Egypt. Acad. J. Biolog. Sci., 5(1): 31-44 (2012) **A. Entomology**
Email: egyptianacademic@yahoo.com **ISSN: 1687–8809**
Received: 11/2/2012 **www.eajbs.eg.net**

Molecular Characterization, Full Length Isolation and Phylogenetic Analysis of C-type Lectin Gene from Bacterial-challenged Cotton Leafworm, *Spodoptera littoralis*

AlaaEddeen M. Seufi
Entomology Department, Faculty of Science, Cairo University, Giza, Egypt

ABSTRACT

Experiments were designed to investigate the molecular immune response of *Spodoptera littoralis* larvae against bacterial infection. In addition, sequence and phylogenetic analyses of the involved gene were studied. Using differential display technique, a partial insect lectin gene (*Spli*Lec) was isolated from bacterial-challenged *S. littoralis* haemolymph. Five differentially displayed bands were sequenced. Sequence results revealed that a fragment of 640 bp was amplified within the open reading frame (*orf*) of a lectin gene. This fragment contained the complete 3` end with a poly(A) tail, but it lacks start codon, AUG at its 5` end. Using RACE PCR reaction, 5` end was extended and a final reaction was performed to obtain the full length of the *Spli*Lec. Sequence analyses of the data revealed that *Spli*Lec consists of a single *orf* encoding a deduced polypeptide consisting of a 18-residue signal peptide and a 291-residue mature peptide. *Spli*Lec sequence contained two CRDs: short form CRD_1 and long form CRD_2 stabilized by two and three highly conserved disulfide bonds, respectively. *Spli*Lec shares homology with some dipteran lectins suggesting possible common ancestor. These results suggested an important role of the *Spli*Lec gene in cell adhesion and non-self recognition. It may cooperate with other AMPs in clearance of invaders of *Spodoptera littoralis*.

Keywords: Immune system, *Spodoptera*, C-type lectin, CRD, PAMP.

INTRODUCTION

After pathogens penetrate the insects' structural barriers, they rely solely on an efficient innate immune system which shares many characteristics with the innate immune system of vertebrates. Insect innate immune system comprises both humoral and cellular responses (Pinheiro and Ellar, 2006, Lemaitre and Hoffmann, 2007). Insect humoral defenses include the production of a potent arsenal of antimicrobial peptides (AMPs) (Pinheiro and Ellar, 2006, Lemaitre and Hoffmann, 2007), coagulation, and melanization led by protease cascades (Kanost *et al.*, 2004). Insect cellular defense refers to haemocyte-mediated immune responses, such as phagocytosis, nodulation, and encapsulation (Lavine and Strand, 2002). The encapsulation process involves cell adhesion and melanization (Eslin and Prevost, 2000).

Lectins are an important class of carbohydrate-binding proteins that have several distinct biological activities. They mediate cell adhesion (i.e. bind to microbial surface components), non-self recognition and immuno-protection processes in immune responses (Vasta *et al.,* 1999). They exist in a wide variety of plants, animals, fungi, bacteria and viruses (Sharon, 1977) and play significant role in clearance of invaders, either as cell surface receptors for microbial carbohydrates or as soluble proteins

existing in tissue fluids (Yu and Kanost, 2003). Such proteins are known as pattern recognition receptors (PRPs), because they bind to the pathogen associated molecular patterns (PAMPs) present in the array of carbohydrate components on the surface of microorganisms and consequently, trigger a series of protective immune responses (Medzhitov and Janeway, 2002). Various proteins that display carbohydrate-binding activity in a calcium-dependent manner are classified into the C-type lectin family (Drickamer and Taylor, 1993). They contain C-type carbohydrate-recognition domains (CRDs) or C-type lectin domains (CTLDs) composed of 110–130 amino acid residues in common. These CRDs or CTLDs contain a characteristic double-loop (loop in a loop) stabilized by two or three highly conserved disulfide bonds. The vertebrate C-type lectins are usually multi-domain lectins and they fall into seven groups (I-VII) (Day, 1994). Seven new groups (VIII-XIV) were added in the revised classification in 2002 (Drickamer and Fadden, 2002) and three new groups (XV-XVII) were updated, recently (Zelensky and Gready, 2004). In contrast, the invertebrate C-type lectins are mostly single-domain proteins, but C-type lectins that contain two CRDs are characterized too. Although all C-type lectin CRDs have sequence similarity, they can be divided into two types: a "short form" approximately 115 residues long and a "long form" approximately 130 residues long, which includes two additional disulfide-bonded cysteine residues at the amino terminus (Drickamer and Taylor, 1993, Day, 1994). In recent years, more and more C-type lectins with two tandem CRDs have been identified and characterized from invertebrates, especially from insects (Yu and Kanost, 2000, Yu *et al*., 2005, Tian *et al.*, 2009). Examples of the C-type lectins with two tandem CRDs include the *M. sexta* immunolectins (IML-1, IML-2, IML-3 and IML-4) which serve as humoral PRPs (Kanost *et al.*, 2004), LPS-binding lectins from the silkworm, *Bombyx mori* (Koizumi *et al.*, 1999) and the fall webworm, *Hyphantria cunea* (Shin *et al.*, 2000).

In this paper, the full length cDNA of a C-type lectin with two tandem CRDs from *S. littoralis*, was isolated using differential display and RACE PCR techniques. Sequence characterization and phylogenetic analyses were reported, too.

MATERIALS AND METHODS

Insects and bacterial strains

Laboratory colony of the cotton leafworm, *S. littoralis*, used for our experiments was originally collected from a private okra field at Giza, Egypt in 1995 and maintained in the insectary of the Department of Entomology, Faculty of Science, Cairo University according to the technique described by Levinson and Navon (1969) and kept at 25 °C, 65–70% RH and 14L: 10D photoperiod cycle.

Two gram (+) bacteria, *Staphylococcus aureus* and *Streptococcus sanguinis* and three gram (-) bacteria, *Escherichia coli* (D_{31}), *Proteus vulgaris* and *Klebsiella pneumoniae* were obtained from the Unit for Genetic Engineering and Agricultural Biotechnology, Faculty of Agriculture, Ain Shams University and used for insect immunization. Bacteria were grown in a peptone medium (1%), supplemented with 1% meat extract and 0.5% NaCl, at 37 °C in a rotary shaker.

