MOLECULAR AND CELLULAR BIOLOGY, Dec. 1990, p. 6114–6122
0270-7306/90/126114-09$02.00/0

Vol. 10, No. 12

Analysis of a Gene Cluster for Sarcotoxin II, a Group of Antibacterial Proteins of *Sarcophaga peregrina*

AKIO KANAI AND SHUNJI NATORI*

Faculty of Pharmaceutical Sciences, University of Tokyo, Bunkyo-ku, Tokyo 113, Japan

Received 19 June 1990/Accepted 22 August 1990

Sarcotoxin II is a group of antibacterial proteins of *Sarcophaga peregrina* (flesh fly) with related primary structures. We have cloned three genes in this family. These genes formed a tandem array with about 2-kb intervals, and one of them was present in the opposite strand. The putative amino acid sequences of the proteins encoded by these genes were very similar except for a deletion in one of them. All of the genes were found to be activated transiently in the same way when the larval body wall was injured, suggesting that the encoded proteins are acute-phase-responsive proteins for protecting the insect from bacterial infection.

Several antibacterial proteins have been isolated from the hemolymph of various insects (6, 7, 9, 10, 12, 16, 18, 20). These proteins are thought to protect the insects from infection by pathogens, because they are induced in the hemolymph upon injection of live or dead bacteria. The antibacterial proteins of *Sarcophaga peregrina* (flesh fly) termed sarcotoxins I, II, and III and sapecin have been studied extensively (1, 4, 18, 21). Among these proteins, sarcotoxin II is unique in inducing morphological changes of bacteria. When growing *Escherichia coli* cells were treated with sarcotoxin IIA, they became greatly elongated and spheroplastlike bulges and projections appeared on their surface, suggesting that the main effect of sarcotoxin IIA is inhibition of cell wall synthesis, including septum formation (2). So far, three structurally related proteins belonging to the sarcotoxin II family, termed sarcotoxins IIA, IIB, and IIC, have been purified to homogeneity (1).

Sarcotoxin II is synthesized by the fat body in response to pricking of the larval body wall and is eventually secreted into the hemolymph. Therefore, the stimulus of body injury must be transmitted to the fat body to activate the sarcotoxin II gene. To determine the mechanism of selective gene activation in response to body injury, the structure of the sarcotoxin II gene must be analyzed.

This report describes the cloning and sequencing of a gene cluster containing genes for three proteins of the sarcotoxin II family. All of these genes were found to be activated in the same manner when the larval body wall was pricked. Recently, similar gene clusters for sarcotoxin I, cecropin-type antibacterial proteins of *S. peregrina*, and *Drosophila* cecropins have been reported (14, 15). It is probable that many genes for insect antibacterial proteins exist as clusters and that their expression is induced simultaneously by the same regulatory mechanism to provide a potent defense system consisting of multiple related proteins with different antibacterial specificities.

MATERIALS AND METHODS

Construction of a *Sarcophaga* genomic library. A library was constructed essentially as described by Kaiser and Murray (13). Briefly, DNA isolated from adult flies was partially digested with *Sau*3AI, and the resulting DNA fragments were fractionated by 10 to 40% sucrose density gradient centrifugation. Fragments with molecular lengths of 15 to 20 kb were collected and ligated to *Bam*HI-cleaved λ EMBL3 arms (Amersham). The ligated DNA was packed into λ bacteriophage particles in vitro with a packaging extract (Stratagene). The infectious particles thus obtained (about 2×10^6 PFU/μg of *Sarcophaga* DNA) were used to screen sarcotoxin II genes.

Screening of sarcotoxin II genes. *E. coli* Q359 was infected with the recombinant phages, and 3×10^5 plaques were screened with pTII20, a cDNA clone for the sarcotoxin IIA gene (3). Hybridization was carried out at 42°C for 16 h in a reaction mixture consisting of 50% (vol/vol) formamide, 5× SSC (1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate), 1× Denhardt solution (0.02% [wt/vol] Ficoll 400, 0.02% bovine serum albumin, 0.02% polyvinylpyrrolidone), 20 mM sodium phosphate (pH 6.5), and 100 μg of salmon sperm single-stranded DNA per ml. The filters were washed with 0.1× SSC containing 0.1% sodium dodecyl sulfate for 15 min each time at room temperature and at 42°C. The blot was then exposed to a Kodak XAR-5 film at −70°C with an intensifying screen.

DNA blot hybridization. About 200 ng of DNA from the recombinant phage was digested with *Sal*I, *Eco*RI, or both. The digests were separated by electrophoresis in 0.8% agarose gel and transferred to a nylon filter (Dupont). Hybridization was carried out with nick-translated pTII20 or end-labeled 30-mer oligodeoxyribonucleotides with the sequence TTGTTTGAAAATATGTTTTGAAAACTGATC, which corresponded to the first 30 bases of the 5′ end of the antisense strand of pTII20, in rapid hybridization solution (Amersham) at 65°C for 2 h. The filters were washed with 0.1× SSC containing 0.1% sodium dodecyl sulfate for 15 min each time at room temperature and at 42°C. The blot was then exposed to Kodak XAR-5 film at −70°C with an intensifying screen. Oligodeoxyribonucleotides were synthesized in an Applied Biosystems model 381A DNA synthesizer.

DNA sequencing. To sequence DNA, we first prepared various deletion derivatives of the DNA fragment, using exonuclease III and mung bean nuclease (8) or site-specific restriction enzymes. Each deletion derivative was inserted into a pUC118 or pUC119 vector (25). Single-stranded DNA was prepared by using helper phage M13K07 and sequenced by the dideoxy-chain termination method (22) with Sequenase (United States Biochemical Corp.). The sequences of both strands were determined.

* Corresponding author.

