., 1–1 Suzuki-cho, Kawasaki-ku, Kawasaki 210, Japan

Preliminary results of a two-pronged approach to the delineation of the mode of sweetener-receptor binding are reported. First, the results of an extensive program based on N- and C-terminal extension of potently sweet dipeptides are described. It can be concluded that the receptor has a small space that can accomodate an additional small D-amino acid residue at the site facing the N-terminus of sweet aspartyl dipeptides. It can also be concluded that there probably exists a spatial barrier at the site facing the C-terminal part of the sweet peptides. Second, the primary structure of the sweet protein monellin was unambiguously determined by a combination of solid phase peptide synthesis and comparative tryptic digest mapping of synthetic and natural products. It is proposed that knowledge of monellin structure will allow design of novel peptide sweeteners and will aid in determining the mode of sweetener-receptor binding.

In an attempt to deduce the mode of interaction between sweet-tasting peptides and their receptor, we have been engaged in a study of peptide and protein sweetener analogues. This review summarizes our progress in this program.

Peptide Sweetener Studies

In earlier work (*1*), the relative potencies of the aspartyl dipeptide esters **1-11** given in Table I were rationalized. In this work, the Fischer projection formulae "A" and "B" of Figure 1 were employed. It was found that the most potent dipeptide esters uniformly are of the "A" type while the inactive analogues are of the "B" type of molecular topography. This work clearly showed that activity did

0097–6156/91/0450–0041$06.00/0

Table I. Sweetness Potencies of Dipeptide Esters

	Compound	Sweetness Potency (% Sucrose Reference)[a]	Ref.
1	L-Asp-Gly-OMe	8 (0.6)	*7*
2	L-Asp-D-Ala-OMe	25 (0.6)	*8*
3	L-Asp-L-Ala-OMe	-	*8*
4	L-Asp-D-Ala-OPr	170 (0.6)	*8, 9*
5	L-Asp-D-Abu-OMe	16 (0.6)	*9*
6	L-Asp-L-Abu-OMe	0	*9*
7	L-Asp-D-Val-OMe	30 (0.6)	c
8	L-Asp-L-Phe-OMe	180 (1); 400[b]	*10, 11*
9	L-Asp-D-Phe-OMe	-	*10*
10	DL-Ama-D-Ala-OMe	55 (0.6)	c
11	DL-Ama-L-Phe-OMe	300-400	*12*

[a] Times as potent as sucrose (weight basis, sucrose = 1). -, bitter; 0, tasteless.
[b] Recognition threshold
[c] Y. Ariyoshi, unpublished results

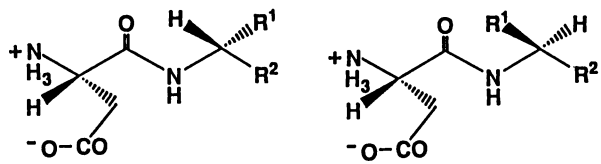

Figure 1. Projection formulae of aspartyl dipeptide esters. R^1 = R^2 = hydrophobic group.

not depend on the L- or D-configuration of the second amino acid ester but rather on the size and shape of the amino acid ester carboalkoxy and sidechain substituents. When R^1 and R^2 are sufficiently dissimilar in size, the sweetness potency is very high.

On the basis of Structure-Activity-Relationships (SAR) and molecular mechanics calculations on sweet-tasting aspartyl dipeptide analogues, the mode of interaction between sweet peptides and the receptor has been hypothesized (*2-7*). These receptor models suggest that, if space remains in the dipeptide ester receptor binding site at the C- or N-terminus of sweet aspartyl dipeptides, then peptides extended at the C- or N-terminus of the sweet peptides may taste sweet. Thus, we undertook a study of the C- and N-terminal extension products of sweet-tasting peptides.

C-Terminal Extension of Sweet Aspartyl Peptides

Extension at the C-Terminus of Aspartyl Dipeptide Esters (Compounds **12-28** in Table II). In order for aspartyl tripeptide esters to be sweet, the second amino acid must have the D-configuration and a small, compact alkyl sidechain. In addition, an L-configuration for the third amino acid is required in order to achieve any significant sweetness potency. Thus, the general structure for sweet aspartyl tripeptide esters can be drawn as in Figure 2, in which R^1 and R^3 are hydrogen atoms, and R^2 and R^4 are small alkyl groups. As suggested by a comparison of the activities of tripeptide esters **12-16** with that of **18** or **20**, the small alkyl groups (R^2 and R^4) may participate in hydrophobic interactions with the receptor thus causing increase in their sweetness potencies.

Extension at the C-Terminus of Aspartyl Tripeptide Esters (Compounds **29-42** in Table II). Tetrapeptides **30-38** were predicted to be weakly sweet, since the C-terminal extension of aspartyl dipeptides significantly decreased their sweetness potencies as described above, although the second and third amino acids have the D- and L-configurations, respectively. As expected, the sweetness potencies of these aspartyl tetrapeptides were extremely low or devoid of sweetness.. The sweetness was accompanied by a bitter or astringent taste.

Extension at the C-Terminus of Aspartyl Tetrapeptides (Compounds **43-49** in Table II). The pentapeptides (i.e., **43-49**) were predicted not to be sweet, since the trend observed with aspartyl tripeptides **12-29** and tetrapeptides **30-42** is a significant decrease in sweetness potency with elongation of the peptide chain. As expected, all the pentapeptides are essentially tasteless or faintly bitter.

