Biochem. J. (1986) 239, 717–722 (Printed in Great Britain)

Molecular cloning of a cDNA and assignment of the *C*-terminal of sarcotoxin IA, a potent antibacterial protein of *Sarcophaga peregrina*

Noriko MATSUMOTO, Masayuki OKADA, Haruo TAKAHASHI, Qu Xian MING,* Yuki NAKAJIMA, Yoshinobu NAKANISHI, Hiroto KOMANO and Shunji NATORI†
Faculty of Pharmaceutical Sciences, University of Tokyo, Bunkyo-ku, Tokyo 113, Japan

A previous paper described the complete amino acid sequences of sarcotoxins IA, IB and IC, which are a group of potent antibacterial proteins with almost identical primary structures produced by *Sarcophaga peregrina* (fleshfly) larvae [Okada & Natori (1985) J. Biol. Chem. **260**, 7174–7177]. The present paper describes the cDNA cloning and complete nucleotide sequencing of a cDNA clone for sarcotoxin IA. The *C*-terminal amino acid residue of sarcotoxin IA deduced from the nucleotide sequence was glycine, whereas it was found to be arginine by amino acid sequencing of purified sarcotoxin IA. Analysis of the elution profiles on h.p.l.c. of the synthetic derivatives of sarcotoxin IA showed that the *C*-terminal amino acid residue of authentic sarcotoxin IA is amidated arginine, which is probably produced by enzymic cleavage of terminal glycine.

INTRODUCTION

In *Sarcophaga peregrina* (fleshfly) humoral antibacterial proteins are induced when the body wall of third-instar larvae or adult flies is injured with a hypodermic needle (Natori, 1977). We have purified a group of these antibacterial proteins named sarcotoxin I from the haemolymph of injured larvae (Okada & Natori, 1983, 1985*a*). Sarcotoxin I is a mixture of three potent bactericidal proteins, named sarcotoxins IA, IB and IC, with almost identical primary structures (Okada & Natori, 1985*a*). We purified sarcotoxins IA, IB and IC to homogeneity, and determined their amino acid sequences (Okada & Natori, 1985*a*). They each consisted of 39 amino acid residues, and differed only in two or three amino acid residues.

The primary target of sarcotoxin I was found to be the cytoplasmic membrane of bacteria (Okada & Natori, 1984, 1985*b*). On treatment with sarcotoxin I, the membrane potential of *Escherichia coli* disappeared almost instantaneously, resulting in loss of active transport of amino acids such as proline and of ATP generation. Since sarcotoxin I is induced in the haemolymph in response to injury of the body wall, its function is thought to be to prevent infection by bacteria entering the body through the damaged body wall. In some insects, antibacterial activity is induced in the haemolymph by injection of live or dead bacteria (Whitcomb *et al.*, 1974; Boman, 1981; Chadwick *et al.*, 1982). Thus such activity may participate in the defence mechanism of insects that have no immune network.

In order to understand the mechanism of the induction of sarcotoxin I in response to injury of the body wall, it is necessary to identify the gene for this protein. Concerning antibacterial proteins of insects, only the cDNAs for cecropins, attacins and lysozymes of *Hyalophora cecropia* (cecropia moth) have so far been cloned (van Hofsten *et al.*, 1985; Kockum *et al.*, 1984; Boman *et al.*, 1985).

The present paper describes the cloning and sequencing of the gene for sarcotoxin IA of *Sarcophaga peregrina*. Although the *C*-terminal amino acid residue of sarcotoxin IA was determined to be arginine, we found a codon for glycine next to one for arginine and then a termination codon in the cDNA for sarcotoxin IA. From comparison of authentic sarcotoxin IA with its synthetic derivatives, we concluded that arginine at the *C*-terminal is amidated in authentic sarcotoxin IA.

MATERIALS AND METHODS

Sarcotoxin I

Sarcotoxin I was purified from the haemolymph of third-instar larvae of *Sarcophaga peregrina* as described previously (Okada & Natori, 1983). Each larva was pricked with a hypodermic needle to induce sarcotoxin I, and the haemolymph was collected 24 h later. About 50 μg of sarcotoxin I was routinely obtained from 15000 larvae. Sarcotoxins IA, IB and IC were separated by h.p.l.c. as described previously (Okada & Natori, 1985*a*).

