Molecular characterization of a c-type lysozyme from the desert locust, Schistocerca gregaria (Orthoptera: Acrididae)

Amr A. Mohamed, Long Zhang, Moataza A. Dorrah, Mohamed Elmogy, Hesham A. Yousef, Taha T.M. Bassal, Bernard Duvic

Accepted Manuscript

Molecular characterization of a c-type lysozyme from the desert locust, *Schistocerca gregaria* (Orthoptera: Acrididae)

Amr A. Mohamed, Long Zhang, Moataza A. Dorrah, Mohamed Elmogy, Hesham A. Yousef, Taha T.M. Bassal, Bernard Duvic

PII: S0145-305X(16)30099-4

DOI: 10.1016/j.dci.2016.03.018

Reference: DCI 2593

To appear in: *Developmental and Comparative Immunology*

Received Date: 27 January 2016

Revised Date: 15 March 2016

Accepted Date: 17 March 2016

Please cite this article as: Mohamed, A.A., Zhang, L., Dorrah, M.A., Elmogy, M., Yousef, H.A., Bassal, T.T.M., Duvic, B., Molecular characterization of a c-type lysozyme from the desert locust, *Schistocerca gregaria* (Orthoptera: Acrididae), *Developmental and Comparative Immunology* (2016), doi: 10.1016/j.dci.2016.03.018.

Molecular characterization of a c-type lysozyme from the desert locust, *Schistocerca gregaria* (Orthoptera: Acrididae)

Amr A. MOHAMED[1,*], Long ZHANG[2], Moataza A. DORRAH[1], Mohamed ELMOGY[1], Hesham A. YOUSEF[1], Taha T.M. BASSAL[1] and Bernard DUVIC[3]

[1] Department of Entomology, Faculty of Science, Cairo University, P. O. Box 12613, Giza, Egypt

[2] Key Lab for Biological Control of the Ministry of Agriculture, Department of Entomology, China Agricultural University, Beijing 100193, P. R. China

[3] Unité DGIMI UMR INRA-UM 1333, Université Montpellier, Place Eugène Bataillon, Montpellier, France

* To whom correspondence should be addressed: mamr@sci.cu.edu.eg

Abstract

Lysozymes are bacteriolytic peptides that are implicated in the insect nonspecific innate immune responses. In this study, a full-length cDNA encoding a c-type lysozyme from *Schistocerca gregaria* (*SgLys*) has been cloned and characterized from the fat body of immune-challenged 5th instar. **The deduced mature lysozyme is 119 amino acid residues in length, has a calculated molecular mass of 13.4 kDa and an isoelectric point of 9.2.** *Sg*Lys showed high **identities** with other insect lysozymes, ranging from **41.5%** to **93.3%** by BLASTp search in NCBI. **Eukaryotic *in vitro* expression of the *SgLys* ORF (r*Sg*Lys) with an apparent molecular mass of ~ 16 kDa under SDS-PAGE is close to the calculated molecular weight of the full-length protein**. r*Sg*Lys displayed growth inhibitory activity against Gram-negative and Gram-positive bacteria. 3D structure modeling of *Sg*Lys, based on comparison with that of silkworm lysozyme, and sequence comparison with the helix-loop-helix (α-hairpin) structure of hen egg white lysozyme (HEWL) were employed to interpret the antibacterial potencies. Phylogenetic alignments indicate that *Sg*Lys aligns well with insect c-type lysozymes that expressed principally in fat body and hemocytes and whose role has been defined as immune-related. Western blot analysis showed that *Sg*Lys expression was highest at 6-12 h post-bacterial challenge and subsequently decreased with time. Transcriptional profiles of *SgLys* were determined by semi-quantitative RT-PCR analysis. *SgLys* transcript was upregulated at the highest level in fat body, hemocytes, salivary gland, thoracic muscles, and epidermal tissue. It was expressed in all developmental stages from egg to adult. These data indicate that *Sg*Lys is a predominant acute-phase protein that is expressed and upregulated upon immune challenge.

Keywords

Schistocerca gregaria – C-type lysozyme – cDNA – Innate immunity – Antibacterial activity – Recombinant expression – Transcriptional profiles

Introduction

Insects have an innate immune system that allows them to be protected against a wide range of microbial pathogens. This system is highly developed and comprises cellular and humoral components that can be activated by invasion of pathogens (Tsakas and Marmaras, 2010). The humoral defense response of insects takes effect by over-expressing an array of potent antimicrobial peptides and proteins (AMPs) in order to kill invaders at the time of pathogenic infection. Among the large number of inducible antimicrobial peptides and proteins, lysozyme, a bacteriolytic enzyme, is the most ubiquitous antibacterial factor and is widely distributed in vertebrate and invertebrate animals. In invertebrates, it is well known that the lysozyme expression is regulated and responds to a bacterial challenge (Jiang et al., 2010; Mohamed et al., 2013). Both hemocytes and fat body, primarily the latter, of various insects have been reported to synthesize and release lysozyme to hemolymph (Lemaitre and Hoffmann, 2007; Mohamed et al., 2013; Zachary and Hoffmann, 1984). However, other tissues, such as epidermal tissues, may also participate in lysozyme production (Lee and Brey, 1995; Mulnix and Dunn, 1994).

Since the first insect lysozyme was reported in honey bees (Mohrig and Messner, 1968), more than fifty lysozyme genes have been identified from several insects, including the orthopterans *Locusta migratoria* (Zachary and Hoffmann, 1984), and *Gryllus bimaculatus* (Schneider, 1985), and the lepidopterans including *Spodoptera littoralis* (Jolles et al., 1979), *Hyalophora cecropia* (Engström et al., 1985), *Manduca sexta* (Mulnix and Dunn, 1994), *Bombyx mori* (Abraham et al., 1995), *Trichoplusia ni* (Kang et al., 1996), *Heliothis virescens* (Lockey and Ourth, 1996), *Hyphantria cunea* (Park et al., 1997), *Antheraea mylitta* (Jain et al., 2001), *Samia cynthia ricini* (Fujimoto et al., 2001), *Spodoptera litura* (Kim and Yoe, 2003), *Artogeia rapae* (Bang and Yoe, 2005; Yoe et al., 1996), *Ostrinia furnacalis* (Wang et al., 2009), and *Spodoptera frugiperda* (Chapelle et al., 2009). It has also been reported, by sequence mining, from both the coleopterans *Tribolium castaneum* (Altincicek et

al., 2008) and *Sitophilus zeamais* (Anselme et al., 2008) **and by immunity-related transcriptome analysis from *Harmonia axyridis* (Vilcinskas et al., 2013)**, and the hymenopteran, *Camponotus floridanus* and *Harpegnathos saltator* (Bonasio et al., 2010). In dipterans, the overwhelming majority of the studied lysozymes, isolated gene and/or cDNA structure were the digestive or involved in blood feeding (Cançado et al., 2008) whereas, a basic lysozyme with muramidase activity, the Lys c-1, is expressed in most tissues of adult mosquitoes and upregulated during bacterial infections (Kajla et al., 2010). In Diptera, the lysozyme genes cluster is a multi-gene family which resulted in a large expansion of the c-type lysozyme gene family in this insect order. For example, in the mosquito *A. gambiae* there are eight c-type lysozyme genes (*Lys c1-8*) (Li et al., 2005; Waterhouse et al., 2007). **Likewise, in Lepidoptera, four c-type lysozyme homologues have been predicted in *Galleria mellonella* (Vogel et al., 2011) and two lysozymes have been identified in the *Manduca sexta* genome (He et al., 2015).**

The survey on insect lysozymes sequenced to date shows that most of them are belonging to holometabolous ones, except for the isopteran *Reticulitermes speratus*, and the hemipterans *Rhodnius prolixus* (Ursic-Bedoya et al., 2008), *Triatoma brasiliensis* (Araujo et al., 2006) and *T. infestans* (Balczun et al., 2008). Interestingly, lysozymes of hemimetabolous insects have digestive function and **almost** belong to c-type group. **However, the genome of the pea aphid *Acyrthosiphon pisum* contains three i-type (invertebrate) lysozymes encoding genes (Gerardo et al. 2010).** To date, the isolation and/or molecular cloning of *Schistocerca gregaria* lysozyme (*SgLys*) has not been reported.

The desert locust, *Schistocerca gregaria* (Forskål) (Orthoptera: Acrididae) remains the most feared of all locusts and is one of the most notorious insects in the world (Lecoq, 2004). Desert locusts, and possibly other hemimetabolous insects, differ in their biological and developmental aspects from holometabolous insects, and therefore, different defense mechanisms against

pathogens may be expected. Accordingly, there is a longstanding interest in understanding the molecular mechanisms of the innate response of the hemimetabolous insects to microorganisms.

In order to understand the antimicrobial mechanism of *S. gregaria* lysozyme, we recently purified and characterized this antimicrobial molecule from the plasma of this insect (Elmogy et al., 2015; Mohamed et al., 2013). In the present work, we report, for the first time, the isolation and cloning of a complete cDNA encoding *S. gregaria* lysozyme (*Sg*Lys) and demonstrate the protein's primary, secondary and three dimensional structures. The directional cloning, *in vitro* expression and activity of r*Sg*Lys were also examined. In addition, *Sg*Lys was compared with other insect and non insect lysozymes and a molecular phylogenetic tree was constructed. The protein synthesis of the mature form by western blot and levels of lysozyme transcript in different tissues, and at different developmental stages post bacterial challenge were analyzed.

Materials and Methods

Insect rearing

The desert locust, *Schistocerca gregaria* (Forskål), was from a well-established laboratory colony at the Entomology Department, Faculty of Science, Cairo University, Egypt. Detailed descriptions of *S. gregaria* colony and the rearing techniques are given by Tanaka and Maeno (2008) and Mohamed et al. (2013).

cDNA cloning

Total RNA was extracted from the fat body tissue of *E. coli*-immunized 5th instar *S. gregaria* using TRIzol reagent (Invitrogen, Carlsbad, CA, USA) according to the supplier's instructions. Genomic DNA contamination was removed by Dnase I (Fermentas Life Sciences, Burlington, ON, Canada) treatment. Poly(A)$^{+}$RNA was purified from total RNA by oligo-dT cellulose (Amersham-Pharmacia Biotech, Buckinghamshire, UK) according to the manufacturer's instructions. Extracted mRNA was used to synthesize cDNA first strand with 200 units of RevertAid™ M-MuLV Reverse Transcriptase

from the RevertAid™ First Strand cDNA Synthesis Kit (Fermentas Life Sciences) and an oligo (dT) primer [a cDNA synthesis primer (10 µM)] with the dNTP mixture (10 mM). The degenerate primers Lys-F1 and Lys-F2, designed from a highly conserved amino acid sequence (DYGL/IFQI) found in insect lysozymes (Fujita et al., 2002; Regel et al., 1998), were first used in PCR to get a partial sequence. All primers used in this study are listed in Supplementary Table S1.

