

## A Novel Member of Lebocin Gene Family from the Silkworm, *Bombyx mori*<sup>1</sup>

Seiichi Furukawa,\* Kiyoko Taniai,† Jun Ishibashi,† Seiichi Hara,‡  
Toshio Shono,\* and Minoru Yamakawa†<sup>2</sup>

\*Laboratory of Applied Zoology, Institute of Agriculture and Forestry, University of Tsukuba, Tsukuba, Ibaraki 305, Japan; †Laboratory of Biological Defense, National Institute of Sericultural and Entomological Science, Tsukuba, Ibaraki 305, Japan; and ‡Noda Institute of Scientific Research, Noda, Chiba 278, Japan

Received August 8, 1997

**We screened genomic clones encoding lebocin, an antibacterial peptide from the silkworm, *Bombyx mori*. Two positive clones were obtained and their nucleotide sequences indicated that they contain no introns. The deduced amino acid sequences revealed that one clone (*Leb* 3) encoded lebocin 3 and another (*Leb* 4) is a new member of the lebocin gene family. Gene expression of both *Leb* 3 and *Leb* 4 was shown to be induced by lipopolysaccharide and to occur tissue-specifically in the fat body and hemocytes. Our results suggest that lebocin as well as cecropin forms a multiple gene family in *B. mori*.** © 1997 Academic Press

Insects and other invertebrate organisms do not possess antibodies found in vertebrates, although proteins having immunoglobulin-like domains have been reported (1-4). However, insects and other arthropods have developed unique and efficient defense systems to protect themselves against invading pathogens and parasites (5, 6). For example, quick induction of synthesis and secretion of self-defense molecules like antibacterial peptides is one of these defense mechanisms.

Many antibacterial peptides have been isolated from different insect species (6) and can be classified into five major groups; cecropins, insect defensins, attacin-like proteins, proline-rich peptides and lysozymes (5). Proline-rich Lebocins were isolated from immunized hemolymph of the silkworm, *Bombyx mori* and found to consist of 32 amino acids (7). Unique threonine residues (see Fig. 4, panel A) in lebocin 1, 2 and 3 are

O-glycosylated and this modification is important for expression of antibacterial activity (7). Lebocin 1 and 2 have identical amino acid sequences but the length of the sugar chains on their threonine residues differs (7). Incubation of lebocins with a liposome preparation caused leakage of entrapped glucose under conditions of low-ionic-strength, suggesting that a target for these peptides is the bacterial membrane (7). The primary structure and antibacterial character of lebocins resemble those of abaecin, an antibacterial peptide of the honeybee, although abaecin is not O-glycosylated.

Recently, a cDNA encoding lebocin 1 and 2 (lebocin 1/2) was isolated from a fat body cDNA library of *B. mori* larvae immunized with *Escherichia coli* (8). Northern blot analysis using this cDNA as a probe showed that gene expression was inducible by injection of bacteria, occurred tissue-specifically in the fat body and continued at least 48 h post-infection.

In order to elucidate the expression mechanisms of lebocin genes, we screened a genomic library of *B. mori* using the cDNA as a probe and obtained two positive clones. Nucleotide sequencing of these clones revealed that one clone encodes lebocin 3 and another one encodes a novel member of the lebocin family, designated lebocin 4. Deduced amino acid sequence (mature portion) of lebocin 4 indicated it to be 96.9% and 87.5% identical to those of lebocin 1/2, and lebocin 3, respectively. Genomic clones of lebocin 3 (*Leb* 3) and lebocin 4 (*Leb* 4) appeared to be intronless, contrary to other *B. mori* antibacterial protein genes such as cecropin B. Expression of both genes was found to be inducible by lipopolysaccharide (LPS) and occurred tissue-specifically as observed in the lebocin 1/2 gene. Our results suggest that lebocin also forms a multiple gene family as observed in cecropins.

### MATERIALS AND METHODS

*Insect.* Silkworms, *Bombyx mori* (Tokai × Asahi) were reared on an artificial diet (Nihonnosanko) at 25 °C.

<sup>1</sup> The sequence data reported in this paper have been deposited with the DDBJ, EMBL, and GenBank Data Libraries under accession numbers AB003035 and AB003036, for lebocin 3 and lebocin 4 genes, respectively.

<sup>2</sup> Corresponding author. Fax: +81-298-38-6028. E-mail: yamakawa@nises.affrc.go.jp.