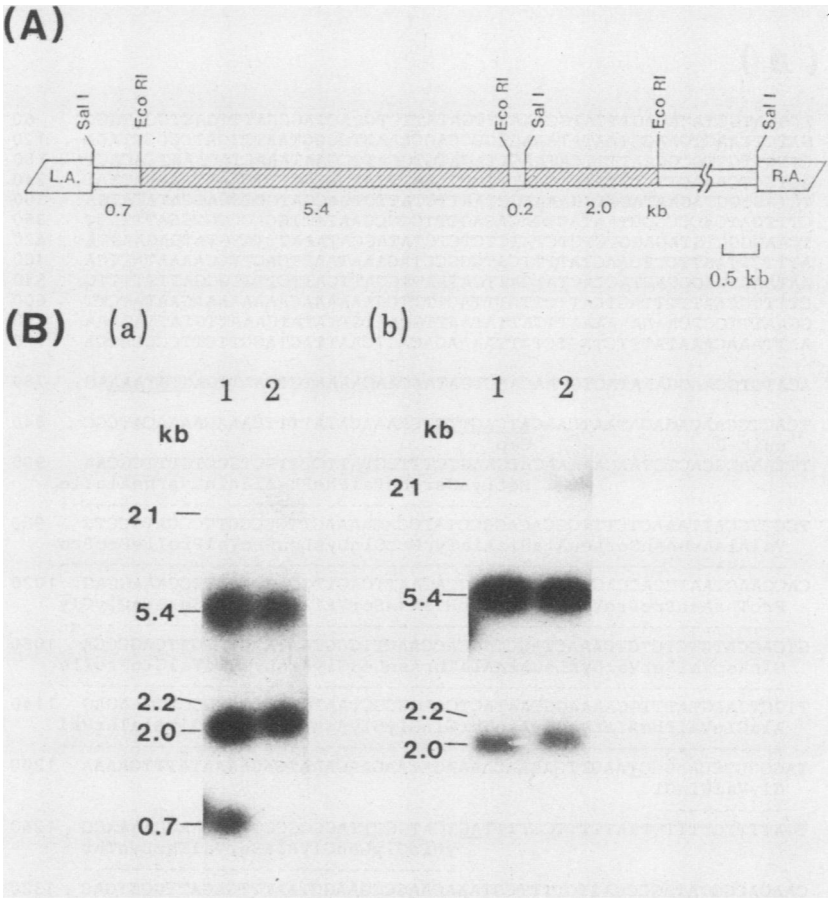

FIG. 1. Restriction map and DNA blot hybridization of the sarcotoxin II genomic clone. (A) Restriction sites of *Sal*I and *Eco*RI located on a 19-kb insert of λII(2). Hatched regions were hybridizable with the first 30 bases of the 5′ end of sarcotoxin IIA cDNA. L.A. and R.A. indicate the left and right arms of the λ phage vector. (B) Hybridization. λII(2) DNA was digested with *Sal*I and *Eco*RI (lane 1) or *Eco*RI alone (lane 2), and the resulting fragments were hybridized with pTII20 (a) or the 5′-end probe of pTII20 (b).

RNase mapping. To determine the cap site and estimate the relative amount of mRNA corresponding to each gene, we performed RNase mapping. For this procedure, we first subcloned the 5′ upstream fragment of each gene, probably containing the cap site, into the Bluescript SK$^+$ vector. Then we synthesized an antisense RNA probe in the presence of [α-^{32}P]ATP (3,000 Ci/mmol), using T7 or T3 RNA polymerase (Stratagene).

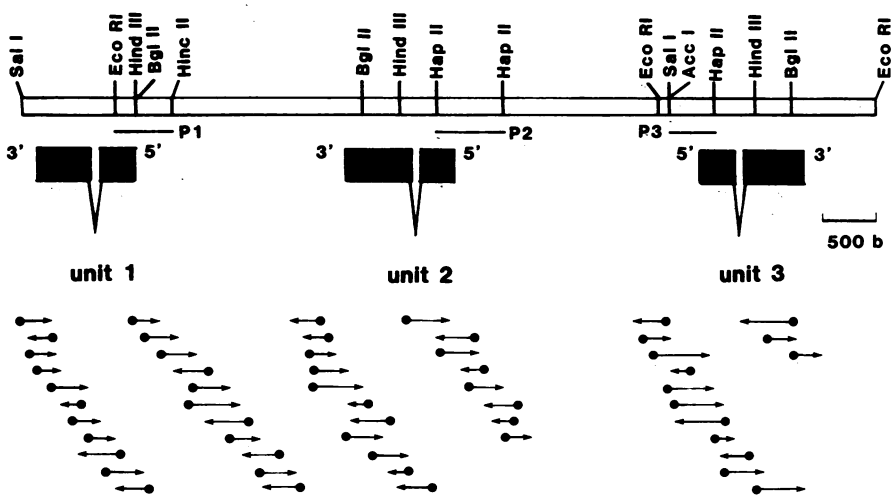

FIG. 2. Diagram of the 8.3-kb *Sal*I-*Eco*RI fragment and sequencing strategy. The directions and extents of sequencing are shown by arrows. Boxes indicate the genes of proteins belonging to the sarcotoxin II family. P1, P2, and P3 represent the fragments used for riboprobe synthesis (see Fig. 5).

(a)

