Journal of Biochemistry and Molecular Biology, Vol. 31, No. 6, November 1998, pp. 536–541

JOURNAL OF
Biochemistry
and
Molecular Biology
© BSRK & Springer-Verlag 1998

Molecular Cloning and Structural Analysis of the Antibacterial Gene from the Common Cutworm, *Spodoptera litura*

Duk-Hwa Nam and Gun-Sik Tae*
Department of Biological Sciences, Dankook University, Cheonan 330–714, Korea

Received 2 June 1998, Accepted 4 July 1998

The cDNA clone encoding the antibacterial peptide (SL-1) was isolated from the fat body of the common cutworm, *Spodoptera litura*, immunized with *E. coli* K12. The primary structure analysis revealed that its deduced amino acid sequence showed the characteristics of the cecropin family antibacterial peptides and that the amino acid residues highly conserved in the antibacterial peptides from moths and flies were also conserved, implying that SL-1 was a cecropin-like, and especially cecropin B-like, peptide. The predicted secondary structure of the mature SL-1 consists of three domains: (i) an amphiphilic α-helical domain (Ile-4 to Gly-18); (ii) the hinge region (Gly-23 and Pro-24); and (iii) a hydrophobic domain (Ala-25 to Ile-38).

Keywords: Antibacterial peptide, Cecropin, RT-PCR, *Spodoptera litura.*

Introduction

Insects have highly effective cell-free immune systems which are induced by a bacterial infection. The cecropin is one of the highly potent, inducible antimicrobial peptides that are synthesized in the fat body and secreted to the hemolymph, as part of the humoral immune response in insects (Boman, 1991; Hultmark, 1993). A cecropin was originally purified and sequenced from the Cecropia moth, *Hyalophora cecropia* (Hultmark *et al.*, 1980; Steiner *et al.*, 1981). Other cecropins have also been isolated from various sources, such as the immune hemolymph of other Lepidopteran and Dipterian insects (Hultmark, 1993), the small intestines of pigs (Lee *et al.*, 1989), and the cuticular matrix of the abraded integuments of the silkworm *Bombyx mori* (Brey *et al.*, 1993), implying that cecropins are ubiquitous in the animal kingdom and are expressed by various tissues. The insect cecropins are peptides of 35–39 amino acids which consist of a highly positively-charged N-terminal region and a long hydrophobic stretch of C-terminal region. The secondary structure of cecropin is a helix-bend-helix motif as predicted from the primary sequence analysis and two-dimensional NMR spectroscopy (Holak *et al.*, 1988). The amphipathic α-helical structure of the NH_2-terminus of this cecropin is important for its antibacterial activity, as well as the COOH-terminus ended with an α-amide group. The first two amino acids of cecropins, especially the highly conserved tryptophan in the second position, play an essential role in their functions. The hinge region between the NH_2- and COOH-terminal helices is formed by the Gly and Pro residues, which are conserved in all cecropin sequences (Andreu *et al.*, 1985).

The antibacterial spectra of cecropins are broad and include both Gram-negative and Gram-positive bacteria (Steiner *et al.*, 1981). However, the mode of action of cecropins and other antibacterial peptides is not yet fully understood. Their antibacterial activities appear to be due to their capabilities of membrane permeation rather than the cooperative formation of the transmembrane pores (Gazit *et al.*, 1994).

Recently, two antibacterial peptides, spodopsin Ia and Ib, were isolated and characterized from the immune hemolymph of the sixth instar larvae of the common cutworm, *Spodoptera litura* (Choi, 1996). They showed high homology to cecropins in their amino acid sequences. In this study, the gene responsible for antibacterial activity in the hemolymph of *S. litura* was searched for by use of the RT-PCR technique with the degenerated primers. The whole DNA sequence of a cecropin-like peptide was obtained and its primary and secondary structures were analyzed.

* To whom correspondence should be addressed.
Tel: 82-417-550-3446; Fax: 82-417-551-9229
E-mail: gtae@anseo.dankook.ac.kr

Materials and Methods

Insects The larvae and the pupae of *Spodoptera litura* were reared with a prepared artificial diet in an incubator (Hanback manufacture). Day and night shifts were monitored as a 12-h period. The humidity and temperatures were set at 69% and 26–28°C (night to day), respectively, and controlled as accurately as possible. The last instar larvae were used throughout this experiment.