Insect immunization and haemolymph collection

Bacterial challenge was performed as described by Seufi *et al.* (2011). Haemolymph was collected at 24, 48 and 72 h post-infection (p.i.) at 4 °C (500 µl/ each), containing few crystals of phenylthiourea to prevent melanization. Aliquots of 100 µl each were stored at -80 °C until investigated. Control group was injected with bacteria-free saline solution.

RNA extraction and reverse transcription

Total RNA of the insect haemolymph (300-500 µl) was extracted using RNeasy kit according to the manufacturer's instructions (Qiagen, Germany). Residual genomic DNA was removed from RNA using RNase-free DNase (Ambion, Germany). RNA integrity and purity were justified by examining 260/280 and 260/230 ratios for protein and solvent contamination. Reverse transcription reaction was carried out according to the ABgene protocol (ABgene, Germany). The cDNA was aliquoted and stored at -80 °C until processed.

Differential display using primers corresponding to lectin sequence (DD-PCR)

A total reaction volume of 25 µl containing 2.5 µl PCR buffer, 1.5 mM $MgCl_2$, 200 µM dNTPs, 1 U *Taq* DNA polymerase (AmpliTaq, Perkin-Elmer), 2.5 µl of 10 pmol primer (Table 1) and 2.5 µl of each cDNA was cycled in a DNA thermal cycler (Eppendorf, Mastercycler 384, Germany). The amplification program was one cycle at 94 °C for 5 min (hot start), followed by 40 cycles at 94 °C for 1 min, 40 °C for 1 min and 72 °C for 1 min. The reaction was then incubated at 72 °C for 10 min for final extension. PCR product was visualized on 1.5 % agarose gel and photographed using gel documentation system. For DNA contamination assessment, a no-reverse transcription control reaction was performed.

Primer design and RT-PCR

Five reproducible bacterial-induced bands were eluted, cloned in *PCR-TOPO* vector (Invitrogen, USA) and sequenced using M_{13} universal primer and model 310 automated sequencer (Applied Biosystems, Foster City, CA, USA). Analyses of nucleotide and deduced amino acid sequences were carried out using EditSeq-DNAstar Inc., Expert Sequence Analysis software, Windows 32 Edit Seq 4.00 (1989-1999) and ExPasy database (http://expasy.org/tools/dna.html).

Based on the sequence and alignment data, specific primers ($LecSF_{1,2}$ and $LecSR_{1,2}$) for lectin-related sequences were designed (Table 1) and tried for reverse transcription polymerase chain reaction (RT-PCR). RT-PCR reaction was performed as previously described in this section regarding to the optimum annealing temperature (T_a) for each specific primer set.

Table 1: Key table for the primers used in this study providing their names, origin and sequences.

Primer name	Origin	Sequence (5` -- 3`)
$LecF_1$	Lectin-based	AGTGGTAACAACGCAGAGTACGCGGGGG
$LecR_1$	Lectin-based	ATATTTTAAAATCATCTCGTGTCCGGC
$LecF_2$	Lectin-based	ATGGGATCCAAGCAACAGAG
$LecR_2$	Lectin-based	ATCCTTCAAAGACACAATGTCG
Fwd_1	Lectin-based	CTCACTGTGAATATGAAAGCGGCGA
Fwd_2	Lectin-based	CTGACGTCCACACGTAAAATGATGTGT
Fwd_3	Lectin-based	CCGGTAGTCGTCACGTCGAAATGAAG
Fwd_4	Lectin-based	CCGGAAGCGGAGATGCTTGCCG
$LecSF_1$	Lectin-based	TTGGTGCCTGCACGTGGAGT
$LecSR_1$	Lectin-based	ACATTTTCAGGTTCAGACCCTTCTT
$LecSF_2$	Lectin-based	AGGGCTGCCTGCTCGTGGAGG
$LecSR_2$	Lectin-based	GACTTTTCATCGTACCCTTCCC
LecFLF	Lectin-based	GAATCTTTCAGTATGGAGTTATATGGACTG
LecFLR	Lectin-based	TAGTAATAAT TACATTTTCAGGTTCAG

Positive PCR products were visualized and eluted from the gel using GenClean Kit (Invitrogen Corporation, San Diego, CA, USA) following the manufacturer's instructions.

The purified PCR product (*Spli*Lec) was cloned into *PCR-TOPO* vector with TOPO TA cloning kit (Invitrogen, USA) following the manufacturer's instructions. Ligation mix was used to transform competent *E. coli* strain $TOPO_{10}$ provided with the cloning kit. White colonies were screened using PCR as described earlier in this section. Two positive clones of *Spli*Lec fragment were selected and sequenced (to exclude PCR errors certainly) using their specific forward and reverse primers (Table 1). Sequencing and sequence analyses were performed as described early in this section.

Full-length cDNA isolation of immunolectin gene

Specific primers (sense and antisense) were designed based on the sequence of *Spli*Lec containing 3µ end. The 5µ end fragment was amplified using SMART RACE cDNA Amplification kit (Clontech) following the procedure outlined in the supplied user manual. The amplified 5µ end fragment was purified, cloned into *PCR-TOPO* vector, and sequenced as described early in this section. The sequences of 3′ and 5′ end fragments were aligned and the predicted full-length cDNA was obtained. Thus a pair of primers, LecFLF and LecFLR (Table 1), was designed for the amplification of full-length *Spli*Lec cDNA. PCR was carried out in a total volume of 25 µl reaction solution containing 2.5 µl PCR buffer, 1.5 mM $MgCl_2$, 200 µM dNTPs, 1 U *Taq* DNA polymerase (AmpliTaq, Perkin-Elmer), 2.5 µl of 10 pmol of each primer and 2 µl cDNA using the following protocol: 94 °C for 5 min (hot start) followed by 35 cycles of amplification (94 °C for 1 min, 60 °C for 1 min, 72 °C for 1.5 min) and a final extention step at 72°C for 10 min. Full-length *Spli*Lec was visualized and eluted from the gel using GenClean Kit (Invitrogen Corporation, San Diego, CA, USA) following the manufacturer's instructions.

Nucleotide sequence and phylogenetic analyses

In addition to the above mentioned analyses, ExPasy Proteomics Server (http://expasy.org/tools) was used to calculate physico-chemical parameters of the translated peptide (ProtParam tool). Furthermore, post-translational modifications and topology predictions were investigated using SignalP, NetCGlyc, NetOGlyc, NetGlycate, YinOYang, OGPET**,** NetPhos, NetPhosK, Sulfinator, NetNES, SOSUI and TMpred tools.

