BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS **218,** 6–11 (1996)
Article No. 0002

Structure and Expression of the Tenecin 3 Gene in *Tenebrio molitor*[1]

Young Jae Lee, Tae Jin Chung, Chang Won Park, Yoonsoo Hahn, Jae Hoon Chung,[2] Bok Luel Lee,* Dong Min Han,† Young Hwan Jung,‡ Semi Kim,‡ and Younghoon Lee‡

*Department of Biological Sciences, and ‡Department of Chemistry, Korea Advanced Institute of Science and Technology, Taejon, 305-701, Korea; *College of Pharmacy, Pusan National University, Pusan, 609-735, Korea; and †Department of Molecular Biology, Wonkwang University, Iri, 570-749, Korea*

Received November 24, 1995

A genomic DNA fragment encoding tenecin 3, an antifungal protein, was cloned from the genomic DNA library of *Tenebrio molitor*. The DNA sequence analysis showed that the coding region is divided into two exons by an intron of 49 bp in the middle of the putative leader peptide coding region. Southern blot analysis suggests that the gene is present as a single copy. The transcription initiation site was determined by primer extension analysis and S1 mapping. The TATA box and CCAAT box sequences were found at −30, and −121, respectively, from the transcription initiation site. Tenecin 3 mRNA is abundantly expressed in larvae and adults, while little was detected in RNAs from pupae, suggesting that the expression of the tenecin 3 gene is developmentally regulated.

Insects have evolved to develop antimicrobial defense systems that are induced in hemolymph of infected insects. More than 50 antibacterial proteins have been purified from Hyalophoran, Lepidopteran, and Dipteran species, and their cDNA sequences characterized (reviewed in refs. 1, 2). According to their amino acid sequence similarity, insect antibacterial proteins are classified into four groups: cecropins (3), attacins (4), lysozymes (5), and defensins (6). In contrast to wealthy of information on these proteins, less is known about proteins with antifungal activity. In recent years, four antifungal proteins isolated from insects have been reported to inhibit growth of fungi. These are Antifungal Protein (AFP) (7) from *Sarcophaga peregrina,* holotricin 3 (8) from *Holotrichia diomphalia,* tenecin 3 (9) from *Tenebrio molitor,* and drosomycin (10) from *Drosophila melanogaster.* AFP, tenecin 3, and holotricin 3 share significant homologies with each other, but not with drosomycin.

Tenecin 3 was recently purified as a secreted hemolymph protein with an antifungal activity from *T. molitor* larvae and its cDNA was characterized (9). Tenecin 3 potentially caused inhibition of *Candida albicans* growth. It has distinctive features in its primary structure of 78 amino acids. Glycine, histidine, and glutamine constitute 80% of total residues. It also exhibits a regularity in the amino acid sequences with an 11-times repeated motif of Gly-X-X-Gly, where X can be His, Gln, or Leu. To gain information about the organization and expression of the tenecin 3 gene, in the present work we isolated the tenecin 3 gene. By analyzing the structure and expression of the tenecin 3 gene, we demonstrated that the gene is interrupted with an intron within the putative leader sequence region and that the transcript is abundantly synthesized in larvae and adults.

MATERIALS AND METHODS

Isolation and characterization of the T. molitor genomic clones. A lambda genomic DNA library was constructed in the λ FIXII vector(Stratagene) with Sau3AI-partial digested genomic DNA isolated from *T. molitor* larvae according to the Stratagene procedure. This library was screened with pSEC35, a cDNA clone for tenecin 3 (9), as a probe. Membranes were hybridized at 65°C in 0.5 M sodium phosphate, pH 7.2, 7% SDS, 1 mM EDTA, 1% BSA (fraction V), 100 μg/m*l* denatured

[1] The nucleotide sequence has been submitted to the GenBank with Accession No. U30627.
[2] To whom correspondence should be addressed. Fax: 82-42-869-2610.

0006-291X/96 $12.00

salmon sperm DNA. The membranes were washed twice in 40 mM sodium phosphate, pH 7.2, 5% SDS, 1 mM EDTA, 0.5% BSA (fraction V) and then washed twice in 40 mM sodium phosphate, pH 7.2, 1% SDS, and 1 mM EDTA at 65°C.