The 3'-terminal end of lysozyme cDNA was amplified by 3' RACE (rapid amplification of cDNA ends). The first PCR reaction was amplified with Lys-F1 (10 µM), 2 µL of poly-T primer (5'-GCGAATTCGTCGACAAGC$(T)_{17}$-3') (10 µM) (Fermentas Life Sciences). The cDNA amplification was performed in a Master cycler Thermocycler (Eppendorf, Germany) based on the following program: Initial denaturation of 94°C for 2 min followed by 30 cycles of denaturation at 94°C for 1 min, annealing at 55°C for 1 min and extension at 72°C for 1.5 min. The second PCR (nested) was identical to the first, except that the reaction mixture comprised 1 µL of the first PCR product and 37.0 µL of nuclease free water; also Lys-F2 and T-amp (5'-GCGAATTCGTCGACAAGC-3') primers were used and a final extension was carried out at 72°C for 20 min. Then, 5 µL of the final PCR product was analyzed by 1.5 % agarose gel electrophoresis (Biobasic Inc., Toronto, Canada) and a 450 bp amplified fragment was detected and purified by the PCR purification kit (Norgen Biotek Corporation, Canada) as per manufacturer's instructions. The purified products were cloned into pTZ57R/T vector with the InsTAclone™ PCR Cloning Kit (Fermentas Life Sciences) for sequencing.

The 5'-terminal end of lysozyme cDNA was amplified by 5' RACE (5´ RACE System for Rapid Amplification of cDNA Ends, Version 2.0, Invitrogen) and the gene-specific primers Lys-R1, Lys-R2, and Lys-R3, and AAP (5'-GGCCACGCGTCGACTAGTACGGGATGGGATGGGATG-3'), AUAP (5'-GGCC ACGCGTCGACTAGTAC-3'). The gene specific primers were designed based on the sequence obtained from 3' RACE. 5' RACE-PCR reactions were carried out according to the instructions of the manufacturer. Amplification conditions employed for dC tailed target cDNA were initial denaturation of 94°C for 2 min followed by 30 cycles of denaturation at 94°C for 1 min, annealing at 55°C for 1 min

147 and extension at 72°C for 1.5 min. Nested amplification were carried under the same conditions
148 except that denaturation comprised 40 cycles and a final extension at 72°C for 20 min. The amplified
149 5' RACE-PCR products were cloned into pTZ57R/T vector by InsTAclone™ PCR Cloning Kit (Fermentas
150 Life Sciences. The PCR was carried out same way as the nested amplification with the primers AUAP
151 and Lys-R3 except that 0.5 µL of plasmid was used as template and the final extension was 5 min
152 instead of 20 min. We named *SgLys*, the lysozyme encoded by the cDNA.

153 Finally, the complete (end to end) sequence of *SgLys* was amplified by the specific primers Lys-
154 tF1 and Lys-tR1 designed on the sequences obtained from 3' RACE and 5' RACE. The amplification
155 was performed with the following program: Initial denaturation of 94°C for 2 min followed by 40
156 cycles of denaturation at 94°C for 1 min, annealing at 56^{o}C for 1 min and extension at 72°C for 1 min.
157 Final extension was carried out at 72°C for 20 min. The amplified complete coding sequence of *SgLys*
158 was cloned into pTZ57R/T vector by InsTAclone™ PCR Cloning Kit (Fermentas Life Sciences), and
159 sequenced.

160 All nucleotide sequences were determined by automated DNA sequencer 3037*xl* DNA analyzer
161 from Applied Biosystems using BigDye® Terminator v3.1 cycle sequencing Kit (Applied Biosystems)
162 and T7 as forward primer and M13 as a reverse primer.

163 **Production of recombinant *S. gregaria* lysozyme**

164 The complete sequence encoding *Sg*Lys was amplified by PCR using specific primer pairs
165 designed based on the sequences obtained from 3' RACE and 5' RACE. The restriction sites for EcoRI
166 and NotI were introduced into the forward (SgLysEcoRI) and reverse (SgLysNotI) primers,
167 respectively, as indicated in Supplementary Table S1. The amplification was performed in a Master
168 cycler® Thermocycler (Eppendorf, Germany) using the following program: initial denaturation of 94°C
169 for 2 min followed by 40 cycles of denaturation at 94°C for 1 min, annealing at 56°C for 1 min and
170 extension at 72°C for 1 min. Final extension was carried out at 72°C for 20 min. The PCR product was

digested with EcoRI and NotI and separated by electrophoresis on 1.5 % agarose gel. The ~480 bp fragment was purified using PCR purification kit (Norgen Biotek Corporation, Canada) according to the manufacturer's instructions. The purified PCR product was cloned into EcoRI and NotI digested *Pichia pastori* expression vector pPIC9K (Invitrogen) to create in frame fusion of lysozyme gene with α-factor signal peptide. **The purpose of cloning full-length lysozyme (native *Sg*Lys) in *Pichia* with alfa factor signal sequence (MF-α Prepro) is to improve secretion based on human lysozyme expression studies (Oka et al. 1999)**. Representative clones containing complete coding sequence of lysozyme (pPIC9K-Lys) were confirmed by sequencing using α-factor (5′-TACTATTGCCAGCATTGCTGC-3′) and 3'AOX1 (5′-GCAAATGGCATTCTGACATCC-3′) primers (Supplementary Figure S1). The plasmid pPIC9K-Lys was linearized by digestion with BglII and transformed into *P. pastoris* strain GS115 by spheroplasting method according to Invitrogen Pichia expression kit manual.

To check that the gene of interest was integrated into the *Pichia* genome, genomic DNA isolated from the recombinant *Pichia* GS115 was analyzed by PCR using either α-factor/3'-AOX1 primers or 5'-AOX1 (5′-GACTGGTTCCAATTGACAAGC-3′)/3'-AOX1 primers. The expected size of amplicon for pPIC9K-Lys is ~972bp with α-factor/3'-AOX1 primers and ~675bp with 5'-AOX1/3'-AOX1 primers.

P. pastoris GS115 transformed with pPIC9K-SgLys were grown in BMGY medium at 30°C with shaking at 250 rpm (buffered glycerol complex medium) for 48 h. The cells were harvested by centrifugation at 10,000 for 5 min. To induce expression, the cells were resuspended in 1/5 the original volume of culture with BMMY medium (buffered methanol complex medium) and grown for 4 days with shaking at 30°C. At every 24 h, 100% methanol was added to a final concentration of 0.5% to maintain induction. The induced culture was centrifuged at 10,000 rpm for 30 min and the culture supernatant containing r*Sg*Lys was used (Supplementary Figure S2A).

Antibacterial assay

Activity of r*Sg*Lys was determined by the zone inhibition test as described in Mohamed et al. (2013). In this approach, a suspension of 1 mg lyophilized *Micrococcus luteus* (ATCC4698) cells (Sigma-Aldrich) in 50 mM phosphate buffer pH 8.0 mixed in 1% agarose was poured in Petri dishes. A well was done by puncture with an inverted Pasteur pipette, and known amounts of expressed r*Sg*Lys or, as control, hen egg white lysozyme (HEWL) were applied. Finally, the plates were incubated overnight at 37°C and the activity was recorded by measuring the diameter (mm) of the clear zone. Muramidase activity of r*Sg*Lys was detected (Supplementary Figure S2B).

Sequence analysis

Nucleotide sequence analysis and prediction of the open reading frame were performed with Genetyx software version 7.3.0 (GENETYX, Tokyo, Japan). The nucleotide homology search for the DNA sequence and protein homology search for the translated amino acids were performed by BLAST on the NCBI website (http://blast.ncbi.nlm.nih.gov/Blast.cgi). Sequence alignment analysis was performed with the ClustalW program (Thompson et al., 1994) and BioEdit v7.1.3 program (http://www.mbio.ncsu.edu/bioedit/bioedit.html). Phylogenetic analysis was performed with MEGA6 (http://www.megasoftware.net). The molecular weight and isoelectric point of the protein were estimated with the use of ProtParam tool (http://www.expasy.org/tools/protparam.html). The signal peptide was predicted by SignalP (http://www.cbs.dtu.dk/services/SignalP/). The domain search was performed by the CD-search in NCBI (http://www.ncbi.nlm.nih.gov/structure/cdd/cdd.shtml). For 3D analysis of *Sg*Lys, alignment of mature lysozymes without signal peptides was performed by ClustalW and *Bombyx* lysozyme (*Bm*Lys, PDB1G6A) was used as a suitable template in the homology modeling analysis using CPH models 3.0.

SDS-PAGE and Western blotting

Samples from the hemolymph of *E. coli*-immune challenged *S. gregaria* 5th instar nymphs collected at different time intervals, purified lysozyme from *S. gregaria* plasma (Mohamed et al., 2013) or r*Sg*Lys were subjected to 12% SDS-PAGE, according to Laemmli (1970)(Laemmli, 1970) and stained with Coomassie brilliant blue or transferred from the gel onto polyvinylidene difluoride membrane (Sigma-Aldrich) by electroblotting using semi-dry transfer cell (Bio-Rad) at 15 V for 1 h at room temperature in transfer buffer (25 mM Tris pH 8.3, 192 mM glycine, 20% (v/v) methanol). Non-specific binding sites on the PVDF membrane were blocked in a 5% skim milk in 1x Tris-buffered saline, 0.1% Tween 20 [TBS-T] solution and then incubated with a 1:2000 dilution of a polyclonal antiserum to *Sg*Lys (Elmogy et al., 2015) as primary antibody, in 1% skim milk–TBS–T for 90 min. Blots were washed with TBS–T containing 1% skim milk (TBS-T-M) and incubated with a 1:2000 dilution of a goat anti-rabbit horseradish peroxidase conjugated antibody (Sigma) as secondary antibody for another 90 min. Blots were washed three times in TBS–T-M. Finally, color development was carried out by adding a diaminobenzidine (DAB) solution (5 mg of 3,3'-diaminobenzidine (Sigma) in 100 mL of 50 mM Tris-HCl, pH 7.2) and 2 µL of 30% H_2O_2. Control experiments were performed to check the specificity of the antiserum and that no cross reactive bands were found on the membrane (Supplementary Figure S2A).

Analysis of developmental and tissues expression of *SgLys* transcripts by RT-PCR

Tissue expression patterns of *SgLys* was conducted on total RNA from cuticular epidermis (whole integument), fat body, hemocytes, midgut, salivary glands, and thoracic muscles by semi-quantitative RT-PCR. On the other hand, *SgLys* expression analysis during locust development was performed on total RNA from eggs, fat body tissues of different instars (1st to 5th), and adults (male and female) stages.