## A

```

TTTAGTCTTTACACGAAAATAGTCTTTAAATTCACACCATCAGCACATCTTTAACATTACCTAATGCTGCTCCCGTACTTACATAATTCA 90
ACTCAAGCTATCAACAGCAGTCGTCCACAACAAAGCGGCATCGTAGTCTTAAAGGACCCAAAGTCGGTTATTCATCATGCCCCCTGGTT 180
CGACGGCATCTTACTTCATCATCAGGCCAGCAGCCCTAGCAGCCACGACTGAATCTTCAGCATTTCGGAATTCACATCGCCGATAC 270
GGTAGCTCTGCACCACATCGACTATTCTCGGCAGTACGCGAGCATTCAGCAGCATTTTCGACTCAGCGCAGATTTCGAGCAACACTGGAA 360
AAACGGTAACACGATCTCCACACTCAGATGGGAGTGATGTAAGATGTAGTGATGTAGCTATCATAGAACATCTTTGAATCGCGTTAACGA 450
CATTTTAAAGTAATCTTTGTATACGAGTTTCATACAATAACATTCGACAGTAAGTAGATCTACAAGAAATGATACCTCCTCGGTGGA 540
AATCATCCCTTTTGCTATCCCAAAGCGCCACAGTAGGTATACATACAGTATTTATGTGTGAATTAATTGATTTTTCATTTTATTTT 630
CTTCATTACTGATTAATATATGAAAGTTTGGTCTAATTTAATAGTCTCTATCTATTTTATTATTGTTACCTAGTACTGCAATGT 720
ACCTGGAGCATTTTTTTAGAATTTCCCGATGTGCAAAATGATATACCAGAACAGAGTGCTTTTATCGATAAGATCGTCCATTGTAGAAAT 810
TATCTGCTTATTTACTAACATCTTGTTAGGCATCAAAGTATAATCTTCTCTCATAATTTATATGAAACATTATAATATCTACGTATT 900
GATTGTATAGGTACGTATATGCAACAGTAATGAGGAAACAAATCATGAAAAATGTTTGTAGATGAATCTTTGTCATAATAAGCCAGAC 990
TTCCCGTTATCATTCGTAAATAATATCGCTAAAGCGGAGGGAAGTACCAGACTTTCCTCAAGACTGTGTATAAATACCGATCATTG 1080
TGATAGTTACTCACATTACCGGTGTTCAATCGATACAATCAACATGTACAAGTTTTCAGTATTCAGTTCAGTTCTGGTCTGTTCTTTGC 1170
M Y K F L V F S S V L V L F F A
TCAGGCTTCGTGCCAGAGTTCATCCAGCCGACCTTCAGGCCACCGCCAAACACAGCGCCCGATAACACGTACAGTGCAGCAAGCTGGCCA 1260
Q A S C Q R F I Q P T F R P P P T Q R P I T R T V R Q A G Q
GGAACCGCTATGGCTGATCAAGGTGACAATGTTCTCGTGCGCCAAGTACCGCAGACCATCCGATTCTTCCTTCGAAAATCGACGAGCT 1350
E P L W L Y Q G D N V P R A P S T A D H P I , P S K I D D V
GCAGCTCGATCCAAACCGAAGTATGTTTCGAGTGTACCAATCCAGAAAAAATACGAGGCGTCCATTGAACATTACATCATACAGTTGA 1440
Q L D P N R R Y V R S V T N P E N N E A S I E H S H H T V D
TATTGGACTTGACCGCCGATCGAGAGCCCGTAAACACAGGACCTGCGGTTTTCATCCCTCGAGGGAAATGCCTGTTTCCAACGCT 1530
I G L D Q P I E S H R N T R D L R F L Y P R G K L P V P T L
TCCTCCGTTTAAACCCCAAGCCAATATATATGATATGGGAAACCGTTACCGACGACATCGCTCGAGGATCAAGAAGAAATGCGGCAATA 1620
P P F N P K P I Y I D M G W R Y R R H A S E D Q E E L R Q Y
TAATGAGCACTTTCGATTCGAGGGATATTTTCAAGAAATAGGAAAGTTCAGAAACAAAGATTTCGGAGTGCACTCTTATATTCAT 1710
N E H F L I P R D I F Q E ***
TGAATCGTAATTAAGTTTAAAGATATTTTAAATAGTTCTGCATAAAAAACAAAGTATTTTAAAGTGGCCGTTTCAATTCAAATA 1800
ATGTAAGTACTCGTACGTATGTTTACCATGATAATATCAAAATTAAAAAACTCAACCAATATTGAACATTATGTTTGTGTAATTTGCTT 1890
TGCAACAAATTTCTCTCACCTGTTGTAATAATATTTTATAATTTGCTGTGTTTAAAGTGTAATTAAGTGTTGTAATTAAGGAAAT 1980
CATTTTGAATAATGTTGTTTCTGTTTCTTTTCGCCAACACTACTGTAGTGTATACCAGAGTACATGTCACTCCCACTGCTGAG 2070
ACGAT 2075

