Available online at www.sciencedirect.com
ScienceDirect
Journal of Insect Physiology 54 (2008) 593–603
Journal of Insect Physiology
www.elsevier.com/locate/jinsphys

Identification and characterization of two novel lysozymes from *Rhodnius prolixus*, a vector of Chagas disease

Raul J. Ursic-Bedoya[a,*,1], Hamed Nazzari[a,1,2], Dawn Cooper[a], Omar Triana[b], Marta Wolff[b], Carl Lowenberger[a]

[a]*Department of Biological Sciences, Simon Fraser University, 8888 University Dr., Burnaby, BC, Canada V5A 1S6*
[b]*Instituto de Biología, Universidad de Antioquia, Calle 67 No 53-108, Medellín, Colombia*

Received 30 October 2007; received in revised form 12 December 2007; accepted 14 December 2007

Abstract

Lysozymes have been described in invertebrates as digestive or immune molecules. We report here the characterization of two novel c-type lysozymes, RpLys-A (EU250274) and RpLys-B (EU250275), isolated from the fat body and digestive tract of immune stimulated *Rhodnius prolixus,* a major vector of Chagas disease. Transcriptional profiles indicate that the temporal and spatial expression patterns of these two peptides are very different. RpLys-A is expressed predominantly in the midgut after ingestion of *Trypanosoma cruzi* in a bloodmeal, or after injection of bacteria into the hemocoel. RpLys-B is expressed primarily in the fat body after bacterial injection. Phylogenetic alignments indicate that RpLys-A aligns best with molecules from other hemipterans whose major expression is found in the intestinal tract whereas RpLys-B aligns best with mosquito and tick molecules whose expression is found principally in hemocytes and fat body and whose role has been described as immune-related. These data suggest a differential compartmentalized role of two closely related molecules; one for immunity in the hemocoel and the other for digestion in the midgut.

Keywords: *Rhodnius prolixus*; Lysozyme; Antimicrobial peptides; Chagas disease; Innate immunity; *Trypanosoma cruzi*

1. Introduction

Chagas disease, or American trypanosomiasis, is caused by the parasitic protozoan *Trypanosoma cruzi* that is transmitted to humans principally by hematophagous triatomine insects such as *Rhodnius prolixus*. Chagas disease remains prevalent in many areas of the Americas, ranging from southern Argentina to the southern United States, and afflicts over 17 million people in these locations (Dutra et al., 2005). A number of studies have suggested that inducible immune peptides can limit parasite development in vectors (Jaynes et al., 1988; Rodriguez et al., 1995; Lowenberger et al., 1996, 1999; Boisbouvier et al., 1998; Possani et al., 1998; Shahabuddin et al., 1998; Lowenberger, 2001; Vizioli et al., 2001). The majority of these studies examined associations in which the parasites make direct contact with hemolymph factors as they move from their site of development to the salivary glands for the subsequent transmission to vertebrates (Lowenberger et al., 1999). *T. cruzi*, however, never leaves the intestinal tract and is voided in the feces during blood ingestion, and therefore has no direct contact with hemolymph factors (Azambuja and Garcia, 1987; Lopez et al., 2003). This inefficient transmission of *T. cruzi* may be an evolutionary adaptation by the parasite to avoid contact with lethal components of the innate immune response of the vector (Lopez et al., 2003). This concept is supported by studies in which *T. cruzi* was killed and cleared after injection into the hemocoel of *R. prolixus* (Azambuja and Garcia, 1987) and by para-transgenesis studies which demonstrated the susceptibility

*Corresponding author. Tel.: +1 778 782 4391; fax: +1 778 782 3496.
E-mail address: rursicbe@sfu.ca (R.J. Ursic-Bedoya).
[1]The authors have contributed equally to this work and should be considered as co-first authors.
[2]Present address: Department of Cellular and Physiological Sciences, University of British Columbia, 2350 Health Sciences Mall, Vancouver, Canada V6T 1Z3.

0022-1910/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jinsphys.2007.12.009

of *T. cruzi*, *in vivo*, to cecropin, a small immune peptide (Durvasula et al., 1997).

Because previous studies had indicated the presence of a ~15 kDa protein in the hemolymph of *R. prolixus* following immune activation with antibacterial activity similar to other insect lysozymes (Lopez, Wolff, Triana, and Lowenberger unpublished), and because lysozyme-like activity had been reported in this species (Ribeiro and Pereira, 1984; Azambuja and Garcia, 1987) we sought to identify cDNAs encoding lysozymes from this vector species. Lysozymes hydrolyze the 1, 4-β-linkage between *N*-acetylmuramic acid and *N*-acetylglucosamine of the cell wall peptidoglycans of bacteria (Grunclova et al., 2003). As such, lysozymes may function in a digestive role for insects that ingest large numbers of bacteria (Regel et al., 1998), as immune related molecules to prevent colonization of the hemocoel by pathogens, and in some insects different isoforms of lysozymes may serve both functions (Ursic-Bedoya et al., 2005).

In *Drosophila melanogaster*, lysozymes are found extensively in the gastrointestinal tract and are involved in digestion (Kylsten et al., 1992). In many Lepidoptera and nematoceran Diptera, lysozymes are found in the hemolymph but not in the gut (Lemos and Terra, 1991). The recruitment of lysozymes as digestive enzymes, and their adaptation to an acidic midgut, may have occurred after the divergence of Cyclorrhapha from the Nematocera (Hultmark, 1996). Lysozymes may be expressed constitutively to regulate gut flora and help initiate the rapid immune response of the insect. Previous studies have shown that peptidoglycan fragments, produced by the enzymatic action of lysozymes on bacterial cell walls, are very potent inducers of the fat body response (Dunn et al., 1985). However, the kissing bugs that transmit *T. cruzi* (such as *R. prolixus*) are a much more ancient insect family and the role lysozymes may play in digestion, immunity or both functions is unknown. We report here the isolation and characterization of two chicken type (c-type) lysozymes isolated from *R. prolixus* and their temporal and spatial expression in response to bloodfeeding, immune activation with bacteria, and the ingestion of a bloodmeal containing the human parasite, *T. cruzi*.

2. Materials and methods

2.1. Insect maintenance, immune activation, and exposure to T. cruzi

A colony of *R. prolixus* has been maintained at the Institute of Biology, Universidad de Antioquia, Medellín Colombia for over 10 years and at Simon Fraser University, British Columbia, Canada for 5 years. The bacteria used for immune activation were grown and maintained as described previously (Lowenberger et al., 1996). Briefly, *Escherichia coli* and *Micrococcus luteus* were grown in Luria-Bertani's rich nutrient medium (LB medium) overnight at 37 °C while shaking at 350 rpm. Cultures were combined, pelleted by centrifugation, and insect minuten pins (0.10 mm diameter) were dipped into the pellet and inserted directly into the hemocoel of adult insects as described (Lopez et al., 2003). Sterilized pins were inserted into the hemocoel of control insects. For detecting peptide expression in response to ingested parasites, insects were fed on mice infected, or not, with the HA strain of *T. cruzi* at the Institute of Biology, Universidad de Antioquia, Colombia.