Table II. Sweetness Potencies of C-Terminus Elongation Peptides

	Compound	Sweetness Potency[a]	Ref.
12	L-Asp-Gly-Gly-OMe	0	*13*
13	L-Asp-Gly-L-Ala-OMe	2	*13*
14	L-Asp-Gly-D-Ala-OMe	1	*13*
15	L-Asp-Gly-L-Val-OMe	-	*13*
16	L-Asp-D-Ala-Gly-OMe	3	*13*
17	L-Asp-D-Val-Gly-OMe	0	*13*
18	L-Asp-D-Ala-L-Ala-OMe	50	*13*
19	L-Asp-D-Ala-D-Ala-OMe	5	*13*
20	L-Asp-D-Ala-L-Val-OMe	50	*13*
21	L-Asp-D-Ala-L-Phe-OMe	-	*13*
22	L-Asp-D-Val-L-Ala-OMe	25	*13*
23	L-Asp-D-Val-D-Ala-OMe	4	*13*
24	L-Asp-D-Val-L-Val-OMe	30	*13*
25	L-Asp-D-Leu-L-Ala-OMe	-	*13*
26	L-Asp-L-Ala-L-Ala-OMe	-	*13*
27	L-Asp-L-Ala-L-Val-OMe	0	*13*
28	L-Asp-L-Val-L-Val-OMe	+/- (4)	*13*
29	L-Asp-Gly-Gly-Gly-OMe	0	*14*
30	L-Asp-D-Ala-L-Ala-L-Ala-OMe	0.5	*14*
31	L-Asp-D-Ala-L-Ala-L-Val-OMe	-	*14*
32	L-Asp-D-Ala-L-Ala-L-Leu-OMe	-	*14*
33	L-Asp-D-Ala-L-Val-L-Ala-OMe	2	*14*
34	L-Asp-D-Ala-L-Val-L-Val-OMe	-	*14*
35	L-Asp-D-Val-L-Ala-L-Ala-OMe	+	*14*
36	L-Asp-D-Val-L-Ala-L-Val-OMe	-	*14*
37	L-Asp-D-Val-L-Val-L-Ala-OMe	+/- (5)	*14*
38	L-Asp-D-Val-L-Val-L-Val-OMe	-	*14*
39	L-Asp-L-Ala-L-Ala-L-Ala-OMe	-	*14*
40	L-Asp-L-Ala-L-Val-L-Ala-OMe	0	*14*
41	L-Asp-L-Val-L-Ala-L-Ala-OMe	0	*14*
42	L-Asp-L-Val-L-Val-L-Ala-OMe	0	*14*
43	L-Asp-D-Ala-L-Ala-L-Ala-L-Ala-OMe	0	*15*
44	L-Asp-D-Ala-L-Ala-L-Val-L-Ala-OMe	0	*15*
45	L-Asp-D-Ala-L-Val-L-Ala-L-Ala-OMe	-	*15*
46	L-Asp-D-Ala-L-Val-L-Val-L-Ala-OMe	0	*15*
47	L-Asp-D-Ala-L-Val-L-Ala-L-Leu-OMe	0	*15*
48	L-Asp-D-Val-L-Ala-L-Val-L-Ala-OMe	0	*15*
49	L-Asp-L-Ala-L-Val-L-Val-L-Ala-OMe	0	*15*

[a] Times as potent as sucrose (weight basis, 0.6% sucrose = 1). +, faintly sweet; -, bitter; 0, tasteless; +/-, bitter-sweet, the number in parentheses is the sweetness value.

The results described above suggest to us that the receptor site may exist in the form of a deep pocket with critical binding sites deep inside. Thus, as an example for the case of aspartyl tripeptide esters, we propose that the mode of interaction with the receptor may be as is illustrated in Figure 3. In the case of the sweet-tasting dipeptide ester aspartame (L-Asp-L-Phe-OMe), the methyl ester group corresponds to R^2 and will interact with the receptor through a hydrophobic interaction, and the benzyl moiety corresponding to -CONHCR^3(R^4)COOR^5 will interact with the receptor at the complement of R^4 through a hydrophobic interaction. It is apparent that these hydrophobic interactions enhance sweetness potency. It is our view that increases in the length of the peptide chain cause increased difficulty in fitting the proposed deep receptor pocket. If true, the concomitant decrease in binding affinity to the receptor would result in diminished sweetness potency. In conclusion, it is clear that the receptor can accommodate an additional small L-amino acid ester at the site facing the C-terminus of sweet aspartyl dipeptides as shown in Figure 3.

We hypothesize that there exists a spatial barrier at the site facing the C-terminus of the sweet-tasting peptides. This proposal is based on the observation that, while most of the aspartyl dipeptide analogues described are potently sweet, potency diminishes sharply on elongation at the C-terminus, finally reaching zero for the cases of the pentapeptides. This spatial barrier is the same as the spatial wall suggested by Iwamura (*5*), based on a quantitative analysis of the SAR of an extensive series of sweet-tasting dipeptides.