```

**FIG. 1.** Nucleotide sequences of two lebocin genomic clones. The nucleotide sequences of FS-1 (A) and FS-2 (B) clones along with the deduced amino acid sequences are shown. The nucleotide numbers are indicated in the right margin of the panels. TATA box and CAAT box are shown by single underline. Putative LPS responsive element and GATA motif are indicated by bold and thin boxes, respectively. IL-6 responsive element is shown by double underline. Arrows reveal the direction of LPS responsive element and GATA motif. The polyadenylation signal (AATAAA) is indicated by a bold wavy line. Stars denote the translation stop codon. Amino acid sequence deduced from the nucleotide sequence is expressed by a single letter and mature (FS-1) and putative mature (FS-2) peptides are shown in bold letters. A probable signal peptide and a putative prosegment are as described elsewhere (Refer to reference 8).

**Screening of lebocin genes.** We screened  $1 \times 10^5$  plaques of *B. mori* genomic library constructed with EMBL 3 (10) using a cDNA encoding lebocin 1/2 as a probe. To prepare a probe the lebocin cDNA was amplified by polymerase chain reactions (PCR) using the primers, 5'-ACAAGCTGGCCAGGAACCGCTATGG-3' (5'-forward primer) and 5'-ACTTAAGTAATTACGATTCAATG-3' (3'-reverse primer). The resultant 482 bp DNA fragment was labeled with [ $\alpha$ - $^{32}$ P] dCTP (ICN) using a "DNA labeling kit" (Nippon Gene). Five

positive clones were obtained in the first screening. Screening conditions were as described previously (11). These clones were further screened with the same probe and two clones remained positive.

**Southern blot analysis.** Phage DNA was extracted from the two positive clones, designated FS-1 and FS-2. DNA sample from FS-1 was separately digested with *Apa*I, *Hind* III, *Kpn*I, *Sac*I, *Sac*II and *Xba*I. The resultant DNA fragments were electrophoretically separated on

B

```

CCTTTAGAGGCAGTGAATCCGTCATTCCCTAACCATGACGTAATTCGCGACTTTCGGGTACGATGCGAGGGCTTCAGGCAGGTCTCTTGC 90
ACTAACAGAATATCTATTAGATTGTCGCGGAGAAATTCGTATATTTATCGCGTTGCCGCGAGTCCGTTAGCATGTAGAAATGTTAAC 180
TTACGGGAATATGGTTTTCTCTACCGTTTCGCGCCATTGATTACCGGTAAGATTTATAACGTCAGGATACGACGATCGTAAACGTCCA 270
TGAGGTCAAAAGCGCGCGAGCCGCTGTTCGGGGGCGGGTCACTAGTAATTTATATGAAACATTGTAATACTATGTATTACTAATAC 360
TATGTATTGATTATTGTATTAGGTACGTATAGCGACAGTAATGAGGAAACAAATCATGAGAATTTTTCGGGTGAACCTTTGTGTCAATG 450
ATGAATACTGACTTCCCTTATCAATTCGTAATAATTAATATCGCTAAAGCGGATCGAAGTACAGACTTTCCTCAAGACAGTGTATAAA 540
TACCGATCATTGTGATAGTTACTCTCATTACCGTGCATCAATCGATACAATCAACATGTACAAGTTTGTAGTATTCAGTTTCAGTTCTGGT 630
M Y K F L V F S S V L V
GCTCTCTTTGCTCAGGCTTCGTGCGAGAGGTTCATCCAGCCGACCTACAGGCCACCAACGACCGCGCCGATAATACGTACAGCGCG 720
L F F A Q A S C Q R F I Q P T Y R P P P T R R P I I R T A R
ACAAGCTGGCGAGGAACCGCTATGGCTGTATCAAGGTGACAATATTCCTCGTGCGCCGAGTACTGCAGACCATCGATTCTTCCTCGAA 810
Q A G Q E P L W L Y Q G D N I P R A P S T A D H P I L P S K
AATCGACGATGTGAAGCTGATCCAAACAGAAGGTATGTTGCGAGTGTACCAATCCAGAAATAACGAGGCATCCATTGAATCTTCCCA 900
I D D V K L D P N R R Y V R S V T N P E N N E A S I E S S H
TCATACAGTTGATATTGGACTTGACCGCGCGATCGAGAGCCACCGTAACACAAAGGACCTGCGGTTTGGAAACCTCGAGAGAAACTGCC 990
H T V D I G L D R P I E S H R N T R D L R F W N P R E K L P
TCTTCCACGCTTCTCCGTTTAAACCCCAAGCCAATATATATTGATATGGGAAATCGTTACCGACGACATGCATCGGACGATCAAGAAGA 1080
L P T L P P F N P K P I Y I D M G N R Y R R H A S D D Q E E
ATTGCGGCATCATATAGACACTTTCTGATTCGAGGGATATTTTGCAAGACTAGGAAAAGTTCTCGAAACAAAAGATTTCGGAGTGCAC 1170
L R H H N E H F L I P R D I L Q D ***
TCTTTTATTTCATGAATCGTAATTACTTAAGCTTACTTGATATTTTAAATAGTTCTGCATTAAAAATACAAAATATTTTAAAGTGCCCGTG 1260
TCAATTCAAAATATGTAACTACTCGTACATGTATATCATGATAATATCAAAATTAATAAATCTCAACAAACATAGAATTCATATTTGT 1350
TTATTTGCTTTGCAACAAATCACCTATTGTAAAAATATTTTATAATTTGTGCTGTATATTTAAAGAAATACTTTGTGAATATTGATGTT 1440
TCTGTTTTCTTTTCGCCAACACTATACTGTAGTGTGACACTACACGTCCTCCACACGATTCTAATGACACCTCATATAATTGGAATC 1530
ATTCAACATGC 1541