2.2. Tissue collection

We isolated the intestinal tracts or fat body tissue for RNA extraction from bacteria-inoculated, bloodfed, or *T. cruzi*-exposed *R. prolixus* adults at various times after treatment. Tissues from bacteria-inoculated insects were collected 8 and 24 h post inoculation. Tissues were collected from bloodfed or *T. cruzi*-exposed insects at 0, 2, 7 and 14 days post feeding which represents different developmental stages and location of the parasite in the insect.

2.3. RNA isolation and cDNA synthesis

Total RNA was extracted from the selected tissues of immune activated and naive insects at various time points after inoculation or bloodfeeding using TRI REAGENT (Molecular Research Center, Cincinnati, OH) following manufacturer's instructions. RNA was quantified using a Biophotometer (Eppendorf Germany) and 2.5 µg of total RNA was used for reverse transcription as described previously (Lowenberger et al., 1999) using an oligo dT primer (MG) with a 5′ extension (5′-CGGGCAGTGAG-CAACG(T_{12})-3′). Degenerate forward primers (5′-GAY-AAYGGNYTNTTYCARAT-3′ and 5′-GGNGGNCC-NAAYAARAAYGGN-3′) were designed against a partial protein sequence obtained previously (Lopez, Wolff, Triana, and Lowenberger unpublished) and conserved regions of other insect lysozymes. These primers were used with the MG primer in a PCR reaction with the conditions: 95 °C (3 min), and 30 cycles of 95 °C (10 s), 53 °C (10 s), 72 °C (2 min) on an Idaho Technologies Rapid Cycler (Salt Lake City, UT). The products of these reactions were size-fractioned on a 1.2% low melting point agarose gel. Bands of predicted size were excised from the gel, heated to 65 °C, and cloned directly into pGEM-T vector (Promega, Madison, WI) using the manufacturer's protocols. Transformations using XL1-Blue cells, and blue-white screening of presumed transformants were done following manufacturer's protocols. Selected colonies were grown overnight in 5 ml LB medium containing ampicillin (100 µg/ml) and purified using the Wizard Plus Minipreps DNA Purification system (Promega, Madison, WI). Sequencing of these clones was carried out on an ABI-310 automated sequencer using Big Dye chemistry. Sequences were compared with available sequences in the NCBI database. Two unique sequences that aligned well with insect lysozymes were obtained and specific primers for each cDNA sequence

were designed. The 5′ end of each sequence was obtained using specific reverse primers in a RACE reaction (Marathon cDNA synthesis kit; Clontech, Palo Alto CA). PCR was done under the conditions: 95 °C (3 min), and 30 cycles of 95 °C (1 min), 60 °C (30 s), 72 °C (1 min) and the resulting products separated, ligated, screened, and sequenced as described above. SeqMan (DNA STAR, Madison, WI) was used to align overlapping sequences of our two clones. Specific primers then were designed to amplify each of the two-cDNA *R. prolixus* sequences: A and B, with no cross-amplification.

2.4. Genomic DNA extraction

Genomic DNA was isolated from five starved adult insects. Insects were ground in a glass tissue grinder with 1.5 mL of fresh DNA extraction buffer (EB) (0.5% SDS, 0.2 M NaCl, 25 mM EDTA, 10 mM Tris pH 8) and 1.5 mL of phenol. After incubation at room temperature for 15 min, the homogenate was transferred to a 15 mL Corex tube and centrifuged at $8800 \times g$ for 20 min at 4 °C on an Allegra 64R (Beckman Coulter, USA) centrifuge. The aqueous phase was transferred to a new tube to which an equal volume of phenol:chloroform (1:1) was added. The mixture was homogenized by vortexing and centrifuged at $6650 \times g$ for 20 min at 4 °C. The supernatant was transferred to a new tube, mixed with an equal volume of chloroform, and centrifuged at $6650 \times g$ for 15 min at 4 °C. The resulting supernatant was transferred to a new tube and 1:10 volume of 4 M ammonium acetate plus 2.5 volumes of 95% ethanol were added, and the mixture was stored at −20 °C for 1 h. The DNA was pelleted by centrifugation at $6650 \times g$ for 30 min at 4 °C. The resulting pellet was dissolved in 100 uL of EB buffer (10 mM Tris–HCl pH 8.5), treated with RNAse A (50 mg at 37 °C for 30 min), and the DNA was further extracted with 100 uL of phenol:chloroform:iso-amylalcohol (25:25:1), washed with 95% ethanol, dried, resuspended in 100 uL of EB buffer and quantified using a Biophotometer (Eppendorf, Germany).

2.5. Target gene identification

Inverse PCR (iPCR) (Triglia, 2000) was used to amplify regions of genomic DNA upstream of the coding region of our genes to identify potential transcription factor binding sites. We used NEBcutter v2.0 (http://tools.neb.com/NEBcutter2/index.php) restriction digest analysis to identify restriction enzymes that would digest the cDNA Lys 1A gene within the first 500 bp of the initial methionine of the coding region. The enzymes used were: Dpn1, Rsa1 and EcoRV (New England Biolabs, USA). One microgram of genomic DNA was digested separately with 10 U of each restriction enzyme in an air incubator at 37 °C for 3–6 h. Restriction enzymes were heat inactivated according to manufacturer's instructions or the digested DNA was isolated by a phenol:chloroform extraction. Approximately 200 ng of each digested genomic DNA were self-ligated with 12 U of T4 Ligase (Promega, USA) at 16 °C for 16 h in 100 uL reactions. Two microliters of the ligation reaction were used in a PCR reaction using iProof DNA polymerase (Bio-Rad, USA). The inverse oriented primers used were: F: 5′-CCAACTACGACGGAAGCTATGATAATGGA-3′ and R: 5′CTAGTGAACACCCTAGCTTGTGTGGC-3′. Amplicons obtained were cloned into pGem- T-easy and transformed into *E. coli* JM109 and sequenced as described above.

In addition to the molecular approach we also used a bioinformatic approach to confirm our findings. We searched for contigs from the recently released trace data from the *R. prolixus* genome sequencing project using Mega Blast searches (http://www.ncbi.nlm.nih.gov/blast/mmtrace.shtml) using the first 200 nucleotides of the open reading frame from our cDNA sequences. Contigs containing the identified genomic clones and the remaining regions of the open reading frames were constructed using the SeqManII module of DNAstar software with loose assembling parameters to accommodate large gaps corresponding to introns. Putative transcription binding sites were identified using Alibaba 2.1 software (Grabe, 2002) using lazy restriction parameters. Alibaba predicts transcription factor binding sites by context dependent matrices generated from TRANSFAC 4.0 public sites.

2.6. Sequence identity analysis

Multiple sequence alignments of *R. prolixus* lysozymes and other invertebrate lysozymes were carried out using MegAlign (DNA Star, Madison, WI) using the Clustal W method with PAM 250 matrix. Prediction of the signal peptide was performed using SignalP v.3.0. Theoretical isoelectric points (pI) and molecular weights were determined using Expasy ProtParam program.