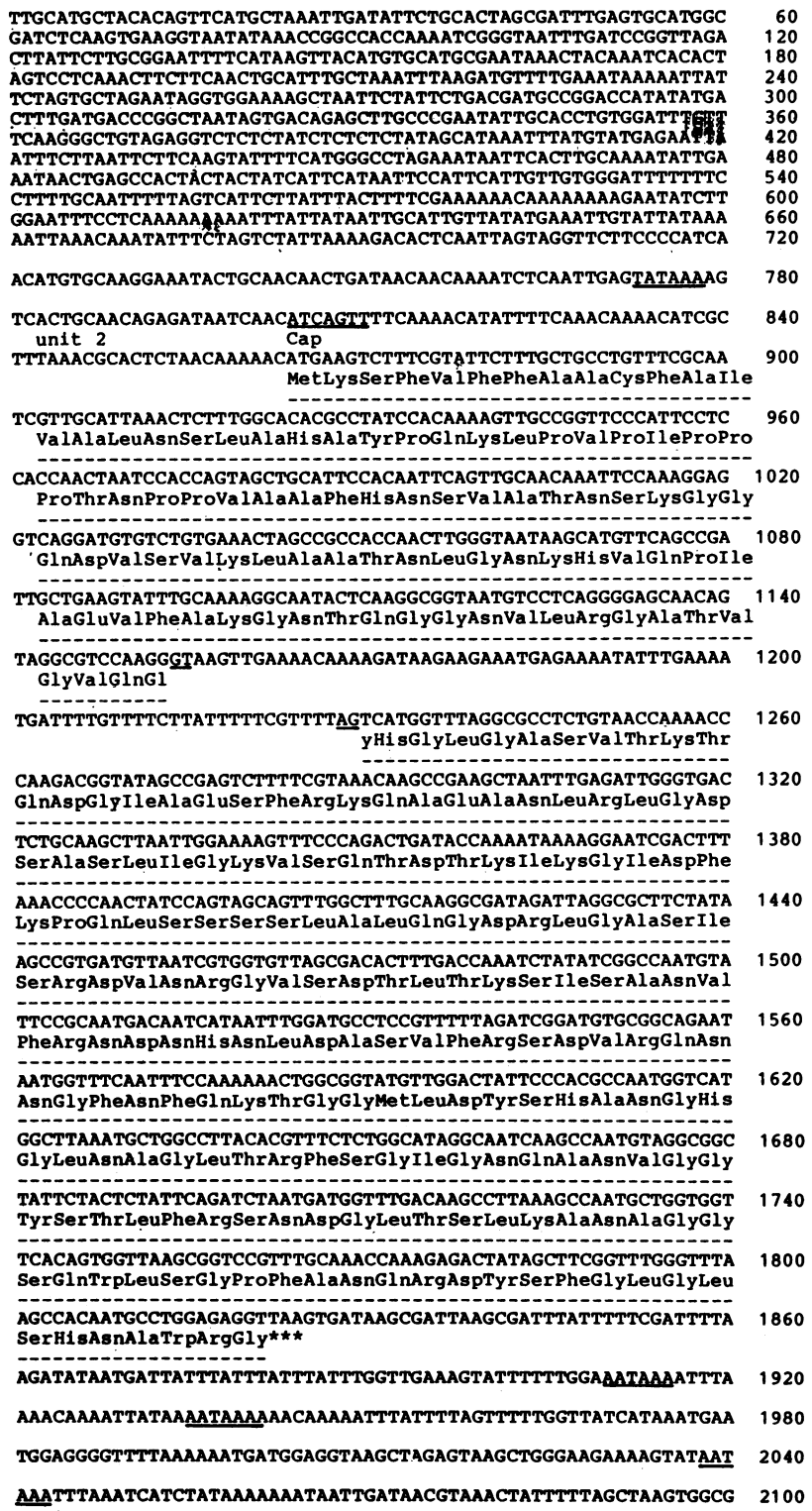

FIG. 3. Sequence of the sarcotoxin II locus. (a) Sequence of the fragment containing units 1 and 2. (b) Sequence of the fragment containing unit 3. Coding regions are underlined with dotted lines, and the corresponding amino acid sequences are shown. The sequences homologous to the TATA box, insect-specific cap site, and poly(A) addition signals are underlined.