Chemically synthesized sarcotoxin IA and its derivatives

Since sarcotoxin IA consists of 39 amino acid residues and its molecular mass is thus less than 4 kDa, it was possible to synthesize this protein and its derivatives chemically in a fully automated peptide synthesizer (Applied Biosystems 430A). Synthetic sarcotoxin IA and its derivatives with different *C*-termini were supplied by Wakunaga Pharmaceutical Co. (Hiroshima, Japan). The primary structure determined by direct protein sequencing of sarcotoxin IA is (in the one-letter notation)

GWLKKIGKKIERVGQHTRDATIQG LGIAQQAANVAATAR

Abbreviation used: SSC, standard saline citrate.

* Present address, Shanghai Institute of Biochemistry, Academia Sinica, 320 Yue-Yang Road, Shanghai, China.
† To whom correspondence should be addressed.

and its *C*-terminal is arginine. The two derivatives used had the same sequence as sarcotoxin IA except that the *C*-terminal arginine residue was amidated in one (sarcotoxin IA-NH_2) and glycine was additionally linked to the arginine residue in the other (sarcotoxin IA-Gly).

Analysis of sarcotoxin IA and its derivatives by h.p.l.c.

Sarcotoxin IA and its derivatives were dissolved in 10 mM-sodium phosphate buffer, pH 6.0, at an appropriate concentration. A 100 μl sample of each was applied to a reverse-phase h.p.l.c. column of Synchropak RP-P (C_{18}) (250 mm × 4.1 mm) connected to a Gilson h.p.l.c. system, and eluted with a linear gradient of 15–40% (v/v) solution B [0.05% (v/v) trifluoroacetic acid in acetonitrile] in solution A [0.05% (v/v) trifluoroacetic acid in water] at a flow rate of 2 ml/min. The absorbance at 280 nm was monitored.

Synthetic oligodeoxyribonucleotides used as probes

Oligodeoxyribonucleotides were designed to fit a partial amino acid sequence of sarcotoxin I, Lys-Lys-Ile-Glu-Arg, corresponding to positions 8–12 from the *N*-terminal. Since there are six possible codons for arginine, two series of mixtures of 24 tetradecamers,

```
        A
C G C T C G A T C T T C T T
    T   T   T   T
```

and

```
        A
C T C T C G A T C T T C T T
    T   T   T   T
```

from the 5′- to the 3′-end, were synthesized by a modified triester method (Hirose *et al.*, 1978). These oligodeoxyribonucleotides represent all possible complementary sequences corresponding to the above peptide except for the third nucleotide residues of the arginine codon.

Cloning procedure

Third-instar larvae of *Sarcophaga peregrina* were each given an injection of 2×10^5 sheep erythrocytes to induce sarcotoxin I efficiently. The fat-body was collected 6 h later, and polyadenylated RNA was extracted. A cDNA library was constructed by the method of Okayama & Berg (1982), by using 10 μg of polyadenylated RNA and 3 μg of vector/primer DNA. *Escherichia coli* HB101 transformed by the method of Morrison (1979) was selected for ampicillin-resistance as described by Hanahan & Meselson (1980) and screened by hybridization with mixtures of 24 synthetic oligodeoxyribonucleotides at 36 °C.

Other analytical procedures

Synthetic oligodeoxyribonucleotides were labelled as described by Kakidani *et al.* (1982). Nick translation with [α-^{32}P]dCTP was performed by the method of Weinstock *et al.* (1978). DNA sequences were determined by the method of Maxam & Gilbert (1980) and of Sanger *et al.* (1977). RNA blot hybridization was performed in 50% (v/v) formamide/5 × SSC (1 × SSC = 0.15 M-NaCl/0.015 M-sodium citrate buffer, pH 7.4)/5 × Denhardt's solution [1 × Denhart's solution = 0.02% (w/v)/Ficoll-400/0.02% bovine serum albumin/0.02% polyvinylpyrrolidone-40]/sonicated denatured salmon sperm DNA solution (50 μg/ml) for 18 h at 42 °C, and blots were washed successively by incubation in 2 × SSC for 10 min at room temperature and in 0.1 × SSC for 30 min at 42 °C (Thomas, 1980).