2.7. Transcriptional profile using quantitative RT-PCR

We have generated a battery of cDNAs from various tissues of *R. prolixus* adults after different immune stimulations. We used these cDNAs in a real-time quantitative-PCR (Q-PCR) analysis using a Rotor-Gene 3000 (Corbett Research, Sydney, Australia) to compare expression patterns of transcripts for both identified cDNAs. We constructed standard curves with known concentrations of purified lysozyme cDNAs. A 150 bp fragment of *R. prolixus* β-actin gene was used to normalize cDNA samples. Standard 25 μl PCR reactions containing 1.0 μl of SYBR-Green (Molecular Probes, Sigma, St. Louis, MO) and primers that distinguished between the two lysozyme sequences (RpLys-A-forward: 5′-ATGAAAGCTGTTTTCTTACTGGC-3′ and reverse: 5′-AAAGCAAACGTTGATATCTGGTA-3′ and RpLys-B-forward: 5′- ATGATTGCAAATCTAGTTTTAACACTATTGC-3′ and reverse: 5′- TTAACAAACCAATGGAGGCAAC-3′) were used in a PCR program of 95 °C (3 min) and 35 cycles of 95 °C (10 s), 58 °C for (10 s), and 72 °C (30 s).

Quantification, melt curve analysis and sample comparison were done with the Rotor-Gene version 5 software (Corbett Research). Three independently synthesized cDNAs for each time point were evaluated in these studies and each cDNA was analyzed at least 5 times to detect levels of the lysozyme and β-actin sequences.

3. Results

3.1. Sequence analysis

The *R. prolixus* lysozyme cDNAs contain deduced open reading frames of 417 and 414 nucleotides for RpLys-A and RpLys-B, respectively, that encode proteins of 139 and 138 residues with predicted sizes of 15.8 and 15.1 kDa, respectively. The coding region contains a stop codon and a 3′ untranslated region (UTR) of 96 and 100 nucleotides for RpLys-A and RpLys-B, respectively, a putative polyadenylation consensus signal (AATAAA) and 15 and 19 additional nucleotides before the poly-A tail for RpLys-A and -B, respectively (Fig. 1). Each deduced protein sequence contains a putative signal peptide comprising the first 18 residues that terminates with an alanine residue (Ala18) based on SignalP v3.0 Expasy tools (Bendtsen et al., 2004). The calculated theoretical pI values are 8.5 and 6.84 for the active

A

```
CGCCCGGGCAGGTCATTCGCAAA
 M   K   A   V   F   L   L   A   I   F   A   L   L   G   A   T   Q   A ⇩R   V
ATG AAA GCT GTT TTC TTA CTG GCT ATT TTC GCC CTN CTT GGC GCC ACA CAA GCT AGG GTG

 F   T   R   C   G   L   A   R   E   L   A   R   Q   G   L   P   R   H   D   L
TTC ACT AGA TGC GGT CTA GCG CGG GAA TTG GCT AGG CAA GGA CTT CCA CGC CAC GAT TTG

 A   N   W   V   C   L   I   E   A  E  S   G   R   N   T   R   A   R   G   G
GCT AAT TGG GTA TGC CTG ATT GAA GCA GAA AGT GGC AGA AAT ACC AGA GCC AGA GGT GGC

 P   N   Y   D   G   S   Y  D  N   G   L   F   Q   I   N   D   R   I   W   C
CCC AAC TAC GAC GGA AGC TAT GAT AAT GGA CTA TTC CAG ATC AAT GAT AGA ATT TGG TGT

 M   N   G   R   P   G   H   A   C   H   V   R   C   E   D   L   R   T   D   D
ATG AAC GGT AGA CCT GGA CAT GCT TGC CAC GTC AGA TGT GAA GAT TTA AGA ACA GAC GAC

 I   R   A   S   V   R   C   A   V   Q   I   K   Q   Q   Q   G   W   S   A   W
ATC AGA GCC AGT GTG AGG TGT GCT GTC CAA ATC AAA CAG CAA CAG GGC TGG TCA GCT TGG

 Y   G   W   Q   Y   H   C   R   G   R   P   L   P   D   I   N   V   C   F   *
TAC GGT TGG CAG TAC CAT TGC AGA GGT CGT CCC CTA CCA GAT ATC AAC GTT TGC TTT TAA

AGCAATCTAATTATTTTAAACTATTAACAAATTTTGTTAATTGTTTATTAGTATGCTTTGATTGTAATTGAGACTATTTG
taaacattacccatcgaataaaatattattgaatttcAn
```

B

```
CGCCCGGGCAGGTCTGATCATATACGGAC
 M   I   A   N   L   V   L   T   L   L   L   L   F   T   V   S   S   A ⇩K   V   F   T
ATG ATT GCA AAT CTA GTT TTA ACA CTA TTG CTG CTG TTT ACT GTC AGT TCA GCC AAA GTG TTC ACC

 D   C   E   L   A   N   V   L   E   N   A   G   F   P   K   D   Q   L   K   D   W   I
GAT TGT GAA CTG GCA AAT GTA TTG GAA AAT GCT GGA TTC CCA AAA GAT CAA CTA AAA GAC TGG ATT

 C   L   A   K   A  E  S   S   L   N   T   T   A   V   G   G   P   N   K   N   G   S
TGT CTA GCT AAA GCA GAA AGC TCA CTG AAC ACC ACG GCC GTC GGA GGA CCA AAT AAA AAT GGA AGC

 Y  D  Y   G   L   F   Q   I   N   D   H   I   W   C   D   P   E   K   R   G   G   D
TAT GAT TAT GGT TTA TTT CAG ATA AAC GAT CAT ATA TGG TGT GAT CCA GAA AAA AGA GGA GGT GAT

 C   N   V   K   C   S   D   L   V   L   E   D   D   I   G   P   S   M   N   C   A   K
TGT AAT GTG AAA TGT TCA GAT CTT GTT CTT GAA GAT GAC ATT GGA CCC AGT ATG AAT TGT GCA AAA

 I   V   Y   K   V   Q   G   F   K   A   W   N   G   L   D   K   K   C   K   G   K   K
ATA GTT TAT AAA GTT CAA GGA TTC AAG GCA TGG AAT GGT TGG ATC AAG AAA TGT AAG GGC AAA AAG

 L   P   P   L   V   C   *
TTG CCT CCA TTG GTT TGT TAA

AAAGAAGAAGAATAAGAAGAAGGAAAGCTAGTTCCTAAAATCATTGTTTATaaatggtcaaattgtaaaagaatttgttatttgcaata
attttagaataaattaaactgaatttattaata
```

Fig. 1. cDNA and translated amino acid sequences of *Rhodnius prolixus* lysozyme A (A) and B (B). The termination codons are marked with an asterisk. The bold amino acid sequence and arrowhead indicate the predicted signal sequence and the N-terminus of the mature protein, respectively. The putative polyadenylation signal is double underlined. The nucleotide sequences used as primer sites for real-time Q-PCR are underlined. The putative active site residues are in bold and italic (Glu50 and Asp68).

regions of RpLys-A and RpLys-B, respectively (Expasy ProtParam).