Hemolymph collection and antibacterial activity assay The larva was injected with 2×10^7 *Escherichia coli* K12 into the thorax. Hemolymph was collected after 1 d, in ice-cooled tubes containing a few crystals of phenylthiourea, by cutting off the abdominal legs. The hemolymph was centrifuged at $10{,}000 \times g$ for 10 min to remove hemocytes and cell debris. The supernatant was used immediately to assay the antibacterial activity by measuring the clear growth inhibition zone on 1% thin agarose plate prepared with 0.1 M phosphate buffer, pH 6.4 (Yoe *et al.*, 1996). Three-mm diameter wells were prepared on the plates into which 3 μl of the supernatant was added per well. The plate was incubated at 37°C for 24 h and the diameter of the cleared zone was measured.

cDNA construction and polymerase chain reaction Total RNA was isolated from the fat body of the *E. coli*-challenged *S. litura* using the RNeasy prep. kit (Qiagen, Valencia, USA). The frozen fat body was ground in liquid nitrogen to a fine powder. The quality of the total RNA was monitored based on the intactness of rRNAs separated on 0.8% agarose gel. A complementary DNA (cDNA) was synthesized from 100 ng of the total RNA using AMV reverse transcriptase (Promega, Madison, USA) with a *Xho*I linker/oligo dT primer (Stratagene, La Jolla, USA) and was extracted with phenol/chloroform (Sigma, St. Louis, USA). A PCR was performed using cDNA as a template with one of the degenerated SL primers (forward primer) for cecropin genes and a GA primer (reverse primer) (Table 1). The SL primers were designed for the cecropin genes. Thirty-five cycles of PCR (each of 1 min at 95°C for denaturation of the cDNA/mRNA heteroduplex followed by 1.5 min at 47°C for the hybridization and 2 min at 72°C for chain elongation) were performed in a reaction medium including 0.8 mM dNTP, 1 M forward primer, 1 M reverse primer, and 2 U of *Taq* DNA polymerase. The PCR products were analyzed on 0.8% agarose gel.

Cloning of the RT-PCR product The band between 0.3 kb and 0.6 kb on 0.8% agarose gel was excised and the DNA was extracted with a Gene clean kit (Bio 101). The extracted DNA was ligated to pGEM T-vector (Promega) with T4 DNA ligase (Promega) in the presence of 10 mM DTT, 30 mM Tris-HCl, pH 7.8, 10 mM $MgCl_2$, and 1 mM ATP and the ligation mixture was transformed into JM109. A plasmid was extracted from a white colony grown in the presence of X-gal and further analyzed by the restriction digestion mapping. The fact that the plasmid with an insert has a larger molecular weight than a plasmid without an insert can be used to confirm the presence of an insert on agarose gel.

DNA sequence analysis The plasmid with a RT-PCR amplified DNA insert was extracted on a large scale with the Midiprep kit (Qiagen). The DNA sequence analysis of an insert was performed, using SP6 and T7 primer sites which were located at both ends of the pGEM T-vector multicloning site, with the automated DNA sequence analyzer (Applied Biosystem, Model 373A) in the Basic Science Research Center. The DNA sequence data obtained were blasted to NCBI (National Center for Biotechnology Information) through the internet to estimate the degree of homology to the known antibacterial peptides. The amino acid sequence alignment was performed with the shareware program, SeqPup.

Results and Discussion

Antibacterial activity measurement Hemolymph from uninjected last instar larva of *S. litura* contained a very low antibacterial activity. The injection of *E. coli* K12 into the thorax, however, increased the antibacterial activity of hemolymph rapidly, which started to increase 12 h after the induction and reached the maximum level after 24 h (data not shown). Since antibacterial peptides are known to be expressed in the fat body and secreted into hemolymph where they show their activities, the period of time required for the fat body to reach the maximum level of transcription might be different from that to reach the maximum level of the antibacterial activity in the hemolymph. In this experiment, total RNAs were extracted from the fat body of *S. litura* 8 h after the injection of *E. coli* K12 and used as a template for RT-PCR.