N-Terminal Extension of Sweet Aspartyl and Aminomalonyl Peptides (*16*)

Extension at the Free α-Amino Group of Sweet Aspartyl Dipeptides (Compounds **50-59** in Table III). It was found that extension of sweet aspartyl dipeptide esters by adding a small D-amino acid residue generally gave sweet compounds, although this alteration also significantly decreased sweetness potency. The D-configuration of a newly introduced amino acid appears to be essential for activity as can be seen by comparison of diastereoisomeric pairs **50/51** and **58/59**. Further extension at the N-terminus of the extended sweet tripeptides **50** and **58** resulted in the nonsweet compounds **60** and **61**, respectively.

Extension at the Free α-Amino Group of Sweet Aspartyl Tripeptides (Compounds **62-65** in Table III). The N-terminal extension of the sweet aspartyl tripeptides **18** and **24** gave the faintly sweet compounds **62** and **63** and the nonsweet compounds **64** and **65**. This may be explained by a poor fit of the tetrapeptides into the proposed narrow receptor pocket.

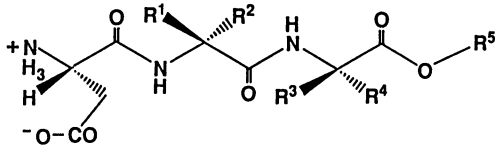

Figure 2. Sweet aspartyl tripeptide esters. $R^1 = R^3 = H$; $R^2 = R^4$ = small alkyl group.

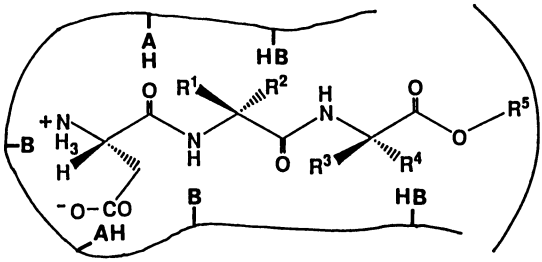

Figure 3. Schematic drawing of the mode of interaction between the sweet aspartyl tripeptides and the receptor. $R^1 = R^3 = H$; $R^2 = R^4 = R^5$ = small alkyl group; AH, Proton donor; B, Proton acceptor; HB, Hydrophobic binding site.

Table III. Sweetness Potencies of N-Terminus Elongation Peptides

	Compound	Sweetness Potency[a]	Ref.
50	D-Ala-L-Asp-D-Ala-OMe	12	*16*
51	L-Ala-L-Asp-D-Ala-OMe	+	
52	D-Val-L-Asp-D-Ala-OMe	+	*16*
53	D-Ala-L-Asp-D-Ala-OPr	30	
54	D-Pro-L-Asp-D-Ala-OPr	12	*16*
55	D-Ala-L-Asp-D-Val-OMe	1.5	
56	D-Val-L-Asp-D-Val-OMe	0	
57	D-Pro-L-Asp-D-Val-OMe	0	*16*
58	D-Ala-L-Asp-L-Phe-OMe	170	*16*
59	L-Ala-L-Asp-L-Phe-OMe	1	
60	L-Ala-D-Ala-L-Asp-D-Ala-OMe	Δ	*16*
61	L-Ala-D-Ala-L-Asp-L-Phe-OMe	-	*16*
62	D-Ala-L-Asp-D-Ala-L-Ala-OMe	+	*16*
63	D-Ala-L-Asp-D-Val-L-Val-OMe	1	
64	Gly-L-Asp-D-Val-L-Val-OMe	0	*16*
65	D-Val-L-Asp-D-Val-L-Val-OMe	0	
66	D-Ala-DL-Ama-D-Ala-OMe	-	*16*
67	D-Ala-DL-Ama-L-Phe-OMe	-	*16*
68	L-Ala-DL-Ama-L-Phe-OMe	-	
69	L-Asp(-Gly-OH)-L-Phe-OMe	0	*16*
70	L-Asp(-D-Ala-OH)-L-Phe-OMe	0	*16*
71	L-Asp(-L-Ala-OH)-L-Phe-OMe	0	*16*

[a] Times as potent as sucrose (weight basis, 0.6% sucrose = 1). +, faintly sweet (<1); Δ, astringent; 0, tasteless; -, bitter.

Extension at the Free α-Amino Group of Sweet Aminomalonyl Dipeptides (Compounds **66-68** in Table III).
The N-terminal extension of the sweet aminomalonyl dipeptides **10** and **11** led to a complete loss of the sweet taste (cf. **66-68**). This result is surprising in view of the close structural similarity of **66-68** to the tripeptides **50, 58** and **59.** The complete absence of sweetness in compounds **66, 67** and **68** may be a consequence of distortion of an essential A-H/B binding system (A-H: NH_3^+, B: COO^-) which has been proposed to be of critical dimensions (*17*). Slight changes in the distance between the A-H and B moieties may have rendered the peptides inaccessible to the narrow receptor pocket.