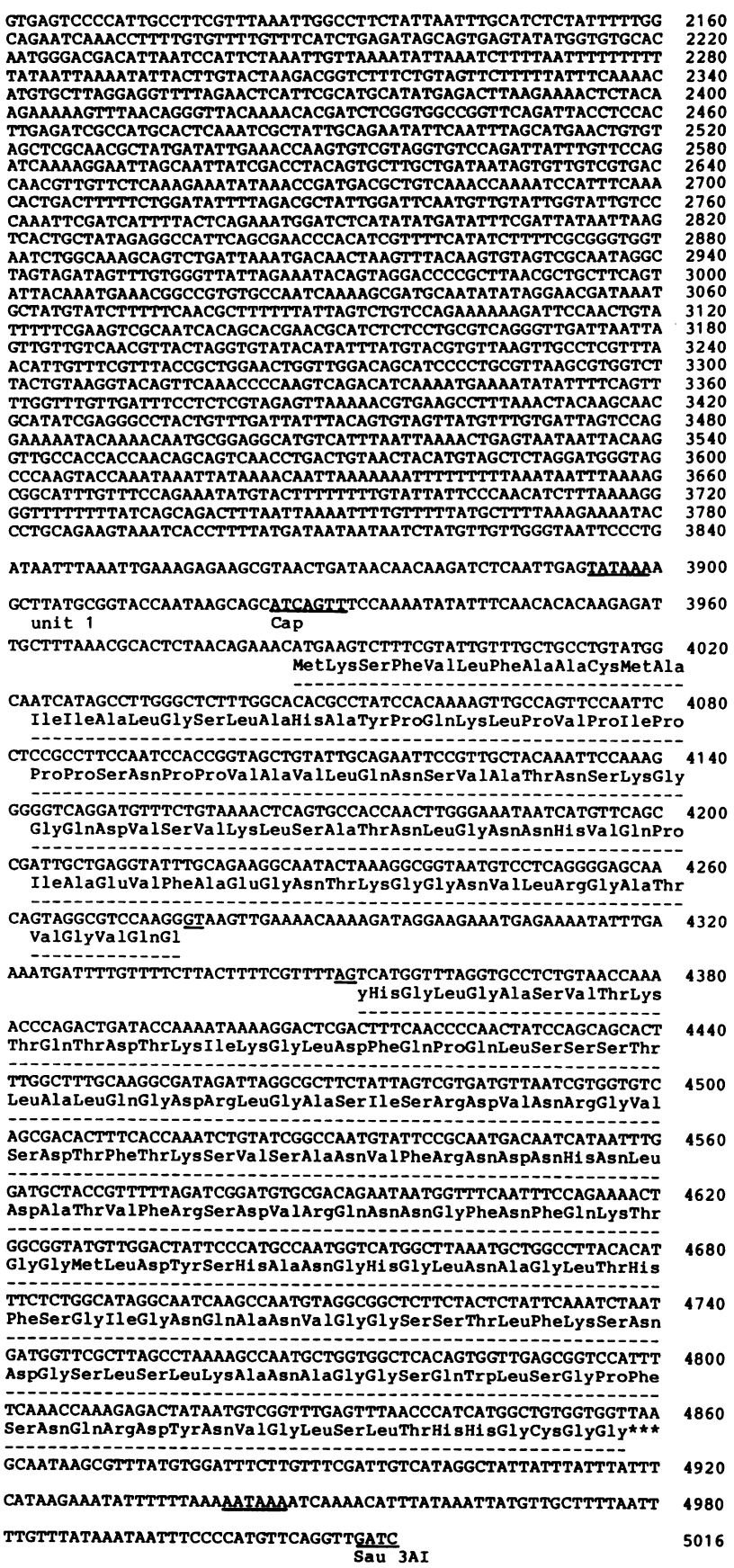

FIG. 3—*Continued.*

(b)

```
TGTAGTAATTGGGCGTTTATGTGGCCATAAGGCGATTTAGTTGATCTTCGACTTTCAGTC     60
GGCTTTTACTGGTCACTAGCTAAATAAACATCCCTGTCTTAAAGCAGATTCTAAATAGGT    120
TTCAAAGTCACCATCACGTAGGGCAGTGCGTTTAAAGGTGTCAGACATAGCGGTTATTGG    180
GTCATTCGATTTAGATACACTCTCTGTGTAGCCATACTCTGTTAGGGAAATGCCGTCTAA    240
AAATCGACAACAAATCTGTACAAAGCCAAACGCATTTCATAGCTGTGTTCACTATATTTA    300
TTTTCATATGTATGCTTTAGGGAAAGTACAAAACAATGTTGGTGGCTTAAAACTGAATTC    360
TGATCTTTTGCAAGGTTGCCAGTACTTGCAAACGTAATTTTTTCTAATCCGCAATCGAAA    420
TCTAAATGCAGTCGACCTAATTTTTCTGGTATGACTTAGGAAATGAATTGTAAAACAATT    480
AAATTTGTTTCCAGAATCATATTTGTACATTTTTGTTGTTTATTTTGTATTATTCCCAGC    540
GTCTTTAAAGGGGATTTTTGTTTTTGTATGCAGACTTTAATTAAAAATTTGTTTTTATGC    600
TTTTGAAGACGCTTACTTACGCTGTAGAAGTAAATCCCCTTTTATGATAATAATTTCTTT    660

GTTGTTGGGTTATTCCAGTAACTGATAACAACAAAATCTCAATGGAGTATAAAAGTCACT    720

GCAACAGAGATAATCAACATCAGTTTTCAAAACATATTTTCAAACAAAACATCGTTTTAA    780
  unit 3          Cap
ACGCACTCTAACAAAAACATGAAGTCTTTCGTATTGTTTGCTGCCTGTATGGCAATCGTT    840
                  MetLysSerPheValLeuPheAlaAlaCysMetAlaIleVal
                  ----------------------------------------
GCATTAAGCTCTTTGGCACACGCCTATCCACAAAAGTTGCCGGTTCCAATTCCTCCACCA    900
AlaLeuSerSerLeuAlaHisAlaTyrProGlnLysLeuProValProIleProProPro
------------------------------------------------------------
ACTAATCCACCAGTAGCTGCATTCCACAATTCAGTTGCAACAAATTCCAAAGGAGGTCAG    960
ThrAsnProProValAlaAlaPheHisAsnSerValAlaThrAsnSerLysGlyGlyGln
------------------------------------------------------------
GATGTGTCTGTGAAACTAGCCGCCACCAACTTGGGTAATAAGCATGTTCAGCCGATTGCT   1020
AspValSerValLysLeuAlaAlaThrAsnLeuGlyAsnLysHisValGlnProIleAla
------------------------------------------------------------
GAAGTATTTGCAGAAGGCAATACTAAAGGCGGTAATGTCATCAGGGGAGCAACAGTAGGC   1080
GluValPheAlaGluGlyAsnThrLysGlyGlyAsnValIleArgGlyAlaThrValGly
------------------------------------------------------------
GTCCAAGGGTAAGTTAAAAATAAAAGATAAGAAGAAATTAGAAAATATTTGAAAATGATT   1140
ValGlnGl
--------
TTGTTTTCTTATTTTTTGTTTTAGTCATGGTTTAGGCGCCTCTGTAACCAAAAGCGGAAA   1200
                      yHisGlyLeuGlyAlaSerValThrLysSerGlyAsn
                      -------------------------------------
CGGTATAGCCGAGTCTTTTCGTAAACAAGCCGAAGCCAATTTGAGATTGGGTGACTCTGC   1260
 GlyIleAlaGluSerPheArgLysGlnAlaGluAlaAsnLeuArgLeuGlyAspSerAla
 ------------------------------------------------------------
AAGCTTAATTGGAAAAGTTTCCCAGACTGATACCAAAATAAAAGGAATCGACTTTAAACC   1320
 SerLeuIleGlyLysValSerGlnThrAspThrLysIleLysGlyIleAspPheLysPro
 ------------------------------------------------------------
CCAACTATCCAGCAGCAGTTTGGCTTTGCAAGGCGATAGATTAGGCGCTTCTATAAGCCG   1380
 GlnLeuSerSerSerSerLeuAlaLeuGlnGlyAspArgLeuGlyAlaSerIleSerArg
 ------------------------------------------------------------
TGATGTTAATCGTGGTGTTAGCGACACTTTAACCAAATCTATATCGGCCAATGTATTCCG   1440
 AspValAsnArgGlyValSerAspThrLeuThrLysSerIleSerAlaAsnValPheArg
 ------------------------------------------------------------
CAATGACAATCATAATTTGGATGCCTCCGTTTTTAGATCGGATGTGCGGCAGAATAATGG   1500
 AsnAspAsnHisAsnLeuAspAlaSerValPheArgSerAspValArgGlnAsnAsnGly
 ------------------------------------------------------------
TTTCAATTTCCAGAAAACTGGCGGTATGTTGGATTATTCCCACGCCAATGGTCATGGCTT   1560
 PheAsnPheGlnLysThrGlyGlyMetLeuAspTyrSerHisAlaAsnGlyHisGlyLeu
 ------------------------------------------------------------
AAATGCTGGCCTTACACGTTTCTCTGGCATAGGCAATCAAGCCAATGTAGGCGGCTATTC   1620
 AsnAlaGlyLeuThrArgPheSerGlyIleGlyAsnGlnAlaAsnValGlyGlyTyrSer
 ------------------------------------------------------------
TACTCTATTCAGATCTAATGATGGTTTGACTAGCCTTAAAGCCAATGCTGGTGGTTCACA   1680
 ThrLeuPheArgSerAsnAspGlyLeuThrSerLeuLysAlaAsnAlaGlyGlySerGln
 ------------------------------------------------------------
ATGGTTGAGCGGTCCGTTTGCAAACCAAAGAGACTATAGCTTCGGTTTGGGTTTAAGCCA   1740
 TrpLeuSerGlyProPheAlaAsnGlnArgAspTyrSerPheGlyLeuGlyLeuSerHis
 ------------------------------------------------------------
TAATGCCTGGAGAGGTTAATTGATAAGTTATAAGCGATTTTTTCGGTTTTATTACATTAT   1800
 AsnAlaTrpArgGly***
 ---------------
TTATTTATTCGACTAGAGGTATTTTTTAGAAATAAAATTTAACACAAAATTATAAAATAA   1860

AATTTTGTTTTAGTTTTTGCTAGAGGGGTTTTATAAGAAATAATGGAAACTGATAACATT   1920
```