3.2. Identification of upstream promoter sites

iPCR successfully amplified only one amplicon that contained a segment of the RpLys-A gene (data not shown). This fragment contained only 60 bp of upstream sequence, but nonetheless contained a putative NF-κB site. Initially we could not find the sequence of RpLys-A in the trace data files but as the *R. prolixus* genome is completed and annotated we were able to identify a genomic clone containing 170 bp of upstream sequence. Data mining in this manner produced a strong match for Rp*Lys-B* and we obtained a 415 bp sequence upstream of the start ATG codon. Analysis with Alibaba 2.1 detected several regions identified as potential transcription factor binding sites. Of particular interest were GATA-1 and NF-κB sites (Table 1) present in the upstream region of both genes. We cannot eliminate the possibility of other binding factors further upstream.

3.3. Multiple alignments and phylogenetic analysis

Multiple protein sequence alignment with lysozymes from selected organisms indicates RpLys-A and RpLys-B share significant identity with other insect c-type lysozymes. All lysozymes documented here have the conserved Glu50 Asp68 in the active site and eight structural cysteine residues. There is a conserved active functional domain region (FQIND) found in the vast majority of insect c-type lysozymes (Fig. 2). It is apparent from the alignment that while there is significant conservation of regions and motifs within the active proteins, there is apparently no such conservation in the signal peptide region. Similarly there does not appear to be a conserved pattern within the signal peptides of proteins identified in closely related species or in a specific tissue (e.g. fat body). Therefore, subsequent analysis only included the sequences of the active proteins from which the signal peptides had been removed.

RpLys-A shares only 49% identity with RpLys-B. RpLys-A shared the greatest identity (79%, 61%, and 78%) with lysozymes from closely related organisms; *T. infestans* lysozyme -1 (Kollien et al., 2003), *T. infestans* lysozyme-2 (Balczun et al., 2008), and *T. brasiliensis* (Araujo et al., 2006), respectively. RpLys-B shares 53, 46, and 52% identity, respectively, with these same molecules. Therefore, the two lysozymes identified from *R. prolixus* are very different from each other. Comparison of the RpLys-A and -B sequences with other insect lysozymes using CLUSTAL W (v.3.2.2) indicated a shared identity with *Ae. aegypti-A* (41%), *Ae. aegypti-B* (41%), *Ae. aegypti-S* (34%), *Ae. albopictus* (38%), *An. gambiae* (40%), *D. andersoni* (41%), *D. variabilis* (41%) *D. melanogaster B* (45%), *D. melanogaster* D (45%), *D. melanogaster* P (42%), *H. cecropia* (48%), *H. virescens* (45%), *M. domestica* (44%), *S. cynthia* (51%), *T. ni* (48%) (Fig. 2).

A comparison of selected invertebrate lysozymes was performed at the amino acid level (Fig. 3) using only the active regions of the proteins. The cladogram (Fig. 3) shows a general separation of lysozymes based on function: lysozymes described as having a principal role in immune function are separate from those whose function has been described mainly as digestive. The sequence of the termite, *R. speratus*, is significantly different from all of the other sequences used here, and is appropriately on its own branch. Lysozymes from the Lepidoptera, mosquitoes, and ticks (*Dermacentor* sp.), whose molecules have been described as having more of an immune function, group together. The lysozymes from the flies, *Triatoma* sp. and a tick (*O. moubata*), whose function has been described as digestive, group in another major branch. Our molecule, RpLys-A, whose expression is greatest in the intestine groups with the branch of triatome lysozymes found in the clade that contains digestive lysozymes whereas RpLys-B, found in the fat body, is found in the clade of molecules whose function has been described as immune related, including molecules from distantly related Lepidoptera, Diptera, and ticks.

3.4. Induction of lysozyme genes

Real-time Q-PCR was used to compare expression patterns of RpLys-A and RpLys-B in different tissues, and at various time points, after inoculation with bacteria or ingestion of a bloodmeal containing *T. cruzi*. Constitutively expressed transcripts were found in fat body tissues

Table 1
Putative transcription factor binding sites for *Rhodnius prolixus* lysozymes

Gene	NF kappaB	GATA-1	Clone
Lys 1A	GGAACTTTCAA ATTAGGAAATAC (−64)	TGTTTCAGATC (−115) CTTATATTTCT (−42)	iPCR Dpn1 NADD–aee07e10.b1
Lys 1B	TAGGAAATGAC (−181)	TTTGAGCAGAA (−356) TTATTATTTTT (−302)	NAAX-ady62g11.g1

Putative transcription binding sites were identified using Alibaba 2.1 software (Grabe, 2002) using lazy restriction parameters. Alibaba predicts transcription factor binding sites by context dependent matrices generated from TRANSFAC 4.0 public sites. Alibaba 2.1 software is freely available at http://www.gene-regulation.com/pub/programs.html#alibaba2. Location of the putative binding site is indicated between parentheses relative to the methionine start codon. The clone indicators refer to the trace data files available at the *Rhodnius prolixus* trace data archives.
Ursic-Bedoya et al. (2005).

