Tetrahedron Vol. 43, No. 20, pp. 4513 to 4518, 1987
Printed in Great Britain.

0040-4020/87 $3.00+.00

STRUCTURE DETERMINATION OF LEPIDOPTERAN C, SELF-DEFENCE SUBSTANCE PRODUCED BY SILKWORM *(BOMBYX MORI)*

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

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

(*Received in Japan* 2 *July* 1987)

Lepidopteran C is a minor component in a group of lepidopterans which are self-defence substances produced in haemolymph of silkworm. Its structure was determined by the Edman degradation of the whole molecule as well as the fragments obtained by digestion of *Staphylococcus aureus* V8 protease, proline specific endopeptidase, and trypsin. Although lepidopteran C is similar to A in amino acid sequence, 11 amino acids of A are substituted in the sequence of C.

Insects do not possess the immune system based on antigen-antibody reaction which is an important self-defence mechanism in vertebrates. When silkworm was vaccinated with killed *E. coli*, the antibacterial peptide called lepidopteran was produced in body fluid. It seems to play an important role in a self-defence mechanism of insects. Lepidopteran is a group of basic peptides, A, B, and C which are all composed of 35 amino acids.[1,2] The structures A and B were already determined[1,2] and the total synthesis of A was also achieved which established the structure of A unequivocally.[3]

Lepidopteran C was isolated as a minor congener accompanied by A and B. The antibacterial activities of three congeners were measured. Since lepidopteran C showed almost the same antibacterial spectrum as those of A and B as shown in Table 2, C was considered to play the same role as A and B in the self-defence system of insects. High performance liquid chromatography (HPLC) profile of a mixture of lepidopteran A, B, and C, and amino acid analysis of C (Table 1) were reported in the previous paper.[2] The sequence of lepidopteran C was determined by the Edman degradation for the whole molecule and its fragments obtained by digestion of *Staphylococcus aureus* V8 protease, proline specific endopeptidase, and trypsin. The C-terminal structure was determined to be an amide form by comparison of C-terminal dipeptide fragment obtained from the natural substance with the synthetic H-Ser-Leu-OH and H-Ser-Leu-NH_2. Consequently, the structure of lepidopteran C was elucidated to be the structure as shown in Fig. 2 which contained one δ-hydroxylysine (Hyl) residue at 21st position as the same as that of lepidopteran B.

Experimental

All melting points are uncorrected. Amino acid analysis was carried out by Hitachi KLA-5 or Hitachi 835 amino acid analyzer. NMR spectrum was measured with JEOL JNM-PMX60 spectrometer, and optical rotation was obtained by a Perkin- Elmer 141 polarimeter. Fast atom bombardment mass spectrum (FAB-MS) was obtained with the Matsuda type mass spectrometer of Osaka University.

Separation of lepidopteran congeners Lepidopteran C was separated by HPLC using the same procedure as reported.[2]

Determination of peptide sequence The Edman degradation was performed by the method of G. E. Tarr,[4] and the phenylthiohydantoin (PTH) derivative of Hyl was prepared by the procedure of P. Edman and A. Henschen.[5] Dansyl-Edman method was performed by the procedure of W. R. Gray.[6] PTH amino acid was identified by HPLC (Nucleosil $5C_{18}$, 4 x 250 mm, 42 % acetonitrile - 0.01 M sodium acetate buffer, pH 4.5).[7]

Staphylococcus aureus V8 protease digestion of lepidopteran C The solution of lepidopteran C (70 nmol) and *Staphylococcus aureus* V8 protease (3.5 nmol) in 0.05 M phosphate buffer (pH 7.80) (400 µl) was incubated at 37°C for 20 h. The reaction mixture was separated by HPLC (Nucleosil 300-$7C_{18}$, 4 x 250 mm, 1 % - 50 % acetonitrile - 0.1 % trifluoroacetic acid (TFA)) to give three peptide fragments SP-1, SP-2, and SP-3. Yield, SP-1 (43 nmol, 62 %), SP-2 (45 nmol, 65 %), SP-3 (43 nmol, 62 %).