FIG. 3—*Continued.*

To determine the cap site, each riboprobe (about 10^6 cpm) was hybridized at 42°C for 18 h with 40 μg of total RNA prepared from the fat body of injured third-instar larvae (24). Then the mixture was treated with RNases A and T_1 for 30 min at 30°C as described by Little and Jackson (17). The resulting digestion products were separated by electrophoresis in a 6% acrylamide–8 M urea gel and autoradiographed, and the size of protected RNA containing the cap site was determined. To determine the relative amount of mRNA, we scanned the intensity of the band on the autoradiogram obtained by RNase mapping.

Nucleotide sequence accession number. The nucleotide sequence data reported will appear in the DDBJ, EMBL, and GenBank nucleotide sequence data bases under acces-

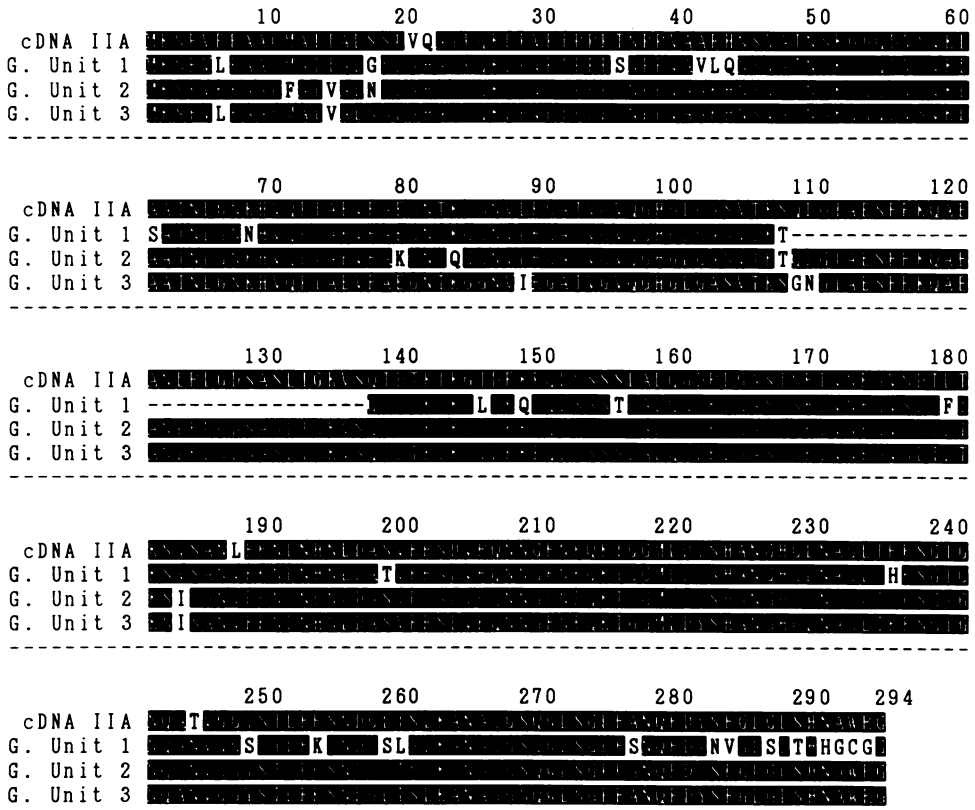

FIG. 4. Deduced amino acid sequences of the proteins encoded by cloned genes compared with that of sarcotoxin IIA. Amino acid residues identical to those of sarcotoxin IIA are shown in black boxes. A gap is introduced in unit 1 to obtain maximal sequence homology.

sion numbers D90153 (sarcotoxin II [TOIIU1&2]) and D90154 (sarcotoxin II [TOIIU3]).

RESULTS

Isolation and characterization of a genomic clone of sarcotoxin II. About 3×10^5 plaques of a *Sarcophaga* genomic library were screened with sarcotoxin IIA cDNA (pTII20) as a probe, and two hybridization-positive clones were obtained. We analyzed one of these clones, termed λII(2), which gave an intense signal. This clone contained a 19-kb insert, the restriction map of which was determined by digestions with *Sal*I and *Eco*RI to be as shown in Fig. 1A. To locate the sarcotoxin IIA gene in this insert, we performed Southern blotting hybridization with two probes: nick-translated pTII20 and end-labeled 30-mer oligodeoxyribonucleotides corresponding to the first 30 bases of the 5′-end of sarcotoxin IIA cDNA (3). Three hybridization-positive bands were obtained when the *Sal*I-*Eco*RI digest was probed with pTII20 (Fig. 1B). Of these fragments, those of 5.4 and 2.0 kb hybridized with the 5′-end probe, indicating that these fragments contained independent 5′ ends. Their locations are shown by hatched boxes in Fig. 1A.