Moreover, Phylogenetic analyses of the nucleotide sequence and its deduced amino acids were done using Mega4. Poorly aligned positions and divergent sequences were eliminated manually. Multiple alignment of available published lectin-related nucleotide sequences was done before phylogenetic analyses to approximate sequence lengths manually. 100% homologous sequences of the same species with different accession numbers were represented by only one sequence. The cloned DNA fragment was deposited in GenBank under the HQ603826 accession number.

RESULTS

Differential display using primers corresponding to well known lectins

Differential display technique was used to characterize the genetic variation (at RNA level) between bacterial-challenged and control cotton leafworm, *S. littoralis*. Fig. (1) shows the results of differentially displayed cDNAs of bacterial-challenged and control insects using 8 primers corresponding to previously characterized lectins (Table 1). Haemolymph samples were differentially displayed at 24, 48 and/ or 72 h p.i. with *S. aureus, S. sanguinis*, *E. coli, P. vulgaris* and *K. pneumoniae* bacterial strains. It was observed that *S. aureus*-challenged insects died 24 h p.i., *E. coli*-

challenged insects died 48 h p.i. and *S. sanguinis*-challenged insects died 72 h p.i. All insects died before sampling in the case of *P. vulgaris* and *K. pneumoniae*. Differential display results revealed that the average number of bands per sample was 4.3 bands for each amplification reaction.

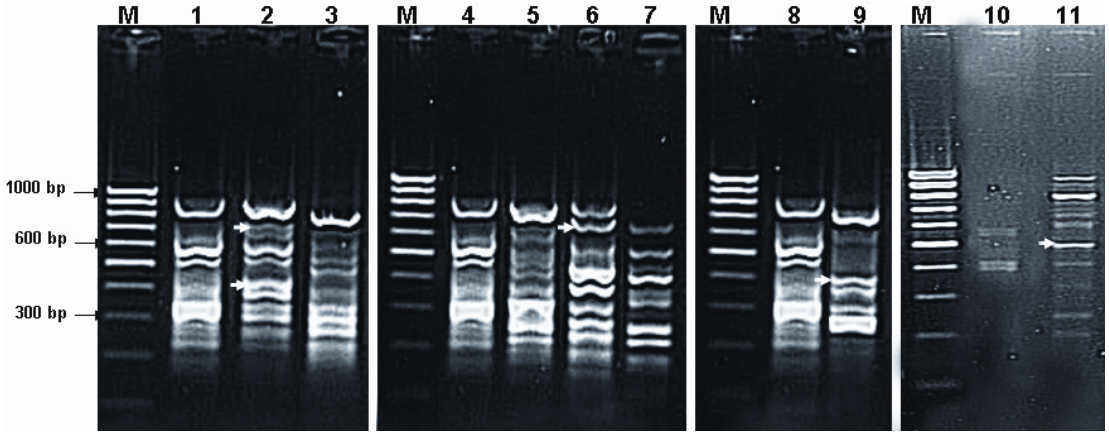

Fig. 1: Representative 1.5% agarose gels of the DD-PCR patterns generated from control and *S. aureus*, *E. coli* and *S. sanguinis*-challenged haemolymph samples using 8 primers corresponding to well known lectin genes. Lane M: DNA marker 100 bp Ladder, lanes 1, 4, 8 and 10: controls of different treatments, lanes 9 and 11: 24 h post-infection with *S. aureus*, lanes 2, 3 and 5, 6: 24 and 48 h post-infection with *E. coli* and lane 7: 72 h post-infection with *S. sanguinis*. Arrows refer to differentially displayed sequenced bands.

The total number of bands (transcripts) resolved in 1.5 % agarose gel for both control and challenged insects was 124 (molecular size ranged from >1300 to ~80 bp). Forty seven polymorphic bands (37.9 %) were differentially displayed with 6 of the used primers. Five reproducible, infection-induced bands were cloned and sequenced using M_{13} universal primer. Analyses of the results revealed that a fragment of 640 bp was amplified within the open reading frame (*orf*) of a lectin gene. This fragment contained the complete 3` end with a poly(A) tail, but it was not complete at the 5` end (lacking starting codon, AUG at its 5` end).

RT-PCR amplification and cloning of the lectin gene

To obtain the full-length sequence, the 5µ end of the cDNA was amplified using RACE PCR method, purified, cloned and sequenced. The full-length sequence of *Spli*Lec cDNA was amplified using LecFLF and LecFLR. RT-PCR was optimized for the primer set and successfully amplified ≈1150 bp fragment (Fig. 2).

The positive PCR product was visualized, eluted and cloned into *PCR-TOPO* vector (Fig. 2, lane 2). Using PCR screening method, the clone *PCR-TOPOSpli*Lec was tested as positive (Fig. 2, lane 4). Two positive clones of *Spli*Lec fragment were selected and sequenced (to exclude PCR errors certainly) using LecFLF and LecFLR primers (Table 1).

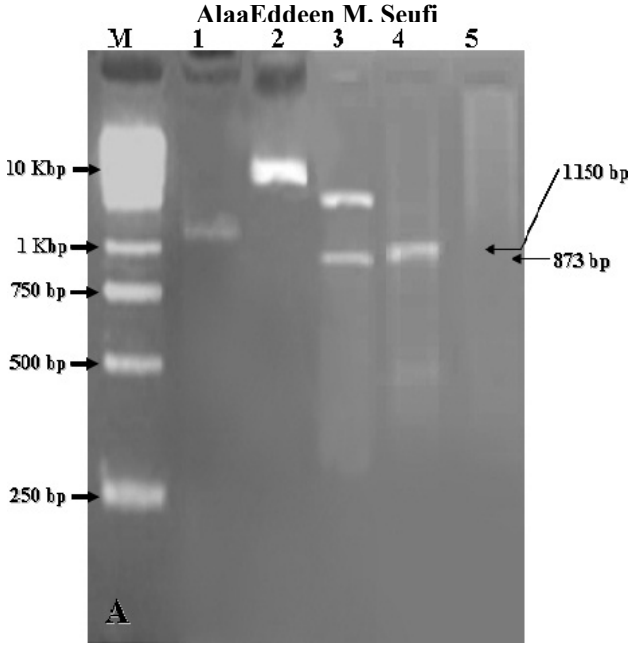

Fig. 2: Agarose gel electrophoresis showing positive PCR representing the full length *Spli*Lec (1150 bp), clone *PCR-Spli*Lec, clone *PCR-Spli*Lec after insert release with *Eco*RI, and PCR confirmation. Lanes 1, 2, 3, 4 and 5 show *Spli*Lec PCR product (1150 bp), *E. coli* harbouring *PCR-Spli*Lec, *PCR-Spli*Lec after digestion with *Eco*RI, positive control (*Spli*Lec amplified from the cDNA), and negative control (PCR mix without cDNA). Lane M: DNA Marker. The size of the bands is shown in bp.