RESULTS

Isolation and characterization of a cDNA clone for sarcotoxin IA

Since the complete amino acid sequences of sarcotoxins IA, IB and IC have been determined, we designed and synthesized oligodeoxyribonucleotides corresponding to Lys-Lys-Ile-Glu-Arg, which is a common sequence for all three proteins. Using these synthetic oligodeoxyribonucleotides as probes, we isolated 20 candidates for hybridization-positive clones by screening about 50000 transformants derived from a cDNA library for fat-body polyadenylated RNA. For enrichment of mRNA for sarcotoxin I, 2×10^5 sheep erythrocytes were injected into each larva and the fat-body was collected 6 h later for isolation of polyadenylated RNA for construction of the cDNA library. Injection of sheep erythrocytes greatly enhanced the induction of haemolymph proteins such as *Sarcophaga* lectin and sarcotoxin I, which participate in the defence mechanism of *Sarcophaga peregrina* (Komano & Natori, 1985). Of the 20 clones, we selected one clone, pTO19, which hybridized intensively with the synthetic oligodeoxyribonucleotide probes, and determined the nucleotide sequence of the insert.

The complete nucleotide sequence and the strategy for sequencing are shown in Fig. 1. For sequencing, we first determined 264 bases from the 5′-end and 253 bases including a 101-base poly(A) sequence from the 3′-end of the insert by the method of Maxam & Gilbert (1980). Then we recloned the insert into pUC8 and determined 354 bases from the 5′-end by the dideoxy method of Sanger *et al.* (1977) to confirm the sequence, as shown in Fig. 1(*a*). The sequences determined by the two methods agreed well with each other.

As is evident from Fig. 1(*b*), clone pTO19 contained an insert consisting of 428 nucleotide residues including a 101-nucleotide 3′-end poly(A) sequence. An open reading frame of 63 amino acid residues was assigned in the insert, and this sequence was found to coincide with that of sarcotoxin IA, assuming a putative signal sequence of 23 amino acid residues. Therefore we concluded that pTO19 is a cDNA clone of sarcotoxin IA.

The nucleotide sequence of clone pTO19 showed that there is a glycine codon between the arginine codon and the termination codon, indicating that the *C*-terminal amino acid residue of sarcotoxin IA is glycine, although by amino acid sequencing of purified sarcotoxin IA we found that it was arginine (Okada & Natori, 1985*a*). A consensus sequence for a poly(A) addition signal of AATAAA (Proudfoot & Brownlee, 1974) is located nine bases upstream of the poly(A)-addition site.

Assignment of the *C*-terminal amino acid residue of sarcotoxin IA

Previously we determined the complete amino acid sequence of sarcotoxin IA, which consists of 39 amino acid residues (Okada & Natori, 1985*a*). Although this sequence coincided with the amino acid sequence deduced from the nucleotide sequence of clone pTO19, the latter contained an additional glycine codon at the

