Tetrahedron Vol. 42, No. 3, pp. 829 to 834, 1986
Printed in Great Britain.

0040-4020/86 $3.00 + .00

STRUCTURE DETERMINATION OF LEPIDOPTERAN, SELF-DEFENSE SUBSTANCE PRODUCED BY SILKWORM.

TADASHI TESHIMA, YASUYUKI UEKI, TAKAHISA NAKAI, and TETSUO SHIBA*

Department of Chemistry, Faculty of Science, Osaka University,
Toyonaka, Osaka 560, Japan

MIKIO KIKUCHI

Research Laboratory of Fundamental Biological Science,
Hiratsuka, Kanagawa 253, Japan

(*Received in Japan* 16 *October* 1985)

The structure of lepidopteran was determined by sequential analysis using the Edman method for the whole molecule as well as several fragment peptides obtained by enzymatic and chemical cleavages. Consequently, lepidopteran was found to be a mixture of congeners composed of 35 amino acids possessing one replaceable amino acid residue of lysine with δ-hydroxylysine.

The immune system of insects lacks the antigen-antibody reaction which is a basic mechanism in a humoral immune system of vertebrates. We isolated the antibacterial peptide from body fluid of silkworm which was vaccinated with killed E. coli. This compound calling lepidopteran[1] seems to play an important role in a self-defense mechanism of insects. Lepidopteran was characterized to be a peptide composed of 35 amino acids including δ-hydroxylysine (Hyl) as one unusual amino acid.[1,2]

The sequence of lepidopteran was determined by the Edman degradation for the whole molecule and its fragments obtained by cyanogen bromide degradation and trypsin digestion. From these results, the sequences of 1-21, 22-33, and 34-35 were able to be deduced. In order to determine the connection between 21 to 22 as well as that between 33 to 34, the connecting peptides were attempted to be obtained by trypsin digestion after citraconic anhydride treatment of lepidopteran and its partial hydrolysis. The sequential analyses of these peptides let us confirm the connection mode mentioned above. Consequently we could determine the structure of lepidopteran as a mixture of congeners, A with Lys[3], B with Hyl at 21st position in the sequence as shown in Fig. 2.

Experimental

Separation of lepidopteran congeners. Lepidopteran[1,2] was separated by HPLC using Nucleosil 300-7C_{18} column eluted gradiently with 25% to 35% acetonitrile containing 0.1 % TFA as shown in Fig. 1.

Determination of peptide sequence. Edman degradation and the synthesis of PTH derivative of Hyl were carried out by the procedure of P. Edman and A. Henschen[4], and PTH amino acids were identified by HPLC (Nucleosil $5C_{18}$, 4 x 250 mm, 42% acetonitrile - 0.01 M sodium acetate buffer, pH 4.5)[5]. Dansyl-Edman method was performed by the procedure of W. R. Gray[6].

Cyanogen bromide treatment of lepidopteran. Lepidopteran (1.9 mg, 489 nmol) and cyanogen bromide (5.6 mg, 53 µmol) were dissolved in 70% formic acid and kept at room temperature for 1 day. Cyanogen bromide (10 mg, 95 µmol) was added to the solution and allowed to react for additional 1 day. After addition of water (180 µl), the reaction mixture was lyophilized. The residue was separated by CM-Sephadex C-25 column by gradient elution with 0 to 1 M sodium chloride. The fraction with Gly at N-terminus was desalted in Sephadex G-10 column by elution with 4% acetic acid to obtain the peptide (CB-1). This was used for Edman degradation.

Trypsin digestions of lepidopteran A and B. In a solution of lepidopteran A (100 nmol) in 0.2 M ammonium acetate buffer (pH 8.0), trypsin treated with tosylphenylalanine chloromethyl ketone[7] (1.5 nmol) was added and kept at 25°C for 6 h. After lyophilization, the residue was separated by HPLC (Nucleosil 300-$7C_{18}$, 6 x 250 mm, 1% - 50% acetonitrile containing 0.1% TFA). Lepidopteran B was treated with trypsin by the same manner.