Proline specific endopeptidase treatment of SP-3 Peptide fragment SP-3 (40 nmol) was digested with proline specific endopeptidase (0.013 nmol) in 0.2 M ammonium acetate buffer (pH 8.0) (160 µl) at 25°C for 100 min. The reaction mixture was separated by HPLC (Nucleosil 300-$7C_{18}$, 4 x 250 mm, 1 % - 50 % acetonitrile - 0.1 % TFA) to give peptide fragments, SP-3-P-1 and SP-3-P-2. Yield, SP-3-P-1 (31 nmol, 79 %), SP-3-P-2 (18 nmol, 45 %).

Trypsin digestion of lepidopteran C To a solution of lepidopteran C (105 nmol) in 0.2 M ammonium acetate buffer (pH 8.0) (70 µl) was added trypsin treated with tosylphenylalanine chloromethyl ketone[7] (0.10 nmol), 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). Further separations of T-5 from T-6 and T-3 from T-7 were carried out by HPLC(Cosmosil $5C_{18}$, 4 x 250 mm, 6% acetonitrile - 0.1 % TFA, and Nucleosil 300-$7C_{18}$, 4 x 250 mm, 10 % acetonitrile - 0.1 % TFA, respectively). Yield; T-1 (55 nmol, 52%), T-2 (23 nmol, 22 %), T-3 (22 nmol, 21 %), T-4 (78 nmol, 74 %), T-5 (20 nmol, 19 %),T-6 (19 nmol, 18 %), T-7 (22 nmol, 21 %), T-8 (49 nmol, 47 %), T-9 (67 nmol, 64 %).

Table 1. Amino Acid Compositions of Lepidopteran C and Peptide Fragments Obtained by Enzymatic Cleavage.

	Lepidopteran C[a)]	SP-1[b)]	SP-2	SP-3[a)]	SP-3-P-1	SP-3-P-2
Asp	1.99(2)	-	2.08(2)	-	-	-
Ser	0.88(1)	-	-	0.91(1)	-	0.90(1)
Glu	2.11(2)	0.78(1)	-	1.11(1)	-	1.02(1)
Pro	1.04(1)	-	-	0.98(1)	0.95(1)	-
Gly	4.00(4)	-	1.00(1)	3.00(3)	2.00(2)	1.00(1)
Ala	4.06(4)	-	-	4.03(4)	1.07(1)	2.77(3)
Val	2.55(3)	-	1.49(2)	0.74(1)	-	0.72(1)
Ile	3.43(4)	1.00(1)	-	2.69(3)	0.82(1)	1.57(2)
Leu	2.90(3)	1.12(1)	-	2.01(2)	0.93(1)	0.95(1)
Phe	0.91(1)	1.01(1)	-	-	-	-
Hyl	0.95(1)	-	-	1.12(1)	0.96(1)	-
Lys	4.81(5)	2.99(3)	0.92(1)	1.11(1)	-	0.91(1)
Trp	1.21(1)[b)]	1.12(1)	-	-	-	-
Arg	2.60(3)	1.03(1)	1.65(2)	-	-	-

Hydrolysis conditions (in a sealed tube):
6M HCl at 110°C for 24 h,
a) 6M HCl at 110°C for 90 h,
b) 6M HCl containing 4% thioglycolic acid at 110°C for 24 h.

Syntheses of H-Ser-Leu-OH and H-Ser-Leu-NH_2•HCl

Boc-Ser(O^tBu)-Leu-OEt To a solution of Boc-Ser(O^tBu)-OH•DCHA (884 mg, 2.0 mmol) and H-Leu-OEt•HCl (391 mg, 2.0 mmol) in N,N-dimethylformamide (8 ml) were added 1-hydroxybenzotriazole (405 mg, 3.0 mmol) and dicyclohexylcarbodiimide (495 mg, 2.4 mmol) at 0°C and then the mixture was stirred at room temperature for 3 d. After the insoluble material was filtered off, the filtrate was concentrated *in vacuo*. The solution of the residue in ethyl acetate was washed with sat. $NaHCO_3$, 10 % citric acid and brine. The organic layer was dried over $MgSO_4$ and concentrated *in vacuo*. The residue was purified by silica-gel column chromatography using benzene - ethyl acetate as a developing solvent (9 : 1). Yield 548 mg (68 %). For elemental analysis the product was recrystallized from hexane. mp 80.5 - 81.5°C, $[\alpha]_D^{29.5}$ +22.0°(c 1.04, DMF). Found: C, 59.87; H, 9.56; N, 6.86 %, Calcd for $C_{20}H_{38}N_2O_6$: C, 59.68; H, 9.52; N, 6.96 %.