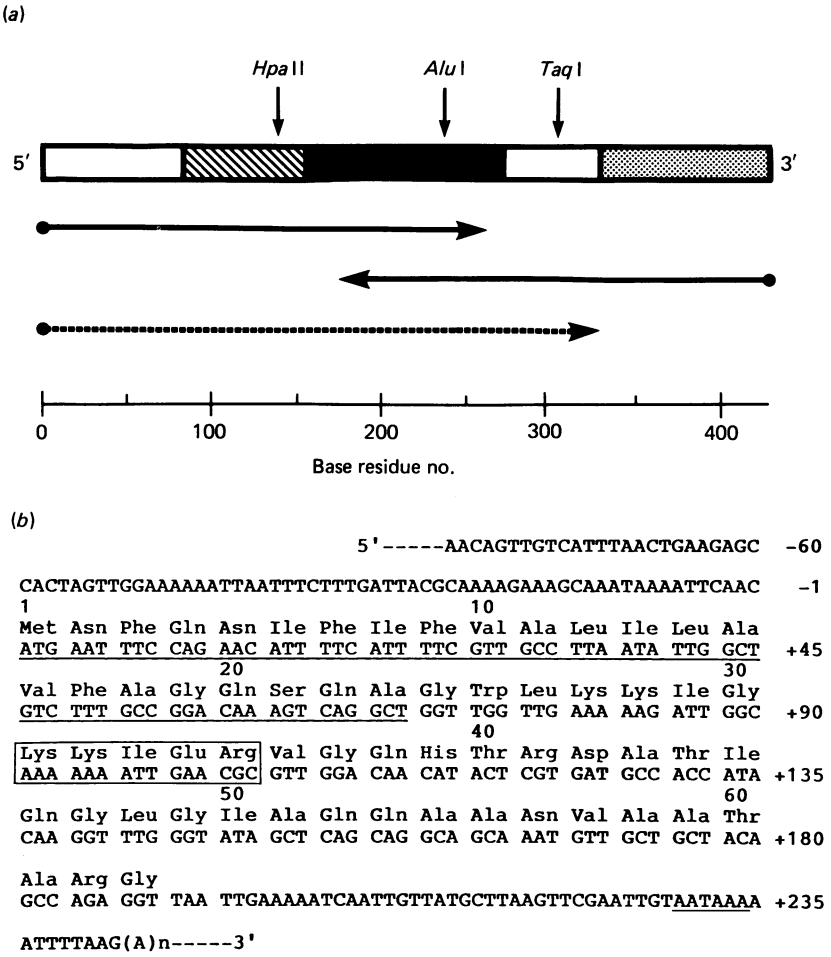


Fig. 1. (*a*) Strategy for sequencing cloned cDNA pTO19 and (*b*) nucleotide sequence of the cloned cDNA encoding sarcotoxin IA

(*a*) Strategy for sequencing cloned cDNA pTO19. The restriction map shows restriction-endonuclease sites, identified by numbers (bottom scale), starting from the 5'-terminal nucleotide generated by cleavage. The poly(dG)–poly(dC) tails are not included in the restriction map. The protein coding region is indicated by a solid black box, the putative signal peptide by a hatched box, and the poly(A) region by a stippled box. The two upper horizontal arrows show the directions of sequencing by the method of Maxam & Gilbert (1980), and the bottom horizontal arrow shows the direction for sequencing by the dideoxy method by Sanger *et al.* (1977). (*b*) Nucleotide sequence of the cloned cDNA encoding sarcotoxin IA. The amino acid sequence of sarcotoxin IA is shown above the nucleotide sequence. Numbers of nucleotides are given at the right of each line. The amino acid sequence used to synthesize oligodeoxyribonucleotides is boxed. The putative signal sequence and the poly(A)addition signal are underlined.

C-terminal of an open reading frame. Therefore, to determine the *C*-terminal amino acid residue of sarcotoxin IA, we synthesized two possible proteins, with arginine (synthetic sarcotoxin IA) and with glycine in addition to arginine (sarcotoxin IA-Gly) at their *C*-terminal, and compared their elution profiles on h.p.l.c. with that of native sarcotoxin IA. As shown in Figs. 2(*a*), 2(*b*) and 2(*d*), under the conditions used synthetic sarcotoxin IA and sarcotoxin IA-Gly were eluted in exactly the same position, but this was slightly before that of authentic sarcotoxin IA. This was confirmed by h.p.l.c. of mixtures of these proteins and authentic sarcotoxin IA (Figs. 2*g* and 2*h*). Thus the *C*-terminal of authentic sarcotoxin IA is not free unmodified arginine or unmodified glycine.

Another possibility is that authentic sarcotoxin IA has a *C*-terminal arginine residue that is amidated, since the *C*-termini of many physiologically active peptides are amidated. Therefore we synthesized the amidated form of sarcotoxin IA (sarcotoxin IA-NH_2) and compared its elution profile on h.p.l.c. with that of authentic sarcotoxin IA. As is evident from Figs. 2(*d*) and 2(*c*), authentic sarcotoxin IA and sarcotoxin IA-NH_2 were eluted in the same position, and a mixture of these two proteins could not be separated by h.p.l.c. (Fig. 2*i*). These results strongly suggest that sarcotoxin IA-NH_2 and authentic sarcotoxin IA have the same structure. Probably, the *C*-terminal glycine residue of sarcotoxin IA is cleaved off enzymically, resulting in an amidated arginine residue during the post-translational processing of sarcotoxin IA.