Trypsin digestion of lepidopteran after citraconylation. To a solution of lepidopteran (1.09 mg, 242 nmol) in 50 mM N-ethylmorpholinium acetate buffer (pH 8.2) (200 µl) containing 0.5 % sodium dodecyl sulfate, citraconic anhydride (9 µl) was added portionwise. The reaction mixture was kept at pH 8 by addition of 1 M sodium hydroxide for 1 h. The reaction mixture was desalted through Sephadex G-25 column eluted with 0.5% ammonium hydrogencarbonate buffer (pH 8.0). The eluate was lyophilized to white powder whose aqueous solution was digested by trypsin and then separated by the same manner as that of lepidopteran to give the peptide fragment C-T-1.

Edman degradation of the peptide C-T-1. At the first cycle of Edman degradation of C-T-1, 3-sulfophenylisothiocyanate was used instead of phenylisothiocyanate. The rest of the procedure was the same as in the normal Edman degradation.

Partial hydrolysis of lepidopteran. Lepidopteran (50 nmol) was dissolved in c.HCl (50 µl) and kept at 25°C for 60 h. After the reaction mixture was concentrated in vacuo, the residue was separated by HPLC (Cosmosil $5C_{18}$, 4 x 125 mm, 0.1% aqueous TFA - 30% acetonitrile containing 0.1% TFA).

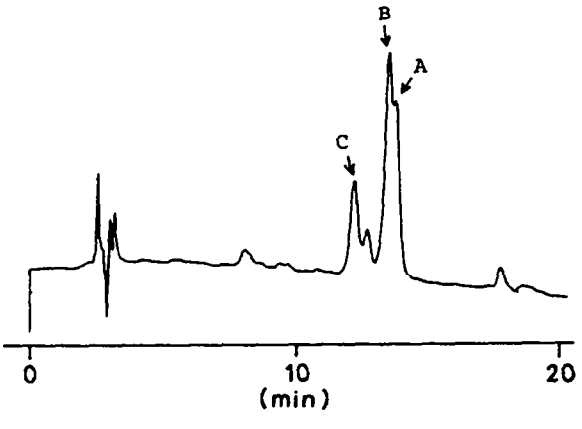

Fig. 1. HPLC pattern of lepidopteran. column: Nucleosil 300-$7C_{18}$, 6 x 250 mm, eluant: A) 25% CH_3CN in 0.1% TFA, B) 35% CH_3CN in 0.1% TFA, gradient A to B (20 min), flow rate :2 ml/min, detection: 220 nm.

Table 1. Amino acid analyses of lepidopteran A, B, and C

	lepidopteran A	B	C
Asp	2.00(2)	2.04(2)	1.99(2)
Ser	0.85(1)	0.89(1)	0.88(1)
Glu	2.01(2)	2.15(2)	2.11(2)
Pro	0.86(1)	1.05(1)	1.04(1)
Gly	4.00(4)	4.00(4)	4.00(4)
Ala	3.95(4)	4.06(4)	4.06(4)
Val	1.60(2)	1.67(2)	2.55(3)
Met	0.87(1)	0.88(1)	-
Ile	5.15(6)	5.15(6)	3.43(4)
Leu	0.96(1)	1.06(1)	2.90(3)
Phe	1.06(1)	0.99(1)	0.91(1)
Hyl	-	0.82(1)	0.95(1)
Lys	5.59(6)	5.04(5)	4.81(5)
Trp*	0.67(1)	1.18(1)	1.21(1)
Arg	2.56(3)	2.69(3)	2.60(3)

Samples were hydrolyzed with 6 M HCl at 110°C for 90 h.
* Hydrolyses were carried out for 20 h.

Synthesis of H-Ala-Ile-NH_2•HCl.