Reverse transcriptase-polymerase chain reaction The intactness of the mRNA extracted from the fat body of the *E. coli* challenged *S. litura* could be essential to make a high quality of cDNA, which could be shown indirectly by the presence of intact rRNA bands in 0.8% agarose gel (Fig. 1). In lanes 2 and 3, 28S and 18S rRNAs appeared as 2 distinct bands and there seemed to be no smeared bands nearby, indicating that mRNAs were intact. The oligo dT/*Xho*I linker primer with a GA extension (Table 1) was used to make cDNA as in Fig. 2. Since almost all of mRNAs from eukaryotic cells have the polyadenylate tails, the dT portion of an oligodT/*Xho*I linker primer binds to it, and the reverse transcriptase is able to make cDNA from the 3′-OH of the primer, resulting in the formation of a DNA–RNA heteroduplex.

The polymerase chain reaction (PCR) was performed with a GA primer and one of SL primers using cDNA as a template. The cDNA was prepared from the DNA–RNA heteroduplex by heat denaturation. In this experiment, the use of a GA primer was planned to increase the specificity of PCR, since the GA primer could bind only on cDNA, and not on mRNA nor on the chromosomal DNA. The degenerated SL primers (1–3) were designed on the basis of highly conserved DNA sequences of the antibacterial peptide genes from moths. The SL primers 1 and 2 were designed to bind the presequence of antibacterial genes, whereas the primer 3 was to bind the mature sequence. The

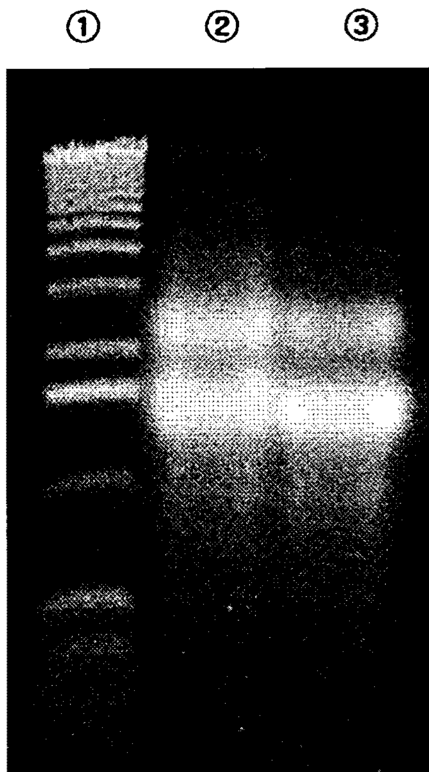

Fig. 1. Intactness of total RNA isolated from the *E. coli* challenged fat body of *S. litura*. The total RNA was isolated from the fat body of *S. litura* with RNeasy spin column (Qiagen) after being challenged by *E. coli*. Lane 1 is the 1 kb ladder size marker (BRL). Lanes 2 and 3 are the total RNAs isolated from the normal fat body and from the *E. coli* induced fat body, respectively.

Table 1. The nucleotide sequences of the SL primers. Three different kinds of the degenerated SL primers were designed to amplify the genes for antimicrobial peptides from cDNA of *S. litura*. The degenerated bases are represented by two or three letters instead of one. The SL primers have been designed on the basis of DNA sequences of the regions that show high degree of homologies among *H. cecropia, B. mori,* and *M. sexta*.

Primer	Nucleotide Sequence (5′ to 3′)
Oligo dT/*Xho*I linker primer	GAGAGAGAGAGAGAGAGAGA ACTAGTCTCGAGTTTTTTTT TTTTTTTTTT
GA primer	GAGAGAGAGAACTAGTCTCG
SL primer 1	TG G G C A C C A GATC T T CTTCG CT A A T T T
SL primer 2	AC A C C AT T T C TTCGT TTCGC CT T T G
SL primer 3	A AA T GTG TTCAAGA T GAGAAAA TGG G GT G ACC