Briefly, total RNA (3 µg extracted from each tissue) was used to synthesize cDNA with the RevertAid™ first Strand cDNA Synthesis Kit and an oligo (dT) primer (Fermentas Life Sciences). **The**

total RNA preparations were treated with DNase I (RNase-free, Thermo Fisher Scientific) in order to digest possible contaminating genomic DNA. The RT-PCR primers used were Lys-tF1 and Lys-tR2 (Supplementary Table S1) which gave a PCR product of 479 bp. The PCR conditions consisted of an initial denaturation at 94°C for 2 min, followed by 30 cycles of denaturation at 94°C for 1 min, annealing at 56°C for 1 min and extension at 72°C for 1 min, and a final extension at 72°C for 5 min.

The expression of the housekeeping gene *β*-actin (Genbank accession number: HQ851398) was used as an internal control with the forward primer 5'-GTGGGGCGCCCCAGGCACCA-3' and the reverse primer 5'-CTCCTTAATGTCACGCACGATTTC-3' and under the same PCR conditions as above. PCR products from three parallel experiments were separated on a 1.5% agarose gel. The gels were stained with ethydium bromide and the quantitation of RT-PCR produced band intensity was carried out in Image J software (http://rsbweb.nih.gov/ij/download.html). Finally, the *SgLys*/*β-actin* ratios were calculated.

Results and discussion

Structure of the *SgLys* gene

The full-length cDNA sequence was obtained by combining sequences of the 3'-RACE and 5'-RACE. The complete (end to end) nucleotides sequence and the deduced amino acids of *S. gregaria* lysozyme are presented in Figure 1. The results show that the cDNA is consisted of 659 bp. It contains untranslated sequences (UTR) in the 5' region (26 bp upstream from start codon) and in the 3' region downstream of stop codon at positions 453-659. The 3'UTR is ended with an 18 poly(A) tail. The 5'UTR has moderate G+C content (65%), while the 3' UTR has low G+C content (41%). The sequence flanking the putative initiator codon follows the rule of eukaryotic initiation sites (Kozak, 1981). Hence, the first ATG at positions 27–29 is most probably the initiation codon. The TGA at positions 450–452 was assigned as the termination codon, and thus an open reading frame (ORF) of 423 bp was predicted (Figure 1). The sequence contained a possible polyadenylation signal (AATAAA) at the

3' end (double underlined in Figure 1). The obtained cDNA sequence of *S. gregaria* lysozyme was designated as *Sg*Lys, submitted to NCBI GenBank and allotted the accession number JQ012999.

Characterization of the predicted *S. gregaria* c-type lysozyme

The encoded amino acid sequence of *Sg*Lys was also deduced with the Genetyx-Win version 7.0.3 software. The *Sg*Lys gene contains a 423 bp ORF with a possible initiation codon at position 27-29 of the nucleotide sequence which confirms our analysis above. The ORF encodes a 141 amino acid sequence composed of 22 residues at the N-terminal side predicted as a signal peptide (bold letters in Figure 1) and 119 residues as a mature peptide. The predicted 22 residues of the signal peptide are mostly hydrophobic amino acids, a feature common to the signal peptides of both eukaryotes and prokaryotes (von Heijne, 1990). Based on these predictions, the molecular weight and isoelectric point of the full-length *Sg*Lys were estimated to be 15,702 Da and 9.09, respectively, **whereas the molecular mass and Ip of mature lysozyme were 13.4 kDa and 9.2, respectively.** Also, the partial N-terminal amino acid sequence of the biochemically purified lysozyme (Elmogy et al., 2015), **KLQR*EIVSALKRHGITSDLRNWV*LVESESGGRTDKRGPRNKNGSY obtained by Edman degradation, is closely matches the deduced amino acid sequence of the mature protein, based on the cDNA data** (underlined in Figure 1). **In Elmogy et al. (2015) some amino acids could not be determined during the Edman degradation and therefore were supposed to be cysteines (* in the above sequence).**

BLASTp search on NCBI (http://blast.ncbi.nlm.nih.gov/) using *Sg*Lys deduced amino acids sequence as query showed that the *Sg*Lys has moderate homology (from 43% to 93% identities) to several reported insects and ticks lysozymes. It has 62% amino acid identity with *Papilio xuthus* (Li et al., 2015), 60% with *Manduca sexta* (Rosenthal and Dahlman, 1991), 59% with *Simulium nigrimanum*, 57% with *G. mellonella*, *Samia cynthia ricini* (Fujimoto et al., 2001), *Hyalophora cecropia* (Engström et al., 1985), *Ostrinia nubilalis* (Khajuria et al., 2011), and 55% with *Bombyx mori* (Matsuura et al., 2002)

and *O. furnicalis*. On the other hand, lysozymes from Orthoptera are highly conserved since the percent identity of *Sg*Lys with that of *Locusta migratoria* reaches 93%.

Alignment of the mature *Sg*Lys amino acid sequence with nineteen mature lysozymes from different insects and ticks was done using Clustal W and is shown in Supplementary Figure S3. According to the alignment, the *Sg*Lys has the 20 amino acid residues (asterisks in Supplementary Figure S3) highly conserved in the vast majority of c-type lysozymes (Fujita et al., 2002; Grunclová et al., 2003; Ursic-Bedoya et al., 2008) including the 8 cysteine residues, Cys^{6}, Cys^{26}, Cys^{61}, Cys^{70}, Cys^{74}, Cys^{88}, Cys^{108} and Cys^{119} (full circles in Supplementary Figure S3). The high conservation of the eight cysteine residues indicates their importance in the formation of disulfide bridges necessary for the stabilization of the molecule. Moreover, the alignment of the two Orthopteran lysozymes presented in Figure 2 shows that *Sg*Lys possesses all of the features of a c-type lysozyme with the two residues, Glu^{31} and Asp^{49} of the catalytic site (arrow heads in Figure 2 and Supplementary Figure S3), the 11 residues Ser^{30}, Glu^{31}, Gly^{33}, Gly^{40}, Gln^{54}, Asn^{56}, Tyr^{59}, Ile^{92}, Arg^{95}, Gly^{100}, and Trp^{101} that compose the lysozyme catalytic cleft (Figure 2, squares) and the residues Asn^{79}, Asp^{84} and Asp^{85} of the Ca^{2+} binding site (Figure 2, asterisks), a characteristic of most lactalbumin and few lysozymes (McKenzie, 1996). These three features form the conserved domain called LYZ1. Together with the two catalytic residues, Glu^{31} and Asp^{49} of the *Sg*Lys, the 8 cysteines are fundamental for the three dimensional structure and the biological activity of the lysozymes (Prager and Jolles, 1996). One major difference in the active site residues in terms of hydrophobicity between *Sg*Lys and the other c-type lysozymes occurs at Gly^{100} (hydropathy index: -0.4) of *Sg*Lys, which is Ala (hydropathy index: 1.3) in all other lysozymes but that of *Reticulitermes speratus*.

The active site of HEWL (Kumagai et al., 1992) consists of 6 subsites from A to F which are able to bind 6 sugar residues (Maenaka et al., 1994; Maenaka et al., 1995). These subsites are present in the *Sg*Lys as Ser^{30}, Gly^{33}, Tyr^{59}, Trp^{60}, Arg^{95}, and Trp^{101} (Figure 2) and are believed to interact with sugar rings placed in the subsites E, F, B, C, A and D, respectively. However, the *Sg*Lys shows some

amino acids substitutions in these subsites when compare to HEWL. Indeed, the HEWL subsites have Phe^{44}, Asn^{37}, Trp^{62}, Trp^{63}, Asp^{101} and Trp^{108}, respectively. The substitutions of the conserved residues with others were supposed to modulate the enzymatic activity (Jolles and Jolles, 1984). For example, when Asn^{37} and Trp^{62} of HEWL were replaced by site-directed mutagenesis with Gly and Tyr, respectively, an enhanced bacteriolytic activity was observed (Kumagai et al., 1987; Kumagai and Miura, 1989). As a matter of fact, such substitutions occur naturally in *Sg*Lys, since the sequence of *S. gregaria* contains a Gly in position 33 and a Tyr in position 59, and indeed, *Sg*Lys exhibits about 2-3 fold higher bacteriolytic activity than HEWL (Mohamed et al., 2013). This observation is in agreement with those reported in human and rat (Mulvey et al., 1973; Mulvey et al., 1974) showing the presence of Gly and Tyr of subsites F and B, respectively, and that these lysozymes exhibit higher bacteriolytic activity than HEWL (4-fold in the case of human lysozyme).

α-helical hairpin structures were reported in c-type lysozymes of chicken, equine, and human by computer-assisted structural analysis (Ibrahim et al., 2001; Nakajima and Kikuchi, 2007), and of *B. mori* (Matsuura et al., 2002), *O. furnicalis* (Wang et al., 2009) as well as in bactericidal and cytolytic pore-forming peptides (Andreu et al., 1985; Boman et al., 1991; Chen et al., 1988; Engelman and Steitz, 1981; Song et al., 1994; Song et al., 1991; Srisailam et al., 2000). They vary in length, conformation, net charge and in its distribution, and hydrophobicity, whose characteristics seem to be main factors leading to variations in their activities (Ibrahim et al., 2001). The helix-loop-helix (HLH) of HEWL comprises the amino acid residues 87-114, of which residues 87-100, 101-106 and 107-114 form the N-terminal helix (H1), the loop (L), and the C-terminal helix (H2), respectively (Figure 3A). When the amino acid sequence of *Sg*Lys was aligned with the corresponding sequence of the HEWL HLH, homologies of 78%, 91% and 85% were revealed (Figure 3B). Moreover, comparison of a generated three dimensional structures of *Sg*Lys with that of *Bm*Lys (Matsuura et al., 2002), which shows 66.6% identity of active site residues with *Sg*Lys (Figure 3C), reveals the presence of an α-helical hairpin motif formed of two α-helices joined by a loop stabilized by glycine (Figure 3D).

Pellegrini et al. (1997) reported that a peptide composed of residues 98-112 of HEWL, released from α-helical hairpin motif by digestion with clostripain (an ArgC-endoproteinase), was shown to have antimicrobial activity against both Gram-positive and Gram-negative bacteria and that the replacement of Asn^{106} with a positively charged Arg generated an increased antimicrobial activity of this peptide. As a matter of fact, an arginine residue is naturally present in *Sg*Lys (as Arg^{99})(Figure 3B) and may also contribute in the observed elevation of *Sg*Lys antimicrobial activity over that of HEWL (Mohamed et al., 2013).