From these results, we concluded that only 8.3 kb of the 19-kb insert adjacent to the left arm of the λ phage vector (*Sal*I-*Eco*RI fragment) hybridized with pTII20 and that this region contained at least two transcription units of proteins of the sarcotoxin II family. To identify these genes, we sequenced most of the *Sal*I-*Eco*RI fragment and found three genes. Two genes (transcription units 1 and 2) were located in the 5.4-kb *Eco*RI-*Eco*RI fragment, and the third gene (transcription unit 3) was located in the 2.0-kb *Sal*I-*Eco*RI fragment in the opposite direction. The sequence strategy and the array of genes are shown in Fig. 2, and the sequences of these fragments are shown in Fig. 3. Each transcription unit was about 1 kb, and the coding sequence was interrupted by a short intron of 76 bases at a conserved site in all three genes. The exon-intron boundaries of these genes were 5′-GT–AG-3′, according to the consensus sequence of splicing (19). The space between each transcription unit was about 2 kb. Therefore, the genes for the proteins of the sarcotoxin II family are probably present in tandem array with a space of about 2 kb.

These genes encoded very similar proteins. The putative amino acid sequences encoded in units 1 to 3 were compared with the amino acid sequence of sarcotoxin IIA (3). None of them was identical with sarcotoxin IIA (Fig. 4), but the sequence homology among these proteins was very high, suggesting that these genes originated by duplication of a common ancestral gene. The unit 1 protein had a deletion of 29 amino acid residues, corresponding to positions 108 to 136. This region may not be related to the antibacterial activity of these proteins.

Each gene contained TATAAA and ATCAGTT sequences in the region upstream of the first Met codon. The former sequence is believed to be a TATA box, and the latter sequence is an insect-specific consensus sequence in which the 5′-end A is designated the cap site (11). This conclusion was also confirmed by RNase mapping as described below. In fact, the first seven bases of sarcotoxin IIA cDNA are ATCAGTT. A poly(A) addition signal of AATAAA was located in the 3′-end nontranslated region. In units 2 and 3,

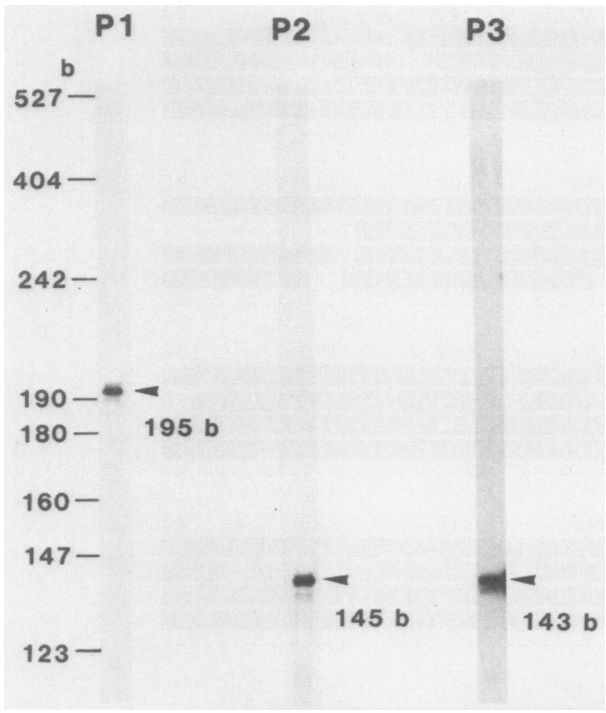

FIG. 5. Detection of mRNA derived from each gene by RNase mapping. RNA was prepared from the fat body of third-instar larvae 3 h after pricking of the body wall and hybridized with labeled riboprobes corresponding to P1, P2, and P3 in Fig. 2. The protected RNA was then analyzed by electrophoresis after digestion with RNases A and T_1. Gels were calibrated with DNA fragments obtained by digesting pBR322 with *Hpa*II. Arrowheads indicate protected RNA derived from the riboprobes shown at the top of each lane.

multiple poly(A) addition signals were identified, suggesting the presence of the same mRNA with different sizes.

Identification of three mRNAs corresponding to cloned genes and their expression. Using RNase mapping, we examined whether the proteins of these three genes belonging to the sarcotoxin II family were expressed in response to injury of the larval body wall, as is the product of the sarcotoxin IIA gene. As templates for RNA synthesis, we prepared DNA fragments containing the 5′ upstream regions of three genes, termed P1 (*Eco*RI-*Hin*cII), P2 (*Hpa*II-*Hpa*II), and P3 (*Acc*I-*Hpa*II) (Fig. 2), and subcloned them into the Bluescript SK^+ vector. These DNA fragments are expected to contain the cap sites of the genes. We then synthesized antisense RNA probes, using T3 or T7 RNA polymerase. Each riboprobe was hybridized with RNA prepared from injured larvae and digested with nucleases, and the protected region of the probe RNA was investigated by electrophoresis.

RNAs with molecular sizes of 195, 145, and 143 bases, corresponding to fragments P1, P2, and P3, respectively, were detected (Fig. 5). These results indicate that the three genes are in fact expressed in injured larvae and that the protected RNA corresponds to the region from the cap site to the restriction enzyme site of each gene. For instance, in transcription unit 1, we concluded that the cap site is located 195 bases upstream of the *Eco*RI site of fragment P1. This nucleotide coincided with the first A of the insect-specific consensus sequence ATCAGTT located at nucleotide positions 3925 to 3931 (Fig. 3a). In the same way, we concluded that the cap sites of transcription units 2 and 3 are also the first A of ATCAGTT of these genes (Fig. 3a and b, respectively).

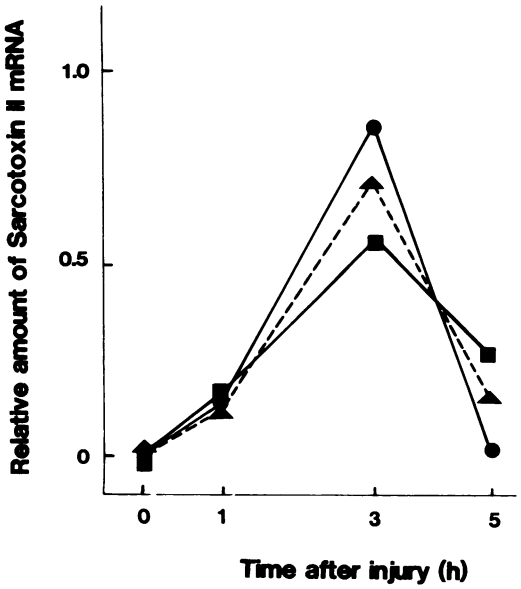

FIG. 6. Time course of expression of the genes after injury of the larval body wall. RNA was extracted from the larval fat body at the times indicated after pricking of the body wall, and the transcript of each gene was examined by RNase mapping with riboprobes. The relative densities of the bands of protected RNA on the autoradiograms are plotted against time after pricking of the larval body wall, taking the total density obtained with each probe as 1. The riboprobes used were P1 (●), P2 (▲), and P3 (■).