Positive clones were tested by PCR reactions. Each 1 μl liquid lysate was diluted with 100 μl of PCR buffer containing 100 pmol each of sense and antisense primers and 2 units of Taq DNA polymerase (Promega). The sense primer (GCG-GATCCTCTGGGTGGTCACCAAACCG) pairs with the nucleotides downstream of the putative leader sequence and a 5′ -GC tail followed by the BamHI recognition site, whereas the antisense primer (GCGGATCCAGGTCCATGCTGCCC-GTGTT) pairs with the nucleotides upstream of the termination codon and contains a 5′-GC tail followed by BamHI recognition site. PCR conditions were 94°C for 60 seconds, 42°C for 60 seconds, and 72°C for 60 seconds for 33 cycles. An about 4 kb EcoRI fragment from the insert DNA of a positive clones was subcloned into the pGEM-3Zf(+) vector and more than 2 kb region of the fragment was sequenced. The DNA sequencing was carried out by the dideoxy chain termination methods (11) using a Sequenase Version 2.0 kit (USB).

Primer extension analysis. Primer ten-p (GAATGTTTTCATTTTGGGATG) was end-labeled with [γ-^{32}P]ATP according to Sambrook *et. al.* (12). The labeled primer was added to larval total RNA, heated to 90°C for 10 min, and quenched on ice. The primer was annealed for 12 hr at 30°C in 30 μl of hybridization buffer (12 mM Tris-HCl, pH 7.0, 0.56 M NaCl, 80% formamide). The RNA/primer hybrid was ethanol-precipitated and dissolved in 25 μl of reverse transcription buffer (50 mM Tris-HCl, pH 8.3, 75 mM KCl, 3 mM $MgCl_2$, 10 mM DTT, 0.5 mM dNTPs) containing 2 units of RNasin (Promega). M-MLV Reverse Transcriptase (BRL) was added, and the incubation was continued for 90 min at 42°C, followed by a 30 min incubation in the presence of RNase A (0.2 μl of a 5 mg/ml stock solution). The extension products were then purified by phenol-chloroform extraction prior to electrophoresis on a sequencing polyacrylamide gel.

S1 mapping. Primer ten-p was end-labeled with [γ- ^{32}P]ATP according to Sambrook *et. al.* (12). Klenow fragment was used to extend the end-labeled 21mer ten-p across the 0.9 kb NdeI - BglII fragment of tenecin 3 genomic DNA to generate the 0.2 kb NdeI - ten-p fragment. Sixty microgram larval total RNA was coprecipitated with hybridization probe, dissolved in hybridization mix (80% deionized formamide, 40 mM PIPES (pH 6.4), 400 mM NaCl, 1 mM EDTA (pH 8.0)), and incubated for 12 hr at 25°C. This mixture was treated with S1 nuclease (BRL) and electrophoresed in a 6% polyacrylamide gel containing 7 M urea as described by Sambrook *et. al.* (12).

Southern blot and Northern blot analysis. Ten micrograms of *T. molitor* genomic DNA were digested with EcoRI HindIII, or PstI, electrophoresed on a 0.8% agarose gel, and transferred to Hybond N^+membrane (Amersham).

A total RNA isolation kit (Promega) was used to purify total RNAs from larvae, pupae, and adults. Twenty micrograms of RNA were electrophoresed in a 1% agarose gel and transferred to Hybond N^+ membrane (Amersham).

Hybridization was carried out with the [α- ^{32}P]dCTP-random primed tenecin 3 cDNA (9) in 0.5 M sodium phosphate, pH 7.2, 7% SDS, 1 mM EDTA, 100 μg/ml denatured salmon sperm DNA at 65°C. Membranes were washed twice in 40 mM sodium phosphate, pH 7.2, 1% SDS for 15 min each at 65°C.

RESULTS AND DISCUSSIONS

Structure of the tenecin 3 gene. We have previously isolated cDNA clones for tenecin 3, an antifungal protein isolated from *T. molitor* larvae (9). In order to obtain information about the organization of the tenecin 3 gene, we performed Southern blot hybridization using the cDNA clone. Total *T. molitor* DNA digested with EcoRI, HindIII, or PstI was electrophoresed, transferred to a membrane, and hybridized with the tenecin 3 cDNA probe. As shown in Figure 1, single hybridization signals at about 4 kb, about 1 kb and >10 kb were detected in the restriction fragments by EcoRI, HindIII, and PstI, respectively, suggesting that the tenecin 3 gene is a single copy gene. The gene for AFP, an antifungal protein closely related to tenecin 3, is also known to be a single copy gene (7). Therefore, the organizations of the antifungal protein genes are different from those of the genes for antibacterial proteins which are reported to be clustered in a genomic region (1, 13, 14).