Extension at the Free β-Carboxyl Group of a Sweet Aspartyl Dipeptide (Compounds **69-71** in Table III). The peptides **69-71** which are extended at the β-carboxyl group of compound **8** are devoid of sweetness. As already described, in order for the tripeptide esters (Figure 2) to be sweet, small alkyl groups (R^2 and R^4) must be placed below the projection plane when the backbone chain of the peptide is projected on the plane of the paper. Exemplary is L-Asp-D-Ala-L-Ala-OMe (**18,** R^1 = R^3 = H, R^2 = R^4 = R^5 = Me) which is 50 times as potent as sucrose, whereas L-Asp-D-Ala-D-Ala-OMe (**19,** R^1 = R^4 = H, R^2 = R^3 = R^5 = Me is only 5 times as potent as sucrose, and L-Asp-L-Ala-L-Ala-OMe (**26,** R^2 = R^3 = H, R^1 = R^4 = R^5 = Me) is bitter. This SAR suggests that a hydrophobic area involving interactions which enhance sweetness potency may reside below the plane.

In the above modifications (Table III), in order for the peptides to be sweet, the first amino acid must be a small D-amino acid, and the third amino acid must fit the model for the sweet-tasting aspartyl dipeptide esters illustrated in Figure 1. Thus, the general structure for the sweet peptides can be drawn as Figure 4, in which small alkyl groups (R^2 and R^3) should be placed below the plane when the backbone chain of the peptides is projected on the plane of the paper. Exemplary of this requirement is D-Ala-L-Asp-L-Phe-OMe (**58**, R^1 = H; R^2 = Me; R^3 = COOMe; R^4 = $CH_2C_6H_5$) which is 170 times as potent as sucrose, whereas L-Ala-L-Asp-L-Phe-OMe (**59**, R^1 = Me; R^2 = H; R^3 = COOMe; R^4 = $CH_2C_6H_5$) is only equivalent to sucrose in potency.

From the results described above, it is suggested that the receptor site for sweet peptides may have a small residual space at a site facing the N-terminus. This restricted space can accommodate an additional small D-amino acid residue. The receptor peptide-binding site appears to be so spatially restricted that very subtle changes of structure strongly affect binding. Thus, only a limited number of peptides such as aspartyl di- and tri-peptide analogues, aminomalonyl dipeptide esters, and the above-

mentioned D-AA-L-Asp-AA-OMe can fit into the receptor and elicit a sweet taste.

Protein Sweetener Studies

Synthesis of the Sweet Protein Monellin (*18*). In attempts to deduce the mode of interaction between the sweet-tasting peptides and the receptor, a number of groups have carried out calculations and made spectroscopic measurements in order to determine the conformational energy minima of aspartame under the assumption that the "active" conformation is one of the low energy conformers. The efforts by Temussi and co-workers (*19*), van der Heijden and co-workers (*20*) and Iwamura (*5*) are illustrative. However, it is not possible to determine the active conformation of the sweet-tasting peptides by these methods since peptides can assume a multitude of low energy conformations in water. We, therefore, began efforts to employ sweet proteins as tools for elucidating the mode of peptide and/or protein sweetener/receptor interaction, since proteins are generally of much reduced conformational mobility. Inherent in this approach are two assumptions. First, it is assumed that the protein and peptide sweeteners bind with equivalent, if not identical, functional groups to the same locus on the same sweetener receptor. Second, we assume that knowledge of the three-dimensional structures of the protein sweeteners will allow unambiguous identification of these equivalent groups. As the first step to study the SAR of sweet proteins, we undertook the synthesis and crystallization of the sweet protein monellin.

Monellin has been isolated from the fruit of the West African plant, *Dioscoreophyllum cumminsii* (Stapf) Diels, by Morris and Cagan (*21*) and by van der Wel (*22*). It consists of two noncovalently associated polypeptide chains, the A-chain of 44 amino acid residues and the B-chain of 50 residues. Monellin is approximately 3000 times (*21*) as potent as sucrose on a weight basis, while neither of the individual A- or B-chain subunits is sweet (*23*). This indicates that the native conformation of monellin is important for the sweet taste. Two different primary structures have been reported for each of the A- and B-chains (*23-25*) as shown in Figure 5. Thus, there are four possible combinations for constructing the monellin structure. The discrepancies lie in positions 22, 25 and 26 of the A-chain, and 49 and 50 of the B-chain. We first synthesized the structure proposed by Frank and Zuber (*24*).

The A- and B-chains were synthesized by the stepwise solid-phase method using Fmoc protection (*26*). The peptide synthesis was performed manually with a semi-automated peptide synthesizer, Labortec SP 640 (for the A-chain), and with a manual shaker in a reaction vessel (for the B-chain). A preliminary synthesis of the B-chain revealed that the sulfur atom of the Met residue was susceptible to autoxidation to form the corresponding

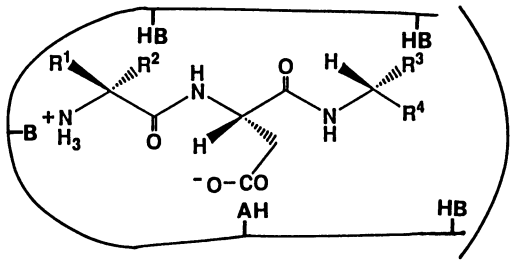

Figure 4. Schematic drawing of the mode of interaction between the sweet peptides and the receptor. R^1 = H; R^2 = R^3 = small hydrophobic group; R^3 and R^4 are hydrophobic groups where $R^3 \leq R^4$ in size; AH, Proton donor; B, Proton acceptor; HB, Hydrophobic binding site.