α-Helical hairpin of HEWL (Ibrahim et al., 2001) and of antimicrobial peptides (Hung et al., 1999; Srisailam et al., 2000; Wang et al., 1999) are active against both Gram-negative and Gram-positive bacteria by channel formation in the bacterial membrane. *Sg*Lys has two tryptophan residues Trp^{101} and Trp^{104} homologous to Trp^{108} and Trp^{111} of HEWL which are located in the C-terminal helix. When either Trp^{108} or Trp^{111} of HEWL was replaced with the less hydrophobic tyrosine, the antimicrobial activity was reduced (Pellegrini et al., 1997). In this respect, Ibrahim et al. (2001) suggested that the amphipathic N-terminal α-helix could account for membrane insertion that places the C-terminal α-helix (containing the tryptophan residues) close to the polar membrane interface. This could be supported by the higher hydrophobicity of the N-terminal α-helix over that of the C-terminal one. The Trp residues are responsible for inducing functionally active conformation of the channel due to their specific preference for the interface (Cross et al., 1999; Hu et al., 1993; Kelkar and Chattopadhyay, 2007). The time-dependent rate of passage of the cytoplasmic β-galactosidase into the incubation medium (Elmogy et al., 2015) indicates a proceeding process of the inner membrane permeabilization by the *Sg*Lys which is most probably due to the referred insertion ability and pore formation.

_Sg_Lys domains and signatures analysis

The predicted *Sg*Lys mature protein was subjected to analysis using ScanProsite (release 20.79) to identify protein domains, families, and functional sites as well as associated patterns and profiles. The scan results indicated the existence of α-lactalbumin/lysozyme C family profile (residues 1–119) and Alpha-lactalbumin/lysozyme C signature with the following consensus pattern C-x(3)-C-x(2)-[LMF]-x(3)-[DEN]-[LI]-x(5)-C (70-88). The scan also determined the positions of the four disulfide bridges to exist between the cysteine residues 6-119, 26-108, 61-74 and 70-88.

Phylogenetic analysis

The phylogenetic relationship of the *Sg*Lys with c-type lysozymes from different insect orders and ticks was inferred using the Neighbor-joining method (Saitou and Nei, 1987) to determine the closest associations among these lysozymes. The analysis conducted with MEGA6 was performed on amino acid sequences of mature lysozymes and HEWL was used as out-group. The data presented in Figure 4 show a general separation of lysozymes into two main groups likely based on their function. Lysozymes from *Lepidoptera* and *Nematocera* (including mosquitoes and black flies) of immune-related function are grouped together in a large clade while those from bugs and cyclorrhaphan flies, which are of digestive function and/or midgut origin, group together under two branches in another clade. In many nematocerans, lysozymes are found in the hemolymph but not in the gut (Lemos and Terra, 1991), and thus the recruitment of lysozymes as digestive enzymes, and their adaptation to an acidic midgut, may have occurred after the divergence of *Cyclorrhapha* from *Nematocera*. The mosquito lysozymes are expressed in the salivary gland (Kang et al., 1996; Moreira-Ferro et al., 1998) and it has been suggested that they might serve to control the microbial population ingested with a sugar meal (Rossignol and Lueders, 1986). This apparent dichotomy between recruitment of lysozyme for a digestive function (mainly in Cyclorrhapha flies) and as an immune peptide (Lepidoptera), was previously reported (Ursic-Bedoya et al., 2005; Ursic-Bedoya et al., 2008).

Tick lysozymes of immune-responsive function like those of the hard ticks *Amblyomma maculatum*, *Dermacentor andersoni*, and *Haemaphysalis longicornis* (Simser et al., 2004; Tanaka et al., 2010) are grouped together in a distinct clade whereas the lysozyme of the soft tick *Ornithodoros moubata* whose function has been described as digestive (Grunclová et al., 2003) separates in another major branch. Lysozymes from the coleopterans *Sitophilus zeamais* (Anselme et al., 2008) and *Tribolium castaneum* (ACV32411) are grouped together in their own clade. **In these latters, the multi-lysozyme domains, with the so-called H-branch of lysozymes share a histidine residue (H) which replaces the conserved tyrosine residue (Y) adjacent to the catalytic aspartic acid residue (D) in the majority of known c-type lysozymes (Beckert et al., 2015).**

The *Sg*Lys, whose activity is greatest in the fat body (Mohamed et al., 2013), is found clustered with the lysozyme of the other orthopteran, *Locusta migratoria*, in a clade of molecules whose function has been described as immune-related, including from distantly related *Lepidoptera* and *Diptera*. It is grouped with the subtree of the hymenopteran lysozymes, but in a separate branch, and both are diverging from the lepidopteran clade.

Expression of *Sg*Lys protein

We have previously reported that a bacterial challenge with *E. coli* induced an increased lysozyme activity in the hemolymph of 5th instars larvae of *S. gregaria* and a maximum of activity was found 6 h after the injection of the bacteria (Mohamed et al., 2013). In the present study, by western blot analysis and the use of a polyclonal antiserum directed towards *Sg*Lys (Elmogy et al., 2015 and Supplementary Figure S2), a single protein band with a molecular weight of about 15.7 kDa was stained in the hemolymph of control larvae (not injected with *E. coli*)(Figure 5) which size corresponds to the lysozyme with its signal peptide still attached. **We suggest that this apparent molecular mass of around 16 kDa is likely due to an unusual migration of the protein upon reduction of disulphide bridges rather than of the non-cleavage of the signal peptide. A**

409 **phenomenon that was reported before by Koganesawa et al. (2001) for the recombinant *B. mori***
410 **lysozyme**. As a matter of fact, Liu et al. (2004) reported the presence of two forms of lysozymes, one
411 of 14.4 kDa and the other of 16.4 kDa, in salivary gland homogenates of *Helicoverpa zea*. These
412 authors showed that the lower molecular weight form was the processed mature lysozyme while the
413 higher molecular weight peptide was considered as the predicted unprocessed, unsecreted form of
414 lysozyme which might be the case in our analysis for *Sg*Lys. Figure 5 also shows a remarkable change
415 in the levels of *Sg*Lys protein in the hemolymph of 5th instars *S. gregaria* following injection with *E.*
416 *coli*. The maximal protein level was detected 6 h post injection which is in agreement with our
417 previous observation (Mohamed et al., 2013).

418 **Developmental and tissue-specific transcriptional expression of *SgLys***

419 Transcriptional expression of *SgLys* at different developmental stages of *S. gregaria* was
420 examined by semi-quantitative RT-PCR (semiQ RT-PCR) on total RNA from eggs or RNA from fat body
421 for the different instars and the adults. The results are presented in Figure 6. They show a
422 progressive and strong *SgLys* expression throughout the insect development with a gradual increase,
423 starting from the 2nd instar up to the adult. Therefore, *SgLys* appears to be developmentally
424 regulated, as this is the case for *M. sexta* (Russell and Dunn, 1991) and *H. armigera* (Zhang et al.,
425 2009). In *Aedes aegypti*, the expression of LysE in the larval and pupal stages, as well as defensin
426 expression in pupae, was supposed to prevent bacteria-induced septicaemia within these stages
427 (Ursic-Bedoya et al., 2005). In *D. melanogaster*, LysX is only expressed in early pupae and was
428 proposed to play a similar protective role (Daffre et al., 1994).

429 Transcriptional expression of *SgLys* was analyzed on different tissues by semiQ RT-PCR (Figure
430 7). A constitutive expression of *SgLys* was found in epidermis, fat body and salivary glands but could
431 not be detected in hemocytes and thoracic muscles. After injection of *E. coli*, *SgLys* transcripts were
432 detected in all tissues but midgut with higher levels in epidermis, fat body and salivary glands.

Therefore, it appears that a bacterial challenge triggers an up-regulation of *SgLys* in a tissue-specific pattern. In the fat body, *E. coli* elicited a 2-fold increase in *SgLys* transcription 6 h post-injection. Consequently, the fat body, salivary gland, epidermis and hemocytes seem to be capable of discriminating the microbial stress factor. Substantial expression of *SgLys* in thoracic muscles was observed which may be due to a contamination of this sample with infiltrated hemocytes in this tissue at the site of injection. This phenomenon was reported during the course of lysozyme gene expression in *M. sexta*, following treatment with peptidoglycans (Mulnix and Dunn, 1994).

The expression profile of *SgLys* in the different immune-related tissues reported here is similar to that of other recognized immune lysozymes. Generally lysozymes are rapidly expressed in response to microbial infection in specific tissues, mainly fat body and hemocytes (Bulet and Stocklin, 2005; Callewaert and Michiels, 2010; Fujita, 2004; Hultmark, 1996). In *M. sexta*, low levels of transcripts in fat body of non-infected larvae increased rapidly after treatment with peptidoglycans and remained elevated over several days (Mulnix and Dunn, 1994). These authors also reported a very low lysozyme transcript level in salivary glands of this insect. In contrast, *H. zea* salivary gland lysozyme is not only present at high transcription levels but is also one of the most abundant proteins in saliva (Liu et al., 2004).

Whereas typical c-type immune-related lysozymes are usually basic enzymes with Ip around 8 or higher, the recruitment of lysozymes for digestive purposes like the *Drosophila* midgut lysozymes is much related to their acidic Ip and/or pH optima of gut lumen (Hultmark, 1996). In insects, lysozymes have a major role in immune responses to pathogens whereas other reports have identified them from larval and adult digestive tracts and from salivary glands, suggesting their role in digestion. **Taking together, c-type lysozymes have been postulated to display immunity- or digestion-related functions. For instance, bacterial-challenge of *H. axyridis* boosted the basic c-lys4 (Ip = 8.18) gene expression 8-fold in the gut, whereas the acidic c-lys3 (Ip= 5.46) gene was expressed at comparable levels in both naïve and challenged beetles (Beckert et al. 2015). These**

458 **two *H. axyridis* c-type lysozymes possessing muramidase activity, may help to control the gut flora,**
459 **and thus are likely to play an important role in the innate immune system.**

460 The data presented here suggest a defensive role of the lysozyme produced by fat body,
461 salivary glands, thoracic muscles, hemocytes and epidermis in *S. gregaria*. *SgLys* is constitutively
462 expressed and is up-regulated after bacterial challenge. It is expressed throughout all developmental
463 stages, suggesting a key role for this gene in *S. gregaria* innate immunity. This is especially true in
464 orthopterans when taking into account that no antimicrobial peptides, such as those inducible after
465 the activation of the two signaling pathways Toll and Imd, have been reported yet in these insects
466 with the exception of locustin (Accession number P83428, Bulet et al., 2002, unpublished).

Acknowledgments

The authors would like to thank Dr Venkatramana Krishna, Bhat Biotech's Vice President R&D (India) (currently postdoctoral researcher at University of Minnesota Twin Cities, USA) for his technical support and enthusiastic work. We are also indebted to the late Dr J. Nagaraju - Khorana Chair at Centre for DNA Fingerprinting and Diagnostics (CDFD), Hyderabad, India for his help. His passing has inflicted a deep loss on both the personal and professional fronts.

References

Abraham, E.G., Nagaraju, J., Salunke, D., Gupta, H.M., Datta, R.K., 1995. Purification and partial characterization of an induced antibacterial protein in the silkworm, *Bombyx mori*. J. Invertebr. Pathol. 65, 17-24.