Approximately 3.0×10^5 plaques of a genomic DNA library, made from larval DNA, were screened with the tenecin 3 cDNA as a probe. Three hybridization-positive clones (λTEN3-1, λTEN3-2, and λTEN3-3) were obtained. The presence of the tenecin 3 gene in these clones was confirmed by PCR with the tenecin 3-specific primers as described in Materials and Methods (data not shown). Since the 4 kb EcoRI DNA fragment had a hybridization signal in the Southern blot analysis, a 4 kb EcoRI DNA fragment from the insert DNA in one of the positive clones, λTEN3-1, was subcloned to generate pTEN3-1. The insert DNA in pTEN3-1 also had a 1 kb HindIII fragment DNA as would be expected from the Southern blot analysis. Therefore, the 1 kb (exactly 1070 bp) HindIII DNA fragment was initially sequenced and then its franking sequences were determined. As shown in Figure 2, the tenecin 3 coding sequence was located within the 1 kb HindIII DNA

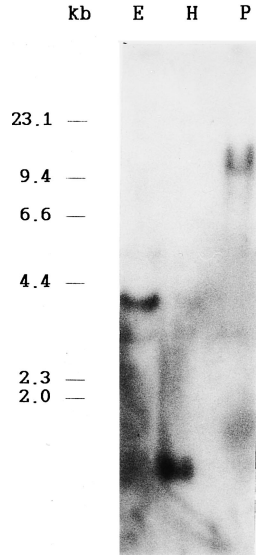

FIG. 1. Southern blot analysis of *T. molitor* genomic DNA. Ten micrograms of *T. molitor* genomic DNA were digested with EcoRI (E), HindIII (H), and PstI (P), respectively. The digested samples were electrophoresed, transferred to a nylon membrane, and hybridized with the ^{32}P-labeled tenecin 3 cDNA probe. The size markers are indicated in kb.

fragment. The nucleotide sequence of tenecin 3 genomic DNA agreed with that of cDNA (9), except that the genomic DNA contained a short (49 bp) intron in the middle of the putative leader peptide coding sequence (Figure 2). Insertion of an intron within the putative leader peptide coding sequence is rarely found in other insect genes (15, 16). The exon/intron boundary sequences, ATTCGTA---TTCAGGT, are in agreement with the consensus sequences (bold letters) at borders of introns of eucaryotic genes. As far as we know, the sequence for the tenecin 3 gene is the first reported genomic DNA sequence for antifungal proteins in insects.

Expression of the tenecin 3 gene. We have used primer extension analysis and S1 mapping to determine the initiation site of transcription in the tenecin 3 gene. In the primer extension analysis, we used the ten-p primer (21mer) which is complementary to the end of exon I (Figure 2). The primer was labeled, hybridized to larval total RNA, and then extended by reverse transcriptase. Figure 3a shows that an extended DNA fragment was 47 bp long. This result confirmed by the S1 mapping. In this experiment the 0.2 kb NdeI - ten-p fragment of tenecin 3 genomic DNA was hybridized with larval total RNA and then digested with S1 nuclease. Figure 3b shows a region of 46-47 protected bases. From these results the most likely initiation site is A at the position +1 in Figure 2. However, an insect-specific cap site (17) is not found in this region. The 5′ flanking region contains a TATAAA sequence at position −30 which is homologous to the TATA box consensus (18). The sequence CCAAT found at position −121 is the best match for the CCAAT box sequence (19).

The genes for insect antibacterial proteins have an upstream sequences related to the NF-κB-binding site (2, 14, 20–23). Phorbol myristate acetate (PMA) and lipopolysaccharide (LPS), which are often used as inducers of genes regulated by NF-κB, can also induce several antibacterial protein genes of insects (14, 22). It was, therefore, proposed that insect immune genes are regulated by the interaction between a transcriptional factor and a sequence related to the NF-κB-binding site (2, 14). However, a similar upstream sequence was not found within the upstream region (from −1 up to −849) of the tenecin 3 gene although a region (from + 851 to + 860) with a single mismatch to the consensus (2) was found within the downstream region. This difference suggests that expression of the tenecin 3 gene might be regulated in a different way from that of the antibacterial