```

FIG. 1—Continued

0.8% agarose S gel (Nippon Gene) and transferred onto a Gene Screen Plus membrane (Dupont). The membrane was hybridized with the same probe as described above, which was labeled with DIG-conjugated dUTP (Boehringer Mannheim). Among the positive DNA fragments, the shortest 4.3 kbp *Xba*I fragment was subcloned into a pBluescript II (SK<sup>+</sup>) vector (Stratagene). FS-2 DNA was similarly digested with *Apa*I, *Kpn*I, *Sac*II and *Xba*I, and Southern blotting was performed under the same conditions as FS-1. The shortest positive *Sac*I fragment (3.2 kbp) was subcloned into the same vector as described above.

**Nucleotide sequencing.** The 4.3 kbp *Xba*I fragment from FS-1 clone was further digested with *Xho*I and the resultant fragments (1 kbp and 3 kbp) were subcloned into pBluescript II (SK<sup>+</sup>) vector. Similarly, the 3.2 kbp *Sac*I fragment from FS-2 clone was also digested with *Xho*I and two DNA fragments (0.8 kbp and 2.4 kbp) were subcloned into the same vector described above. The nucleotide sequence of the 4 subcloned DNA fragments was determined by a dye terminator cycle sequencing method using a DNA sequencer (ABI 373A).

**Northern blot analysis.** Total RNA was extracted from fat bodies, hemocytes, mid guts, the Malpighian tubules and silk glands from the 5th instar larvae (4 days) non-immunized or immunized with LPS. Extracted RNA samples were electrophoresed on 1.2% agarose containing 6.6% formaldehyde (8). The RNA was transferred onto a Gene Screen Plus membrane (Dupont) and hybridized with probes for FS-1 (838 bp fragment from nucleotide number 1124 to 1962 of the genomic clone) and FS-2 (820 bp fragment from 596 to 1416 of

the genomic clone) which were labeled with DIG-conjugated dUTP (Boehringer Mannheim). Prehybridization and hybridization were carried out overnight at 68 °C for 2 h and 68 °C, respectively. Other conditions were as described previously (11).

**RT-PCR.** PCR based on cDNA (40 ng) synthesized with extracted RNA samples was performed under the following conditions; 94 °C for 1 min, 58 °C for 2 min and 72 °C for 3 min (30 cycles). The following primers were used; 5'-CCCGATAACACGTACAGT-3' (5'-forward primer) and 5'-CAACTGTATGATGTGAATG-3' (3'-reverse primer) for FS-1, and 5'-CCCGATAATACGTACAGC-3' (5'-forward primer) and 5'-CAACTGTATGATGGGAAGA-3' (3'-reverse primer) for FS-2. As an internal marker, the following primers based on the homologous region of actin genes from insects such as *B. mori* (12), *Drosophila melanogaster* (accession no. K00667, K00668 and K00669) and *Anopheles gambiae* (13) were synthesized and used. 5'-AGCAGGAGATGGCCACC-3' (5'-forward primer) and 5'-TCCACATCTGCTGGAAGG-3' (3'-reverse primer).

**Phylogenetic analysis.** Computer-aided phylogenetic analysis was performed using Genetyx-Mac software (unweighted pair-group method using arithmetic averages, UPGMA) provided by Software Development.