Identification of sarcotoxin IA mRNA and its induction by injury of the body wall

Using clone pTO19 as a probe, we identified sarcotoxin IA mRNA by Northern-blot hybridization. For this RNA was extracted from fat-body collected from third-instar larvae 6 h after the pricking of their body wall (Takahashi *et al.*, 1985), denatured with glyoxal and subjected to 1.5%-agarose-gel electro-

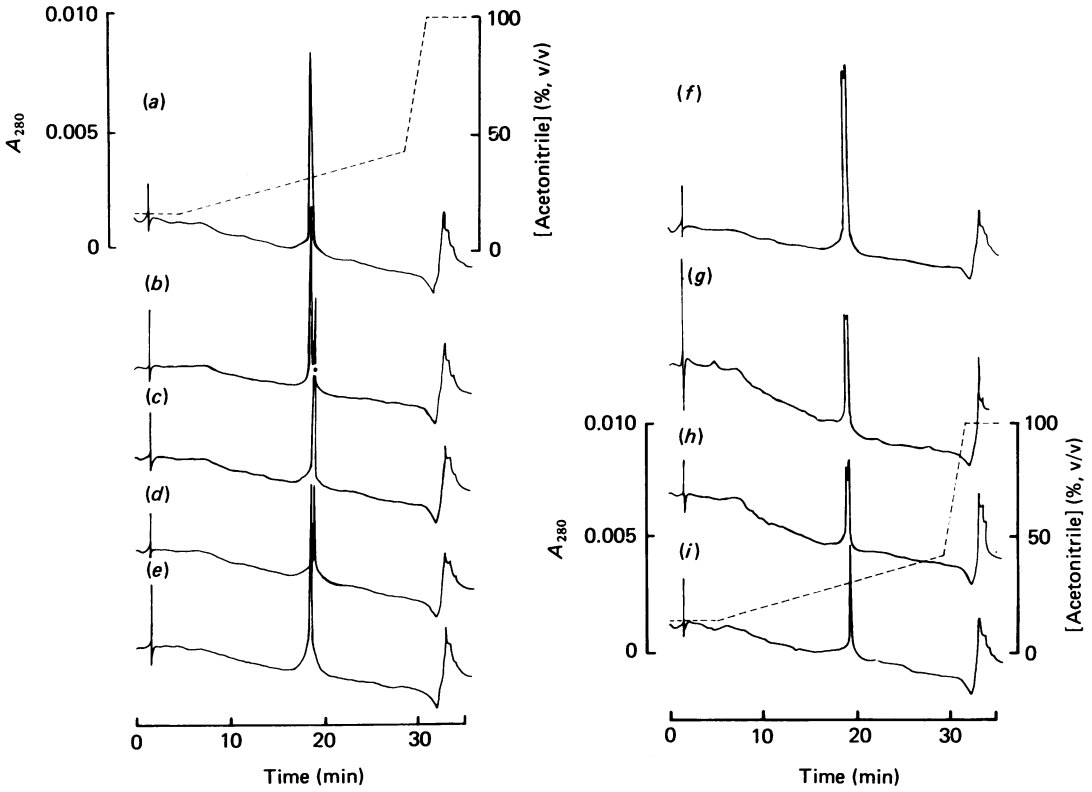
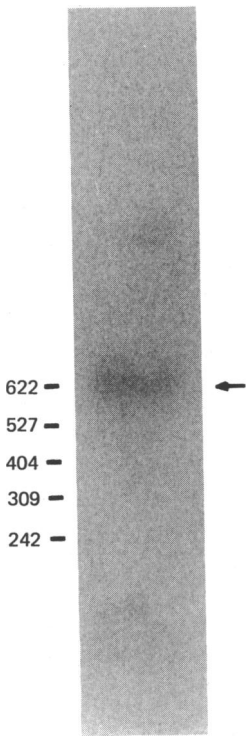


Fig. 2. Reverse-phase h.p.l.c. analysis of authentic and synthetic derivatives of sarcotoxin IA

Protein was analysed in a Gilson h.p.l.c. system. The chromatographic conditions were as follows: column, Synchropack RP-R (C_{18}) (250 mm × 4.1 mm); solution A, 0.05% trifluoroacetic acid in water; solution B, 0.05% trifluoroacetic acid in acetonitrile; gradient, a linear gradient of 15–40% (v/v) solution B in solution A; flow rate, 2 ml/min. The absorbance at 280 nm was monitored. (*a*) Synthetic sarcotoxin IA; (*b*) sarcotoxin IA-Gly; (*c*) sarcotoxin IA-NH_2; (*d*) authentic sarcotoxin IA; (*e*) a mixture of synthetic sarcotoxin IA and sarcotoxin-Gly; (*f*) a mixture of sarcotoxin IA-Gly and sarcotoxin IA-NH_2; (*g*) a mixture of sarcotoxin IA-Gly and authentic sarcotoxin IA; (*h*) a mixture of synthetic sarcotoxin IA and authentic sarcotoxin IA; (*i*) a mixture of sarcotoxin IA-NH_2 and authentic sarcotoxin IA.