gene for the antibacterial peptide was first amplified by PCR with the GA primer and the SL primers 1 or 2. The primer 2 has turned out to work better in this experiment. It might be due to the lower frequency of the degenerated bases in its nucleotide sequence (Table 1). The DNA fragment, generated by the first PCR and separated on agarose gel, was used as the template for the second PCR. Since the binding site of a SL primer 3 was located at the downstream of the gene for the preprosequence of cecropins, PCR with the SL primer 3 could generate the smaller DNA band (data not shown), of which the size was approximately 400 bp. The pGEM T-easy vector (Promega) with the additional dT residues in its *Eco*RV cleavage site, which could increase the blunt end ligation efficiency at least 10 times, was used to clone the RT-PCR product. In the presence of X-gal in a plate, the plasmid was extracted from each of white colonies formed and digested with the restriction enzyme, *Eco*RI. Since pGEM T-vector has two *Eco*RI cleavage sites which are flanking the insert, the *Eco*RI digestion made it possible to confirm the presence and the size of an insert.

Primary structure of SL-1 The cloned antibacterial gene of *S. litura* has an open reading frame of 171 nucleotides corresponding to 56 amino acid residues and the termination codon, TGA (Fig. 3). The peptide, whose amino acid sequence was deduced from the cloned gene, was named as a cecropin-like peptide, SL-1 (hereafter, referred to as SL-1). The amino acid sequence of SL-1 showed the characteristics of the cecropin family proteins, which were composed of the preprosequence and the mature sequence. The SL-1 has the full lengths of the mature and the prosequence, and the presequence without five amino acid residues from the NH_2-terminus. As for the homology of the mature sequence of SL-1 to cecropins, it showed 70%, 73%, and 49% identities to cecropin A, B, and D, respectively. Therefore, SL-1 is more similar to cecropin B. The NH_2-terminal region of SL-1 is more conserved than the COOH-terminal region as in cecropins (Boman *et al.*, 1985), implying that it is important for its antibacterial activity. The second amino acid residue, Trp, of the mature sequence was also conserved in SL-1. The Trp is believed to be essential for its antibacterial function and in both moths and flies it is completely conserved. In addition, Lys (position-6), Glu (position-9), Gly (position-12), Arg (position-16), Ala (position-32), and Gly (position-36), which were all 100% conserved residues in the antibacterial peptides of moths, were also conserved in SL-1 (Fig. 4). The possible splice site, Ile-Glu (positions 8 and 9), is also conserved. The residues, Ala-Gly-Pro-Ala (positions 22 to 25) were also highly conserved in moths except for *M. sexta* (Fig. 4). It showed higher homology to those antibacterial peptides from moths than those from flies (data not shown). In particular, Gly (position-23) and Pro (position-24) are highly conserved in moths, but not in

```
                Exon1          Exon2
        ====[=====>]====[=====>]=====   Genomic DNA

                    | Transcription
                    v

mRNA  ----[==========>]------AAAA(A)n
                      3'-TTTT(T)14GAGCTCTGATCAAGA(GA)9-5'
                            XhoI linker/primer

                    | Reverse Transcriptase
                    v

mRNA  ----[==========>]------AAAA(A)n
cDNA  ====[==========>]======TTTT(T)14GAGCTCTGATCAAGA(GA)9-5'

                    | Heating at 95°C for 2 min
                    v
             1
cDNA  ====[==========>]======TTTT(T)14GAGCTCTGATCAAGA(GA)9-5'
                                                   GA

                    | PCR with primer 1 and GA primer
                    v

          [==========>]======AAAA(A)14CTCGAGACTAGTTCT(CT)9-3'
                             TTTT(T)14GAGCTCTGATCAAGA(GA)9-5'

                    |
                    v

          Subcloning into pGEM T-easy Vector
```