H-Ser-Leu-OH To a solution of Boc-Ser(O^tBu)-Leu-OEt (100 mg, 0.25 mmol) in 95 % EtOH (1 ml) was added 1 M NaOH (300 µl, 0.30 mmol) and the solution was stirred for 2.5 h. The reaction mixture was concentrated *in vacuo* and the residue was dissolved in water. The solution was washed with ethyl acetate and acidified with 1 M HCl. The aqueous layer was extracted with ethyl acetate and the organic layer was washed with water. The organic solution was dried over $MgSO_4$ and concentrated *in vacuo* to give oily residue. The residue was dissolved in TFA (680 µl, 8.8 mmol) and stirred for 2 h at room temperature. The reaction mixture was concentrated *in vacuo* and the solution of the residue was lyophilized. The oily residue was triturated with ether to give white powder. Yield, 47 mg (57 %). mp 166.5-167.5°C, $[\alpha]_D^{17}$ -16.9° (c 1.08, 5% AcOH). Found: C, 49.36; H, 8.27; N, 12.80 %. Calcd for $C_9H_{18}N_2O_4$: C, 49.53; H, 8.31; N, 12.84 %.

H-Ser-Leu-NH_2•HCl A solution of Boc-Ser(O^tBu)-Leu-OEt (435 mg, 1.1 mmol) in methanol (10 ml) was saturated with ammonia at 0°C in a pressure bottle and kept at room temperature for 40 h. The solution was concentrated *in vacuo* to give an oily product. The product thus obtained was dissolved in TFA and stirred for 2 h at room temperature. The residue was dissolved in 4 M HCl in tetrahydrofuran and the solution was concentrated *in vacuo*. The residue in water was lyophilized to give an oily residue. Yield 220 mg (79 %). NMR (δ) (DMSO-d_6): 8.6 (d, NH of Leu), 8.3 (broad s, NH_3^+ of Ser), 7.4 (broad s, $CONH_2$ of Leu), 7.1 (broad s, $CONH_2$ of Leu).

Identification of T-9 with the authentic sample The retention time of T-9 on HPLC (Cosmosil $5C_{18}$, 4 x 125 mm, 4 % acetonitrile - 0.1 % TFA) was identical with that of synthetic H-Ser-Leu-NH_2 but different from that of H-Ser-Leu-OH. The mixture of T-9 and H-Ser-Leu-NH_2 (1:1) showed one peak on HPLC. Retention time: T-9, 7.2 min; H-Ser-Leu-NH_2, 7.2 min; mixture of T-9 and H-Ser-Leu-NH_2, 7.2 min; H-Ser-Leu-OH, 13.3 min.

Table 1. (continued).

	T-1[b]	T-2	T-3	T-4	T-5	T-6	T-7	T-8[a]	T-9
Asp	-	-	-	-	-	1.00(1)	1.00(1)	-	-
Ser	-	-	-	-	-	-	-	-	1.00(1)
Glu	-	-	-	1.03(1)	-	-	-	1.07(1)	-
Pro	-	-	-	-	-	-	-	0.94(1)	-
Gly	-	-	-	-	1.00(1)	-	1.00(1)	2.00(2)	-
Ala	-	-	-	-	-	-	-	3.88(4)	-
Val	-	-	-	-	0.92(1)	0.90(1)	-	0.78(1)	-
Ile	-	-	-	0.84(1)	-	-	0.71(1)	1.65(2)	-
Leu	-	-	1.00(1)	-	-	-	0.96(1)	-	1.15(1)
Phe	-	-	0.91(1)	-	-	-	-	-	-
Hyl	-	-	-	-	-	-	1.06(1)	-	-
Lys	1.00(1)	1.00(1)	1.07(1)	2.00(2)	-	-	-	0.98(1)	-
Trp	0.77(1)	1.10(1)	-	-	-	-	-	-	-
Arg	0.72(1)	-	-	-	1.20(1)	0.86(1)	-	-	-