## RESULTS

### Cloning and Structure of *Lebocin* Genes

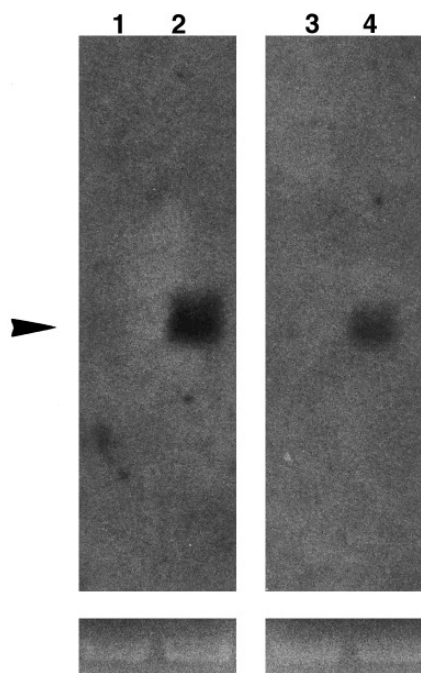
In order to examine gene expression mechanisms of anti-bacterial peptides in insects, we screened a *B. mori* geno-

mic library using a lebocin cDNA as a probe and obtained two distinct positive clones, designated FS-1 and FS-2.

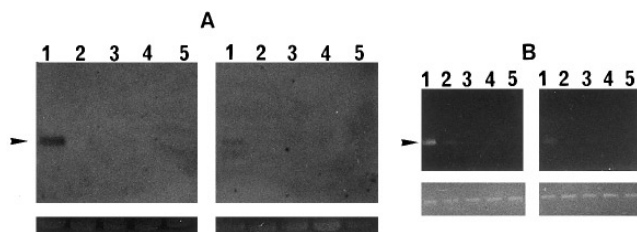
Nucleotide sequencing of FS-1 and FS-2 revealed that these genes do not contain introns (Fig. 1 A and B). Such intron-less genes are also seen in antibacterial peptides such as defensin (14) and dipterocin (15) from *D. melanogaster*. On the contrary, antibacterial peptide genes like cecropin B(16) from *B. mori*, cecropin A, B and D (17, 18) and attacins (19) from the giant silkworm, *Hyalophora cecropia*, cecropin A1, A2 and B (20) from *D. melanogaster* and Sarcotoxin I and II (21, 22) from the flesh fly, *Sarcophaga peregrina*, are known to contain 1-2 introns.

FS-1 and FS-2 had an open reading frame encoding 179 amino acids containing a probable signal peptide (16 amino acids), a putative prosegment (104 amino acids) and a mature peptide (32 amino acids) followed by 27 additional amino acids at its carboxyl-terminus. A poly-adenylation signal, AATAAA, was present at 182 and 180 nucleotides after translation stop codon (TAG) in FS-1 and FS-2, respectively.

Typical TATA, CAAT boxes and GATA motif (23) were present in the 5'-upstream regulatory regions of both genes (Fig. 1A and B). There are putative lipopolysaccharide (LPS) responsive elements in FS-1 and FS-2 (Fig. 1A and B), suggesting that the gene expression is inducible



**FIG. 2.** Inducibility of lebocin gene expression by LPS. Total RNA was extracted from silkworm larvae non-injected (lanes 1 and 3) or injected with LPS (lanes 2 and 4). The RNA sample (7  $\mu$ g) was electrophoresed and blotted onto a membrane. Northern blot analysis was performed on the membrane using PCR fragments of FS-1 (lanes 1 and 2) and FS-2 (lanes 3 and 4) as probes. An arrowhead indicates the position of positive signals. As an internal marker, 28S rRNA is shown in lower panels.



**FIG. 3.** Tissue specific expression of lebocin genes. The sites of lebocin gene expression were analyzed by both Northern blotting (A) using PCR fragments of FS-1 and FS-2 as probes and RT-PCR (B) using specific primers for two lebocin genes. In panel A and B, left and right panels indicate the results of FS-1 and FS-2, respectively. Total RNA was extracted from fat body (lane 1), hemocytes (lane 2), midgut (lane 3), the Malpighian tubule (lane 4) and silk gland (lane 5). The tissues were excised from *B. mori* larvae immunized with LPS. Five  $\mu$ g of RNA sample were used for Northern blot analysis. As an internal marker for Northern blotting, 28S rRNA is shown in lower panels (A). In addition, actin primers were synthesized and used for PCR internal marker as shown in lower panels (B). Arrowheads indicate the position of positive signals for both Northern blotting and PCR. Details of experimental conditions are described in MATERIALS AND METHODS.

by LPS from Gram-negative bacteria. In the upstream region of the CAAT box, an identical interleukin 6 (IL-6) responsive element (TGTAGAAAT), which is present in acute phase immune related genes in the mammalian immune system, was seen in FS-1 (Fig. 1A). However, one nucleotide of this element was lacking (TGTAGAAT) in FS-2 (Fig. 1B). Unlike *B. mori* cecropin B genes and attacin gene, a highly repetitive element, Bm 1(24), was not present in these lebocin genes.