Boc-Ala-Ile-NH_2. To a suspension of H-Ile-NH_2•HCl[8] (1.61 g, 9.63 mmol) in DMF (30 ml) was added triethylamine (1.35 ml, 9.63 mmol) and Boc-Ala-OSu (2.76 g, 9.63 mmol) at 0°C with stirring. The reaction mixture was stirred overnight at room temperature. Crystals separated out were filtered off and the filtrate was concentrated in vacuo. The residue was triturated with ethyl acetate and water to give colorless solid. It was washed with 10% citric acid, sat. $NaHCO_3$, and water. Yield, 2.66 g (91.6%), mp 166-167.5°C, $[\alpha]_D^{21}$ -16.4°(c 1.02, DMF). Found: C, 55.96; H, 9.07; N, 13.75 %. Calcd for $C_{14}H_{27}N_3O_4$: C, 55.79; H, 9.03; N, 13.94 %.

H-Ala-Ile-NH_2•HCl. To a suspension of Boc-Ala-Ile-NH_2 (1.63 g, 5.41 mmol) in THF (12 ml) was added 4.35 M HCl in THF (25 ml) at 0°C with stirring. After clear solution was obtained, crystals were separated out. The reaction mixture was stirred for 3 h and concentrated in vacuo. The residue was washed with ether and recrystallized from methanol. Yield 1.17 g (80.7%), mp 127-129°C (dec), $[\alpha]_D^{21}$ +8.3° (c 1.14, DMSO). Found: C, 43.24; H, 8.75; N, 16.44; Cl, 14.19 %. Calcd for $C_9H_{20}N_3O_2Cl•3/4H_2O$: C, 43.03; H, 8.63; N, 16.73; Cl, 14.11 %.

Elucidation of amide structure at C-terminal of lepidopteran. The peptide fragment T-9 was identified with H-Ala-Ile-NH_2 by HPLC (Cosmosil 5C_{18}, 4 x 125 mm, 4% acetonitrile - 0.1% TFA), but different from H-Ala-Ile-OH[9] in retention time. Retention time: T-9, 4.9 min; H-Ala-Ile-NH_2, 4.9 min; H-Ala-Ile-OH, 11.6 min. The mixture of T-9 and H-Ala-Ile-NH_2 gave single peak at 4.9 min. In addition, PH-1 was identified with synthetic H-Ser-Ala-Lys-Ala-Ile-NH_2[10] by HPLC (the same condition as those of T-9). Retention time: PH-1, 11.0 min; H-Ser-Ala-Lys-Ala-Ile-NH_2, 11.0 min; the mixture gave single peak at 11.0 min.

Results and Discussion

Amino acid composition of lepidopteran A was the same as that of B except Lys and Hyl, and total mole numbers of Lys and Hyl were both six (lepidopteran A: Lys(6); lepidopteran B: Lys(5), Hyl(1)). Moreover, HPLC patterns of tryptic peptides of lepidopteran A and B were also superimposable except T-7a and T-7b as shown in Fig. 3. These results suggested that only one Lys residue of lepidopteran A is replaced with Hyl residue in the sequence of B. Therefore, we

Table 2. Amino acid compositions of peptide fragments obtained by enzymatic and chemical cleavage