Results and Discussion

Although the amino acid composition of lepidopteran C was similar to those of A and B, its difference from A and B was recognized as follows: One Hyl residue was found in lepidopteran C. Met was not detected and the ratio of Ile to Leu (4:3) in lepidopteran C was different from that in A or B (6:1). Molecular weight of lepidopteran C was determined to be 3875±1 (exact mass) from FAB-MS which was in agreement with a value calculated from amino acid residues in Table 2. (Found: M+H, 3876±1. Calcd M+H (Asp(1), Asn(1), Glu(1), Gln(1), amide form at C terminus): 3875.3)

A series of the Edman degradation of the whole molecule of lepidopteran C was first performed to clarify the sequence from the N-terminus to 12th amino acid. Since the 9th amino acid was found to be Glu, lepidopteran was then digested with *Staphylococcus aureus* V8 protease. Three peptides SP-1, SP-2, and SP-3 were obtained. Since SP-3 seemed to be a relatively large peptide and contained one Pro residue, it was further digested with proline specific endopeptidase to give SP-3-P-1 and SP-3-P-2. All sequences of four peptides, SP-1, SP-2, SP-3-P-1, and SP-3-P-2, were determined by the Edman degradation method except for C-terminal part of SP-3-P-2. SP-1 and SP-2 were found to correspond to 1-9 and 10-17 respectively. At 10th cycle of the Edman degradation for SP-3-P-2, dansyl method was used to detect Ser residue, although the C-terminal amino acid at 11th cycle could not be detected. Finally, judging from the specificity of proteases, the sequence of these fragment peptides was presumed to be (SP-1)-(SP-2)-(SP-3-P-1)-(SP-3-P-2). However, it was still uncertain if the C-terminal carboxylic acid was amidated as in lepidopteran A and B.

In order to confirm the sequence of these fragment peptides and to clarify the C-terminal structure, trypsin digestion of lepidopteran C was further applied for the whole molecule. In this experiment, nine fragments were obtained, sequences of which were determined by the Edman or dansyl-Edman method. A fragment T-9 was elucidated to be H-Ser-Leu-NH_2 by comparison with the authentic sample as described in experimental part. Sequences in fragments T-4, T-7, and T-8 were overlapped with those in SP-1, SP-2, SP-3-P-1 and SP-3-P-2 as shown in Fig. 2, and C-terminus was confirmed to be amidated. Thus the whole sequence of lepidopteran C was unequivocally clarified.

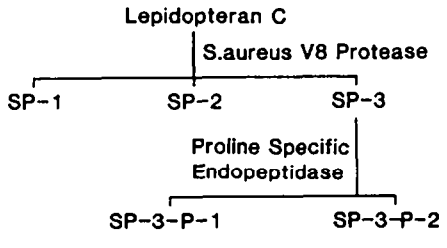

Fig. 1.

Table 2. Antibacterial Activities of Lepidopteran A, B, and C (MIC, μg/ml)

Test organism	Lepidopteran A	B	C
Proteus mirabilis 129	16	16	16
Proteus mirabilis 1287	16	16	16
Escherichia coli NIHJ	4	4	4
Escherichia coli NIHJ JC-2	2	2	2
Salmonella enteritidis IFO 3313	4	8	8
Shigella sonnei EW-33	4	4	4
Pseudomonas aeruginosa	16	8	8
Pseudomonas aeruginosa 99	8	8	8
Enterobacter aerogenes IFO 13534	8	8	8
Staphylococcus aureus FDA 209P	250	250	250
Bacillus subtilis PCI 219	16	16	16
Streptococcus hemolyticus	250	250	125

The amino acid sequences of lepidopteran A, B, and C were compared in Fig. 3. Eleven amino acids of A were substituted in C. However, in the sequence of 10 amino acids from N-terminus of C, only one amino acid residue, i.e., Leu^4, was different from Ile^4 in A, while the three congeners showed almost the same antibacterial activities. (Table 2) Therefore, N-terminal part of lepidopteran may play a major role responsible for common antibacterial activity.