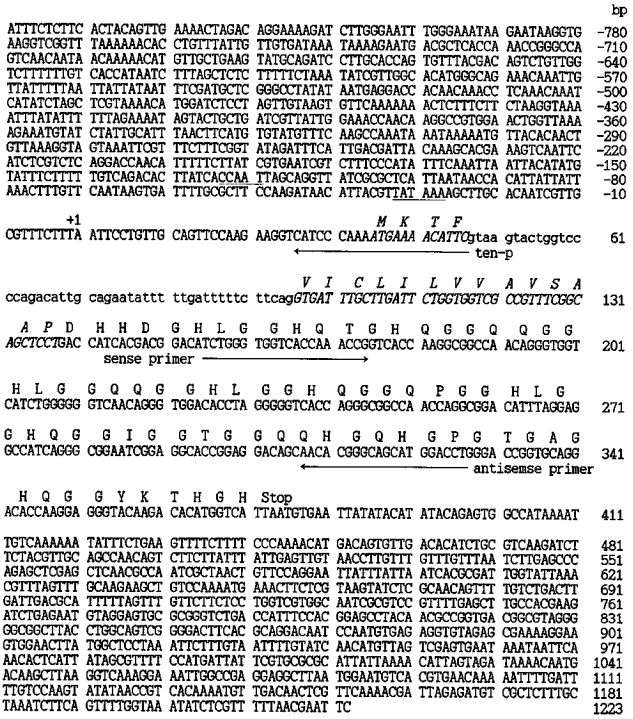

FIG. 2. Nucleotide and deduced amino acid sequences of the tenecin 3 gene. The deduced amino acid sequence is written above the nucleotide sequence. The putative leader sequence is italicized. TATA box and CCAAT box elements are underlined. The transcription initiation site at position +1 was assigned based on the primer extension analysis. Intron sequences are lower-case letters. Primers for PCR, primer extension analysis, and S1 mapping are displayed by arrows. Two HindIII recognition sites were located at positions −26 and +1044, respectively.

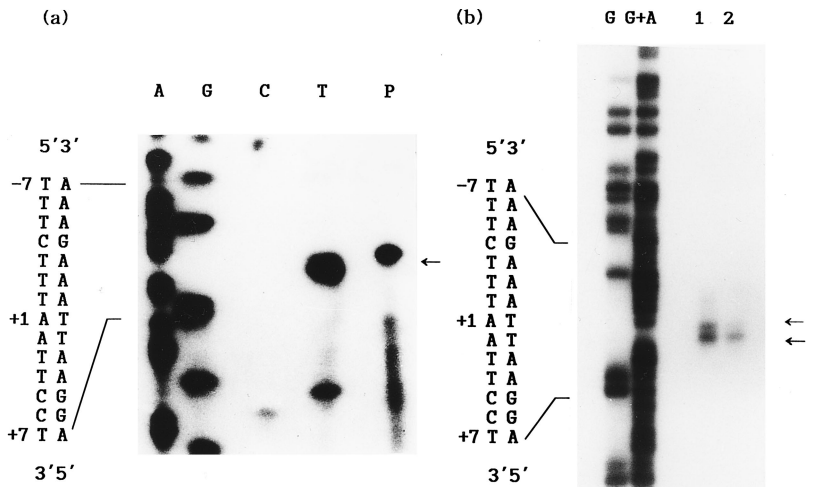

FIG. 3. Mapping of the 5′ end of tenecin 3 mRNA. The 5′ end of tenecin 3 mRNA was determined by primer extension analysis with ten-p primer (a) and S1 mapping (b). (a) The lanes marked as A,G,C, and T show the dideoxy sequencing ladders with the same primer from pTEN3-1 template and are used as size markers. The nucleotide sequence near 5′ ends of tenecin 3 mRNA is shown on the left. The major extension product is indicated by an arrow. (b) Lanes 1 and 2, the hybrid between RNA and the probe treated with 330 units/ml and 660 units/ml of S1 nuclease, respectively; lanes G and G+A, (G)-specific cleavage and (G+A)-specific cleavage fragments of the probe used as size markers, respectively. The nucleotide sequence near the protected fragments is shown on the left. The protected DNA fragments are indicated by arrows.

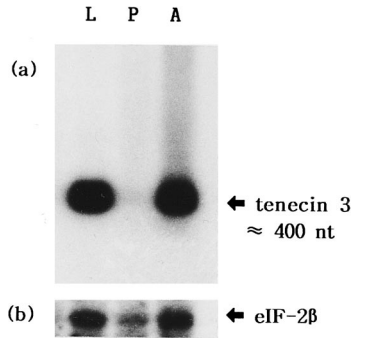

FIG. 4. Developmental analysis of tenecin 3 gene expression. Twenty micrograms of RNAs from larvae (L), pupae (P), and adults (A) was electrophoresed and transferred to nylon membrane. (a) The membrane was hybridized with the ^{32}P-labeled tenecin 3 cDNA probe. Tenecin 3 mRNA of about 400 nucleotides was indicated. (b) The membrane was reprobed using the ^{32}P-labeled eIF-2β coding sequence of *D. melanogaster* which is known to be expressed quasi-constitutively throughout *D. melanogaster* development (24). eIF-2β mRNA was indicated.

genes. This is consistent with the findings that various antibacterial proteins are induced or increased specifically in response to injury and infection (1, 2), whereas tenecin 3 is constitutively present in the hemolymph (9).