	CB-1	T-1	T-2	T-3	T-4	T-5	T-6	T-7a	T-7b	T-8	T-9	C-T-1	PH-1
Asp	1.65(2)	-	-	-	-	-	1.05(1)	0.91(1)	0.89(1)	-	-	1.12(1)	-
Ser	1.08(1)	-	-	-	-	-	-	-	-	1.02(1)	-	0.99(1)	1.12(1)
Glu	1.14(1)	-	-	-	1.16(1)	-	-	-	-	1.15(1)	-	0.92(1)	-
Pro	**(1)	-	-	-	-	-	-	-	-	1.04(1)	-	1.00(1)	-
Gly	4.15(4)	-	-	-	-	1.00(1)	-	1.00(1)	1.00(1)	2.12(2)	-	3.09(3)	-
Ala	4.00(4)	-	-	-	-	-	-	-	-	2.81(3)	1.11(1)	4.00(4)	1.82(2)
Val	1.52(2)	-	-	-	-	-	-	$^{0.44}_{*0.83}$(1)	0.44(1)	0.98(1)	-	1.16(2)	-
Met	-	-	-	-	-	0.82(1)	-	-	-	-	-	-	-
Ile	3.06(4)	-	-	1.00(1)	1.00(1)	-	1.00(1)	$^{0.42}_{*0.80}$(1)	0.39(1)	1.00(1)	1.00(1)	2.40(3)	0.83(1)
Leu	1.10(1)	-	-	-	-	-	-	-	-	0.97(1)	-	0.91(1)	-
Phe	-	-	-	0.86(1)	-	-	-	-	-	-	-	-	-
Hyl	0.62	-	-	-	-	-	-	-	0.77(1)	-	-	-	-
Lys	1.40	1.00(1)	1.00(1)	0.97(1)	1.81(2)	-	-	0.69(1)	-	0.97(1)	-	1.64(2)	1.00(1)
Trp	-	1.04(1)	0.82(1)	-	-	-	-	-	-	-	-	-	-
Arg	1.72(2)	0.85(1)	-	-	-	0.85(1)	0.90(1)	-	-	-	-	-	-

Samples were hydrolyzed with 6M HCl at 110°C for 20 h.
* Hydrolysis was carried out for 90 h.
** The datum was not measured.

first used a mixture of congeners for the structural determination. A series of Edman degradation applied to lepidopteran clarifies the sequence from the N-terminus to 12th amino acid residue as shown in Fig. 1. Since the 11th amino acid was found to be Met, cyanogen bromide degradation was next applied to lepidopteran in order to obtain fragments in appropriate sizes. After separation of peptides obtained, the fragment with Gly at N-terminus (CB-1) was subjected to Edman degradation to clarify the sequence starting from 12th residue resulting in an elucidation of the sequence as far as 20th residue. Consequently a sequence from 1st to 20th residue could be now assigned.

On the other hand, trypsin digestion was applied to lepidopteran A and B separately in order to determine the remaining sequence of the 21st to 35th residues. Their HPLC patterns were compared each other as shown in Fig. 3. The pattern of lepidopteran A digestion was completely the same as that of lepidopteran B except T-7a and T-7b. All fragments from T-1 to T-9 obtained by trypsin digestion were separated by HPLC and each was analyzed by the Edman and the dansyl-Edman methods. The peptides, T-7a and T-7b, were now assigned to Asp^{17}-Gly-Ile-Val-Lys^{21} and Asp^{17}-Gly-Ile-Val-Hyl^{21} respectively, which confirm the fact that lepidopteran A only differ from lepidopteran B at the 21st amino acid residue. Dipeptide (T-9) was deduced to be H-Ala-Ile-OH or H-Ala-Ile-NH_2, because N-terminal amino acid of T-9 was determined to be Ala by dansyl method. The retention time of T-9 was identical with that of synthetic H-Ala-Ile-NH_2 but different from that of H-Ala-Ile-OH as described in the experimental section,