Nucleotide sequence and sequence analyses

Nucleotide sequences of the *Spli*Lec and its deduced amino acid sequence is shown in Fig. (3). A single *orf* encoding a 309-residues polypeptide was detected in the *Spli*Lec sequence. One stop codon was found at the 3` end. The flanking region of the initiation codon ATG is AGTATGGAG, and the length of 5µ untranslated region (UTR) was 60 bp before the start codon ATG. The length of 3µ UTR was 60 bp before the poly (A) track.

The putative polyadenylation sequence AATAAA was located 15 bp downstream from the stop codon (Fig. 3). The identified *Spli*Lec *orf* includes a signal peptide (54 bp), and a mature peptide (873 bp). Analysis of the amino acid sequence deduced from the cDNA indicated that *Spli*Lec is a member of the C-type lectin superfamily. It contains two C-type CRDs, an amino-terminal domain, CRD_1 (residues 1–149), and a carboxyl-terminal domain, CRD_2 (residues 160–301). The deduced *Spli*Lec polypeptide contains 50 strongly basic, 28 strongly acidic, 127 hydrophobic and 104 polar uncharged amino acids.

The calculated molecular masses of the putative *Spli*Lec and its mature peptide are 34.85 and 32.91 KDa, respectively. The theoretical isoelectric points (PIs) were 9.27 and 9.38 for the full length and mature *Spli*Lec peptides, respectively. The net charges at pH 7.0 were 15.9 and 16.9 for the *Spli*Lec and its mature peptide, respectively. Both the full length and the mature *Spli*Lec peptides were classified as unstable (Instability Index (II): 55.81 and 56.95, respectively). Ratios of the hydrophilic residues were calculated as 37 and 38% for the full length and its mature peptides, respectively.

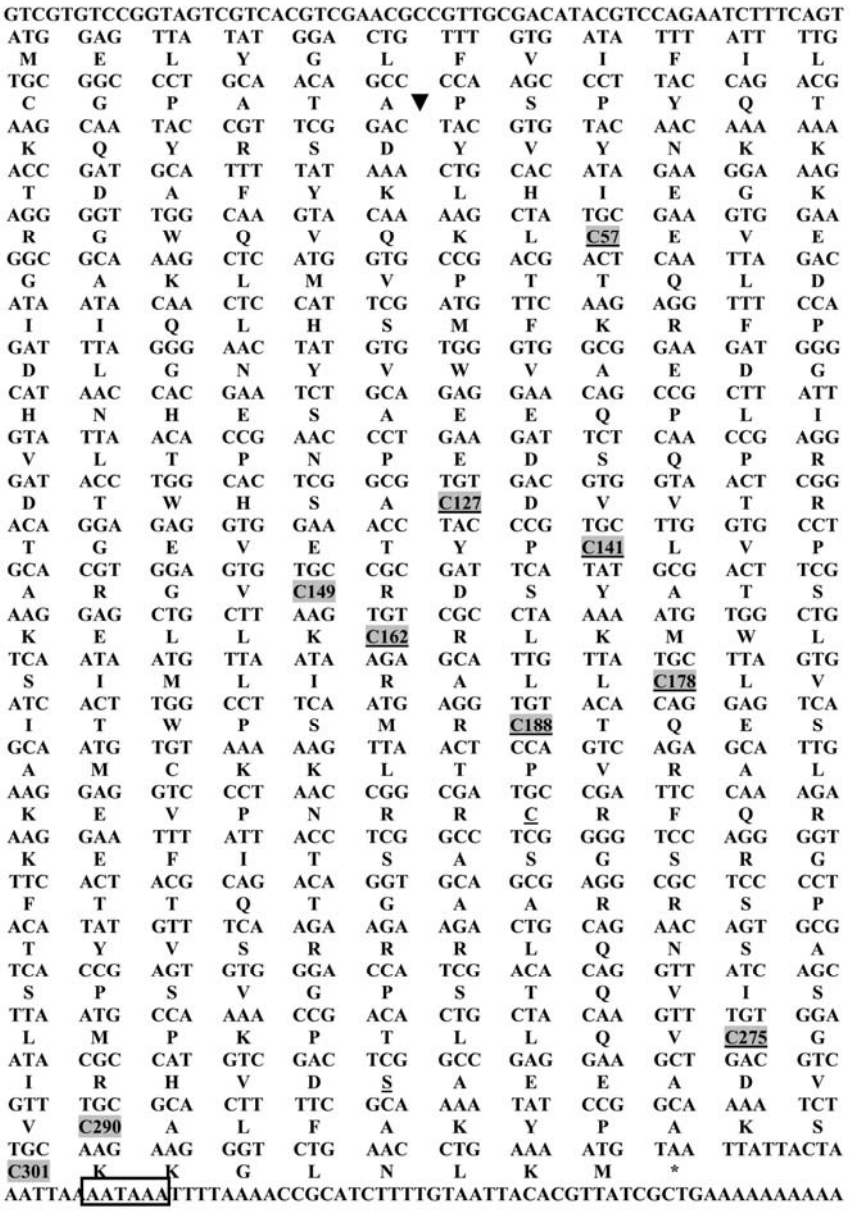

Fig. (3): Nucleotide and corresponding deduced amino acid sequence of *S. littoralis* immunolectin gene (*Spli*Lec). Cleavage site between the signal and mature peptides are indicated by an arrow. Positions of cysteine residues are shaded and numbered. Asterisk indicates the stop codon. Boxed sequence represents the putative polyadenylation signal.

Nucleotide sequence and its deduced amino acid sequence of the *Spli*Lec were blasted with all available sequences in GenBank database. Alignment results revealed that the *Spli*Lec sequence (Acc# HQ603826) has a significant alignment with 9 and 14 published lepidopteran DNA and peptide sequences, respectively. Although the percentage identity ranged from 100% to 69% with IML-A precursor (Acc# AF053131) and IML-3 (Acc# AY768811) of *Manduca sexta*, it did not necessarily mean full consistence, especially when the percentage coverage of the gene was regarded. Some insect lectins covered the forward region of the *Spli*Lec sequence and others covered the backward segment (e.g. *M. sexta* and *Bombyx mori* immunolectins) (Fig. 4 A and B).