(a) REIKGYEYQLYVYASDKLFRADISEDYKTRGRKLLRFNGPVPPP
(b) REIKGYEYGLYVYASDKLFRANISQNYKTRGRKLLRFBGPVPPP

A chain

(a) GEWEIIDIGPFTQNLGKFAVDEENKIGQYGRLTFNKVIRPCMKKTIYENE
(c) GEWEIIDIGPFTQNLGKFAVDEENKIGQYGRLTFNKVIRPCMKKTIYEEN

B chain

Figure 5. Primary structures reported for monellin. For a, b and c, see references *22*, *23*, and *21*, respectively.

sulfoxide. Contamination of the synthetic B-chain by the sulfoxide form complicated the HPLC chromatogram. Therefore, the B-chain was synthesized with Met in the sulfoxide form.

The protection and deprotection schemes for the stepwise solid-phase synthesis of monellin are shown in Figure 6. The peptide was assembled on a p-alkoxybenzyl alcohol resin (*27*). Elongation of the peptide chain was carried out by the DCC/HOBt method (*28*) in CH_2Cl_2-DMF. The Fmoc group was removed with 50% or 20% piperidine in DMF. The coupling reaction and deprotection of the Fmoc group were monitored by the Kaiser test (*29*). When the deprotection was insufficient, the treatment with piperidine was repeated until a clear positive Kaiser test was obtained. When the coupling was incomplete, the reaction was repeated until a negative Kaiser test was obtained. When conversion was incomplete even following repeated coupling, the symmetrical anhydride method was used. If no further improvement could be made by these procedures, the "capping" procedure with acetic anhydride-pyridine was applied to eliminate the formation of deletion peptides.

Treatment of Fmoc-Pro-Pro-OCH_2-C_6H_4-OCH_2-resin with various concentrations of piperidine in DMF resulted in elimination of the dipeptide from the resin support. Therefore, Fmoc-Pro-Pro-Pro-OCH_2-C_6H_4-OCH_2-resin was synthesized by coupling Fmoc-Pro-Pro-OH with H-Pro-OCH_2-C_6H_4-OCH_2-resin.

In the synthesis of the A-chain, monitoring by the Kaiser test indicated that, for most residues, repeated coupling was necessary for quantitative incorporation, and for some residues, a second deprotection was necessary. After the last coupling step, the peptide-resin was treated with piperidine and then with CH_2Cl_2-anisole-thiophenol-TFA, and the resulting peptide was further treated with thiophenol-thioanisole-TFA, and then purified by preparative HPLC to give pure A-chain in an overall yield of 14.1% based on the amine content of the starting amino acid resin.

In the synthesis of the B-chain, either the symmetrical anhydride method or the DCC/HOBt method in NMP was used for amino acid residues for which coupling had been found to be difficult in a preliminary synthesis of the B-chain. After the last coupling step, the peptide-resin was treated with CH_2Cl_2-anisole-m-cresol-1,2-ethanedithiol-TFA, and the resulting peptide was further treated with thioanisole-m-cresol-1,2-ethanedithiol-TFA, and then with TMSBr (*30*). The crude B-chain was purified by HPLC to give the pure B-chain in an overall yield of 5.6% based on the amine content of the starting amino acid resin.

The purity of each peptide was confirmed by analytical HPLC, FAB-MS, quantitative amino acid analysis following hydrolysis in constant-boiling HCl containing 1% phenol at 110°C for 24 and 96 hr, and sequence analysis by automatic Edman degradation. HPLC analysis of each peptide gave satisfactory results. FAB mass spectrometric analysis of the A-chain showed a protonated