Boman et. al., found similar self-defence peptides called cecropin from cecropia moss and proposed the structures.[9-11] However, later on, they revised the structure of cecropin B to be the peptide composed of 35 amino acids as shown in Fig. 3 by results of cDNA which coincides to the length of lepidopteran.[12] Another cecropin, CM_{IV} was also isolated from *Bombyx mori* by Chinese group.[13] Its structure was again similar to that of lepidopteran although seven amino acids of cecropin CM_{IV} were different from lepidopteran A. On the other hand, lepidopteran is the only peptide containing Hyl residue in such self-defence peptides whose structures were clarified so far, although the contribution of hydroxyl group of this unusual amino acid residue to antibacterial activity was not understood. The content of common amino acid residues in all congeners of lepidopterans and cecropins was 80% in N-terminal decapeptide part whereas only 30% in C-terminal decapeptide part as shown in Fig. 3. This fact may indicate that the structure of N-terminal part was much more required for exhibition of antibacterial activity than that of C-terminal as mentioned in the previous paragraph. Evolutional process may be elucidated after more structural features of such insect peptides other than lepidopteran or cecropin were clarified. On the other hand, another antibacterial peptide, sarcotoxin, was isolated from flesh fly[14]. This may belong to a different peptide group of the similar biological significance from lepidopteran in view of its structure.

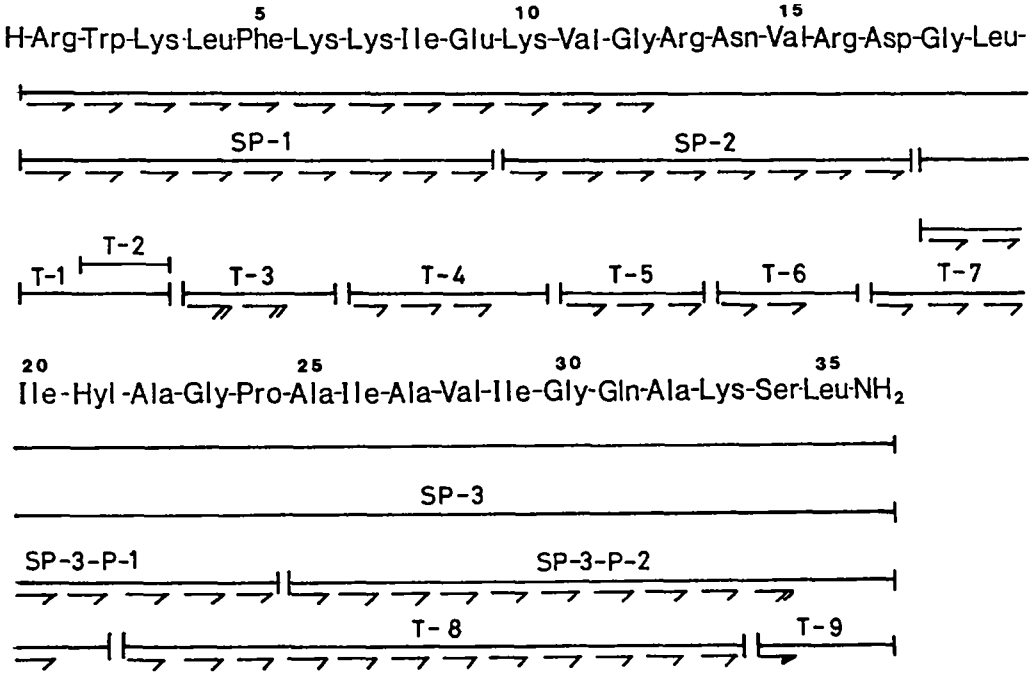

Fig. 2. Amino acid sequence of lepidopteran C. —→: Edman degradation, —//→: dansyl Edman method, —▼:dansyl method.