Next we investigated the expression of these genes after injury of the larval body wall. To do so, we measured the density of each band on RNase mapping of total RNA prepared from larvae at various times after injury of the body wall. The relative amount of mRNA was plotted against the time after body injury. The three genes were expressed in the same way (Fig. 6). No mRNA was detected in RNA prepared from normal larvae, but the amount of mRNA increased with time after body injury to a maximum after 3 h and then decreased rapidly. Thus, expression of these genes was transient, and the mRNAs were rather unstable.

Since these three genes are expressed coordinately in response to body injury, it is expected that they contain a common element that responds to the stimulus of injury. The sequences of the upstream 5′-flanking regions of the three genes were searched by computer with alignment for maximum matching. Although these genes are believed to have originated by duplication of a single ancestral gene, their 5′ upstream sequences are much more diverse than their coding sequences, and many deletions and insertions were detected. However, the nucleotide sequences from positions −36 to −59 were highly conserved in the three genes (Fig. 7). Nothing is known about the biological significance of these sequences, but they may well be regulatory elements controlling expression of the sarcotoxin II genes.

DISCUSSION

We found that three active genes for proteins of the sarcotoxin II family form a tight cluster in about 8 kb of *Sarcophaga* DNA. Therefore, *S. peregrina* has at least four related genes, among which is the sarcotoxin IIA gene, which has not yet been identified. The Southern blotting

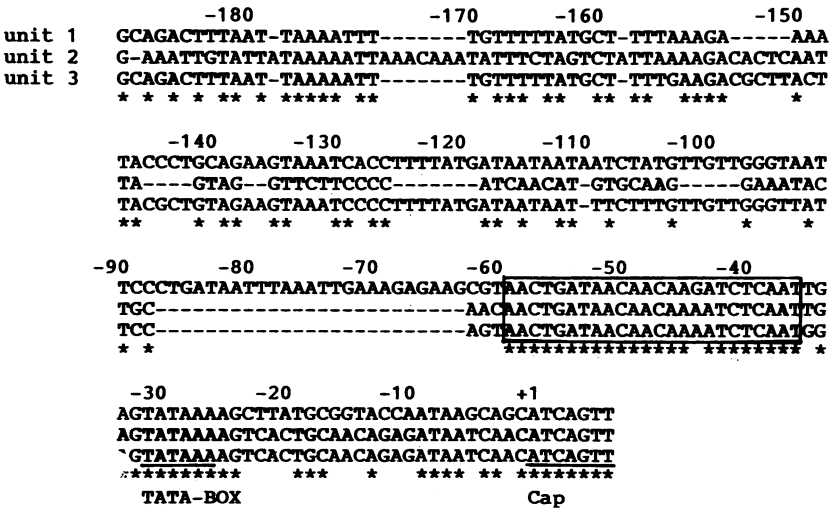

FIG. 7. Comparison of 5'-flanking regions of the three genes of the sarcotoxin II family. About 200 bases of 5'-flanking sequences of the three genes were aligned to give maximum homology by inserting gaps. Asterisks show identical bases. Highly conserved sequences are boxed. The TATA -box and insect-specific consensus sequence containing the cap site are underlined.

patterns of λII(2) DNA and *Sarcophaga* genomic DNA obtained by probing with pTII20 were identical except for an extra positive band in genomic DNA (unpublished data). Therefore, *Sarcophaga* DNA probably contains a few more genes for proteins of the sarcotoxin II family, including sarcotoxin IIA.

An RNase mapping experiment showed that all three genes are active and are activated coordinately in response to injury of the larval body. A similar gene cluster was identified with proteins of the sarcotoxin I family, another group of cecropin-type antibacterial proteins of this insect (14). Moreover, a gene cluster for *Drosophila* cecropins was also cloned recently (15). Therefore, it is likely that many genes of dipteran antibacterial proteins form gene clusters.

We found that the putative amino acid sequences of the three proteins were not identical. For instance, there is deletion of 29 amino acid residues in the unit 1 protein. There are also various differences in individual amino acid residues in these proteins. However, judging from the sequence of sarcotoxin IIA (3), the amino-terminal amino acid residue of each protein is Gln at position 25 (Fig. 4). This Gln may exist as pyroglutamate, as in sarcotoxin IIA.

Why does *S. peregrina* need so many related proteins? Does each protein have a specific function? These genes probably originated from a common ancestral gene by gene duplication, and insects with multiple genes have been selected during evolution. Nothing is known about the functions of these proteins except that they have antibacterial activity against certain gram-negative bacteria (1). If each protein has a slightly different antibacterial spectrum, the presence of multiple proteins should be advantageous for survival of the insect in various pathogenic environments.

Expression of the sarcotoxin II gene is induced upon pricking of the body wall of *Sarcophaga* larvae (1, 3). However, the mRNA seems to be unstable because it disappeared rapidly. As reported previously, sarcotoxin IIA mRNA was detected continuously for at least 24 h after injection of a light suspension of *E. coli* into the larvae (3). The mRNA itself probably is unstable, because the consensus sequence TTATTTAT of unstable mRNA was found in the 3'-nontranslated region of each gene (5, 23). After introduction of foreign substances such as *E. coli* into the body, the mRNA may continue to be synthesized for much longer than after simple pricking of the body wall.

ACKNOWLEDGMENTS

This work was supported by a Grant-in-Aid for Scientific Research from the Ministry of Education, Science and Culture of Japan and in part by a Grant-in-Aid (Bio Media Programs) from the Ministry of Agriculture, Forestry and Fisheries of Japan (BMP 90-IV-2-6).