Fig. 2. Schematic diagram of RT-PCR. The oligo dT/*Xho*I linker primer with the GA extension was used to make cDNA. The dT portion of the primer binds to the poly(dA) tail and the reverse transcriptase were used to make cDNA from the 3'-OH of the oligo dT/*Xho*I linker primer, resulting in the cDNA–mRNA heteroduplex. The polymerase chain reaction was performed with a GA primer and a SL primer 1 h after heat-denaturing the cDNA–mRNA heteroduplex. The DNA fragments for the antibacterial gene, separated on agarose gel, were subcloned to pGEM T-easy vector.

```
1                                                         15
 I   L   S   F   V   F   A   L   V   L   G   M   S   A   V
ATA TTA TCC TTC GTG TTC GCC TTG GTT CTT GGC ATG TCT GCT GTA

16                  → Mature sequence                     30
 S   A   A   P   K   W   K   I   L   K   K   I   E   K   V
TCA GCA GCT CCC AAA TGG AAG ATT CTT AAG AAA ATT GAA AAA GTC

31                             •                          45
 G   R   N   A   R   D   G   I   I   K   A   G   P   A   I
GGA AGG AAC GCC CAT GAT GGT ATT ATC AAG GCG GGA CCA GCG ATA

46                                          56
 Q   V   L   G   Q   A   K   A   I   G   K   *
CAA GTG CTG GGA CAG GCG AAA GCG ATT GGA AAA TGA
```

Fig. 3. Nucleotide and its deduced amino acid sequence of the antibacterial gene cloned from the fat body of *S. litura*. The deduced amino acid sequence contains the full length of the mature sequence as indicated by an arrow as well as the preprosequence without 5 amino acid residues from the NH_2-terminus. The stop codon, TGA, is written in bold letter.

```
B. mori CecA          RWKLFKKIEKVGRNVRDGLIKAGPAIAVIGQAKSL..GK
B. mori CecB          RWKIFKKIEKMGRNIRDGIVKAGPAIEVLGSAKAI..GK
T. ni CecA            RWKFFKKIEKVGQNIRDGIIKAGPAVAVVGQAASIT.GK
H. cecropia CecA      KWKLFKKIEKVGQNIRDGIIKAGPAVAVVGQATQIAKG.
H. cecropia CecB      KWKVFKKIEKMGRNIRNGIVKAGPAIAVLGEAKAL..G.
H. cecropia CecD      .WNPFKELEKVGQRVRDAVISAGPAVATVAQATALAKGK
A. pernyi D           .WNPFKELERAGQRVRDAIISAGPAVATVAQATALAK..
M. sexta Cec D-like   .WNPFKELERAGQRVRDAVISAA.AVATVGQAAAIARGG
S. litura SL1         KWKILKKIEKVGRNARDGIIKAGPAIQVLGQAKAI..GK
```

Fig. 4. Multiple amino acid sequence alignment of the SL-1 peptide with the known antibacterial peptides. The SL-1 peptide from *S. litura* was aligned with the antibacterial peptides from *Bombyx mori, Trichoplusia ni, Hyalophora cecropia,* and *Manduca sexta.* The highly conserved amino acid residues are shaded.