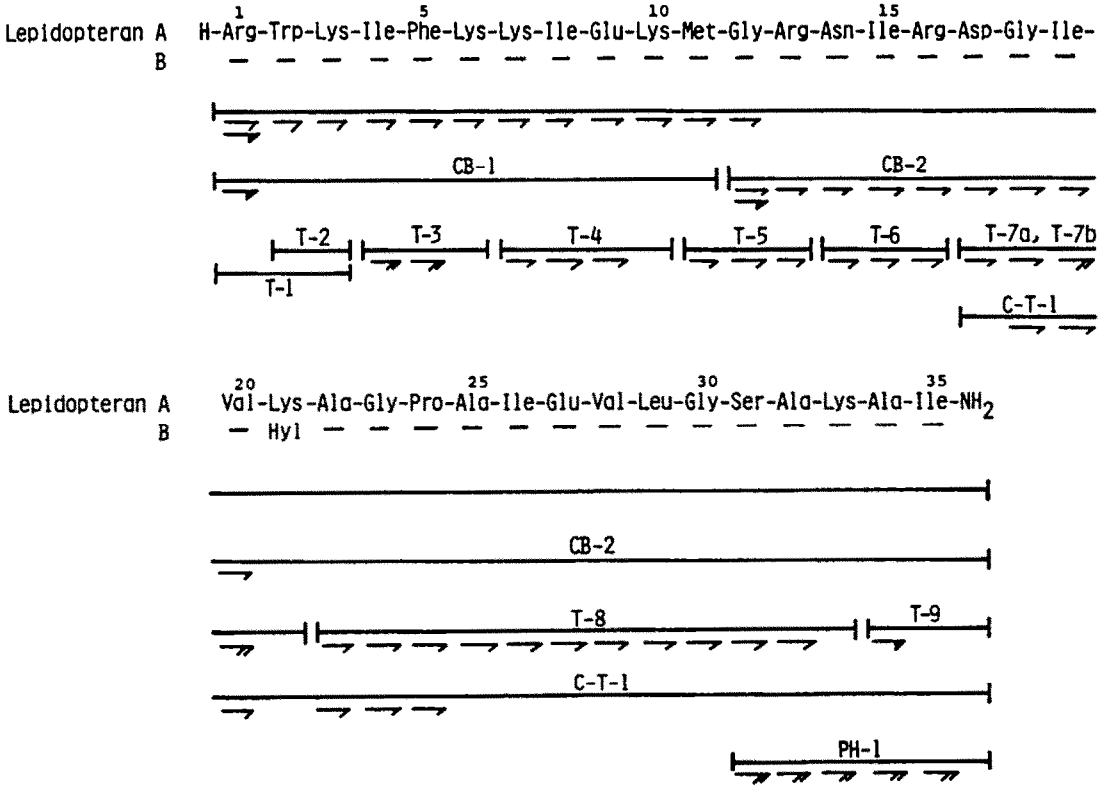

Fig. 2. Amino acid sequence of lepidopteran. CB: cyanogen bromide degradation, T: trypsin digestion, C-T: trypsin digestion after citraconylation, PH: partial hydrolysis, ⟶ :Edman degradation, ⟶ :dansyl Edman method, ⟶ :dansyl-method. Amino acid residues identical to those of lepidopteran A are indicated by dashes.

which confirmed the amide structure of T-9. From the results thus obtained, three partial peptide sequences, 1-21, 22-33, and 34-35, could be now assigned which cover all amino acid residues in the whole molecule of lepidopteran.

In order to confirm the connection of these peptides, trypsin digestion after treatment of citraconic anhydride or simple partial hydrolysis with conc HCl for 24 h was then carried out. The former procedure gave the new peptide fragment (C-T-1) corresponding to 17-35. The Edman degradation of C-T-1 using 3-sulfophenylisothiocyanate gave the sequence from Gly^{18} to Pro^{24}. In addition, the sequence of the peptide (PH-1) obtained by the partial hydrolysis was determined by dansyl-Edman method resulting in a clarification of C-terminal amide structure from 31st to 35th residues which was identified with the synthetic pentapeptide as described in experimental section. Thus, the whole sequences of lepidopteran A and B were determined.

On the other hand, lepidopteran C, seemed to be a similar peptide as judged by amino acid composition, and its structure elucidation will be reported later elsewhere. All these results elucidated that lepidopteran was the mixture of three congeners and that two of them, lepidopteran B and C, contained Hyl besides usual amino acids.

The similar peptide with comparable function named cecropin had been found by Boman et al.[11-13], In comparison of both peptides, lepidopteran A was very similar to cecropin B in the whole sequence except five amino acid residues and C-terminal part. However, characteristic feature in lepidopteran molecule is the presence of Hyl while cecropin was composed of only usual amino acids. Recently, another antibacterial peptide, sarcotoxin, was found in larvae of flesh fly[14], although this sequence is quite different from that of lepidopteran.