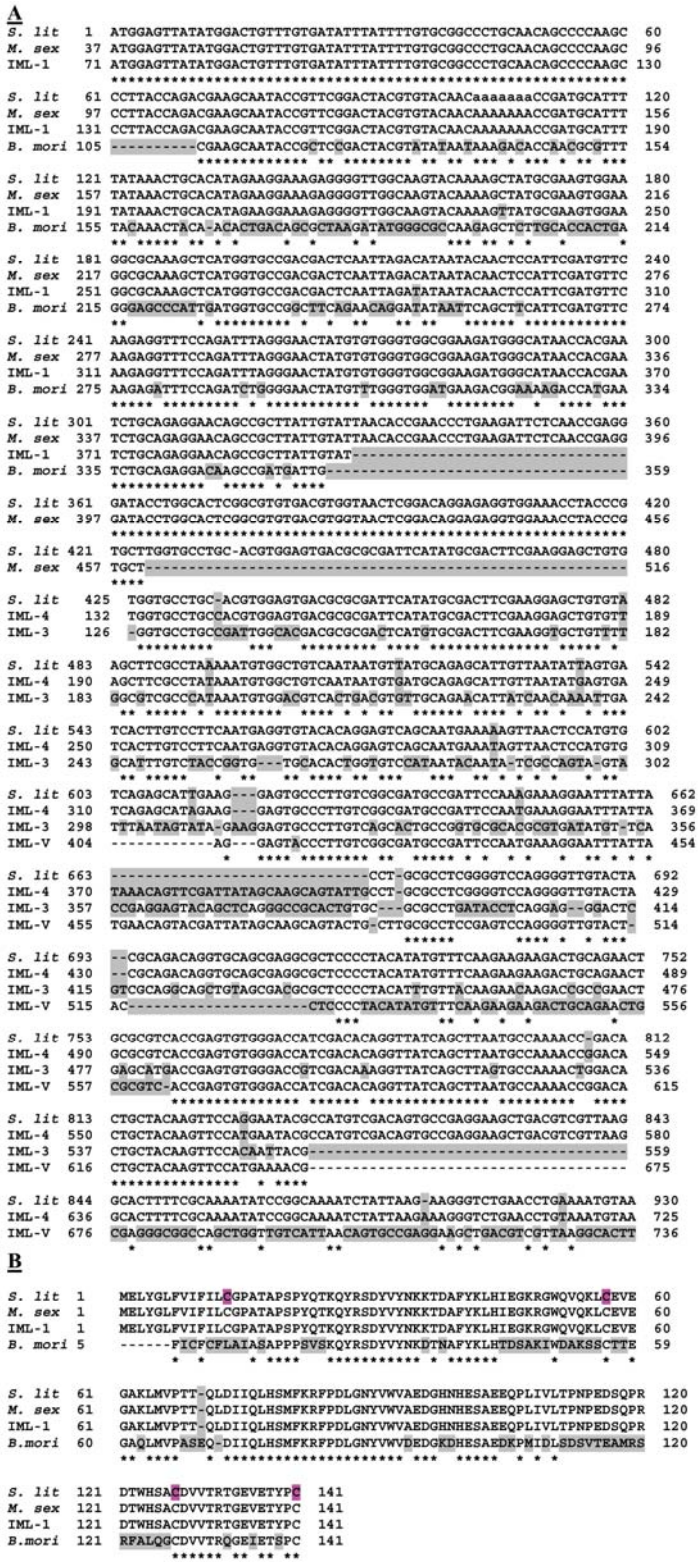

Fig. 4: Comparison of the lectin nucleotide sequence of *S. littoralis* (Acc# HQ603826) with *M. sexta* immunolectin-A precursor, IML-1, IML-2, IML-3, IML-4, IML-V and *B. mori* C-type lectin (Acc# AF053131, GU454799, AF242202, AY768811, AY768812, AM293329 and AY297159, respectively). *S. lit*: *S. littoralis* lectin, *M. sex*: *M. sexta* immunolectin-A, IML-1: *M. sexta* immunolectin-1, IML-2: *M. sexta* immunolectin-2, IML-3: *M. sexta* immunolectin-3, IML-4: *M. sexta* immunolectin-4, IML-V: *M. sexta* immunolectin-V and *B. mori*: *B. mori* C-type lectin. Gaps and different nucleotides are shaded.

Primary, secondary structure analyses, post-translational modifications and topology predictions revealed that amino acid sequence of the putative *Spli*Lec peptide had one signal peptide cleavage site (between positions 18 and 19), one

tyrosine-glycosylated and two tyrosine-sulfated sites at positions 111, 31 and 33, respectively. Fifteen *O*-GlcNAcylated residues (8 Ser and 7 Thr) and six potentially glycated lysines were predicted. Twenty one phosphorylation sites (Ser: 11, Thr: 6 and Tyr: 4) and 44 (24 S, 2 Y and 18 T) kinase specific phosphorylation sites (highest score: 0.82 PKC at position 185) were also predicted. In addition, two transmembrane helices (one primary: 166-182 with outside to inside orientation and one secondary: 3-22 with inside to outside orientation) were predicted.

Phylogenetic analyses of the *Spli*Lec sequence

Phylogenetic analyses of the *Spli*Lec have been performed with the 47 nucleotide seuquence (including 10 insect genera from the order Lepidoptera.) and 14 polypeptides (including 8 insect species: 3 lepidopterans and 5 dipterans). The results of these analyses are shown in Figs. (5 A and B).