LITERATURE CITED

1. **Ando, K., M. Okada, and S. Natori.** 1987. Purification of sarcotoxin II, antibacterial proteins of *Sarcophaga peregrina* (flesh fly) larvae. Biochemistry **26**:226–230.
2. **Ando, K., and S. Natori.** 1988. Inhibitory effect of sarcotoxin IIA, an antibacterial protein of *Sarcophaga peregrina* on growth of *Escherichia coli*. J. Biochem. **103**:735–739.
3. **Ando, K., and S. Natori.** 1988. Molecular cloning, sequencing, and characterization of cDNA for sarcotoxin IIA, an inducible antibacterial protein of *Sarcophaga peregrina* (flesh fly). Biochemistry **27**:1715–1721.
4. **Baba, K., M. Okada, T. Kawano, H. Komano, and S. Natori.** 1987. Purification of sarcotoxin III, a new antibacterial protein of *Sarcophaga peregrina*. J. Biochem. **102**:69–74.
5. **Caput, D., B. Beutler, K. Hartog, R. Thayer, S. Brown-Shimer, and A. Cerami.** 1986. Identification of a common nucleotide sequence in the 3'-untranslated region of mRNA molecules specifying inflammatory mediators. Proc. Natl. Acad. Sci. USA **83**:1670–1674.
6. **Casteels, P., C. Ampe, L. Riviere, J. Van Damme, C. Elicone, M. Fleming, F. Jacobs, and P. Tempst.** 1990. Isolation and characterization of abaecin, a major antibacterial response peptide in the honeybee (*Apis mellifera*). Eur. J. Biochem. **187**:381–386.
7. **Dimarcq, J.-L., E. Kedpi, B. Dunbar, J. Lambert, J.-M. Reichhart, D. Hoffmann, S. M. Rankine, J. E. Fothergill, and J. A. Hoffmann.** 1988. Insect immunity. Purification and characterization of a family of novel inducible antibacterial proteins from immunized larvae of the dipteran *Phormia terranovae* and complete amino-acid sequence of the predominant member, diptericin A. Eur. J. Biochem. **171**:17–22.
8. **Henikoff, S.** 1987. Unidirectional digestion with exonuclease III in DNA sequence analysis. Methods Enzymol. **155**:156–165.
9. **Hultmark, D., Å. Engström, K. Anderson, H. Steiner, H. Ben-**

nich, and H. G. Boman. 1983. Insect immunity. Attacins, a family of antibacterial proteins from *Hyalophora cecropia*. EMBO J. **2**:571–576.
10. **Hultmark, D., Å. Engström, H. Bennich, R. Kapur, and H. G. Boman.** 1982. Insect immunity. Isolation and structure of cecropin D and four minor antibacterial components from cecropia pupae. Eur. J. Biochem. **127**:207–217.
11. **Hultmark, D., R. Klemenz, and W. G. Gehring.** 1986. Translational and transcriptional control elements in the untranslated leader of heat-shock gene *hsp*22. Cell **44**:429–438.
12. **Hultmark, D., H. Steiner, T. Rasumuson, and H. G. Boman.** 1980. Insect immunity. Purification and properties of three inducible bactericidal proteins from hemolymph of immunized pupae of *Hyalophora cecropia*. Eur. J. Biochem. **106**:7–16.
13. **Kaiser, K., and N. E. Murray.** 1985. The use of phage lambda replacement vectors in the construction of representative genomic DNA libraries, p. 1–47. *In* D. M. Glover (ed.), DNA cloning: a practical approach, vol. 1. IRL Press, Oxford.
14. **Kanai, A., and S. Natori.** 1989. Cloning of gene cluster for sarcotoxin I, antibacterial proteins of *Sarcophaga peregrina*. FEBS Lett. **258**:199–202.
15. **Kylsten, P., C. Samakovlis, and D. Hultmark.** 1990. The cecropin locus in *Drosophila*; a compact gene cluster involved in the response to infection. EMBO J. **9**:217–224.
16. **Lambert, J., E. Keppi, J.-L. Dimarcq, C. Wicker, V. A. Dorsselaer, J. A. Hoffmann, J. E. Fothergill, and D. Hoffmann.** 1989. Insect immunity: isolation from immune blood of the dipteran *Phormia terranovae* of two insect antibacterial peptides with sequence homology to rabbit lung macrophage bactericidal peptides. Proc. Natl. Acad. Sci. USA **86**:262–266.
17. **Little, P. F. R., and I. J. Jackson.** 1987. Application of plasmids containing promoters specific for phage-encoded RNA polymerases, p. 1–42. *In* D. M. Glover (ed.) DNA cloning: a practical approach, vol. 3. IRL Press, Oxford.
18. **Matsuyama, K., and S. Natori.** 1988. Purification of three antibacterial proteins from the culture medium of NIH-Sape-4, an embryonic cell line of *Sarcophaga peregrina*. J. Biol. Chem. **263**:17112–17116.
19. **Mount, S. M.** 1982. A catalogue of splice junction sequences. Nucleic Acids Res. **10**:459–472.
20. **Okada, M., and S. Natori.** 1983. Purification and characterization of an antibacterial protein from haemolymph of *Sarcophaga peregrina*. Biochem. J. **211**:727–734.
21. **Okada, M., and S. Natori.** 1985. Primary structure of sarcotoxin I, an antibacterial protein induced in the hemolymph of *Sarcophaga peregrina*. J. Biol. Chem. **260**:7174–7177.
22. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain terminating inhibitor. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
23. **Show, G., and R. Kamen.** 1986. A conserved AU sequence from the 3′ untranslated region of GM-CFS mRNA mediates selective mRNA degradation. Cell **46**:659–667.
24. **Takahashi, H., H. Komano, and S. Natori.** 1986. Expression of the lectin gene in *Sarcophaga peregrina* during normal development and under conditions where the defense mechanism is activated. J. Insect. Physiol. **32**:771–779.
25. **Vieira, J., and J. Messing.** 1982. The pUC plasmids, an M13mp7-derived system for insertion mutagenesis and sequencing with synthetic universal primer. Gene **19**:259–268.