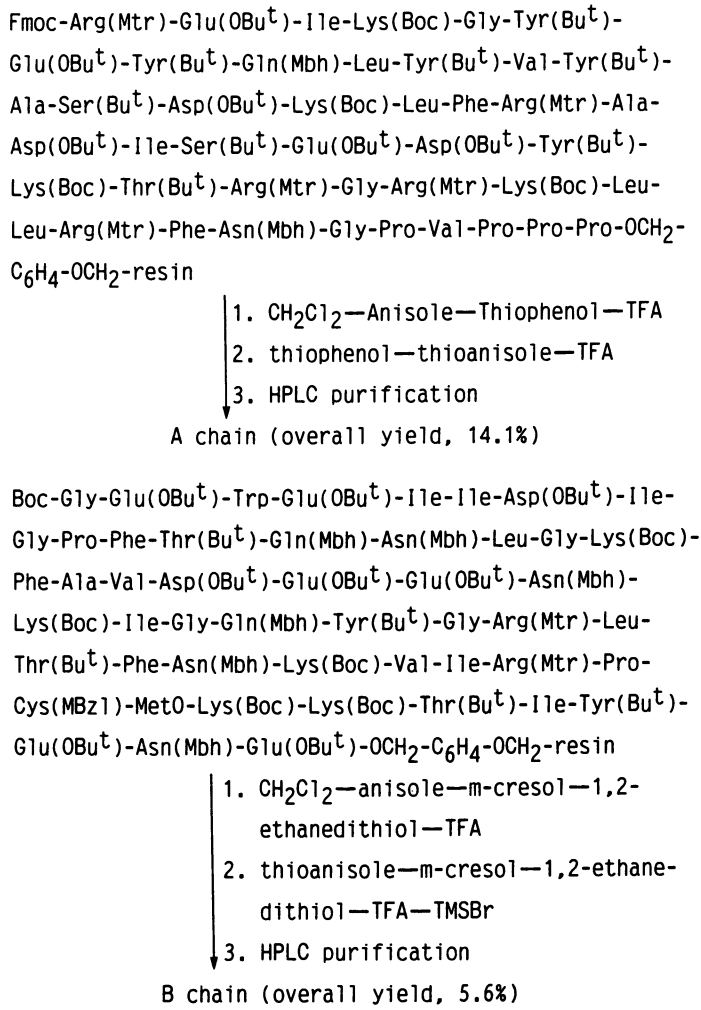

Figure 6. Protection and deprotection schemes for the solid-phase synthesis of monellin.

monoisotopic molecular ion at m/z 5248.8 relative to a calculated value of m/z 5248.8. Similarly, the observed value for the B-chain was m/z 5832.1 vs. a calculated value of m/z 5832.0. Amino acid composition of each peptide gave the expected values. Sequencing of the synthetic A- and B-chains was carried out on an automated sequencer, using the intact peptides and their tryptic peptides, and the results fully supported the expected sequences.

We observed that the individual synthetic A- and B-chain subunits were not sweet when each individual substance was tasted as a powder. Combination of the A and B-chains, and subsequent HPLC purification gave monellin in a yield of 53.9%. Its amino acid analysis gave a satisfactory result. Thus, given the comparable sweetness potency of this synthetic monellin with the natural monellin described above, it appears that the natural sweet conformation was almost completely formed. The synthetic monellin was noted to exhibit a distinct, lingering sweet taste about 4000 times more potent than sucrose on a weight basis (130,000 times/molar basis) relative to 0.6% sucrose as reference.

Interestingly, we noted that, after tasting a 0.6% sucrose solution, the sweetness intensity of a solution of synthetic monellin was significantly enhanced. Moreover, after rinsing the tongue two times with a highly diluted subthreshold tasteless solution (0.075 mg/L) of synthetic monellin, the solution itself tasted sweet. Under these conditions, we estimated synthetic monellin to exhibit a potency of 8000 times that of sucrose (0.6% sucrose reference). A similar phenomenon has been observed for thaumatin, another sweet protein (*31*).

Crystallization of the Synthetic Monellin. Crystallization was carried out by a combination of the methods described by Tomlinson and Kim (*32*), and by Wlodawer and Hodgson (*33*), using the "hanging-drop" vapor diffusion method, in which the synthetic monellin was dissolved in a 14% (w/w) solution of polyethylene glycol and phosphate buffer, and equilibrated with a 28% (w/w) solution of polyethylene glycol and phosphate buffer at 4°C. Crystals obtained by this method are shown in Figure 7.

Comparison of the Synthetic Product with Natural Monellin. It is known that approximately 10% of the A-chain of natural monellin carries an extra Phe residue at the N-terminus (*24*). This peptide is termed Phe-A-chain. Thus, a comparison of the synthetic product with natural monellin was performed after separating these peptide chains. Separation of the A-, Phe-A- and B-chains of natural monellin was readily accomplished by HPLC. Recombination of the separated A-chain (44 residues) with the separated B-chain gave "natural monellin". The synthetic monellin was identical to the reconstituted monellin by HPLC, but not by tryptic mapping. Therefore, a comparison was made between the individual chains. The synthetic A-chain was identical to the natural A-chain (44

residues) by tryptic mapping, but the synthetic B-chain was not. These results indicate that the reported structure for the B-chain differs from that of the natural B-chain. Therefore, we then determined the primary structure of natural monellin.

Determination of the Primary Structure of Natural Monellin (*34*). The sequence of monellin was determined by a combination of mass spectrometry (FAB and ESI), and automatic Edman degradation. The results are described briefly. A sample of monellin was obtained through the courtesy of Dr. J.G. Brand of Monell Chemical Senses Center. Sequencing was performed after separating the A-chain, Phe-A-chain and B-chain. The analytical HPLC of monellin revealed that approximately 10% of the A-chain carried an extra Phe residue at the N-terminus as described by Frank and Zuber (*24*).

Residues Arg^{1} to Phe^{37} of the A-chain were determined by automatic sequencing of the intact A-chain. The rest of the sequence was determined by automatic sequencing of a tryptic peptide after tryptic digestion of the A-chain. The complete amino acid sequence of the A-chain was determined to be as shown in Figure 8. This sequence is identical to that proposed by Frank and Zuber (*24*) and is also supported by our synthetic study described above (*18*).

Similarly, the separated B-chain was subjected to an automatic Edman degradation. By this method, Gly, Glu and Thr residues were identified as the N-terminal amino acids. This indicated that the B-chain was a mixture of three peptides, which we consider to be formed by enzyme action on the native protein. ESI mass spectrometric analysis of the mixture gave the measured mean molecular weights of 5834.3, 5777.1 and 5935.1, corresponding to calculated values of 5834.7, 5777.6 and 5935.8 for the respective intact B-chain subunit, and the des-Gly^{1}-B- and Thr-B-chain subunits. The ratio of these three components was semi-quantitatively estimated from the relative intensities of the peaks to be 100 : 43 : 33, respectively. Attempts to separate the modified peptides from the intact B-chain subunit by HPLC and by FSCE were unsuccessful. Therefore, the sequence was determined by automatic sequencing of tryptic peptides, after tryptic digestion of the mixture. The complete amino acid sequence of the B-chain was determined to be as shown in Figure 9. This sequence is identical to that reported by Bohak and Li (*23*).