Developmental expression pattern of tenecin 3 gene was also investigated. Northern blot analysis was carried out with RNAs isolated from larvae, pupae, and adults (Figure 4). Tenecin 3 RNA was abundantly transcribed in larvae and adults, but little in pupae. This result suggests that the transcription of the tenecin 3 gene is controlled in a developmental stage-specific manner. Admitting that tenecin 3 plays an important role in host defence against fungi, the reduced expression of tenecin 3 at the pupal stage implies that a passive structural barrier can take the place of tenecin 3 in the pupal stage. Alternatively, tenecin 3 may have some other physiological roles during development.

ACKNOWLEDGMENT

This work is supported by a grant (KOSEF 92-50-00-04) from Korea Science and Engineering Foundation.

REFERENCES

1. Kimbrell, D. A. (1991) *BioEssays* **13,** 657–663.
2. Hultmark, D. (1993) *Trends Genet.* **9,** 178–183.
3. Steiner, H., Hultmark, D., Engström, A., Bennich, H., and Boman, H. G. (1981) *Nature* **292,** 246–248.
4. Hultmark, D., Engström, A., Andersson, K., Steiner, H., Bennich, H., and Boman, H. G. (1983) *EMBO J.* **2,** 571–576.
5. Boman, H. G., and Hultmark, D. (1987) *Annu. Rev. Microbiol.* **41,** 103–126.
6. Lambert, J., Keppi, E., Dimarcq, J.-L., Wicker, C., Reichhart, J.-M., Dunbar, B., Lepage, P., Van Dorsselaer, A., Hoffman, J., Fothergill, J., and Hoffman, D. (1989) *Proc. Natl. Acad. Sci. USA* **86,** 262–266.
7. Iijima, R., Kurata, S., and Natori, S. (1993) *J. Biol. Chem.* **268,** 12055–12061.
8. Lee, S. Y., Moon, H. J., Kurata, S., Natori, S., and Lee, B. L. (1995) *Biol. Pharm. Bull.* **18,** 1049–1052.
9. Jung, Y. H., Park, B. Y., Lee, D.-K., Hahn, Y., Chung, J. H., Han, D. M., Moon, H. J., Lee, B. L., and Lee, Y. (1995) *Mol. Cells* **5,** 287–292.
10. Fehlbaum, P., Bulet, P., Michaut, L., Lagueux, M., Broekaert, W. F., Hetru, C., and Hoffmann, J. A. (1994) *J. Biol. Chem.* **269,** 33159–33163.
11. Sanger, F., Nicklen, S., and Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74,** 5463–5467.
12. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) Molecular Cloning 2nd ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.
13. Kanai, A., and Natori, S. (1990) *Mol. Cell. Biol.* **10,** 6114–6122.
14. Sun, S.-C., Lindström, I., Lee, J.-Y., and Faye, I. (1991) *Eur. J. Biochem.* **196,** 247–254.
15. Hibner, B. L., Burke, W. D., and Eickbush, T. H. (1991) *Genetics* **128,** 595–606.
16. Galli, J., and Wieslander, L. (1993) *J. Biol. Chem.* **268,** 11888–11893.
17. Hultmark, D., Klemenz, R., and Gehring, W. J. (1986) *Cell* **44,** 429–438.

18. Breathnach, R., and Chambon, P. (1981) *Annu. Rev. Biochem.* **50,** 349–383.
19. Efstratiadis, A., Posakony, J. W., Maniatis, T., Lawn, R. M., O'Connell, C., Spritz, R. A., DeRiel, J. K., Forget, B. G., Weissman, S. M., Slightom, J. L., Blechl, A. E., Smithies, O., Baralle, F. E., Shoulders, C. C., and Proudfoot, N. J. (1980) *Cell* **21,** 653–668.
20. Lenardo, M. J., and Baltimore, D. (1989) *Cell* **58,** 227–229.
21. Xanthopoulos, K. G., Lee, J.-Y., Gan, R., Kockum, K., Faye, I., and Boman, H. G. (1988) *Eur. J. Biochem.* **172,** 371–376.
22. Sun, S.-C., Åsling, B., and Faye, I. (1991) *J. Biol. Chem.* **266,** 6644–6649.
23. Kylsten, P., Samakovlis, C., and Hultmark, D. (1990) *EMBO J.* **9,** 217–224.
24. Ye, X., and Cavener, D. R. (1994) *Gene* **142,** 271–274.