#### Inducibility of Lebocin Gene Expression by LPS

Putative LPS responsive elements were present in FS-1 and FS-2, therefore, we examined inducibility of expression of these genes by a bacterial cell wall component, LPS. For this purpose, silkworm larvae (5th instar, 4 days) were injected with LPS and Northern blot analysis was performed on RNA samples extracted from fat bodies. As shown in Fig. 2, clear signals were found with FS-1 and FS-2 probes. Non-injected silkworm larvae were similarly investigated as a control and no positive signals were detected using both FS-1 and FS-2 as probes. The results suggest that expression of FS-1 and FS-2 is inducible by LPS.

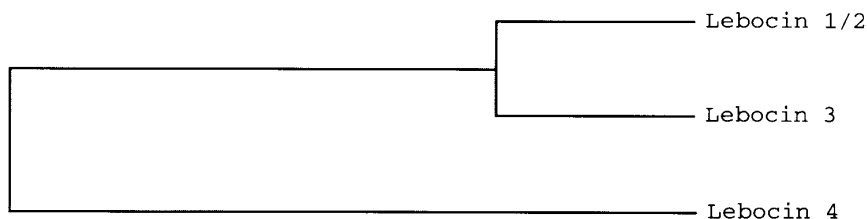
#### Tissue Specificity of Lebocin Gene Expression

Gene expression in different tissues was investigated by Northern blotting with RNA samples from fat bodies, hemocytes, midguts, the Malpighian tubules and silk glands. For this, silkworm larvae were injected with LPS. Results of Northern blot analysis revealed that a clear signal appeared in RNA samples from fat bodies by FS-1 probe and a weaker signal was observed with FS-2 probe (Fig. 3A). RT-PCR was

## A

Lebocin 1/2	1	: MYKFLVFSSVLVLFFAQASCQRFIQPTFRPPPTQRPIIRT
Lebocin 3	1	: MYKFLVFSSVLVLFFAQASCQRFIQPTFRPPPTQRPIIRT
Lebocin 4	1	: MYKFLVFSSVLVLFFAQASCQRFIQPTFRPPPTRRPIIRT
Lebocin 1/2	41	: ARQAGQEPLWLYQGDNPAPSTADHPILPSKIDDVQLDP
Lebocin 3	41	: VRQAGQEPLWLYQGDNPAPSTADHPILPSKIDDVQLDP
Lebocin 4	41	: ARQAGQEPLWLYQGDNPAPSTADHPILPSKIDDVQLDP
Lebocin 1/2	81	: NRRYVRSVTNPENNEASIEHSHHTVDIGLDQPIESHRNTR
Lebocin 3	81	: NRRYVRSVTNPENNEASIEHSHHTVDIGLDQPIESHRNTR
Lebocin 4	81	: NRRYVRSVTNPENNEASIEHSHHTVDIGLDQPIESHRNTR
Lebocin 1/2	121	: <b>DLRFLYPRGKLPVPTPPFPNPKPIYIDMGNRYRRHASDDQ</b>
Lebocin 3	121	: <b>DLRFLYPRGKLPVPTPPFPNPKPIYIDMGNRYRRHASDDQ</b>
Lebocin 4	121	: <b>DLRFWNPREKLPLPTLPPFPNPKPIYIDMGNRYRRHASDDQ</b>
Lebocin 1/2	161	: EELRQYNEHFLIPRDIFQE
Lebocin 3	161	: EELRQYNEHFLIPRDIFQE
Lebocin 4	161	: EELRHHNEHFLIPRDILQD

## B



**FIG. 4.** Sequence identity and phylogenetic relationships of three members of the lebocin family. Amino acid sequences of lebocins 1/2, 3 and 4 were aligned and compared (A). Identical amino acid residues in one letter are shadowed. Bold letters indicate a mature portion of lebocins. Numbers of amino acid residues are shown on the left margin of the panel. The phylogenetic relationship (B) was analyzed by the unweighted pair-group method using arithmetic averages (Software Development).

similarly carried out with the same RNA samples to detect low levels of transcripts in different tissues by highly sensitive methods. Results of RT-PCR indicated that both genes were expressed in hemocytes as well as fat bodies (Fig. 3B).