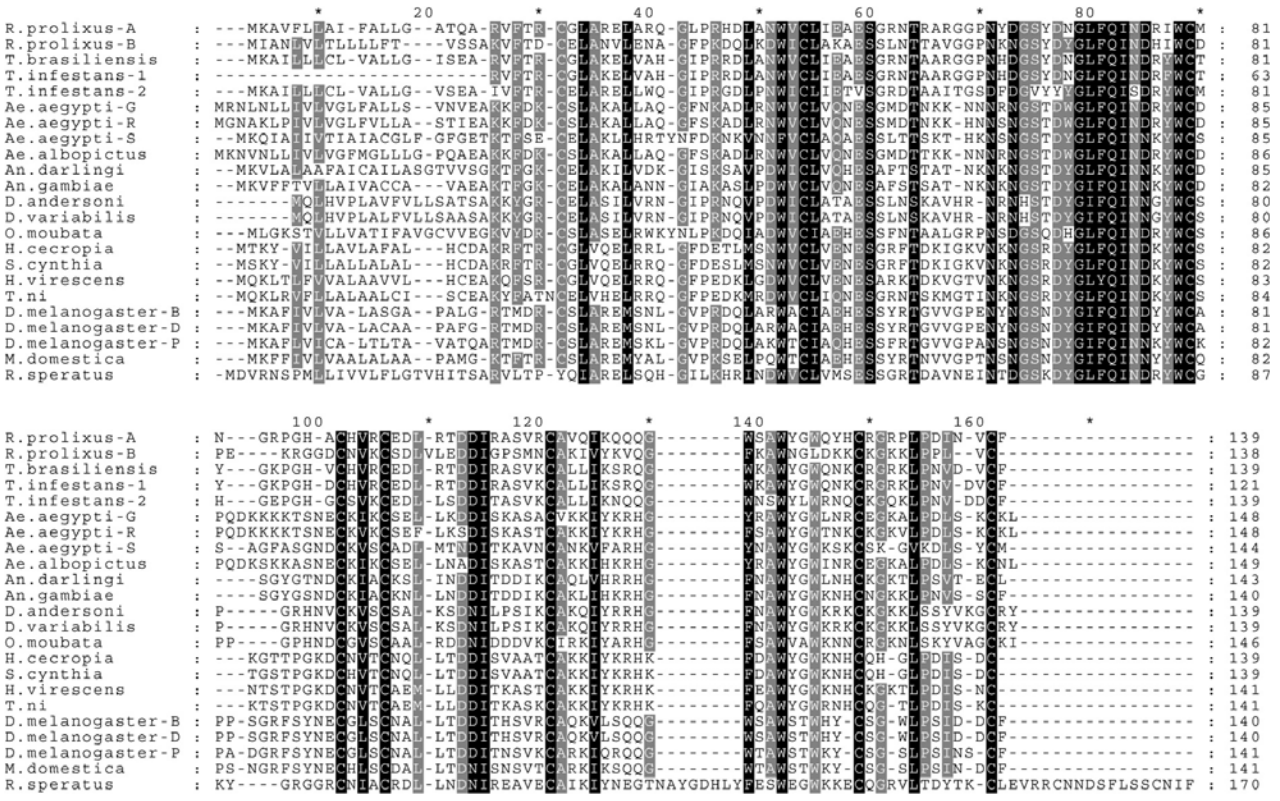

Fig. 2. Multiple sequence alignments of the amino acid sequences of *Rhodnius prolixus* lysozymes and other invertebrate lysozyme precursors using CLUSTAL W (V3.2.2) (DNA STAR, Madison, WI). The shading reflects the level of conservation throughout these molecules. The sequences used in this analysis were: *R. prolixus*-A (genbank:EU250274EU250274), *R. prolixus*-B (EU250275) *A. albopictus*: (AY089957), *Ae. aegypti-A*: (AJ290428), *Ae. aegypti*- B: (AY693973), *Ae. aegypti*-S: (AF466591), *An. darlingi:* (AF003945), *A. gambiae*: (Q17005), *D. andersoni*: (AY207371), *D. melanogaster*- B: (Z22225), *D. melanogaster*- D: (X58382), *D. melanogaster*- P: (X58383), *D. variabilis*: (AY183671), *H. cecropia*: (P05105), *H. virscens*: (U50551), *M. domestica*: (AY344588), *O. moubata:* (AF425264), *S. Cynthia*: (AB048258), *T. infestans* Lys 1: (AY253830), *T. infestans* Lys 2 (sequence obtained from Balczun et al., 2008), *T. brasiliensis* (AAU04569) and *T. ni*: (P50718).

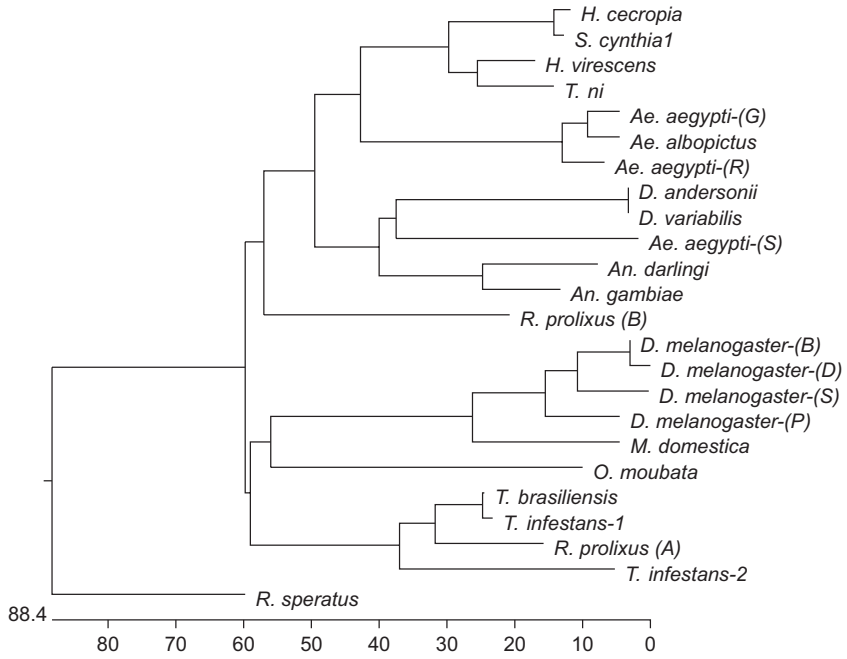

Fig. 3. A cladogram of selected insect lysozyme sequences representing molecules from diverse insects and for which a known function has been described. Alignments were constructed with the Clustal W method with the PAM250 residue weight table [DNA STAR, Madison, WI]) using the active regions of the lysozymes from which the signal peptide regions had been removed.

and the intestinal tract for both lysozyme sequences. After introduction of bacteria into the hemocoel we found a differential induction of transcription of both molecules. RpLys-B increased 12 and 18 fold in fat body tissues at 8 and 24 h post inoculation, respectively, but showed no significant difference in midgut expression at these time points (Fig. 4). In contrast, RpLys-A was up-regulated 24 fold transiently in the midgut of *R. prolixus* 8 h after the inoculation of bacteria into the hemocoel. Transcript levels decreased to baseline amounts in midgut tissues 24 h post inoculation (Fig. 4A). There was no induction of RpLys-A in the fat body 8 and 24 h post inoculation (Fig. 4B).

In bloodfed insects we measured minimal differences in expression of RpLys-A between 0 and 48 h post ingestion of a sterile bloodmeal or a meal containing the parasite *T. cruzi*. Subsequently, RpLys-A transcripts increased >20 fold in midgut and intestinal tissues extracted 7 and 14 days post ingestion of the parasite laden bloodmeal, but no differences were determined after the ingestion of a parasite free bloodmeal (Fig. 5). The presence or absence of *T. cruzi* in the bloodmeal did not produce significant changes in RpLys-B expression (Fig. 5).

4. Discussion

In a previous study (Lopez et al., 2003) we identified, using HPLC, an inducible protein approximately 15 kDa with lysozyme-like activity in the hemolymph of immune activated *R. prolixus*. Lysozyme-like activity had been reported previously in the hemolymph of this vector, (Azambuja and Garcia, 1987) and in intestinal homogenates obtained from adult insects, with 2 peaks 3 days and 3 weeks after feeding (Ribeiro and Pereira, 1984). A similar activity was observed in response to *M. lysodeikticus* (Azambuja and Garcia, 1987) or *T. cruzi* (Mello et al., 1995) injection into the hemolymph but no characterization or sequencing of these proteins was done (Azambuja and Garcia, 1987). Our data corroborate the conclusions of these authors that the effects they observed were likely due to lysozyme expression.