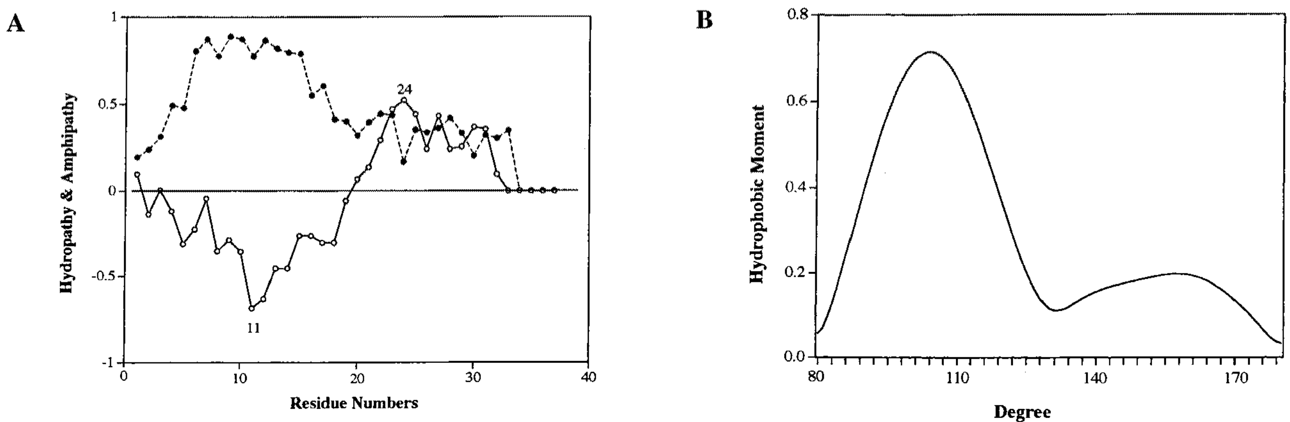

Fig. 5. Hydrophobicity and amphipathicity plots of the mature SL-1 and hydrophobic moment analysis of the amino terminal domain of SL-1. A. Hydrophobicity (o) and amphipathicity (•) of the mature SL-1 was calculated and plotted based on the Eisenberg consensus index and a window size of 11, according to the algorithm discussed by Shiver *et al.* (1989). The total number of amino acid residues analyzed is 39. B. Hydrophobic moment analysis of the amino terminal segment of SL-1 (Ile-4 to Gly-18), analyzed from 80° to 180° with a residue window size of 11. The peak was positioned at 104°.

flies in which the aliphatic amino acids such as Ile and Val are conserved. Pro or Gly is generally considered as helix breakers. The dipeptide, Gly-Pro, is thought to be involved in the formation of the hinge between NH_2- and COOH-terminal helices of SL-1, which may confer the flexibility on two helices.

The amino acid sequence comparison of SL-1 to spodopcin Ia and Ib, another pair of antibacterial peptides whose genes were also cloned from *S. litura* (Choi, 1996), revealed that SL-1 showed more than 73% of homology and it increased up to 95% if the property of the amino acid side chain was considered. The reason why *S. litura* has at least three almost identical genes for the antibacterial activity can be explained by the facts that each peptide may vary its antibacterial spectra and that the presence of bacteria inside insects, regardless of their species, is able to induce the antibacterial peptides in *S. litura* (Choi, 1996). As a part of highly effective cell-free immune systems of insects, the production of antibacterial peptides is likely to be controlled at the transcriptional level.

Predicted secondary structure of SL-1 The hydropathy analysis revealed that the mature sequence of SL-1 antibacterial peptide consisted of two domains: the hydrophilic NH_2-terminal domain and the hydrophobic hinge region and the COOH-terminal domain (Fig. 5A). When the hydrophilicity of the NH_2-terminal region of the mature SL-1 was also calculated with a consensus Einsenburg index for each amino acid residues with a window size of 11 amino acids, it turned out that the amino-terminal domain was not only hydrophilic, but also amphiphilic, in which one or two charged amino acid residues such as Lys and Glu are alternatively positioned among the hydrophobic residues such as Val, Ala, and Ile. (Fig. 3). Based on the hydrophobic moment calculation (Shiver *et al.*, 1989), the angle between the amino acid side chains of the NH_2-terminal domain from Ile-4 to Gly-18 is approximately 104° (Fig. 5B), implying that the NH_2-terminal domain forms an amphiphilic α-helix, in which the charged residues, especially positively-charged residues, are facing one side of an α-helix, and the hydrophobic residues are on the other side, as shown in Fig. 6.

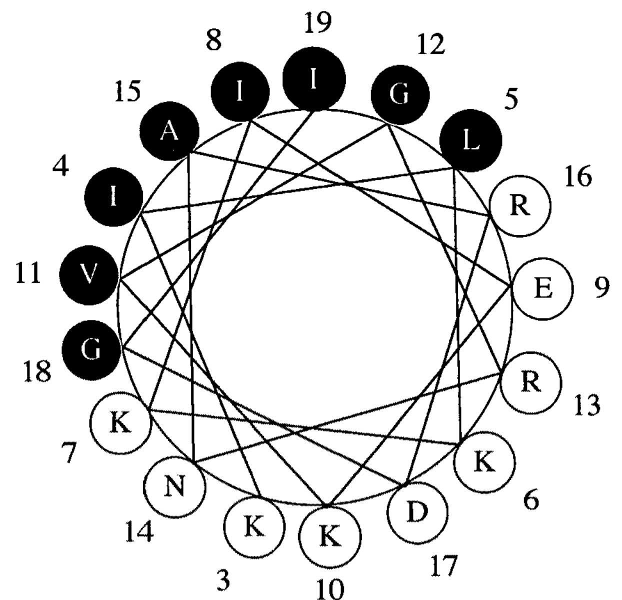

Fig. 6. The helical wheel representation of the NH_2-termainal domain of the mature SL-1. On the assumption that the segment from Ile-4 to Gly-18 forms an α-helix, the side chain of each amino acid residue is protruded outward and the angle between them is 104°. The hydrophobic and hydrophilic amino acids residues are represented by shaded and opened circles, respectively, with the name inside. The residue numbers are given outside of the wheel.