## DISCUSSION

Among numerous insect antibacterial peptides, lebocins are classified into a specific group that is characteristic for a proline-rich sequence (25). This proline-rich group contains abaecin (26) and apidaecin I and II (27) from the honey bee, *Apis mellifera*, drosocin (28) and metchnikowin (29) from *D. melanogaster*, pyrrhocoricin (30) from the sap-sucking bug, *Pyrrhocoris apterus*, metalnukowin I, IIA, IIB and III (31) from a bug, *Palomena prasina* and lebocin 1, 2 and 3 (7) from the silkworm, *B. mori*. Of these antibacterial peptides, drosocin, pyrrhocoricin and lebocin 1, 2 and 3 are known to carry an O-glycosylated substitution on the threonine

residue (7, 28, 30). To date, there is a report concerning structural analysis of a gene encoding unique O-glycosylated antibacterial peptide, drosocin (32). Cloning studies of the *Drosophila* gene indicated that it is a single and intronless gene (32).

We cloned lebocin genes and determined their nucleotide sequences. Amino acid sequences deduced from nucleotide sequences of FS-1 and FS-2 were compared to those of lebocin 1/2 and lebocin 3 (Fig. 4A). Results revealed that FS-1 encodes lebocin 3 and a deduced amino acid sequence from FS-2 indicated that it has 96.9% and 87.5% identity to sequences of the mature portion of lebocin 1/2 and lebocin 3, respectively. This suggests that FS-2 is a novel member of the lebocin gene family. Based on these results, we propose to designate lebocin gene 3 (*Leb 3*) and 4 (*Leb 4*) from FS-1 and FS-2 genomic clones, respectively. Overall identity of the deduced amino acid sequence from *Leb 4* shows 91.1% and 89.6% to those of lebocin 1/2 and lebocin 3, respectively, suggesting that these lebocin genes have

a very close structural relationship and have evolved from a common ancestral gene. A computer-aided analysis of the phylogenetic relationship of lebecin family exhibited that an original ancestral gene first duplicated into two groups, namely, a common gene for lebecin 1/2 and 3, and lebecin 4 gene, and then the former group was further divided into two groups (lebecin 1/2 and lebecin 3) (Fig. 4B).

As mentioned above, one of the unique characteristics of lebecins is that they have a sugar chain on a threonine residue (7). Our previous results indicated that the sugar chain plays an important role in expressing antibacterial activity (7). Interestingly, a possible O-glycosylation site (Pro-Thr/Ser-Xaa-Xaa-Pro) (33) is well conserved in lebecin 4 as well as in other peptides of lebecin family (Fig. 4A), suggesting that lebecin 4 is also an O-glycosylated antibacterial peptide.

Results of Northern blotting and RT-PCR demonstrated that expression of *Leb 3* and *Leb 4* in specific tissues is inducible by LPS. The results strongly suggest that these genes are involved in the insect immune system. Concerning gene expression, a question arises why is *Leb 4* expression weaker than in *Leb 3*. Comparison of nucleotide sequences of 5'-upstream regulatory regions in *Leb 3* and *Leb 4* show they contain TATA box, CAAT box, GATA motif and putative LPS responsive element. However, IL-6 responsive element in *Leb 4* is different from that of *Leb 3*. In addition, a number of the LPS responsive elements and GATA motif is also different in *Leb 3* and *Leb 4*. Although there is no direct evidence, it is speculated that these differences might reflect the level of gene expression. There may also be yet unknown regulatory elements such as enhancers in these genes. Details of expression mechanisms of lebecin genes and screening of the gene encoding lebecin 1/2 still remain to be explored.

## ACKNOWLEDGMENTS

We thank Dr. D. Taylor for critically reading the manuscript. This work was supported by Enhancement of Center of Excellence, Special Coordination Funds for Promoting Science and Technology, Science and Technology Agency, Japan and in part by Grants-in-Aid (Bio Media Program, BMP 97-V-1-5-2 and Glyco-technology Program) of the Ministry of Agriculture, Forestry and Fisheries, Japan.