Altincicek, B., Knorr, E., Vilcinskas, A., 2008. Beetle immunity: Identification of immune-inducible genes from the model insect *Tribolium castaneum*. Dev. Comp. Immunol. 32, 585-595.

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

Anselme, C., Perez-Brocal, V., Vallier, A., Vincent-Monegat, C., Charif, D., Latorre, A., Moya, A., Heddi, A., 2008. Identification of the weevil immune genes and their expression in the bacteriome tissue. BMC Biol. 6, 43.

Araujo, C.A., Waniek, P.J., Stock, P., Mayer, C., Jansen, A.M., Schaub, G.A., 2006. Sequence characterization and expression patterns of defensin and lysozyme encoding genes from the gut of the reduviid bug *Triatoma brasiliensis*. Insect Biochem. Mol. Biol. 36, 547-560.

Balczun, C., Knorr, E., Topal, H., Meiser, C.K., Kollien, A.H., Schaub, G.A., 2008. Sequence characterization of an unusual lysozyme gene expressed in the intestinal tract of the reduviid bug *Triatoma infestans* (*Insecta*). Parasitol. Res. 102, 229-232.

Bang, I.S., Yoe, S.M., 2005. Purification and cDNA cloning of lysozyme II from cabbage butterfly, *Artogeia rapae* larvae. Entomol. Res. 35, 207-211.

Beckert, A., Wiesner, J., Baumann, A., Pöppel, A.K., Vogel, H., Vilcinskas, A., 2015. Two c-type lysozymes boost the innate immunity system of the invasive ladybird *Harmonia axyridis*. Dev. Comp. Immunol. 49, 303-312.

Boman, H.G., Faye, I., Gudmundsson, G.H., Lee, J.Y., Lidholm, D.A., 1991. Cell-free immunity in *Cecropia*. A model system for antibacterial proteins. Eur. J. Biochem. 201, 23-31.

Bonasio, R., Zhang, G., Ye, C., Mutti, N.S., Fang, X., Qin, N., Donahue, G., Yang, P., Li, Q., Li, C., Zhang, P., Huang, Z., Berger, S.L., Reinberg, D., Wang, J., Liebig, J., 2010. Genomic comparison of the ants *Camponotus floridanus* and *Harpegnathos saltator*. Science 329, 1068-1071.

Bulet, P., Stocklin, R., 2005. Insect antimicrobial peptides: structures, properties and gene regulation. Protein Pept. Lett. 12, 3-11.

Callewaert, L., Michiels, C.W., 2010. Lysozymes in the animal kingdom. J. Biosci. 35, 127-160.

Cançado, F.C., Chimoy Effio, P., Terra, W.R., Marana, S.R., 2008. Cloning, purification and comparative characterization of two digestive lysozymes from *Musca domestica* larvae. Braz. J. Med. Biol. Res. 41, 969-977.

Chapelle, M., Girard, P.A., Cousserans, F., Volkoff, N.A., Duvic, B., 2009. Lysozymes and lysozyme-like proteins from the fall armyworm, *Spodoptera frugiperda*. Mol. Immunol. 47, 261-269.

Chen, H.C., Brown, J.H., Morell, J.L., Huang, C.M., 1988. Synthetic magainin analogues with improved antimicrobial activity. FEBS lett. 236, 462-466.

Cross, T.A., Arseniev, A., Cornell, B.A., Davis, J.H., Killian, J.A., Koeppe, R.E., 2nd, Nicholson, L.K., Separovic, F., Wallace, B.A., 1999. Gramicidin channel controversy--revisited. Nat. Struct. Biol. 6, 610-611.

Daffre, S., Kylsten, P., Samakovlis, C., Hultmark, D., 1994. The lysozyme locus in *Drosophila melanogaster*: an expanded gene family adapted for expression in the digestive tract. Mol. Gen. Genet. 242, 152-162.

Elmogy, M., Bassal, T.T., Yousef, H.A., Dorrah, M.A., Mohamed, A.A., Duvic, B., 2015. Isolation, characterization, kinetics, and enzymatic and nonenzymatic microbicidal activities of a novel c-type lysozyme from plasma of *Schistocerca gregaria* (*Orthoptera*: *Acrididae*). J. Insect Sci. 15, 57

Engelman, D.M., Steitz, T.A., 1981. The spontaneous insertion of proteins into and across membranes: the helical hairpin hypothesis. Cell 23, 411-422.

Engström, A., Xanthopoulos, K.G., Boman, H.G., Bennich, H., 1985. Amino acid and cDNA sequences of lysozyme from *Hyalophora cecropia*. EMBO J. 4, 2119-2122.

Fujimoto, S., Toshimori-Tsuda, I., Kishimoto, K., Yamano, Y., Morishima, I., 2001. Protein purification, cDNA cloning and gene expression of lysozyme from eri-silkworm, *Samia cynthia ricini*. Comp. Biochem. Physiol. B Biochem. Mol. Biol. 128, 709-718.

Fujita, A., 2004. Lysozymes in insects: what role do they play in nitrogen metabolism ? Physiol. Entomol. 29, 305-310.

Fujita, A., Minamoto, T., Shimizu, I., Abe, T., 2002. Molecular cloning of lysozyme-encoding cDNAs expressed in the salivary gland of a wood-feeding termite, *Reticulitermes speratus*. Insect Biochem. Mol. Biol. 32, 1615-1624.

Gerardo, N.M., Altincicek, B., Anselme, C., Atamian, H., Barribeau, S.M., de Vos, M., Duncan, E.J., Evans, J.D., Gabaldón, T., Ghanim, M., Heddi, A., Kaloshian, I., Latorre, A., Moya, A., Nakabachi, A., Parker, B.J., Pérez-Brocal, V., Pignatelli, M., Rahbé, Y., Ramsey, J.S., Spragg, C.J., Tamames, J., Tamarit, D., Tamborindeguy, C., Vincent-Monegat, C., Vilcinskas A., 2010. Immunity and other defenses in pea aphids, *Acyrthosiphon pisum*. Genome Biol. 11(2):R21.

Grunclová, L., Fouquier, H., Hypsa, V., Kopácek, P., 2003. Lysozyme from the gut of the soft tick *Ornithodoros moubata*: the sequence, phylogeny and post-feeding regulation. Dev. Comp. Immunol. 27, 651-660.

He, Y., Cao, X., Li, K., Hu, Y., Chen, Y., Blissard, G., Kanost, M., Jiang, H. 2015. A genome-wide analysis of antimicrobial effector genes and their transcription patterns in *Manduca sexta*. Insect Biochem. Mol. Biol. 62, 23-37.

Hu, W., Lee, K.C., Cross, T.A., 1993. Tryptophans in membrane proteins: indole ring orientations and functional implications in the gramicidin channel. Biochemistry 32, 7035-7047.

Hultmark, D., 1996. Insect lysozymes. In: *Lysozymes: Model Enzymes in Biochemistry and Physiology* (Jollès, P., Ed.). Birkhäuser, Basel. 75, 87-102.

Hung, S.C., Wang, W., Chan, S.I., Chen, H.M., 1999. Membrane lysis by the antibacterial peptides cecropins B1 and B3: A spin-label electron spin resonance study on phospholipid bilayers. Biophys. J. 77, 3120-3133.

Ibrahim, H.R., Thomas, U., Pellegrini, A., 2001. A helix-loop-helix peptide at the upper lip of the active site cleft of lysozyme confers potent antimicrobial activity with membrane permeabilization action. J. Biol. Chem. 276, 43767-43774.

Jain, D., Nair, D.T., Swaminathan, G.J., Abraham, E.G., Nagaraju, J., Salunke, D.M., 2001. Structure of the induced antibacterial protein from tasar silkworm, *Antheraea mylitta*. Implications to molecular evolution. J. Biol. Chem. 276, 41377-41382.

Jiang, H., Vilcinskas, A., Kanost, M.R., 2010. Immunity in lepidopteran insects. Adv. Exp. Med. Biol. 708, 181-204.

Jolles, J., Schoentgen, F., Croizier, G., Croizier, L., Jolles, P., 1979. Insect lysozymes from three species of Lepidoptera: their structural relatedness to the C (chicken) type lysozyme. J. Mol. Evol. 14, 267-271.

Jolles, P., Jolles, J., 1984. What's new in lysozyme research? Always a model system, today as yesterday. Mol. Cell. Biochem. 63, 165-189.

Kajla, M.K., Andreeva, O., Gilbreath, T.M.I., Paskewitz, S.M., 2010. Characterization of expression, activity and role in antibacterial immunity of *Anopheles gambiae* lysozyme c-1. Comp. Biochem. Physiol. B Biochem. Mol. Biol. 155, 201-209.

Kang, D., Liu, G., Gunne, H., Steiner, H., 1996. PCR differential display of immune gene expression in *Trichoplusia ni*. Insect Biochem. Mol. Biol. 26, 177-184.

Kelkar, D.A., Chattopadhyay, A., 2007. The gramicidin ion channel: a model membrane protein. Biochim. Biophys. Acta 1768, 2011-2025.

Khajuria, C., Buschman, L.L., Chen, M.S., Zurek, L., Zhu, K.Y., 2011. Characterization of six antibacterial response genes from the European corn borer (*Ostrinia nubilalis*) larval gut and their expression in response to bacterial challenge. J. Insect Physiol. 57, 345-355.

Kim, Y.S., Yoe, S.M., 2003. Purification of lysozyme from hemolymph of tabacco cutworm, *Spodoptera litura*. J. Entomol. 33, 287-291.

Koganesawa, N., Aizawa, T., Masaki, K., Matsuura, A., Nimori, T., Bando, H., Kawano, K., Nitta, K., 2001. Construction of an expression system of insect lysozyme lacking thermal stability: the effect of selection of signal sequence on level of expression in the *Pichia pastoris* expression system. Protein Eng. 14, 705-710.

Kozak, M., 1981. Possible role of flanking nucleotides in recognition of the AUG initiator codon by eukaryotic ribosomes. Nucleic Acids Res. 9, 5233-5252.

Kumagai, I., Kojima, S., Tamaki, E., Miura, K., 1987. Conversion of Trp 62 of hen egg-white lysozyme to Tyr by site-directed mutagenesis. J. Biochem. 102, 733-740.

Kumagai, I., Miura, K., 1989. Enhanced bacteriolytic activity of hen egg-white lysozyme due to conversion of Trp62 to other aromatic amino acid residues. J. Biochem. 105, 946-948.

Kumagai, I., Sunada, F., Takeda, S., Miura, K., 1992. Redesign of the substrate-binding site of hen egg white lysozyme based on the molecular evolution of C-type lysozymes. J. Biol. Chem. 267, 4608-4612.

Laemmli, U.K., 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227, 680-685.

Lecoq, M., 2004. Vers une solution durable au problème du criquet pélerin ? Sécheresse 15, 217-224.