We have synthesized monellin, the structure of which was determined as is described here, as well as analogues. The results of this effort will be comprehensively described elsewhere. Upon solution of the crystal structure of monellin and its analogues, we expect that the synthesis of oligopeptides as described in the first part of this review will lead to delineation of the receptor binding site of monellin. In addition, we expect these studies to lead to the

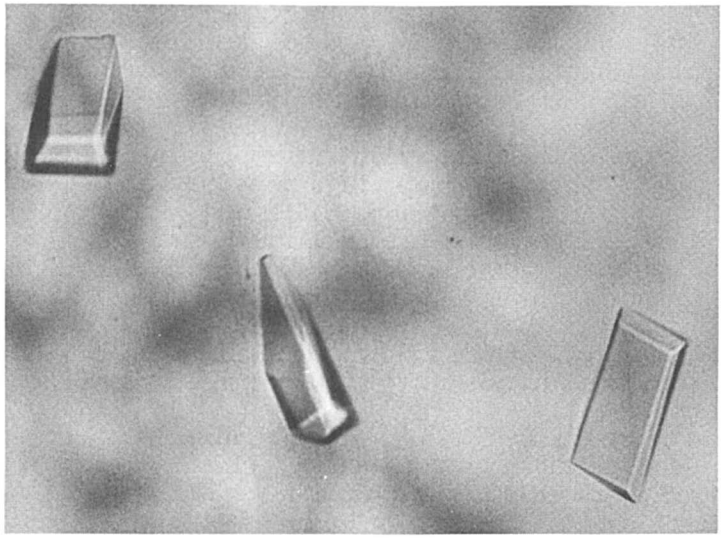

Figure 7. Crystals of synthetic monellin.

```
          1                   5                  10
(H-Phe)-Arg-Glu-Ile-Lys-Gly-Tyr-Glu-Tyr-Gln-Leu-Tyr-Val-

         15                  20                  25
Tyr-Ala-Ser-Asp-Lys-Leu-Phe-Arg-Ala-Asp-Ile-Ser-Glu-Asp-

         30                  35                  40
Tyr-Lys-Thr-Arg-Gly-Arg-Lys-Leu-Leu-Arg-Phe-Asn-Gly-Pro-

         44
Val-Pro-Pro-Pro-OH
```

Figure 8. The complete amino acid sequence of the A-chain of monellin, approximately 10% of which carries an extra Phe residue at the N-terminus as described by Frank and Zuber (*22*).

```
          1                   5                  10
(H-Thr)-Gly-Glu-Trp-Glu-Ile-Ile-Asp-Ile-Gly-Pro-Phe-Thr-

         15                  20                  25
Gln-Asn-Leu-Gly-Lys-Phe-Ala-Val-Asp-Glu-Glu-Asn-Lys-Ile-

         30                  35                  40
Gly-Gln-Tyr-Gly-Arg-Leu-Thr-Phe-Asn-Lys-Val-Ile-Arg-Pro-

         45                  50
Cys-Met-Lys-Lys-Thr-Ile-Tyr-Glu-Glu-Asn-OH
```

Figure 9. The complete amino acid sequence of the B-chain of monellin, approximately 19% of which carries an extra Thr residue at the N-terminus, and approximately 24% of which lacks the N-terminal Gly residue.

design of novel sweeteners and the elucidation of the mode of interaction between sweet compounds and their receptor(s).

Acknowledgments. We thank Dr. J.G. Brand of Monell Chemical Senses Center for sending us the sample of monellin, Dr. S. Ikeda of Kokusan Chemical Works Ltd. for providing us with Fmoc-Pro-Pro-OH, Mr. M. Kaise and Mr. S. Fukushima of Applied Biosystems Japan for FSCE, Dr. B. N. Green of VG Masslab Ltd. and Dr. A. Shibata of Jasco International for measurements of ESI mass spectra, Dr. K. Hirayama, Ms. S. Akashi and Ms. M. Furuya for measurements of FAB mass spectra, Mr. S. Ozawa, Mr. T. Seino and Ms. M. Fuchiagami for amino acid analyses, and Mr. K. Iijima for elemental analyses.

Abbreviations follow the recommendations of the IUPAC-IUB Commission on Biochemical Nomenclature in *Eur. J. Biochem.* **1984**, *138*, 9-37. Other abbreviations used: OPr, propyloxy; Ama, aminomalonyl; HPLC, high-performance liquid chromatography; Mtr, 4-methoxy-2,3,6-trimethylbenzenesulfonyl; MBzl, p-methoxybenzyl; Mbh, 4,4'-dimethoxybenzhydryl; DMF, N,N-dimethyformamide; DCC, dicyclohexylcarbodiimide; HOBt, 1-hydroxybenzotriazole; TFA, trifluoroacetic acid; TMSBr, trimethylsilyl bromide; FAB-MS, fast atom bombardment mass spectrometry; ESI-MS, electrospray ionization mass spectrometry; FSCE, free solution capillary electrophoresis.