## REFERENCES

- Boman, H. G., Faye, I., Gudmundsson, G. H., Lee, J.-Y., and Lindholm, D.-A. (1991) *Eur. J. Biochem.* **201**, 23–31.
- Hoffmann, J. A., Hetru, C., and Reichhart, J.-M. (1993) *FEBS Lett.* **325**, 63–66.
- Sun, S.-C., Lindström, I., Boman, H. G., Faye, I., and Schmidt, O. (1990) *Science* **250**, 1729–1732.
- Ladendorff, N. E., and Kanost, M. R. (1991) *Arch. Insect Biochem. Physiol.* **18**, 285–300.
- Hultmark, D. (1993) *Trend. Genet.* **9**, 178–183.
- Kimbrell, D. A. (1991) *BioEssays* **13**, 657–663.
- Hara, S., and Yamakawa, M. (1995) *Biochem. J.* **310**, 651–656.
- Chowdhury, S., Taniai, K., Hara, S., Kadono-Okuda, K., Kato, Y., Yamamoto, M., Xu, J., Debnath, N. C., Choi, H. K., Miyano-shita, A., Sugiyama, M., Asaoka, A., and Yamakawa, M. (1995) *Biochem. Biophys. Res. Comm.* **214**, 271–278.
- Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- Fujii, T., Sakurai, H., Izumi, S., and Tomino, S. (1989) *J. Biol. Chem.* **264**, 11020–11025.
- Kato, Y., Taniai, K., Hirochika, H., and Yamakawa, M. (1993) *Insect Biochem. Mol. Biol.* **23**, 285–290.
- Mounier, N., Gaillard, J., and Prudhomme, J. C. (1987) *Nucleic Acids Res.* **15**, 2781.
- Andersen, J. F., Utermohlen, J. G., and Feyereisen, R. (1994) *Biochemistry* **33**, 2171–2177.
- Dimarcq, J.-L., Hoffman, D., Meister, M., Bulet, P., Lanot, R., Reichhart, J.-M., and Hoffmann, J. A. (1994) *Eur. J. Biochem.* **221**, 201–209.
- Reichhart, J.-M., Meister, M., Dimarcq, J.-L., Zachary, D., Hoffmann, D., Ruiz, C., Richards, G., and Hoffmann, J. A. (1992) *EMBO J.* **11**, 1469–1477.
- Taniai, K., Kadono-Okuda, K., Kato, Y., Yamamoto, M., Shimabukuro, M., Chowdhury, S., Xu, J., Kotani, E., Tomino, S., and Yamakawa, M. (1995) *Gene* **163**, 215–219.
- Xanthopoulos, K. G., Lee, J.-Y., Gan, R., Kockum, K., Faye, I., and Boman, H. G. (1988) *Eur. J. Biochem.* **172**, 371–376.
- Gudmundsson, G. H., Lindholm, D.-A., Åsling, B., Gan, R., and Boman, H. G. (1991) *J. Biol. Chem.* **266**, 11510–11517.
- Sun, S.-C., Lindström, I., Lee, J.-Y., and Faye, I. (1991) *Eur. J. Biochem.* **196**, 247–254.
- Kylsten, P., Samakovlis, C., and Hultmark, D. (1990) *EMBO J.* **9**, 217–224.
- Kanai, A., and Natori, S. (1989) *FEBS Lett.* **258**, 199–202.
- Kanai, A., and Natori, S. (1990) *Mol. Cell. Biol.* **10**, 6114–6122.
- Kadalayil, L., Petersen, U.-M., and Engström, Y. (1997) *Nucleic Acids Res.* **25**, 1233–1239.
- Adams, D. S., Eickbush, T. H., Herrea, R. J., and Lizardi, P. M. (1986) *J. Mol. Biol.* **187**, 465–478.
- Wicker, C., Reichhart, J.-M., Hoffmann, D., Hultmark, D., Samakovlis, C., and Hoffmann, J. A. (1990) *J. Biol. Chem.* **265**, 22493–22498.
- Casteels, P., Ampe, C., Riviere, L., Damme, J. V., Elicone, C., Fleming, M., Jacobs, F., and Tempst, P. (1990) *Eur. J. Biochem.* **187**, 381–386.
- Casteels, P., Ampe, C., Jacobs, F., Vaeck, M., and Tempst, P. (1989) *EMBO J.* **8**, 2387–2391.
- Bulet, P., Dimarcq, J.-L., Hetru, C., Laguex, M., Charlet, M., Hegy, G., Van Dorsselaer, A., and Hoffmann, J. A. (1993) *J. Biol. Chem.* **268**, 14893–14897.
- Levashina, E. A., Ohresser, S., Bulet, P., Reichhart, J.-M., Hetru, C., and Hoffmann, J. A. (1995) *Eur. J. Biochem.* **233**, 694–700.
- Cociancick, S., Dupont, A., Hegy, G., Lanot, R., Holder, F., Hetru, C., Hoffmann, J. A., and Bulet, P. (1994) *Biochem. J.* **300**, 567–575.
- Chernysh, S., Cociancick, S., Briand, J.-P., Hetru, C., and Bulet, P. (1996) *J. Insect Physiol.* **42**, 81–89.
- Charlet, M., Laguex, M., Reichhart, J. M., Hoffmann, D., Braun, A., and Meister, M. (1996) *Eur. J. Biochem.* **241**, 699–706.
- Wilson, I. B. H., Gavel, Y., and von Heijne, G. (1991) *Biochem. J.* **275**, 529–534.