Published in final edited form as:

Insect Biochem Mol Biol. 2007 July ; 37(7): 702–712. doi:10.1016/j.ibmb.2007.04.004.

The sialotranscriptome of the blood-sucking bug *Triatoma brasiliensis* (Hemiptera, Triatominae)

Adriana Santos^a, José Marcos C. Ribeiro^b, Michael J. Lehane^c, Nelder Figueiredo Gontijo^a, Artur Botelho Veloso^a, Mauricio R.V. Sant'Anna^c, Ricardo Nascimento Araujo^a, Edmundo C. Grisard^d, and Marcos Horácio Pereira^{a,*}

^aLaboratório de Fisiologia de Insetos Hematófagos, Departamento de Parasitologia/ICB, Universidade Federal de Minas Gerais – Caixa Postal 486, 31270-901, Belo Horizonte, MG, Brazil

^bSection of Vector Biology, Laboratory of Malaria and Vector Research, National Institute of Allergy and Infectious Diseases, National Institutes of Health, 12735 Twinbrook Parkway, Room 2E32D, Rockville, MD 20852, USA

^cLiverpool School of Tropical Medicine, Pembroke Place, Liverpool, L3 5QA, United Kingdom

^dLaboratório de Protozoologia, MIP, CCB, Universidade Federal de Santa Catarina, Caixa postal 476, 88040-900, Florianópolis, SC, Brazil

Abstract

Triatoma brasiliensis is the most important autochthon vector of *Trypanosoma cruzi* in Brazil, where it is widely distributed in the semiarid areas of the Northeast. In order to advance the knowledge of the salivary biomolecules of Triatominae, a salivary gland cDNA library of *T. brasiliensis* was mass sequenced and analyzed. Polypeptides were sequenced by HPLC/Edman degradation experiments. 1,712 cDNA sequences were obtained and grouped in 786 clusters. The housekeeping category had 24.4% and 17.8% of the clusters and sequences, respectively. The putatively secreted category contained 47.1% of the clusters and 68.2% of the sequences. Finally, 28.5% of the clusters, containing 14% of all sequences, were classified as unknown. The sialoma of *T. brasiliensis* showed a high amount and great variety of different lipocalins (93.8% of secreted proteins). Remarkably, a great number of serine proteases that were not observed in previous blood-sucking sialotranscriptomes were found. Nine Kazal peptides were identified, among them one with high homology to the tabanid vasodilator vasotab, suggesting that the *Triatoma* vasodilator could be a Kazal protein.

Keywords

Saliva; Transcriptome; Hematophagy; Salivary proteins; *Triatoma brasiliensis*

1. Introduction

Triatoma brasiliensis is the most important autochthonous vector of *Trypanosoma cruzi* in Brazil, where it is widely distributed in the semiarid areas of the Northeast (Costa et al.,

* Corresponding author. Tel.: +55 31 3499-2835, fax: +55 31 3499-2970. E-mail address: marcoshp@icb.ufmg.br.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting galley proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

2003; Dias et al., 2000; Silveira et al., 1984). This species is able to colonise houses and peridomestic areas, and is also widely distributed in sylvatic habitats where it is mainly found among rock piles associated with various species of vertebrates. This increases its importance as a vector, because it can re-colonise domestic habitats after it has been eliminated through insecticide spraying (Alencar, 1987; Diotaiuti et al., 2000).

Blood feeders have evolved a wide set of pharmacologically active molecules to counteract host defense systems (haemostasis, inflammation, immune response) in the feeding site (Andrade et al., 2005; Ribeiro and Francischetti, 2003). Several biomolecules have already been described in triatomine bug saliva, including anticoagulants (Hellman and Hawkins, 1964; 1965; Pereira et al., 1996; Ribeiro et al., 1998;), vasodilators (Ribeiro et al., 1990; Ribeiro et al., 1993; Ribeiro and Nussenzeig, 1993), antihistamines (Ribeiro & Walker, 1994), sialidase (Amino et al., 1998), sodium channel blockers (Dan et al., 1999), immunosuppressors (Kalvachova et al., 1999), pore formers (Amino et al., 2002), complement inhibitor systems (Cavalcante et al., 2003) and inhibitors of platelet aggregation induced by collagen (Noeske-Jungblut et al., 1994; Ribeiro and Garcia, 1981), ADP (Ribeiro & Garcia, 1980; Sarkis et al., 1986), arachidonic acid (Ribeiro & Sarkis, 1982), thrombin (Francischetti et al., 2000; Noeske-Jungblut et al., 1995) or PAF (Golodne et al., 2003). Few of these activities have been reported from *T. brasiliensis*. Sant'Anna et al. (2002), using the suppression subtractive hybridization (SSH) technique, identified six full-length differentially expressed cDNAs from *T. brasiliensis*. Among them, three had their activities inferred by similarity with other sequences in GenBank, but the others had no obvious orthologues.