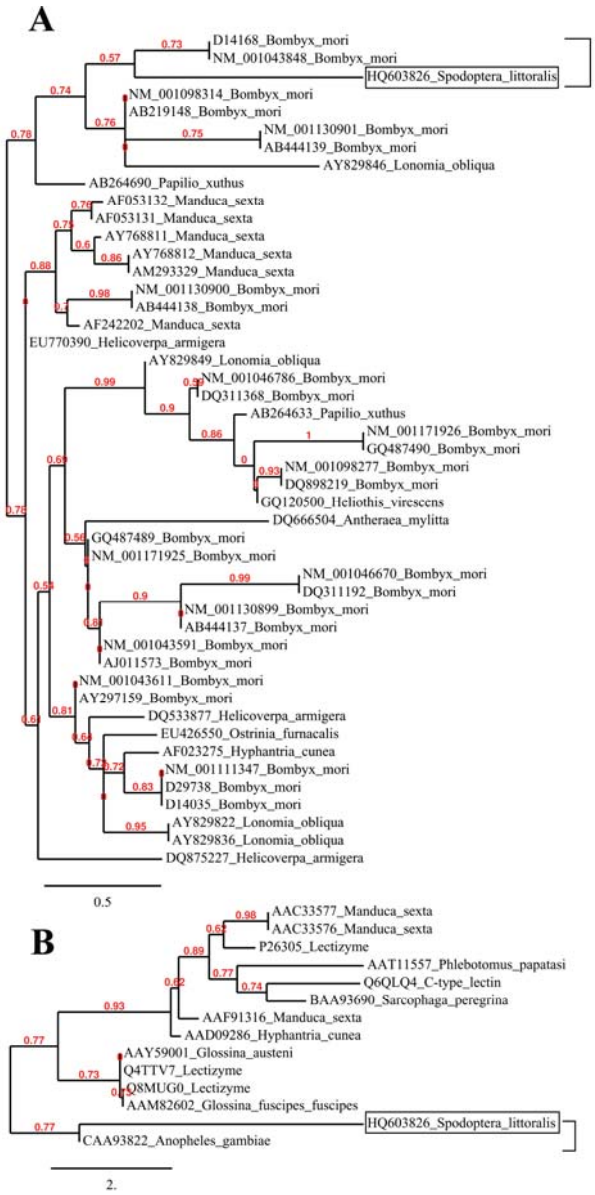

Fig. 5: Phylogenetic analysis of *Spli*Lec nucleotide and deduced amino acid sequences compared to 46 and 13 sequences registered in NCBI. Phylogenetic trees were generated from 47 and 14 lectin-related sequences by neighbor-joining distance analysis using Mega4 software. Full sequence names and accession numbers are included in the tree.

Phylogenetic trees were generated by neighbor-joining distance analyses with maximum sequence difference 1.0. The nucleotide topology shows two distinct lineages including 9 (6 phylogenetic groups) and 38 (24 phylogenetic groups) lectin-related sequences, respectively. *Spli*Lec seuquence (Acc# HQ603826) was clustered in a monophyletic sister clade with 2 *B. mori* lectins (Acc# NM_001043848 and D14168) (Fig. 5A). The polypeptide topology shows two distinct lineages including 12 (7 phylogenetic groups) and 2 (1 phylogenetic group) lectin peptides, respectively. However in this case, *Spli*Lec polypeptide was clustered with *Anopheles gambiae* lectin (Acc# CAA93822) in the same lineage (Fig. 5B).

DISCUSSION AND CONCLUSION

In the present study, the common bands revealed by DD-PCR in both control and challenged samples may represent the house-keeping genes. Some bands were recorded in control insects and disappeared in challenged ones (genes were turned off). On the other hand, many bands were induced as a result of bacterial-challenge at different time intervals post-infection. DD-PCR technique is considered a powerful genetic screening tool for complicated dynamic tissue processes, to detect and compare altered gene expression in eukaryotic cells, to screen and to characterize differentially expressed mRNAs, because it allows for simultaneous amplification of multiple arbitrary transcripts (Santana *et al.*, 2006). Many publications described the enhancement of the insect immune system and induction of lectins due to stress and/ or infection (*e.g.* Basseri, 2002, Yu *et al*., 2005, Wang *et al.,* 2010, Seufi *et al.*, 2011). Lectins were isolated from six insect orders: Lepidoptera (*e.g.* Koizumi *et al.*, 1999, Shin *et al.*, 2000, Yu and Kanost, 2003, Kanost *et al.*, 2004, Yu *et al.*, 2005, Tian *et al.*, 2009, Wang *et al.*, 2010), Diptera (*e.g.* Haq *et al.*, 1996, Fujita *et al.*, 1998, Lemaitre and Hoffmann, 2007), Coleoptera (*e.g.* Umetsu *et al.*, 1984), Hemiptera (*e.g.* Mello *et al*., 1999, Pinheiro and Ellar, 2006), Orthoptera (*e.g.* Bradley *et al.,*1989) and Dictyoptera (*e.g.* Chen *et al.*, 1993). As the C-type lectins are important molecules in the innate immune systems, we isolated the full-length cDNA of *S. littoralis* lectin (*Spli*Lec) which shares typical features of the C-type insect lectins.

The full-length cDNA of the *Spli*Lec was 1150 bp, a size very similar to that of *M. sexta* IML-2 (Yu and Kanost, 2000), and contained a 927 bp *orf* encoding 309 amino acids. The flanking region of the *Spli*Lec initiation codon ATG keeps the adenine nucleotide at position -3 which is a universal feature in all the eukaryote genes (Zhu *et al*., 2000). Like many insect lectins, the *Spli*Lec was predicted to have a 18-residue secretion signal peptide and a 291-residue mature protein. The deduced amino acid sequence of *M. sexta* IML-2 was reported to contain a 19-residue secretion signal peptide and a 308-residue mature protein (Yu and Kanost, 2000). *Sarcophaga* C-type lectin was predicted to have a 23-residue secretion signal peptide and a 150-residue mature protein (Fujita *et al*., 1998). Signal secretion peptides were reported to be 21-, 21- and 26-residue in the cases of *Drosophila* DL1, DL2 and DL3 C-type lectins, respectively (Tanji *et al.,* 2006). It is notable that the *Spli*Lec gene also shares homology to many C-type insect lectins and it consists of two CRDs: the amino-terminal CRD_1 is short form, with two intramolecular disulfide bonds (Cys^{57}–Cys^{127} and Cys^{141}–Cys^{149}) and the carboxyl-terminal CRD_2 is long form, with three intramolecular disulfide bonds (one additional disulphide bond near the amino terminus: Cys^{162}-Cys^{178}, Cys^{188}-Cys^{275} and Cys^{290}-Cys^{301}). This feature of the *Spli*Lec is similar to the two immulectins of *M. sexta* (IML-1 and IML-2) (Yu *et al.,* 1999),

LPS-binding proteins of the silkworm, *B. mori* (Koizumi *et al.*, 1999) and the putative lectin of the fall webworm, *H. cunea* (Shin *et al.*, 2000).