The positively-charged face of an α-helix of the SL-1 amino-terminal domain could bind phospholipid membranes as monomers lying on the surface, as shown in a spectrofluorometric study of the cecropin B (Gazit *et al.*, 1994), and the hydrophobic COOH-terminal domain could get inserted into the membranes by contacting the hydrophobic residues with the aliphatic phase of the membrane. However, the detailed mechanism is not yet clear. In order to gain insight into the structure and function relationships of the antibacterial peptide in *S. litura*, a three-dimensional structure determination of SL-1 will be necessary in the future, in addition to the primary structure deduced from the gene we have cloned.

Acknowledgments This work was supported by a research grant from the Korean Ministry of Education (BSRI-96-4443).

References

Andreu, D., Merrifield, R. B., Steiner, H. and Boman, H. G. (1985) N-terminal analogs of cecropin A: synthesis, antibacterial activity and conformational properties. *Biochemistry* **24**, 1683–1688.

Boman, H. G. (1991) Antibacterial peptides: key components needed in immunity. *Cell* **65**, 205–207.

Boman, H. G., Faye, I., Hofsten, P. V., Kockum, K., Lee, J.-Y., Xanthopoulos, K. G., Bennich, H., Engström, Merrifield, R. B. and Andreu, D. (1985) On the primary structures of lysozyme, cecropins and attacins from *Hyalophora cecropia*. *Dev. Comp. Immunol.* **9**, 551–558.

Brey, P. T., Lee, W.-J., Yamakawa, M., Koizumi, Y., Perrot, S., Francois, M. and Ashida, M. (1993) Role of the integument in insect immunity: Epicuticular abrasion and induction of cecropin synthesis in cuticular epithelial cells. *Proc. Natl. Acad. Sci. USA* **90**, 6275–6279.

Choi, C. S. (1996) Purification and cDNA sequences of antibacterial peptides, spodopsins Ia and Ib from the haemolymph of *Spodoptera litura*. Ph.D. thesis, Korea University.

Gazit, E., Lee, W.-J., Brey, P. T. and Shai, Y. (1994) Mode of action of the antibacterial cecropin B2: A spectrofluorometric study. *Biochemistry* **33**, 10681–10692.

Holak, T. A., Engström, Ä., Kraulis, P. J., Lindeberg, G., Bennich, H., Jones, T. A., Gronenborn, A. M. and Clore, G. M. (1988) The solution conformation of the antibacterial peptide cecropin A: a nuclear magnetic resonance and dynamical simulated annealing study. *Biochemistry* **27**, 7620–7629.

Hultmark, D. (1993) Immune reactions in *Drosophila* and other insects: a model for innate immunity. *Trends Genet.* **9**, 178–183.

Hultmark, D., Steiner, H., Rasmuson, T. and Boman, H. G. (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.

Lee, J.-Y., Boman, A., Chauanxin, S., Andersson, M., Jornvall, H., Mutt, V. and Boman, H. G. (1989) Antibacterial peptides from pig intestine: Isolation of a mammalian cecropin. *Proc. Natl. Acad. Sci. USA* **86**, 9159–9162.

Shiver, J. W., Peterson, A. A., Widger, W. R., Furbacher, P. N. and Cramer, W. A. (1989) Prediction of bilayer spanning domains of hydrophobic and amphipathic membrane proteins: Application to the cytochrome *b* and colicin families. *Methods Enzymol.* **172**, 439–461.

Steiner, H., Hultmark, D., Engström, Ä., Bennich, H. and Boman, H. G. (1981) Sequence and specificity of two antibacterial proteins involved in insect immunity. *Nature* **292**, 246–248.

Yoe, S. M., Bang, I. S., Kang C. S. and Kim J. J. (1996) Purification and characterization of two lysozymes from larval haemolymph of cabbage butterfly, *Artogeia rapae*. *Mol. Cells* **6**, 609–614.