phoresis (McMaster & Carmichael, 1977). After electrophoresis, RNA was transferred to a nitrocellulose paper (Thomas, 1980) and hybridized with nick-translated clone pTO19. The autoradiograph in Fig. 3 shows a single broad band of material with a molecular length of about 620 bases, which was about 200 bases larger than the size of the insert of clone pTO19. A signal observed above that band is thought to be a background, since it did not cross the lane. Probably, clone pTO19 hybridizes with mRNAs for sarcotoxins IB and IC in addition to that for sarcotoxin IA, because the base sequences of the three mRNAs are believed to be very similar. Therefore it is likely that the band detected by Northern-blot hybridization is that of a mixture of these three mRNAs rather than that of only mRNA for sarcotoxin IA.

Fig. 3. Determination of the molecular size of sarcotoxin IA mRNA

Total mRNA (65 μg) extracted from the fat-body of third-instar larvae 3 h after the pricking of their body wall was subjected to electrophoresis on 1.5% agarose gel by the method of McMaster & Carmichael (1977), blotted on to nitrocellulose paper, and hybridized with nick-translated clone pTO19. *Hpa*II-digested pBR322 was co-electrophoresed and was used as molecular-size standard. Sizes are shown in numbers of bases.

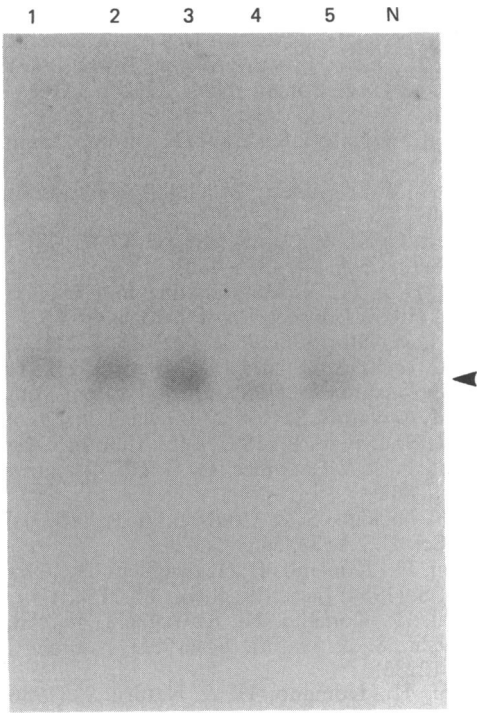


Fig. 4. Expression of the sarcotoxin IA gene after injury of the body wall of third-instar larvae

Fat-bodies were excised at the indicated times after the pricking of the body wall of larvae, their RNA was extracted and the amount of sarcotoxin IA mRNA was examined by Northern-blot hybridization. Lanes 1–5 show the profiles of RNAs prepared 1.5, 3, 6, 12 and 24 h respectively after injury of the body wall. Lane N shows RNA from normal larvae. Each lane contained 15 μg of RNA.

Next, we examined the synthesis of sarcotoxin I mRNA in the fat-body with time after injury of the body wall. For this, fat-body was collected at various times after the pricking of the body wall, and its RNA was extracted, denatured with 2.2 M-formaldehyde, and analysed by Northern-blot hybridization (Goldberg, 1980). As shown in Fig. 4, no mRNA for sarcotoxin I was detected in the fat-body of normal larvae (lane N), indicating that transcription of the sarcotoxin I gene is shut off in normal third-instar larvae. However, mRNA was detected after 1.5 h and subsequently increased in amount to a maximum after 6 h, and then it decreased gradually (lanes 1–5). These results clearly indicate that the synthesis of sarcotoxin I in response to injury of the body wall is due to activation of specific transcription of the sarcotoxin I genes, not to cleavage of precursor proteins present *in situ*.