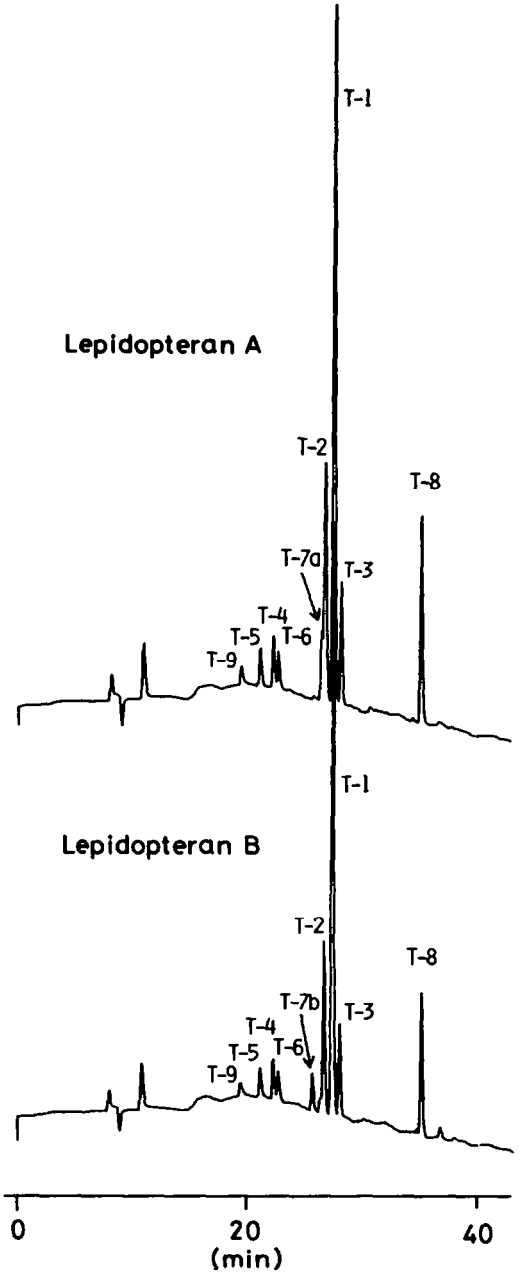

Fig. 3. HPLC patterns of tryptic peptides of lepidopteran A and B. colomn: Nucleosil 300-7C_{18}, 6 x 250 mm, eluant: A) 1% CH_3CN in 0.1% TFA, B) 50% CH_3CN in 0.1% TFA, gradient A to B (49 min), flow rate: 0.7 ml/min, detection: 220 nm.

From these findings it may be suggested that an immune system of insect though of different species has very similar mechanism which produces the comparable peptide with defensive activity against an attack of microbial invaders.