Lee, W.J., Brey, P.T., 1995. Isolation and characterization of the lysozyme-encoding gene from the silkworm *Bombyx mori*. Gene 161, 199-203.

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

Lemos, F.J., Terra, W.R., 1991. Digestion of bacteria and the role of midgut lysozyme in some insect larvae. Comp. Biochem. Physiol. B Biochem. Mol. Biol. 100, 265-268.

Li, B., Calvo, E., Marinotti, O., James, A.A., Paskewitz, S.M., 2005. Characterization of the c-type lysozyme gene family in *Anopheles gambiae*. Gene 360, 131-139.

Li, X., Fan, D., Zhang, W., Liu, G., Zhang, L., Zhao, L., Fang, X., Chen, L., Dong, Y., Chen, Y., Ding, Y., Zhao, R., Feng, M., Zhu, Y., Feng, Y., Jiang, X., Zhu, D., Xiang, H., Feng, X., Li, S., Wang, J., Zhang, G., Kronforst, M.R., Wang, W., 2015. Outbred genome sequencing and CRISPR/Cas9 gene editing in butterflies. Nat. Commun. 6, 8212.

Liu, F., Cui, L., Cox-Foster, D., Felton, G.W., 2004. Characterization of a salivary lysozyme in larval *Helicoverpa zea*. J. Chem. Ecol. 30, 2439-2457.

Lockey, T.D., Ourth, D.D., 1996. Purification and characterization of lysozyme from hemolymph of *Heliothis virescens* larvae. Biochem. Biophys. Res. Commun. 220, 502-508.

Maenaka, K., Kawai, G., Watanabe, K., Sunada, F., Kumagai, I., 1994. Functional and structural role of a tryptophan generally observed in protein-carbohydrate interaction. TRP-62 of hen egg white lysozyme. J. Biol. Chem. 269, 7070-7075.

Maenaka, K., Matsushima, M., Song, H., Sunada, F., Watanabe, K., Kumagai, I., 1995. Dissection of protein-carbohydrate interactions in mutant hen egg-white lysozyme complexes and their hydrolytic activity. J. Mol. Biol. 247, 281-293.

Matsuura, A., Yao, M., Aizawa, T., Koganesawa, N., Masaki, K., Miyazawa, M., Demura, M., Tanaka, I., Kawano, K., Nitta, K., 2002. Structural analysis of an insect lysozyme exhibiting catalytic efficiency at low temperatures. Biochemistry 41, 12086-12092.

McKenzie, H.A., 1996. Alpha-lactalbumins and lysozymes as a model. In: *Lysozymes: Model Enzymes in Biochemistry and Physiology* (Jollès, P., Ed.). Birkhäuser, Basel. 75, 365-409.

Mohamed, A.A., Elmogy, M., Dorrah, M.A., Yousef, H.A., Bassal, T.T., 2013. Antimicrobial activity of lysozyme in the desert locust, *Schistocerca gregaria* (*Orthoptera*: *Acrididae*). Eur. J. Entomol. 110, 559-565.

Mohrig, W., Messner, B., 1968. Lysozyme as antibacterial agent in honey and bees venom. Acta Biol. Med. Ger. 21, 85-95.

Moreira-Ferro, C.K., Daffre, S., James, A.A., Marinotti, O., 1998. A lysozyme in the salivary glands of the malaria vector *Anopheles darlingi*. Insect Mol. Biol. 7, 257-264.

Mulnix, A.B., Dunn, P.E., 1994. Structure and induction of a lysozyme gene from the tobacco hornworm, *Manduca sexta*. Insect Biochem. Mol. Biol. 24, 271-281.

Mulvey, R.S., Gaultieri, R.J., Beychok, S., 1973. Spectral properties of human lysozyme and its inhibitor complexes. Fluorescence and difference spectra. Biochemistry 12, 2683-2690.

Mulvey, R.S., Gualtieri, R.J., Beychok, S., 1974. Composition, fluorescence, and circular dichroism of rat lysozyme. Biochemistry 13, 782-787.

Nakajima, S., Kikuchi, T., 2007. Analysis of the differences in the folding mechanisms of c-type lysozymes based on contact maps constructed with interresidue average distances. J. Mol. Model. 13, 587-594.

Oka, C., Tanaka, M., Muraki, M., Harata, K., Suzuki, K., Jigami, Y., 1999. Human lysozyme secretion increased by alpha-factor pro-sequence in *Pichia pastoris*. Biosci. Biotechnol. Biochem. 63, 1977-1983.

Park, H.Y., Park, S.S., Shin, S.W., Park, D.S., Kim, M.G., Oh, H.W., Joo, C.K., 1997. Protein purification and nucleotide sequence of a lysozyme from the bacteria-induced larvae of the fall webworm, *Hyphantria cunea*. Arch. Insect Biochem. Physiol. 35, 335-345.

Pellegrini, A., Thomas, U., Bramaz, N., Klauser, S., Hunziker, P., von Fellenberg, R., 1997. Identification and isolation of a bactericidal domain in chicken egg white lysozyme. J. Appl. Microbiol. 82, 372-378.

Prager, E.M., Jolles, P., 1996. Animal lysozymes c and g: an overview. In: *Lysozymes: Model Enzymes in Biochemistry and Physiology* (Jollès, P., Ed.). Birkhäuser, Basel. 75, 9-31.

Regel, R., Matioli, S.R., Terra, W.R., 1998. Molecular adaptation of *Drosophila melanogaster* lysozymes to a digestive function. Insect Biochem. Mol. Biol. 28, 309-319.

Rosenthal, G.A., Dahlman, D.L., 1991. Studies of L-canavanine incorporation into insectan lysozyme. J. Biol. Chem. 266, 15684-15687.

Rossignol, P.A., Lueders, A.M., 1986. Bacteriolytic factor in the salivary glands of *Aedes aegypti*. Comp. Biochem. Physiol. B Biochem. Mol. Biol. 83, 819-822.

Russell, V.W., Dunn, P.E., 1991. Lysozyme in the midgut of *Manduca sexta* during metamorphosis. Arch. Insect Biochem. Physiol. 17, 67-80.

Saitou, N., Nei, M., 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. 4, 406-425.

Schneider, P.M., 1985. Purification and properties of three lysozymes from hemolymph of the cricket, *Grillus bimaculatus* (De Geer). Insect Biochem. 15, 463-470.

Simser, J.A., Macaluso, K.R., Mulenga, A., Azad, A.F., 2004. Immune-responsive lysozymes from hemocytes of the American dog tick, *Dermacentor variabilis* and an embryonic cell line of the Rocky Mountain wood tick, *D. andersoni*. Insect Biochem. Mol. Biol. 34, 1235-1246.

Song, H., Inaka, K., Maenaka, K., Matsushima, M., 1994. Structural changes of active site cleft and different saccharide binding modes in human lysozyme co-crystallized with hexa-N-acetyl-chitohexaose at pH 4.0. J. Mol. Biol. 244, 522-540.

Song, H.Y., Cohen, F.S., Cramer, W.A., 1991. Membrane topography of ColE1 gene products: the hydrophobic anchor of the colicin E1 channel is a helical hairpin. J. Bacteriol. 173, 2927-2934.

Srisailam, S., Arunkumar, A.I., Wang, W., Yu, C., Chen, H.M., 2000. Conformational study of a custom antibacterial peptide cecropin B1: implications of the lytic activity. Biochim. Biophys. Acta 1479, 275-285.

Tanaka, S., Maeno, K., 2008. Maternal effects on progeny body size and color in the desert locust, *Schistocerca gregaria*: examination of a current view. J. Insect Physiol. 54, 612-618.

Tanaka, T., Kawano, S., Nakao, S., Umemiya-Shirafuji, R., Rahman, M.M., Boldbaatar, D., Battur, B., Liao, M., Fujisaki, K., 2010. The identification and characterization of lysozyme from the hard tick *Haemaphysalis longicornis*. Ticks Tick Borne Dis. 1, 178-185.

Thompson, J.D., Higgins, D.G., Gibson, T.J., 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 22, 4673-4680.

Tsakas, S., Marmaras, V.J., 2010. Insect immunity and its signalling: an overview. Invertebrate Surviv. J. 7, 228-238.

Ursic-Bedoya, R.J., Mitzey, A.M., Obraztsova, M., Lowenberger, C., 2005. Molecular cloning and transcriptional activation of lysozyme-encoding cDNAs in the mosquito *Aedes aegypti*. Insect Mol. Biol. 14, 89-94.

Ursic-Bedoya, R.J., Nazzari, H., Cooper, D., Triana, O., Wolff, M., Lowenberger, C., 2008. Identification and characterization of two novel lysozymes from *Rhodnius prolixus*, a vector of Chagas disease. J. Insect Physiol. 54, 593-603.

Vilcinskas, A., Mukherjee, K., Vogel, H., 2013. Expansion of the antimicrobial peptide repertoire in the invasive ladybird *Harmonia axyridis*. Proc. Biol. Sci. 280, 20122113.

Vogel, H., Altincicek, B., Glöckner, G., Vilcinskas, A., 2011. A comprehensive transcriptome and immune-gene repertoire of the lepidopteran model host *Galleria mellonella*. BMC Genomics, 12: 308.

von Heijne, G., 1990. The signal peptide. J. Membr. Biol. 115, 195-201.

Wang, W., Smith, D.K., Chen, H.M., 1999. The effect of pH on the structure, binding and model membrane lysis by cecropin B and analogs. Biochim. Biophys. Acta 1473, 418-430.

Wang, W.X., Wang, Y.P., Deng, X.J., Dang, X.L., Tian, J.H., Yi, H.Y., Li, Y.F., He, X.F., Cao, Y., Xia, Q.Y., Lai, R., Wen, S.Y., Paskowitz, S., 2009. Molecular and functional characterization of a c-type lysozyme from the Asian corn borer, *Ostrinia furnacalis*. J. Insect Sci. 9, 17.

Waterhouse, R.M., Kriventseva, E.V., Meister, S., Xi, Z., Alvarez, K.S., Bartholomay, L.C., Barillas-Mury, C., Bian, G., Blandin, S., Christensen, B.M., Dong, Y., Jiang, H., Kanost, M.R., Koutsos, A.C., Levashina, E.A., Li, J., Ligoxygakis, P., Maccallum, R.M., Mayhew, G.F., Mendes, A., Michel, K., Osta, M.A., Paskewitz, S., Shin, S.W., Vlachou, D., Wang, L., Wei, W., Zheng, L., Zou, Z., Severson, D.W., Raikhel, A.S., Kafatos, F.C., Dimopoulos, G., Zdobnov, E.M., Christophides, G.K., 2007. Evolutionary dynamics of immune-related genes and pathways in disease-vector mosquitoes. Science 316, 1738-1743.