The predicted post-translational modifications of the *Spli*Lec protein suggested an important role of the *Spli*Lec protein in modulating a broad range of biological processes in the cell. The predicted O-GlcNAcylation suggested a possible function of the *Spli*Lec protein in macromolecular complex assembly and intracellular transport. Glycosylation and glycation serve for the correct folding and stability of the protein (unglycosylated proteins degrade quickly). Glycosylation of proteins play a role in cell-cell adhesion (a mechanism employed by cells of the immune system), as well (Varki *et al*., 2009). Reversible phosphorylation of proteins (using kinases and phosphatases) is considered an important regulatory mechanism in protein-protein interaction via recognition domains, (i.e. many proteins and receptors are switched "on" or "off" by phosphorylation and dephosphorylation). It also results in a conformational changes in the structure in many peptides, causing them to become activated, deactivated or degraded (Olsen *et al.*, 2006). In addition, many transmembrane proteins (TPs) function as gateways or "loading docks" to deny or permit the transport of specific substances across the biological membranes (to get into or out of the cell by folding up or bending through the membrane).

Reconstruction of the phylogenetic trees of the *Spli*Lec nucleotide sequence and its deduced polypeptide resulted in two different topologies. Both of the two trees clustered *Spli*Lec sequence in two different groups (clustered with *Bombyx* in the case of nucleotide-based tree and with *Anopehles* in the case of amino acid-based tree) indicating the possibility of evolutionary trend between these lectins which might descend from a common ancestor. Grouping of some lepidopteran and dipteran lectins (*e.g. M. sexta* with *Sarcophaga* and *S. littoralis* with *Anopehles*) in one sister clade indicated that they may be homologous or share some similarity. In addition, lepidopteran lectin-like sequences were diverged in many sister clades as amino acids due to the difference in codon usage in different species.

In short, these findings shed a new light on the lectin-mediated immune system. Combination of these findings with that reported by Seufi *et al.* (2009), Seufi *et al.* (2011) and Seufi (2011) suggested that the *Spli*Lec, *Spli*Def and *Spli*Cec peptides with other possible AMPs may constitute the defense network of *S. littoralis* (Lepidoptera) against invading microorganisms.

Conclusively, the current results provide a novel insect lectin gene (*Spli*Lec) with a two tandem CRDs. The *Spli*Lec plays an important immune role in *S. littoralis* by cooperating with other AMPs to clear invading microorganisms. These findings would be helpful in future studies on lectins concerning ELISA, PCR and other related molecular and immunological techniques. Future studies on the carbohydrate-binding and blood group specificities, on the determination of molecular weight and three-dimensional structure of the *Spli*Lec will be needed to provide direct evidences and more understanding of the *Spli*Lec mode of action.

REFERENCES

Basseri, H. (2002): Role of lectins in interaction between parasites and the important insect vectors. *Iranian J. Publ. Health*, 31: 69-74.

Bradley, R.; Stuart, G.; Stiles, B.; Hapner, K. (1989): Grasshopper haemagglutinin: immunochemical localization in haemocytes and investigation of opsonic properties. *J. Insect Physiol.,* 35: 353-361.

Chen, C.; Ratcliffe, N.; Rowley, A. (1993): Detection, isolation and characterization of multiple lectins from the haemolymph of the cockroach *Blaberus discoidalis*. *Biochem. J.,* 294: 181-190.

Day, A. (1994): The C-type carbohydrate recognition domain (CRD) superfamily. *Biochem. Soc. Trans.,* 22: 83-88.

Drickamer, K.; Fadden, A. (2002): Genomic analysis of C-type lectins. *Biochem. Soc. Symp.,* 69: 59-72.

Drickamer, K.; Taylor, M. (1993): Biology of animal lectins. *Annu. Rev. Cell Biol.,* 9: 237–264.

Eslin, P.; Prevost, G. (2000): Racing against host's immunity defenses: a likely strategy for passive evasion of encapsulation in *Asobara tabida* parasitoids. *J. Insect Physiol.,* 46: 1161–1167.

Fujita, Y.; Kurata, S.; Homma, K.; Natori, S. (1998): A novel lectin from *Sarcophaga*: Its purification, characterization and cDNA cloning. *J. Biol. Chem.,* 273: 9667-9672.

Haq, S.; Kubo, T.; Kurata, S.; Kobayashi, A.; Natori, S. (1996): Purification, Characterization, and cDNA Cloning of a Galactose-specific C-Type Lectin from *Drosophila melanogaster*. *J. Biol. Chem.,* 271: 20213–20218.

Kanost, M.; Jiang, H.; Yu, X. (2004): Innate immune responses of a lepidopteran insect, *Manduca sexta*. *Immunol. Rev.,* 198: 97-105.

Koizumi, N.; Imamura, M.; Kadotani, T.; Yaoi, K.; Iwahana, H.; Sato, R. (1999): The lipopolysaccharide-binding protein participating in hemocyte module formation in the silkworm *Bombyx mori* is a novel member of the C-type lectin superfamily with two different tandem carbohydrate-recognition domains. *FEBS Lett.,* 443: 139-143.

Lavine, M.; Strand, M. (2002): Insect haemocytes and their role in immunity. *Insect Biochem. Mol. Biol.,* 32: 1295-1309.

Lemaitre, B.; Hoffmann, J. (2007): The host defense of *Drosophila melanogaster*. *Annu. Rev. Immunol.,* 25: 697-743.

Levinson, Z.; Navon, A. (1969): Ascorbic acid and unsaturated fatty acids in the nutrition of the Egyptian cotton leaf worm. *J. Insect Physiol.,* 15: 591-595.

Medzhitov, R.; Janeway, C. (2002): Decoding the patterns of self and nonself by the innate immune system. *Science,* 296: 298–300.

Mello, C.; Nigan, Y.; Garcia, E.; Azambuja, P.; Newton, R.; Ratcliffe, N. (1999): Studies in a haemolymph lectin isolated from *Rhodnius prolixus* and its interaction with *Trypanosoma rangeli*. *Exp. Parasitol.,* 91: 289-296.

Olsen, J.; Blagoev, B.; Gnad, F.; Macek, B.; Kumar, C.; Mortensen, P.; Mann, M. (2006): Global, *in vivo*, and site-specific phosphorylation dynamics in signaling networks. *Cell,* 127: 635–648.

Pinheiro, V.; Ellar, D. (2006): How to kill a mocking bug? *Cell Microbiol.,* 8: 545-557.

Santana, F.; Nunes, F.; Vieira, C.; Machado, M.; Kerr, W.; Silva Jr. W.; Bonetti, A. (2006): Differentially displayed expressed sequence tags in *Melipona scutellaris* (Hymenoptera, Apidae, Meliponini) development. *Anais da Academia Brasileira de Ciências,* 78: 69-73.