DISCUSSION

This paper describes the isolation and characterization of a cDNA clone for sarcotoxin IA, a potent bactericidal protein produced by *Sarcophaga peregrina* larvae in response to injury of the body wall.

Analysis of a cDNA clone pTO19 showed that sarcotoxin IA consists of 40 amino acid residues, although we previously identified 39 amino acid residues by amino acid sequencing of purified sarcotoxin IA. The difference was in the *C*-terminal amino acid residue: from amino acid sequencing data, we assigned arginine as the *C*-terminal amino acid residue, but in the nucleotide sequence a glycine codon was found between the arginine codon and the termination codon. To determine the reason for this difference, we chemically synthesized three derivatives of sarcotoxin IA: a synthetic sarcotoxin IA with arginine at its *C*-terminal, and sarcotoxin IA-Gly and sarcotoxin IA-NH_2, and compared their elution profiles on h.p.l.c. with that of authentic sarcotoxin IA. Results showed that the *C*-terminal amino acid residue of authentic sarcotoxin IA is amidated arginine, not unmodified arginine. This is reasonable, since the *C*-terminal glycine residue of several other proteins is known to be cleaved off enzymically, with formation of a new amidated *C*-terminal amino acid residue (Glembotski *et al.*, 1984; Kizer *et al.*, 1984). A similar situation was found by analysis of the genes for cecropins, which are antibacterial proteins of *Hyalophora cecropia* (van Hofsten *et al.*, 1985).

In a previous paper we discussed the homology of the amino acid sequences of sarcotoxin I and cecropins (Okada & Natori, 1985*a*). Comparison of the putative signal peptides of sarcotoxin IA and cecropin B (van Hofsten *et al.*, 1985) showed extensive homology (about 70%) between the two in the first 15 (sarcotoxin IA) to 16 (cecropin B) amino acid residues, but homology in the subsequent eight (sarcotoxin IA) to ten (cecropin B) amino acid residues including the cleavage sites was much less (Fig. 5). The homologous regions of signal peptides may be important in the secretion of these antibacterial proteins.

In the present paper we demonstrated that in normal third-instar larvae the sarcotoxin IA gene in the fat-body is shut off, but that it is activated with consequent increase in mRNA with time after injury of the body wall. Probably, the genes for sarcotoxins IB and IC are activated in the same way. These results indicate that the production of sarcotoxin I in response to injury of the body wall is due to activation at the gene level. We have identified several haemolymph proteins induced under the same conditions, including *Sarcophaga* lectin, which like sarcotoxin I is thought to participate in the defence mechanism (Takahashi *et al.*, 1984). By analysis with a cDNA clone we found that induction of the *Sarcophaga* lectin is due to the activation of transcription of the *Sarcophaga* lectin gene (Takahashi *et al.*, 1986). Conceivably on injury of the body wall a common stimulus that is transmitted to fat-body cells in some way activates the genes for both sarcotoxin I and *Sarcophaga* lectin. Therefore it is likely that these genes have

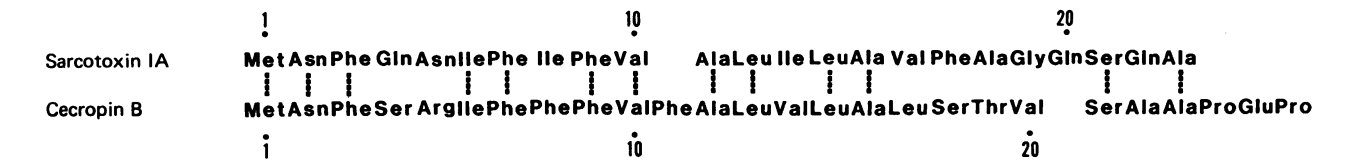


Fig. 5. Comparison of the amino acid sequences of signal peptides of sarcotoxin IA and cecropin B

common sequences or structures in their regulatory regions that respond to such a stimulus. Analyses of genomic clones for these proteins may provide a clue to this problem.

This study was supported by a Grant-in-Aid for Scientific Research from the Ministry of Education, Science and Culture of Japan, a grant from the Mitsubishi Foundation and a Grant-in-Aid from the Tokyo Biochemical Research Foundation. We thank Dr. K. Miyoshi and Dr. T. Yatsunami (Wakanaga Pharmaceutical Co.) for the syntheses of oligodeoxyribonucleotides and derivatives of sarcotoxin IA respectively.