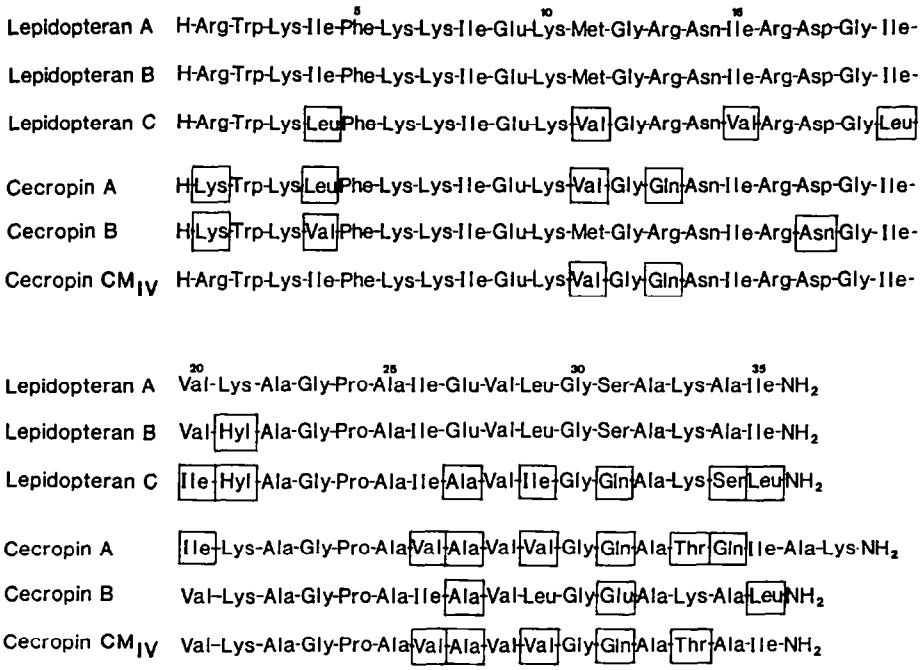

Fig. 3. Amino acid sequences of lepidopterans and cecropins.

Acknowledgment

We would like to thank Prof. Tokuji Ikenaka and Dr. Saburo Hara, Osaka University, for use of an amino acid analyzer, and Dr. Takekiyo Matsuo and Dr. Itsuo Katakuse, Osaka University, for measurement of FAB-MS. We are also grateful to Dr. Mikio Kikuchi, Research Laboratory of Fundamental Biological Science, for measurement of antibacterial activity.

References

1) T. Shiba, Y. Ueki, I. Kubota, T. Teshima, Y. Sugiyama, Y. Oba, and M. Kikuchi, Peptide Chemistry 1983, 209 (1984).
2) T. Teshima, Y. Ueki, T. Nakai, T. Shiba, and M. Kikuchi, Tetrahedron, 42, 829 (1986).
3) T. Nakai, M. Kitazawa, T. Teshima, and T. Shiba, Peptide Chemistry 1985, 207 (1986).
4) G. E. Tarr, Method in Enzymology, 47, 335 (1977).
5) P. Edman and Henschen, in "Protein Sequence Determination" (S. B. Needleman, ed), p 232. Springer-Verlag, Berlin and New York, 1975.
6) W. R. Gray, Method in Enzymology, 25, 333 (1972).
7) C. L. Zimmerman, E. Appella, J. J. Pisano, Anal. Biochem., 77, 569 (1977).
8) S. S. Wang, F. H. Carpenter, J. Biol. Chem., 240, 1619 (1965).
9) H. Steiner, D. Hultmark, Å. Engström, H. Bennich, and H. G. Boman, Nature, 292, 246 (1981).
10) D. Hultmark, Å. Engström, H. Bennich, R. Kapur, and H. G. Boman, Eur. J. Biochem., 127, 207 (1982).
11) X. Qu, H. Steiner, Å. Engström, H. Bennich, and H. G. Boman, Eur. J. Biochem., 127, 219 (1982).
12) P. van Hofsten, I. Faye, K. Kockum, J. -Y. Lee, K. G. Xanthopoulos, I. A. Boman, H. G. Boman, Å. Engström, D. Andreu, and R. B. Merrifield, Proc. Natl. Acad. Sci. USA, 82, 2240 (1985).
13) S. -Q. Zhang, J. -K. Fang, Y. -H. Tang, X. -M. Qu, and C. -W. Chi, Peptide Seminar, Shanghai, Nov. 20, 1985, Shanghai.
14) M. Okada, and S. Natori, J. Biol. Chem., 260, 7174 (1985).