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

Zachary, D., Hoffmann, D., 1984. Lysozyme is stored in the granules of certain hemocyte types in *Locusta*. J. Insect Physiol. 30, 405-411.

Zhang, Y., Huang, J., Zhou, B., Zhang, C., Liu, W., Miao, X., Huang, Y., 2009. Up-regulation of lysozyme gene expression during metamorphosis and immune challenge of the cotton bollworm, *Helicoverpa armigera*. Arch. Insect Biochem. Physiol. 70, 18-29.

Figures legends

Figure 1. cDNA sequence (upper) and predicted amino acid sequence (one letter code, below) of *Sg*Lys. Numerical designation of the nucleotide and deduced amino acid sequences are shown on the left and right, respectively. Assigned initial and terminal codons are in bold-italic. The termination codon is marked with an asterisk. The bold face amino acid sequence indicates the predicted signal peptide whereas the N-terminal amino acid sequence confirmed by the Edman degradation method is single underlined. The putative polyadenylation signal (AATAAA) is bold-double underlined. The cDNA sequence has been deposited in GenBank with the accession number JQ012999.

Figure 2. Alignment of sequences from two orthopterans lysozymes, *Schistocerca gregaria* and *Locusta migratoria*. Amino acids that are different between the two sequences are boxed. The 11 residues that compose the lysozyme catalytic cleft on the conserved LYZ1 domain are indicated with black squares. The two residues, Glu^{31} and Asp^{49}, of the putative lysozyme catalytic site on conserved LYZ1 domain are marked with arrow heads. Asterisks indicate Ca^{2+} binding site on conserved domain LYZ1.

Figure 3. (A) Ribbon representation of the fold of α-helical hairpin, i.e. helix-loop-helix (HLH), motif of hen egg white lysozyme (HEWL). The secondary structural elements are as follows: α-helix 1 (H1) Asp^{87} to Ser^{100} - loop (L) Asp^{101} to Asn^{106} - α-helix 2 (H2) Ala^{107} to Arg^{114} (from Ibrahim et al., 2001). The tightness of the loop between the two helices is constrained by Gly^{102} and Gly^{104}. (B) Pairwise alignment of amino acid sequences of the α-helical hairpin motifs of *SgLys*, *Bm*Lys with HEWL. Large bold letters represent the identical residues while the partially homologous ones are represented by large regular letters. To the right, percent identity of both *Sg*Lys and *Bm*Lys with HEWL. (C) Pairwise alignment of mature lysozymes of *Sg*Lys and *Bm*Lys without signal peptide using clustalW showing *Bm*Lys as a suitable template in CPH models. Black boxes are identical amino acids while grey boxes indicate conserved amino acids. (D) Three dimensional structures of *Sg*Lys and *Bm*Lys (PDB1GD6A) using CPH models-3.0.

Figure 4. Multiple phylogenic analysis. 36 amino acid sequences of mature c-type lysozymes from different insect orders (Coleoptera, Diptera, Hemiptera, Hymenoptera, Isoptera, Lepidoptera and Orthoptera) and ticks (Ixodida) were aligned by ClustalW and a phylogenetic tree was constructed by Neighbor-Joining criteria using MEGA6. Only bootstrap values higher than 50% are indicated for each root. *Gallus gallus* lysozyme was used as the out-group. *Schistocerca gregaria* lysozyme is written bold face letters. The sequences used in this analysis are listed in Supplementary Table S2. Lysozymes with digestive function are underlined while full circles indicate acidic lysozymes.

Figure 5. Kinetics of expression of *Sg*Lys in 5th instar of *Schistocerca gregaria* after bacterial challenge with *Escherichia coli*. The lysozyme (15.7 kDa) was detected by western blot analysis using a polyclonal antiserum directed against *Sg*Lys (Elmogy et al., 2015). Hemolymph proteins were withdrawn at different time intervals post-injection with *E. coli* as indicated. Each lane contains 30 µL of the hemolymph. C: non-immunized hemolymph (control).

Figure 6. Developmental expression of *SgLys*. RNA was extracted from fat body tissues of 1st, 2nd, 3rd, 4th, 5th instars, male and female insects 6 h post-injection with *E. coli* beside egg homogenate. Statistical analysis of 3 parallel RT-PCR results of *SgLys* expression was performed. Error bars represent the SD of 3 independent PCR amplifications and quantifications (n= 3). Columns, with the same letter(s) are homogenous (not significantly different at $p>0.05$), whereas those with different letters are significantly different at $p<0.001$. The relative expression level (was expressed as the ratio of *SgLys* RT-PCR product band intensity to that of the house keeping *β-actin)* calculated from the gel picture using Image J software.

Figure 7. Tissue expression of *SgLys*. Total RNA was extracted from the different tissues 6 h post-injection of bacteria and from native insects. (u) unchallenged larvae and (c) larvae injected with *E. coli*. He, hemocytes; Ep, epidermis; MG, midgut; FB, fat body; SG, salivary glands; TM, thoracic

muscles. To the left, DNA molecular size markers. The results shown are from 3 independent PCR amplifications.

Supplementary Figure S1. Sequence analysis of *SgLys* cloned in *Pichia pastoris* expression vector pPIC9K. **The primers used for the cloning of *SgLys* ORF are double-underlined**. **Bold/**Italic: α-factor primer, Italic**/single-**underlined: 3'AOX1 primer.

Supplementary Figure S2. Production of recombinant *Sg*Lys. (A) SDS-PAGE analysis of culture of *Pichia pastoris* GS115 transformed with pPIC9K-SgLys after 0 h, 24 h, 48 h, 72 h and 96 h of incubation (lanes 1 to 5) and Western blot of the purified r*Sg*Lys immune-stained with anti-*Sg*Lys polyclonal antiserum (lane 6). **Prior to SDS-PAGE, fractions (1-2 ml) were boiled for 5 min in denaturation buffer containing 65 mM Tris-HCl (pH 6.8), 2% (w/v) SDS, 10% (v/v) glycerol, and 5% (v/v) 2-mercaptoethanol.** (B) Zone inhibition assay showing lysozyme muramidase activity of r*Sg*Lys compared to that of hen egg white lysozyme (HEWL). (C) Cross reactivity of anti-*Sg*Lys polyclonal antiserum against the r*Sg*Lys as revealed by Ouchterlony double immune-diffusion assay. Center well (anti-*Sg*Lys), anti-*Sg*Lys polyclonal antiserum. Upper peripheral wells (r*Sg*Lys), purified r*Sg*Lys. Agarose plate was stained with Coomassie brilliant blue.

Supplementary Figure S3. Alignment of the deduced amino acid sequence of *S. gregaria* lysozyme with those of other insects and ticks (Supplementary Table S2). Sequences of mature proteins were aligned using ClustalW. Dashes indicate gaps inserted in the alignment. Numbers of amino acids are indicated to the right of the sequences. Asterisks indicate 20 conserved amino acid residues. The 8 conserved cysteine residues are marked by arrow heads and the two catalytic residues E and D with black squares. Signal peptides were predicted using SignalP 4.0 server and removed. The multiple sequence alignment was edited to define the sets of residues of high homology (shadowed).

Figure 1

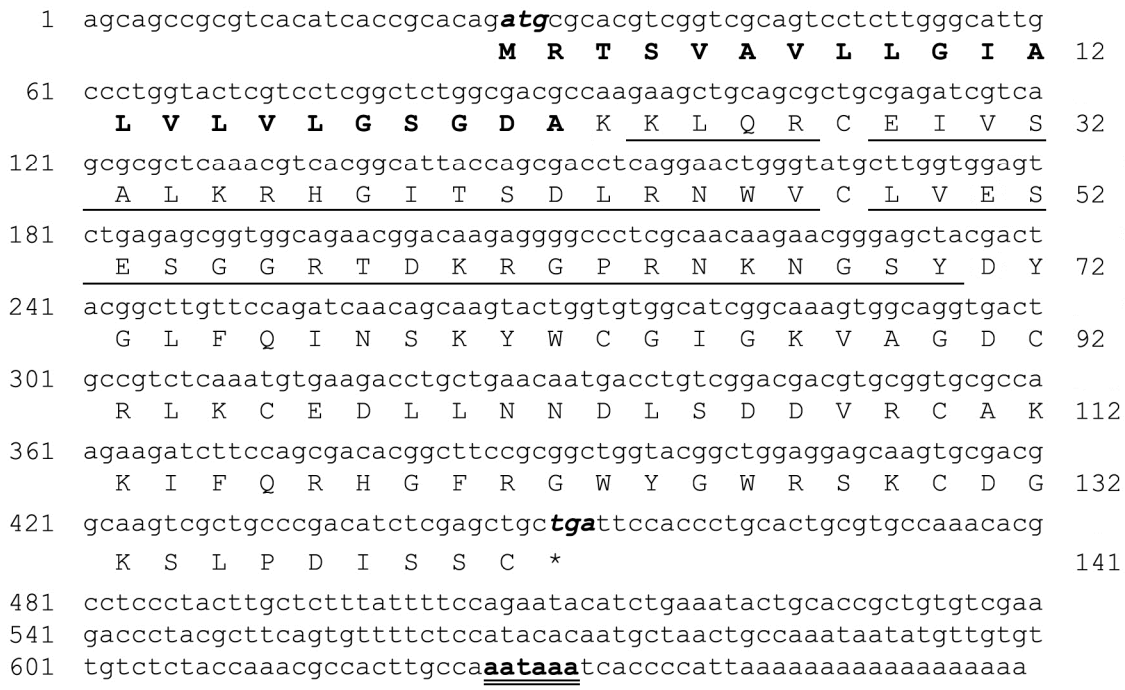


Figure 2

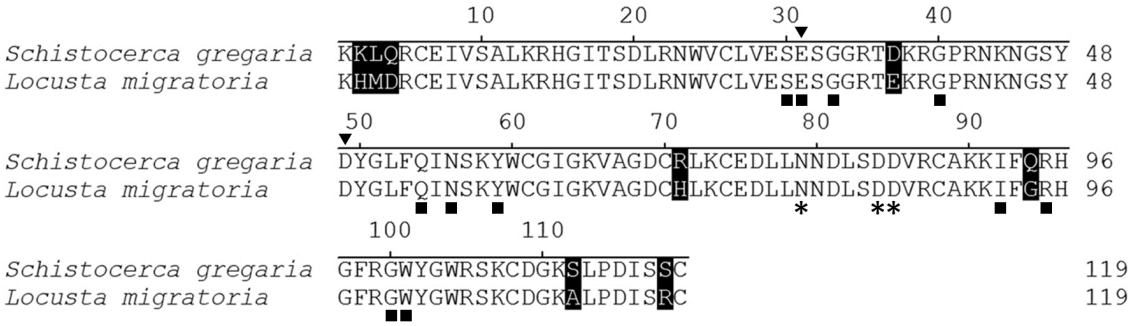
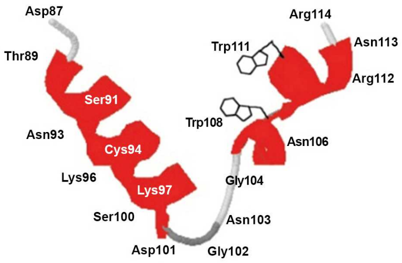