We report here the expression pattern of two novel lysozymes in this vector of Chagas disease. Lysozymes are ubiquitous peptides described from many groups of invertebrates and vertebrates and which have been characterized as immune related molecules (Roxstrom-Lindquist et al., 2004), digestive enzymes (Grunclova et al., 2003), or multifunctional molecules (Li et al., 2005; Ursic-Bedoya et al., 2005). At the protein level, RpLys-A and RpLys-B share the greatest identity with lysozymes isolated from the closely related insects, *T. infestans* (Kollien et al., 2003) and *T. brasiliensis* (Araujo et al., 2006) (Fig. 3). A sequence analysis of the active regions of these lysozymes indicates a general grouping determined more by function (immune or digestion) or possibly location (hemocytes/fat body or digestive tract). RpLys-B, found in the fat body, aligns closest to a group containing immune related lysozymes isolated from the distantly related Lepidoptera or the ticks, *D. andersoni* and *D. variabilis* (Simser et al., 2004), rather than organisms that are more closely related taxonomically. RpLys-A, found mainly in the digestive

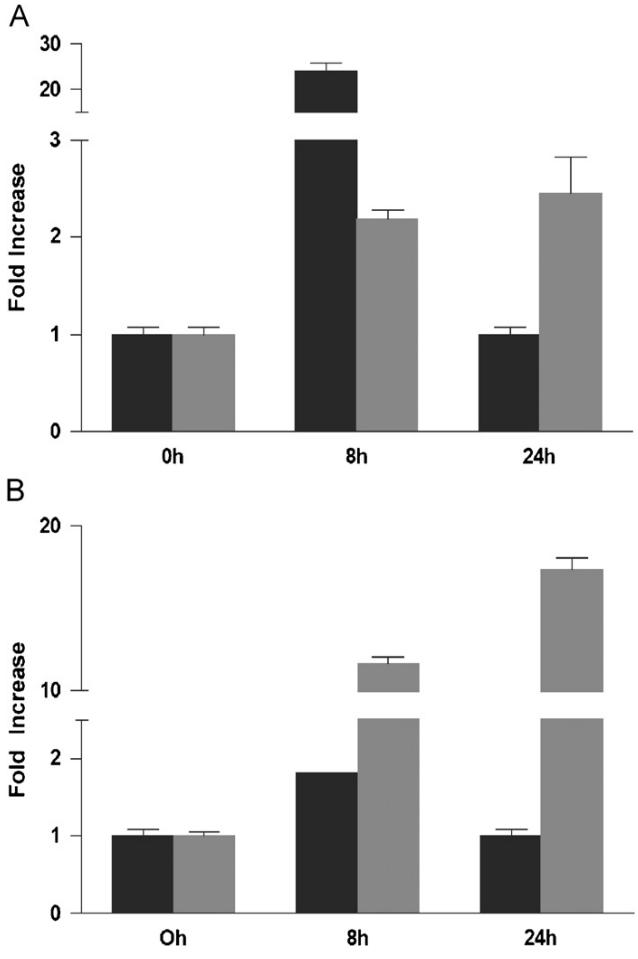

Fig. 4. Real-time quantitative PCR (Q-PCR) profile of *Rhodnius prolixus* lysozymes-A and -B in midgut (4A) and fat body (4B) tissues after bacterial inoculation. Standard curves were established for RpLys-A and RpLys-B and a 150 bp fragment of R. prolixus β-actin. Products from each sample were normalized to actin levels in each cDNA. The levels in control insects that were injected with a sterile needle were arbitrarily designated as 1, and all other levels were expressed as a fold increase over controls. Each bar represents the mean (+SD) fold increase of five replicates from three independently derived cDNAs. Black bars—RpLys-A and gray bars—RpLys-B.

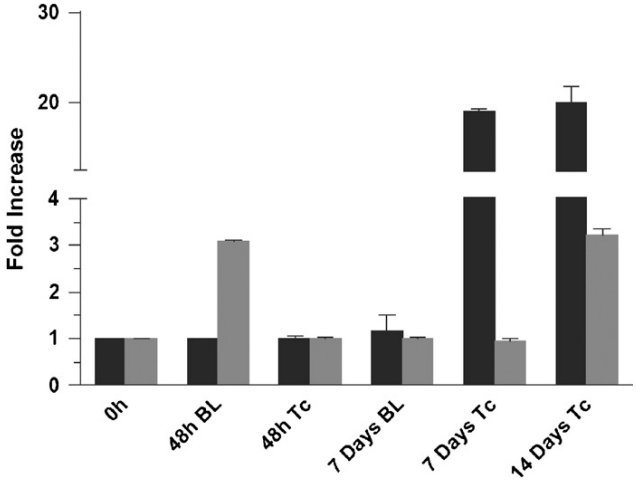

Fig. 5. Real-time quantitative PCR profile of *Rhodnius prolixus* lysozymes-A and -B in midgut and intestinal tissues after the ingestion of a sterile bloodmeal or a bloodmeal containing the parasite *Trypanosoma cruzi*. Standard curves were established for RpLys-A and RpLys-B and a 150 bp fragment of *R. prolixus* β actin. Products from each sample were normalized to actin levels in each cDNA. The levels in naive insects were arbitrarily designated as 1, and all other levels were expressed as a fold increase over controls. Each bar represents the mean (+SD) fold increase of five replicates from three independently derived cDNAs. Black bars—RpLys-A, gray bars—RpLys-B, BL—sterile bloodmeal, Tc—bloodmeal containing *T. cruzi*.

tract of *R. prolixus,* aligns best with the lysozyme found in the digestive tract of the other hemipterans, *T. infestans* and *T. brasiliensis*, and falls within the general grouping of digestive lysozymes. By including sequences from termites and the distantly related ticks in the analysis it appears that the conservation of specific aspects of these very similar molecules is based on their functional role or tissue origin (fat body/hemocytes) rather than solely by taxonomic relatedness. In addition the calculated theoretical pI values of 8.5 and 6.84 for RpLys-A and RpLys-B, respectively, reflect the environment in which these molecules operate. Digestion of proteins and lipids in *R. prolixus* occurs mainly in the posterior midgut (intestine) where cathepsins break down proteins under acidic conditions (Terra, 1990). The anterior midgut, where the bloodmeal is stored and bacterial symbionts reside, has a neutral-basic pH. Lysozymes found in the anterior midgut function in carbohydrate digestion and possibly regulate the proliferation of Gram-negative bacteria in this region. Therefore, RpLys-A may function optimally in the anterior midgut and RpLys-B would operate optimally in the buffered neutral hemolymph. Despite these differences, all invertebrate lysozymes compared in this study share a common theme; eight cysteine residues that form four disulfide bridges, and, with the exception of *T. infestans*-2 (Balczun et al., 2008), all share the conserved catalytic sites of glutamic and aspartic acid residues. This conservation of common structural components in all insect and tick lysozymes suggests a major role for these molecules in invertebrates.