REFERENCES

Boman, H. G. (1981) in Microbial Control of Pests and Plant Diseases (Bruges, H. G., ed.), pp. 769–784, Academic Press, New York

Boman, H. G., Faye, I., van Hofsten, P., Kockum, K., Lee, J.-Y., Xanthopoulos, K. G., Bennich, H., Engström, Å., Merrifield, R. B. & Andreu, D. (1985) Dev. Comp. Immunol. **9**, 551–558

Chadwick, J. S., Deverno, P. J., Chung, K. L. & Aston, W. P. (1982) Dev. Comp. Immunol. **6**, 433–440

Glembotski, C. C., Eipper, B. A. & Mains, R. E. (1984) J. Biol. Chem. **259**, 6385–6392

Goldberg, D. A. (1980) Proc. Natl. Acad. Sci. U.S.A. **77**, 5794–5798

Hanahan, D. & Meselson, M. (1980) Gene **10**, 63–67

Hirose, T., Crea, R. & Itakura, K. (1978) Tetrahedron Lett. **28**, 2449–2452

Kakidani, H., Furutani, Y., Takahashi, H., Noda, M., Morimoto, Y., Hirose, T., Asai, M., Nakanishi, S. & Numa, S. (1982) Nature (London) **298**, 245–249

Received 8 April 1986/22 May 1986; accepted 11 July 1986

Kizer, J. S., Busy, W. H., Jr., Cottle, C. & Youngblood, W. W. (1984) Proc. Natl. Acad. Sci. U.S.A. **81**, 3228–3232

Kockum, K., Faye, I., van Hofsten, P., Lee, J.-Y., Xanthopoulos, K. G. & Boman, H. G. (1984) EMBO J. **3**, 2071–2075

Komano, H. & Natori, S. (1985) Dev. Comp. Immunol. **9**, 31–40

Maxam, A. M. & Gilbert, W. (1980) Methods Enzymol. **65**, 499–560

McMaster, G. K. & Carmichael, G. G. (1977) Proc. Natl. Acad. Sci. U.S.A. **74**, 4835–4838

Morrison, D. A. (1979) Methods Enzymol. **68**, 326–331

Natori, S. (1977) J. Insect. Physiol. **23**, 1169–1173

Okada, M. & Natori, S: (1983) Biochem. J. **211**, 727–734

Okada, M. & Natori, S. (1984) Biochem. J. **222**, 119–124

Okada, M. & Natori, S. (1985*a*) J. Biol. Chem. **260**, 7174–7177

Okada, M. & Natori, S. (1985*b*) Biochem. J. **229**, 453–458

Okayama, H. & Berg, P. (1982) Mol. Cell. Biol. **2**, 161–170

Proudfoot, N. J. & Brownlee, G. G. (1974) Nature (London) **252**, 359–362

Sanger, F. Nicklen, S. & Coulson, A. R. (1977) Proc. Natl. Acad. Sci. U.S.A. **74**, 5463–5467

Takahashi, H., Komano, H., Kawaguchi, N., Obinata, M. & Natori, S. (1984) Insect Biochem. **14**, 713–717

Takahashi, H., Komano, H., Kawaguchi, N., Kitamura, N., Nakanishi, S. & Natori, S. (1985) J. Biol. Chem. **260**, 12228–12233

Takahashi, H., Komano, H. & Natori, S. (1986) J. Insect Physiol., in the press

Thomas, P. S. (1980) Proc. Natl. Acad. Sci. U.S.A. **77**, 5201–5205

van Hofsten, P., Faye, I., Kockum, K., Le, J.-Y., Xanthopoulos, K. G., Boman, I. A., Boman, H. G., Engström, Å., Andreu, D. & Merrifield, R. B. (1985) Proc. Natl. Acad. Sci. U.S.A. **82**, 2240–2243

Weinstock, R., Sweet, R., Weiss, M., Cedar, H. & Axel, R. (1978) Proc. Natl. Acad. Sci. U.S.A. **75**, 1299–1303

Whitcomb, R. F., Shapiro, M. & Granados, R. P. (1974) Physiol. Insecta 2nd Ed. **5**, 447–536