Seufi, A. (2011): Molecular Characterization, cDNA Cloning and Phylogenetic Analysis of Cecropin Gene Isolated from Bacterial-challenged Cotton Leafworm, *Spodoptera littoralis. Egypt. Acad. J. Biolog. Sci. (A- Entomology)*, (2): 65-75.

Seufi, A.; El Bassiony, G.; Ibrahim, S. (2009): Purification and Characterization of Two Antimicrobial Peptides from Bacterial-challenged Haemolymph of *Bombyx mori* Larva. *Egypt. Acad. J. biolog. Sci.,* 2 (2): 231- 243.

Seufi, A.; Hafez, E.; Galal, F. (2011): Identification, phylogenetic analysis and expression profile of an anionic insect defensin gene, with antibacterial activity, from bacterial-challenged cotton leafworm, *Spodoptera littoralis*. *BMC Molec. Biol.,* 12: 47.

Sharon, N. (1977): Lectins. *Sci .Amer.,* 236: 108-119.

Shin, S.; Park, D.; Kim, S.; Park, H. (2000): Two carbohydrate recognition domains of *Hyphantria cunea* lectin bind to bacterial lipopolysaccharides through O specific chain. *FEBS Lett.*, 467: 70-74.

Tanji, T.; Ohashi-Kobayashi, A.; Natori, S. (2006): Participation of a galactose-specific C-type lectin in *Drosophila* immunity. *Biochem. J.,* 396: 127–138.

Tian, Y.; Liu, Y.; Zhao, X.; Wang, J. (2009): Characterization of a C-type lectin from the cotton bollworm, *Helicoverpa armigera*. *Dev. Comp. Immunol.,* 33: 772–779.

Umetsu, K.; Kosaka, S.; Suzuki, T. (1984): Isolation and characterization of a lectin from the beetle *Allomyrina dichotoma. J. Biochem.,* 95: 239- 246.

Varki, A.; Cummings, R.; Esko, J.; Freeze, H.; Stanley, P.; Bertozzi, C.; Hart, G.; Etzler, M. (2009): Essentials of Glycobiology. *(2nd ed.) Cold Spring Harbor Laboratories Press (NY). La Jolla, California, USA.*

Vasta, G.; Quesenberry, M.; Ahmed, H.; O'Leary, N. (1999): Lectins and galectins mediate innate and adaptive immune functions: their roles in the complement activation pathway. *Dev. Comp. Immunol.,* 23: 401-420.

Wang, Q.; Liu, Y.; He, H-J.; Zhao, X-F; Wang, J-X. (2010): Immune responses of *Helicoverpa armigera* to different kinds of pathogens. *BMC Immunol.,* 11: 9.

Yu, X.; Gan, H.; Kanost, M. (1999): Immulectin, an inducible C-type lectin from an insect, *Manduca sexta*, stimulates activation of plasma prophenoloxidase. *Insect Biochem. Mol. Biol.,* 29: 585-597.

Yu, X.; Kanost, M. (2000): Immulectin-2, a lipopolysaccharide-specific lectin from an insect, *Manduca sexta*, is induced in response to Gram-negative bacteria. *J. Biol. Chem.,* 275: 37373-37381.

Yu, X.; Kanost, M. (2003): *Manduca sexta* lipopolysaccharide-specific immulectin-2 protects larvae from bacterial infection. *Dev. Comp. Immunol.,* 27:189–196.

Yu, X.; Tracy, M.; Ling, E.; Scholz, F.; Trenczek, T. (2005): A novel C-type immulectin-3 from *Manduca sexta* is translocated from haemolymph into the cytoplasm of haemocytes. *Insect Biochem. Mol. Biol.,* 35: 285-295.

Zelensky, A.; Gready, J. (2004): C-type lectin-like domains in *Fugu rubripes*. *BMC Genomics,* 5.

Zhu, S.; Li, W.; Zeng ,X.; Liu, H.; Jiang, D. and Mao, X. (2000): Nine novel precursors of *Buthus martensii* scorpion *α*-toxin homologues. *Toxicon,* 38: 1653–1661.

ARABIC SUMMARY

التوصيف الجزيئي و عزل الطول الكامل وتحليل النشوء والتطور لجين اللكتين (نوع-ج) المعزول من دودة ورق القطن، *سبودوبترا ليتوراليس*، المستحثة مناعيا بالبكتيريا

علاء السيوفى
قسم علم الحشرات - كلية العلوم -جامعة القاهرة -القاهرة -مصر

صممت التجارب و أجريت بهدف دراسة الاستجابة المناعية لدودة ورق القطن، *سبودوبترا ليتوراليس*، المستحثة بالعدوى البكتيرية ، وبالإضافة إلى ذلك، تمت دراسة تحليل تسلسل النشوء والتطورللجينات المستحثة. و باستخدام تقنية العرض التباينى، تم اختيار خمس حزم تباينية من دم اليرقات المستحثة مناعيا بالبكتريا، و تم الكشف عن تتابعات قواعدها النيتروجينية، و اكتشف أن أحد هذه الجزيئات يمثل جزءا من جين اللكتين طوله 640 زوجا من القواعد النيتروجينية، و يحتوى على شفرة نهاية الجين فقط دون شفرة البدايه. و صممت البادئات التخصصية لجين اللكتين لاستخدامها فى تقنية التكبير السريع لنهايات الجينات (RACE PCR)، و تم عزل الطول الكامل لهذا الجين، و بعد تحليل تتابعات القواعد النيتروجينية له وجد أنه يمثل التركيب النموذجى لعائلة جين اللكتين (نوع-ج) الحشرى. وقد أوضح استنتاج تسلسل الأحماض الأمينية لهذا الجزئ أن البروبيبتيد يتكون من بروببتيد (18 حمضا أمينيا)، و ببتيد ناضج يتكون من 291 حمضا أمينيا. و قد أظهرت النتائج أنه يحتوى على اثنين مدى تعرف سكرى (CRD). و قد أظهرت الدراسات المقارنة للجين أنه يشبه تركيب جين اللكتين الحشرى لبعض ثنائيات الأجنحة، مما يوحى بوجود سلف مشترك لهما، و قد أكدت هذه النتائج دورا هاما لهذا الجين في الالتصاق والتعرف غير الذاتي، كما أنها قد تتعاون مع الببتيدات الأخرى لمكافحة العدوى البكتيرية لدودة ورق القطن، *سبودوبترا ليتوراليس*.