In *D. melanogaster*, it has been widely documented that promoter sequences in antimicrobial peptide (AMP) genes contain combinations of transcription factor binding sites responsible for their tissue and signal-dependent specificities (Uvell and Engstrom, 2007). Two different classes of transcription factors are implicated in the transcriptional activation of AMPs in the *Drosophila* fat body; the NF-κB factors Dorsal, Dif, and Relish (RHD containing proteins), and the GATA factor Serpent (Senger et al., 2004, 2006). Generally, GATA sites are located within 20 bp of the NF-κB sites in functionally important promoter regions (Kadalayil et al., 1997). Furthermore, regulatory regions for Diptericin and Metchnikowin require GATA sites for their activation in the midgut and a second GATA factor, dGATAe, mediates a Toll-independent immune response

in the midgut (Senger et al., 2006). Although three other lysozyme genes have been characterized molecularly from kissing bugs (Kollien et al., 2003; Araujo et al., 2006; Balczun et al., 2008), no information concerning their regulatory sequences has been published. Our data is in concurrence with information available from higher order dipteran insects. Both of the genes we describe in this paper seem to contain both GATA and NF-κB sites in the promoter's proximal region. No GATA transcription factor homolog has been identified in triatomes to date and our early attempts using homology searches to Drosophila's Serpent did not yield any results. We have, however, identified an EST corresponding to a Dorsal homolog (Ursic-Bedoya and Lowenberger, 2007) and current work is undergoing to characterize this gene. Further work also will focus on mapping regulatory sequences from other immunity genes in *R. prolixus*, which we have identified through a series of SSH libraries (Ursic-Bedoya and Lowenberger, 2007), in order to conduct a comparative analysis among different insect orders.

We detected very low baseline transcripts for both sequences in naive insects. In response to bacteria inoculation in the hemocoel we detected significant increases in fat body transcription for RpLys-B 8 and 24 h after stimulation. This inoculation also produced a spike of expression of RpLys-A in the midgut/intestine, which indicates that these tissues are immune responsive, and that they may be activated by stimuli received elsewhere in the body. These data suggest a systemic and coordinated immune response, possibly mediated by cytokines, in which the stimulation of one region results in an increased expression of peptides in another region as has been reported previously (Lowenberger et al., 1996; Dimopoulos et al., 1997; Ursic-Bedoya et al., 2005).

In bloodfed insects, with or without parasites in the bloodmeal, we found no significant changes in the transcriptional profile for RpLys-B, neither in the midgut nor the fat body, suggesting that the interactions between parasite and digestive tract do not result in a systemic response in the fat body. In addition, because this parasite normally does not cross the digestive tract, there is no direct contact between parasite and hemolymph factors such as hemocytes and fat body cells. However, we do see a significant increase in RpLys-A expression in the digestive tract in response to the presence of the *T. cruzi,* 7 and 14 days post ingestion. This up-regulation is not reflected in insects that ingested a parasite-free bloodmeal, suggesting a specific interaction between *T. cruzi* and midgut and intestinal tissues over the first 14 days of parasite development. Although we have demonstrated an inducible response in *R. prolixus* to *T. cruzi* in the digestive tract, these compounds expressed at normal physiological levels are not lethal to the parasite. *R. prolixus* relies on obligate mutualistic bacterial symbionts in the intestine to provide essential nutrients (Azambuja and Garcia, 1987; Garcia et al., 2007). While hemocoel AMPs kill pathogens, AMPs in the intestine may regulate the proliferation of microbes in the digestive tract but must not eliminate these essential symbionts by expressing high concentrations of AMPs to which the symbionts are susceptible. *T. cruzi* is susceptible to insect AMPs, as demonstrated *in vitro* and *in vivo* (Beard et al., 2001; Durvasula et al., 1997; Lopez et al., 2003) suggested that its exclusive relegation to the intestine may have evolved to permit development and multiplication in an area of low AMP expression.

Molecules such as lysozymes may have evolved to play dual roles in digestion and defense. Recently, a number of studies have examined the expression of lysozymes in *T. brasiliensis* (Araujo et al., 2006) and *T. infestans* (Kollien et al., 2003; Balczun et al., 2008) which showed significant up-regulation in the digestive tract. The molecules from *T. infestans* also showed an up-regulation in response to molting and feeding. These data are supported by previous studies that demonstrated lysozyme-like activity in the intestine (Ribeiro and Pereira, 1984) or after pathogens were injected into the hemocoel (Azambuja and Garcia, 1987; Mello et al., 1995).

Characterization of the seven c-type lysozymes found in *Anopheles gambiae* suggests that lysozymes may have been adopted more recently to play a role in immunity (Li et al., 2005). We have identified two lysozymes that act in a compartmentalized manner to protect the hemocoel or to aid in digestion and maintain intestinal flora at acceptable levels. Our data suggest that lysozymes evolved for important roles in both digestion and immunity in the more 'ancient' hemimetabolous triatominae but also play a role in the 'higher' Diptera and Lepidoptera. To speculate on the evolution of lysozymes and to develop a prediction of which physiological function, digestion or immunity, preceded the other would require a detailed examination of these molecules throughout the invertebrates.

The parasite *T. cruzi* lives in the milieu of the digestive tract and appears not to be affected by normal physiological levels of lysozymes, or other molecules, expressed in these tissues. Further studies will reveal if hemocoel components, such as RpLys-B, can reduce *T. cruzi* viability. Whether the molecules described here originated for a digestive or defense role cannot be determined. Our results lend support for the dual role of lysozymes in invertebrates. The identification of novel immune peptides in vectors such as *R. prolixus* will increase our knowledge of general insect immunity, and the evolutionary origins of these and other immune peptides, and may provide insight into vector-parasite interactions that affect and regulate parasite transmission.

Acknowledgments

We thank L. Lopez and G. Morales for initial studies and M. Griffiths and M. Zakhary for technical expertise. This research was funded, in part, by an NSERC undergraduate student research award to HN, a Michael Smith Foundation senior graduate fellowship to DMC,

the Canada Research Chairs program, CIHR (69558), NSERC (RPG261940), and a Michael Smith Scholar award to CL.

References

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 Biochemistry and Molecular Biology 36, 547–560.

Azambuja, P., Garcia, E.S., 1987. Characterization of inducible lysozyme activity in the hemolymph of *Rhodnius prolixus*. Brazilian Journal of Medical and Biological Research 20, 539–548.

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). Parasitological Research 102, 229–232.

Beard, C.B., Dotson, E.M., Pennington, P.M., Eichler, S., Cordon-Rosales, C., Durvasula, R.V., 2001. Bacterial symbiosis and paratransgenic control of vector-borne Chagas disease. International Journal of Parasitology 31, 621–627.

Bendtsen, J.D., Nielsen, H., von Heijne, G., Brunak, S., 2004. Improved prediction of signal peptides: SignalP 3.0. Journal of Molecular Biology 340, 783–795.

Boisbouvier, J., Prochnicka-Chalufour, A., Nieto, A.R., Torres, J.A., Nanard, N., Rodriguez, M.H., Possani, L.D., Delepierre, M., 1998. Structural information on a cecropin-like synthetic peptide, Shiva-3 toxic to the sporogonic development of *Plasmodium berghei*. European Journal of Biochemistry 257, 263–273.

Dimopoulos, G., Richman, A., Muller, H.M., Kafatos, F.C., 1997. Molecular immune responses of the mosquito *Anopheles gambiae* to bacteria and malaria parasites. Proceedings of the National Academy of Science USA 94, 11508–11513.

Dunn, P.E., Dai, W., Kanost, M.R., Geng, C.X., 1985. Soluble peptidoglycan fragments stimulate antibacterial protein synthesis by fat body from larvae of *Manduca sexta*. Developmental and Comparative Immunology 9, 559–568.