Until now, only the sialome of *Rhodnius prolixus* has been published (Ribeiro et al., 2004), but the evolutionary diversity among triatomines suggests that extrapolating findings from one group to another should be undertaken with caution. So we can see that there are 138 species in the Triatominae placed in six tribes (Triatomini, Rhodniini, Cavernicoloni, Bolboderini, Alberproseniini, Linschosteiini) and 19 genera (Galvão et al., 2003). Schofield (1988) suggested that the Triatominae had different origin from Reduviidae predators that converged to haematophagy habit independently. Several studies showed differences between Rhodniini and Triatomini tribes (Bargues et al., 2000; Catalá, 1997; Dujardin et al., 1999; Garcia and Powell, 1998; Jurberg, 1996; Marcilla et al., 2001; Stothard et al., 1998). Marked differences are also seen between *Rhodnius* and *Triatoma* saliva, in that *Triatoma* lacks nitrophorins, and their apyrase, anticoagulant and vasodilator activities show distinct mechanisms of action (Ribeiro et al., 1998). So a comparative study of a sialome from the Triatomini will be very informative and permit interesting comparisons with that from the Rhodniini already published. Thus, in this study a salivary gland cDNA library of *Triatoma brasiliensis* was mass sequenced and analyzed, and polypeptides were sequenced by HPLC/Edman degradation experiments.

2. Materials and methods

2.1. Insect rearing

T. brasiliensis were captured in Simplício Mendes, Piauí (Northeastern region of Brazil) and reared in the insectary of the Centro de Pesquisas René Rachou - Fiocruz MG, maintained at 28 ± 2 °C and $65 \pm 10\%$ relative humidity. They were kept in cages containing vertical strips of coarse filter paper and fed weekly on chickens.

2.2. Salivary gland cDNA library construction

T. brasiliensis salivary gland mRNA was isolated from 50 salivary gland pairs from starved adult insects using the Dynabeads mRNA[®] DIRECT[™] kit (DYNAL, Great Neck, NY). The PCR-based cDNA library was constructed using a SMART cDNA library construction kit

(BD-Clontech, Palo Alto, CA), according to the manufacturer's instructions. The obtained library was plated by infecting log-phase XL-1 Blue cells (Stratagene, La Jolla, CA, USA). The titer of the cDNA library was 0.492×10^6 pfu/ml, with a recombination efficiency of 87%.

2.3. Sequencing of *T. brasiliensis* cDNA library

T. brasiliensis salivary gland cDNA library was plated to approximately 100 plaques per plate (80mm Petri dish). The plaques were randomly picked and transferred to 1.5 ml centrifuge microtubes containing 100 μ L of distilled water. Five microliters of the phage sample were used as template for a PCR to amplify random cDNAs. The primers TriplEx2-F (5'-CTC CGA GAT CTG GAC GAG C-3') positioned upstream of the cDNA of interest (5' end), and TriplEx2-R (5'-TAA TAC GAC TCA CTA TAG GGC-3') positioned downstream of the cDNA of interest (3' end) were used for the PCR (94°C/4 min followed by 35 cycles of 94°C/1 min, 52.5°C/1 min and 72°C/1.1 min, and a final extension of 72°C/7 min) carried out with the Pht *Taq* polymerase system (Phoneutria, Belo Horizonte, MG, Brazil). Amplified products were visualized by 1.0% agarose gel electrophoresis and cleaned up using the GFX PCR DNA and Gel Band Purification Kit (GE/Amersham Biosciences, Buckinghamshire, UK) or the Wizard® SV Gel and PCR Clean-Up System kit (Promega, Madison, WI, USA). Four microliters of the cleaned PCR product was used as a template for a cycle-sequencing reaction using the DYEnamic ET dye terminator cycle sequencing kit (GE/Amersham Biosciences). The primer Seq.Clontech-F (5'-CTC GGG AAG CGC GCC ATT GTG TTG GT-3') was used for sequencing. Conditions were 94°C/1 min, and 35 cycles of 94°C/30 sec, 51°C/25 sec and 60°C/4 min. After cycle-sequencing the samples, a post-reaction cleanup step, consisted of isopropanol precipitation followed by 70% ethanol wash, was performed. After the supernatant removal, each pellet was dissolved in 6 μ L of MegaBACE loading solution and sequenced on a MegaBACE™ 1000 sequencing instrument (GE/Amersham Biosciences).