Figure 3

(A)



(B)

		% Identity		
		HLH (/27)	H1 (/14)	H2 (/8)
HEWL	DITASVNCAKKIVSDGNGMNAWVAWRNR (90, 100, 110)			
*Sg*Lys	**D**LSdD**VR**C**AKKI**FQr-H**G**FrG**W**yG**WR**SK (90, 100)	41	50	38
*Bm*Lys	**DIT**kAAK**CAKKI**ykr-HrFD**AW**yG**W**K**N**H (90, 100)	44	57	50

(C)

```
S. gregaria   KKLQRCEIVSALKRHGITSD-LRNWVCLVESESGGRTDKR    39
B. mori       KTFTRCGLVHELRKHGFEENLMRNWVCLVEHESSRDTSKT    40

S. gregaria   GPRNKNGSYDYGLFQINSKYWCGIGKVAG-DCRLKCEDLL    78
B. mori       N-TNRNGSKDYGLFQINDRYWCSKGASPGKDCNVKCSDLL    79

S. gregaria   NNDLSDDVRCAKKIFQRHGFRGWYGWRSKCDGKSLPDISSC  119
B. mori       TDDITKAAKCAKKIYKRHRFDAWYGWKNHCQG-SLPDISSC  119
```

(D)

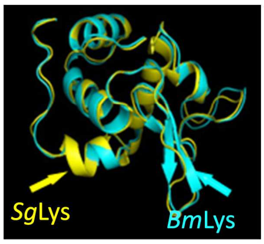


Figure 4

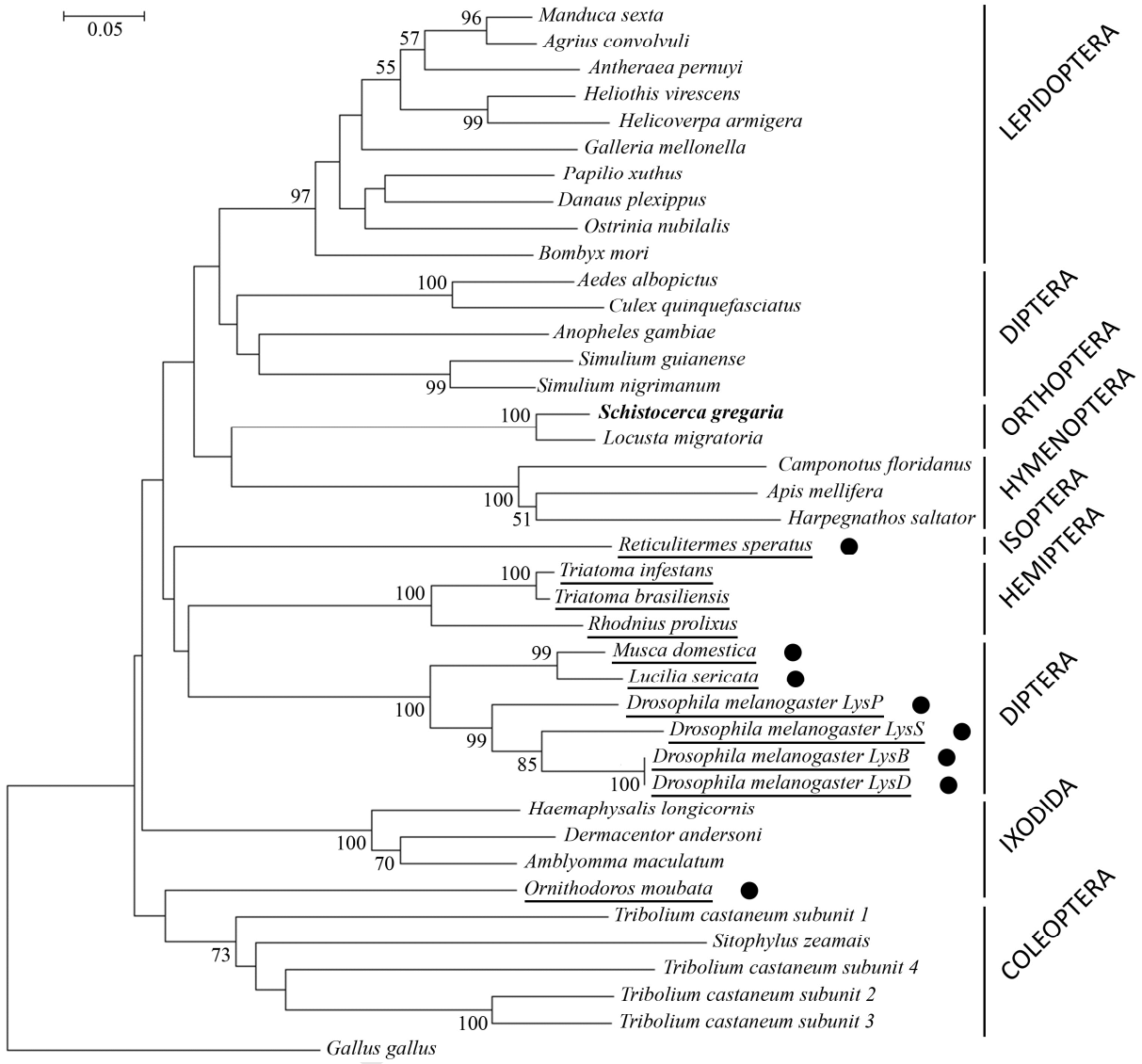


Figure 5

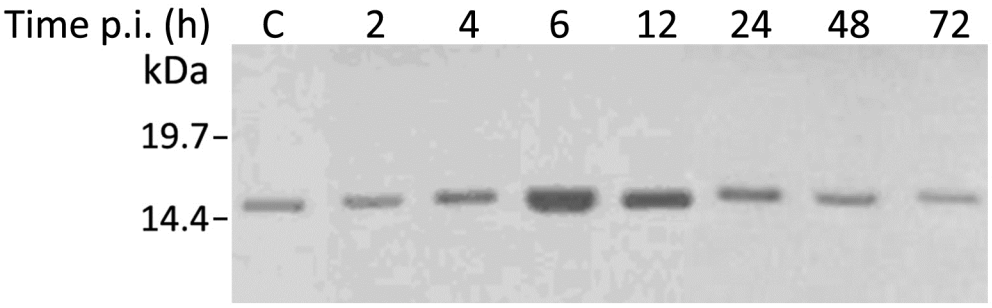


Figure 6

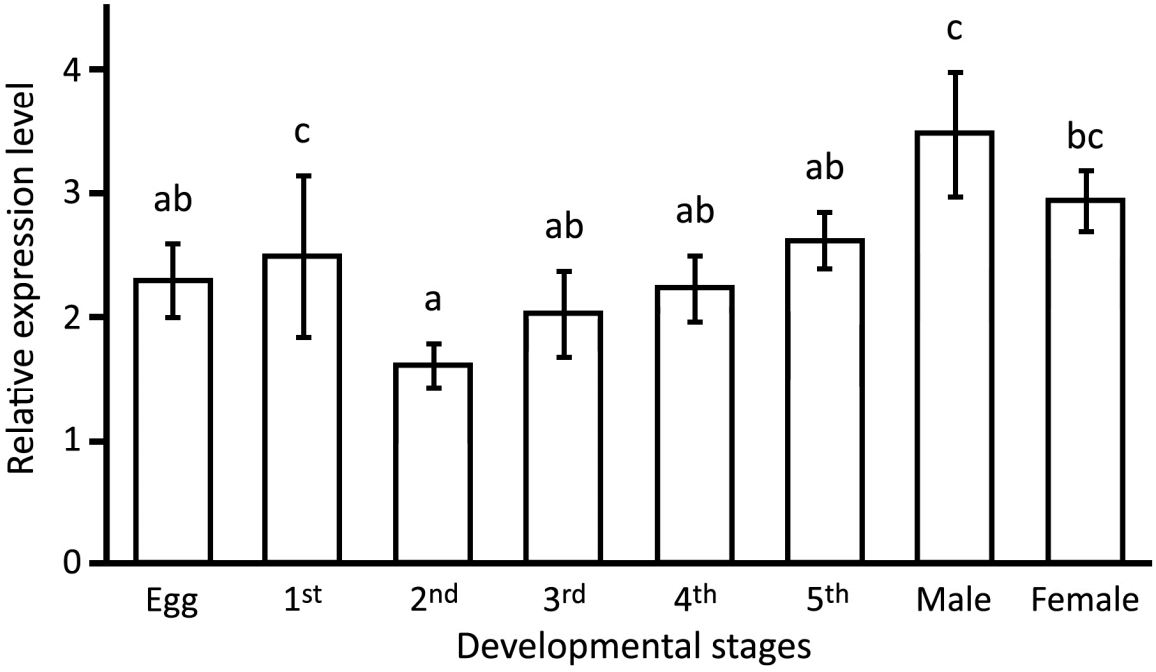
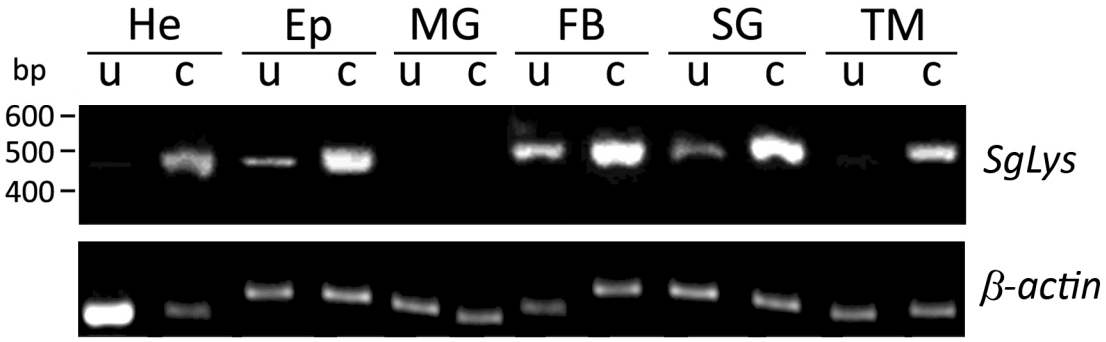


Figure 7



Highlights

- A lysozyme from fat body of the orthopteran *Schistocerca gregaria* was cloned.
- It contains all the features of a c-type lysozyme.
- *SgLys* is expressed at all developmental stages from egg to adults.
- *SgLys* is strongly upregulated in several tissues but midgut after *E. coli* injection.
- *Sg*Lys is likely a predominant acute-phase protein.