Literature Cited

1. Ariyoshi, Y. *Agric. Biol. Chem.* **1976**, *40*, 983-992.
2. Fujino, M.; Wakimasu, M.; Mano. M.; Tanaka, K.; Nakajima, N.; Aoki, H. *Chem. Pharm. Bull.* **1976**, *24*, 2112-2117.
3. Lelj, F.; Tancredi, T.; Temussi, P.A.; Toniolo, C. *J. Am. Chem. Soc.* **1976**, *98*, 6669-6675.
4. van der Heijden, A.; Brussel, L.B.P.; Peer, H.G. *Chem. Senses Flavour.* **1979**, *4*, 141-152.
5. Iwamura, H. *J. Med. Chem.* **1981**, *24*, 572-583.
6. Goodman, M. *Biopolymers.* **1985**, *24*, 137-155.
7. Thomsen, M.W.; Dalton, J.M.; Stewart, C.N. *Chem. Senses* **1988**, *13*, 397-405.
8. Mazur, R.H.; Reuter, J.A.; Swiatek, K.A.; Schlatter, J.M. *J. Med. Chem.* **1973**, *16*, 1284-1287.
9. Ariyoshi, Y.; Yasuda, N.; Yamatani, T. *Bull. Chem. Soc. Jpn.* **1974**, *47*, 326-330.
10. Mazur, R.H.; Schlatter, J.M.; Goldkamp, A.H. *J. Am. Chem. Soc.* **1969**, *91*, 2684-2691.
11. Beck, C.I. In *Symposium: Sweeteners;* Inglett, G.I., Ed.; AVI: Westport, CT., 1974; pp 164-181.
12. Briggs, M.T.; Morley, J.S. Brit. Patent 1 299 265, 1972..
13. Ariyoshi, Y. *Bull. Chem. Soc. Jpn.* **1984**, *57*, 3197-3202.

14. Ariyoshi, Y. *Bull. Chem. Soc. Jpn.* **1985**, *58*, 1727-1730.
15. Ariyoshi, Y. *Bull. Chem. Soc. Jpn.* **1986**, *59*, 1027-1030.
16. Ariyoshi, Y.; Hasegawa, Y.; Nio, N. In *Peptide Chemistry 1986; Proceedings of the 24th Symposium on Peptide Chemistry;* Miyazawa, T., Ed.; Protein Research Foundation: Osaka, 1987, pp 251-256.
17. Shallenberger, R.S.; Acree, T.E.; Lee, C.Y. *Nature.* **1969,** *221*, 555-556.
18. Kohmura, M.; Nio, N.; Ariyoshi, Y. In *Peptide Chemistry 1989, Proceedings of the 27th Symposium on Peptide Chemistry*; Yanaihara, N., Ed.; Protein Research Foundation: Osaka, 1990, pp 175-180.
19. Lelj, F.; Tancredi, T.; Temussi, P.A.; Toniolo, C. *J. Am. Chem. Soc.* **1976**, 98, 6669.
20. van der Heijden, A.; van der Wel, H.; Peer, H.G., *Chem. Senses* **1985**, *10*, 57-72.
21. Morris, J.A.; Cagan, R.H. *Biochim. Biophys. Acta.* **1972**, *261*, 114-122.
22. van der Wel, H. *FEBS Lett.* **1972**, *21*, 88-90.
23. Bohak, Z.; Li, S.-L. *Biochim. Biophys. Acta.* **1976**, *427*, 153-170.
24. Frank, G.; Zuber, H. *Hoppe-Seyler's Z. Physiol. Chem.* **1976**, *357*, 585-592.
25. Hudson, G.; Biemann, K. *Biochem. Biophys. Res. Commun.* **1976**, *71*, 212-220.
26. Carpino, L.A.; Han, G.Y. *J. Am. Chem. Soc.* **1970**, *92*, 5748-5749.
27. Wang, S.-S. *J. Am. Chem . Soc.* **1973,** *95*, 1328-1333.
28. König, W.; Geiger, R. *Chem. Ber.* **1970**, *103*, 2041-2051.
29. Kaiser, E.; Colescott, R.L.; Bossinger, C.D.; Cook, P.I. *Anal. Biochem.* **1970**, *34*, 595-598.
30. Yajima, H.; Fujii, N.; Funakoshi, S.; Watanabe, T.; Murayama, E.; Otaka, A. *Tetrahedron.* **1988**, *44*, 805-819.
31. van der Wel, H.; Loeve, K. *Eur. J. Biochem.* **1972**, *31*, 221-225.
32. Tomlinson, G.E.; Kim, S.-H. *J. Biol. Chem.* **1981**, *256*, 12476-12477.
33. Wlodawer, A.; Hodgson, K.O. *Proc. Nat. Acad. Sci. USA* **1975,** *72*, 398-399.
34. Kohmura, M.; Nio, N.; Ariyoshi, Y. *Abstracts of Papers,* 59th Annual Meeting of the Chemical Society of Japan, Yokohama; Chemical Society of Japan, 1990; p 1127.

RECEIVED August 27, 1990