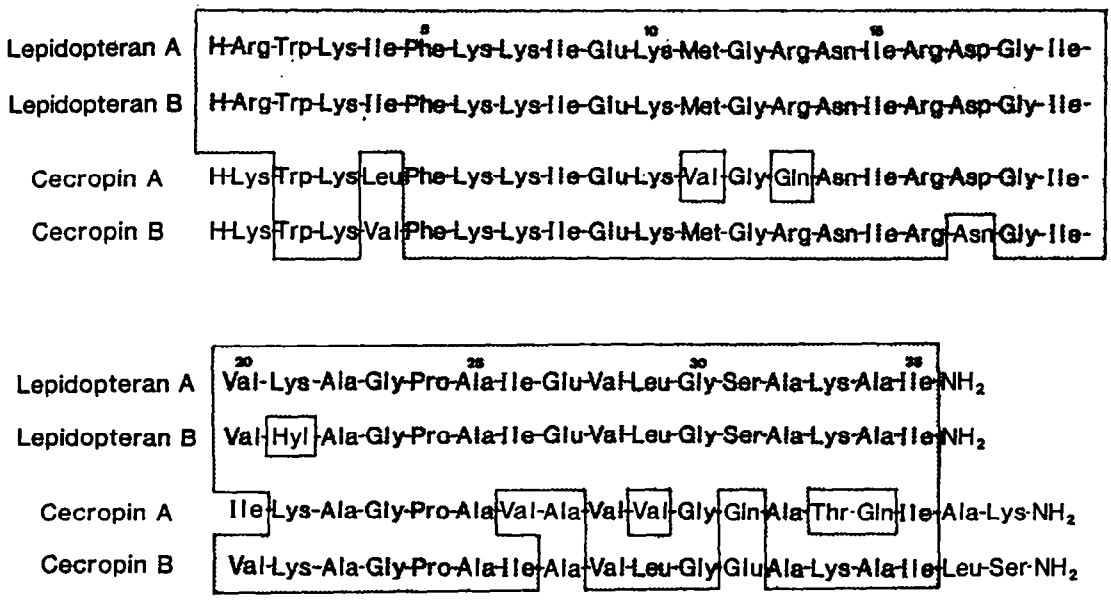

Fig. 4. Amino acid sequences of lepidopterans and cecropins. Tints show the same residues as that of lipidopteran A.

Acknowledgement

We would like to thank Prof. Tokuji Ikenaka and Dr. Saburo Hara, Faculty of Science, Osaka University, for use of the amino acid analyzer, and Dr. Shoichi Kusumoto, Faculty of Science, Osaka University, for giving us the authentic sample.

References

1) This work was presented at the 21st Symposium on Peptide Chemistry, Tsukuba, October 1983, T. Shiba, Y. Ueki, I. Kubota, T. Teshima, Y. Sugiyama, Y. Oba, and M. Kikuchi, Peptide Chemistry 1983, 209 (1984), and also at Akabori Conference, 1st German-Japanese Symposium on Peptide Chemistry, Munich, June 1985.
2) The results concerning isolation as well as characterization of lepidopteran will be reported soon elsewhere.
3) Abbreviations according to IUPAC-IUB commision, Eur. J. Biochem., 138, 9 (1984), are used. Hyl: δ-hydroxylysine, Boc: t-butoxycarbonyl, HOSu: N-hydroxysuccinimide, DMF: N,N-dimethylformamide, THF: tetrahydrofuran, DMSO: dimethylsulfoxide, TFA: trifluoroacetic acid, HPLC: high performance liquid chromatography.
4) P. Edman and A. Henschen, in "Protein Sequence Determination" (S. B. Needleman, ed), p 232. Springer-Verlag, Berlin and New York, 1975.
5) C. L. Zimmerman, E. Appella, J. J. Pisano, Anal. Biochem., 77, 569 (1977).
6) W. R. Gray, Methods in Enzymology, 25, 333 (1972).
7) S. S. Wang, F. H. Carpenter, J. Biol. Chem., 240, 1619 (1965).
8) K. Morihara, T. Oka, and H. Tsuzuki, Arch. Biochem. Biophys., 135, 311 (1969).
9) H-Ala-Ile-OH was the gift from Dr. S. Kusumoto, Faculty of Science, Osaka University.
10) The authentic sample was synthesized as a fragment peptide in the synthetic study of lepidopteran A, which will be reported elsewhere.
11) H. Steiner, D. Hultmark, Å. Engström, H. Bennich, and H. G. Boman, Nature, 292, 246 (1981).
12) D. Hultmark, Å. Engström, H. Bennich, R. Kapur, and H. G. Boman, Eur. J. Biochem., 127, 207 (1982).
13) X. Qu, H. Steiner, Å. Engström, H. Bennich, and H. G. Boman, Eur. J. Biochem., 127, 219 (1982).
14) M. Okada, S. Natori, J. Biol. Chem., 260, 7174 (1985).