Durvasula, R.V., Gumbs, A., Panackal, A., Kruglov, O., Aksoy, S., Merrifield, R.B., Richards, F.F., Beard, C.B., 1997. Prevention of insect-borne disease: an approach using transgenic symbiotic bacteria. Proceedings of the National Academy of Science USA 94, 3274–3278.

Dutra, W.O., Rocha, M.O., Teixeira, M.M., 2005. The clinical immunology of human Chagas disease. Trends in Parasitology 21, 581–587.

Garcia, E.S., Ratcliffe, N.A., Whitten, M.M., Gonzalez, M.S., Azambuja, P., 2007. Exploring the role of insect host factors in the dynamics of *Trypanosoma cruzi–Rhodnius prolixus* interactions. Journal of Insect Physiology 53, 11–21.

Grabe, N., 2002. AliBaba2: context specific identification of transcription factor binding sites. In Silico Biology 2, S1–S15.

Grunclova, L., Fouquier, H., Hypsa, V., Kopacek, P., 2003. Lysozyme from the gut of the soft tick *Ornithodoros moubata*: the sequence, phylogeny and post-feeding regulation. Developmental and Comparative Immunology 27, 651–660.

Hultmark, D., 1996. Insect lysozymes. In: Jollès, P., Birkhäuser, B. (Eds.), Lysozymes—Model Enzymes in Biochemistry and Biology. Verlag, pp. 87–102.

Jaynes, J.M., Burton, C.A., Barr, S.B., Jeffers, G.W., Julian, G.R., White, K.L., Enright, F.M., Klei, T.R., Laine, R.A., 1988. In vitro cytocidal effect of novel lytic peptides on *Plasmodium falciparum* and *Trypanosoma cruzi*. FASEB Journal 2, 2878–2883.

Kadalayil, L., Petersen, U.M., Engstrom, Y., 1997. Adjacent gata and kappa B-like motifs regulate the expression of a *Drosophila* immune gene. Nucleic Acids Research 25, 1233–1239.

Kollien, A.H., Fechner, S., Waniek, P.J., Schaub, G.A., 2003. Isolation and characterization of a cDNA encoding for a lysozyme from the gut of the reduviid bug *Triatoma infestans*. Archives of Insect Biochemistry and Physiology 53, 134–145.

Kylsten, P., Kimbrell, D.A., Daffre, S., Samakovlis, C., Hultmark, D., 1992. The lysozyme locus in *Drosophila melanogaster*: different genes are expressed in midgut and salivary glands. Molecular and General Genetics 232, 335–343.

Lemos, F.J., Terra, W.R., 1991. Digestion of bacteria and the role of midgut lysozyme in some insect larvae. Comparative Biochemistry and Physiology B 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.

Lopez, L., Morales, G., Ursic, R., Wolff, M., Lowenberger, C., 2003. Isolation and characterization of a novel insect defensin from *Rhodnius prolixus*, a vector of Chagas disease. Insect Biochemistry and Molecular Biology 33, 439–447.

Lowenberger, C.A., 2001. Form, function, and phylogenetic relationships of mosquito Immune peptides. In: Beck, G., Sugumaran, M., Cooper, R. (Eds.), Phylogenetic Perspectives on the Vertebrate Immune System. Kluwer Academic, Dordrecht, pp. 113–129.

Lowenberger, C.A., Ferdig, M.T., Bulet, P., Khalili, S., Hoffmann, J.A., Christensen, B.M., 1996. *Aedes aegypti*: induced antibacterial proteins reduce the establishment and development of *Brugia malayi*. Experimental Parasitology 83, 191–201.

Lowenberger, C.A., Kamal, S., Chiles, J., Paskewitz, S., Bulet, P., Hoffmann, J.A., Christensen, B.M., 1999. Mosquito–*Plasmodium* interactions in response to immune activation of the vector. Experimental Parasitology 91, 59–69.

Mello, C.B., Garcia, E.S., Ratcliffe, N.A., Azambuja, P., 1995. *Trypanosoma cruzi* and *Trypanosoma rangeli*: interplay with hemolymph components of *Rhodnius prolixus*. Journal of Invertebrate Pathology 65, 261–268.

Possani, L.D., Zurita, M., Delepierre, M., Hernandez, F.H., Rodriguez, M.H., 1998. From noxiustoxin to Shiva-3, a peptide toxic to the sporogonic development of *Plasmodium berghei*. Toxicon 36, 1683–1692.

Regel, R., Matioli, S.R., Terra, W.R., 1998. Molecular adaptation of *Drosophila melanogaster* lysozymes to a digestive function. Insect Biochemistry and Molecular Biology 28, 309–319.

Ribeiro, J.M., Pereira, M.E., 1984. Midgut glycosidases of *Rhodnius prolixus*. Insect Biochemistry 14, 103–108.

Rodriguez, M.C., Zamudio, F., Torres, J.A., Gonzalez-Ceron, L., Possani, L.D., Rodriguez, M.H., 1995. Effect of a cecropin-like synthetic peptide (Shiva-3) on the sporogonic development of *Plasmodium berghei*. Experimental Parasitology 80, 596–604.

Roxstrom-Lindquist, K., Terenius, O., Faye, I., 2004. Parasite-specific immune response in adult *Drosophila melanogaster*: a genomic study. EMBO Reports 5, 207–212.

Senger, K., Armstrong, G.W., Rowell, W.J., Kwan, J.M., Markstein, M., Levine, M., 2004. Immunity regulatory DNAs share common organizational features in *Drosophila*. Molecular Cell 13, 19–32.

Senger, K., Harris, K., Levine, M., 2006. GATA factors participate in tissue-specific immune responses in *Drosophila* larvae. Proceedings of the National Academy of Science USA 103, 15957–15962.

Shahabuddin, M., Fields, I., Bulet, P., Hoffmann, J.A., Miller, L.H., 1998. *Plasmodium gallinaceum:* differential killing of some mosquito stages of the parasite by insect defensin. Experimental Parasitology 89, 103–112.

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, *Dandersoni*. Insect Biochemistry and Molecular Biology 34, 1235–1246.

Terra, W.R., 1990. Evolution of digestive systems of insects. Annual Review of Entomology 35, 181–200.

Triglia, T., 2000. Inverse PCR (IPCR) for obtaining promoter sequence. Methods in Molecular Biology 130, 79–83.

Ursic-Bedoya, R.J., Lowenberger, C.A., 2007. *Rhodnius prolixus*: identification of immune-related genes up-regulated in response to pathogens and parasites using suppressive subtractive hybridization. Developmental and Comparative Immunology 31, 109–120.

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 Molecular Biology 14, 89–94.

Uvell, H., Engstrom, Y., 2007. A multilayered defense against infection: combinatorial control of insect immune genes. Trends in Genetics 23, 342–349.

Vizioli, J., Richman, A.M., Uttenweiler-Joseph, S., Blass, C., Bulet, P., 2001. The defensin peptide of the malaria vector mosquito *Anopheles gambiae*: antimicrobial activities and expression in adult mosquitoes. Insect Biochemistry and Molecular Biology 31, 241–248.