2.4. Chromatography

Approximately 20 μ L of saliva from starved adult insects were chromatographed according Ribeiro et al. (2004). Briefly, experiments used 0.24 ml bed volume columns of strong cation (Mono-S) and strong anion (Mono-Q) ion exchangers obtained from Amersham Biosciences (Piscataway, NJ). To elute the proteins of interest, the ion-exchange columns were submitted to gradients of NaCl (0 to 1 M). For the cation exchange column, the buffer used was 50 mM sodium acetate at pH 5.0 and for the anion exchange, 50 mM Tris-Cl at pH 8.0. Fractions of interest had 40 μ L removed and diluted with an equal volume of 20% methanol containing 0.4% tri-fluoroacetic acid (TFA) and were applied to a ProSorb cartridge (Perkin Elmer, Foster City, CA) previously treated with 10 μ L of methanol. After absorption of the solution through the polyvinylidene difluoride (PVDF) membrane, the cartridge was washed three times with the same volume of 10% methanol containing 0.1% TFA.

2.5. Bioinformatics

Bioinformatics procedures were as by Francischetti et al. (2002) and Valenzuela et al. (2002), except that the clustering of the cDNA sequences was accomplished using the CAP program (Huang, 1992) after initial clustering of the database following a BLASTN (Altschul et al., 1997) of the database against itself and reading the output to join those sequences that had at least 95 identical residues in a window of 100 residues. Accession numbers for the National Center for Biology Information (NCBI) databases are given as gi|XXXX, where XXXX is the accession number. Signal peptide predictions were done with the SignalP program (Nielsen et al., 1997). Transmembrane helices were predicted with the TMHMM program (Sonnhammer et al., 1998). Sequence alignments and phylogenetic tree

analysis used the ClustalW package (Thompson et al., 1994). Phylogenetic trees were constructed by the neighbor-joining method (Saitou and Nei, 1987). Boot strapping of phylogenetic trees was done with the Clustal package for 1000 trials. Phylogenetic trees and dendograms were formatted with TreeView (Page, 1996) using the ClustalW output. The electronic version of the complete tables in Microsoft Excel format with hyperlinks to web-based databases and to BLAST results is available at http://www.ncbi.nlm.nih.gov/projects/omes/T_brasiliensis_sialome

2.6. Supplemental material

Supplemental Tables S1 and S2 and figures are hyperlinked throughout the paper to the NCBI pages, http://www.ncbi.nlm.nih.gov/projects/omes/T_brasiliensis_sialome/Sup_tab1/TB-Sup-table1.xls and http://www.ncbi.nlm.nih.gov/projects/omes/T_brasiliensis_sialome/Sup_tab2/TB-Sup-table2.xls respectively.

3. Results and Discussion

3.1. cDNA library characteristics

To obtain an insight into the salivary transcriptome of *T. brasiliensis*, we randomly sequenced 1,712 cDNA clones from a salivary gland cDNA library from this insect and assembled a clusterised database (Supplemental Table S1), yielding 786 clusters of related sequences, 666 of which contained only a single EST. The consensus sequence of each cluster is designated either a contig (deriving from two or more sequences) or a singleton (deriving from a single sequence). In this paper, for simplicity sake, we will use the denomination contig to address sequences deriving both from consensus sequences and from singletons.

Using the BLAST package of programs (Altschul et al., 1997), we compared the sequence of each cluster in the database with the nonredundant protein and nucleotide sets of the NCBI and the gene ontology database (Ashburner et al., 2000; Hvidsten et al., 2001; Lewis et al., 2000). The translated sequences were also screened with RPSBLAST for protein motifs of the combined set of Pfam (Bateman et al., 2000) and SMART (Schultz et al., 2000) databases (also known as the Conserved Domains Database – CDD). Finally, we submitted all translated sequences (starting with a Met) to the SignalP server (Nielsen et al., 1997) to detect the presence of signal peptides indicative of secretion. The EST assembly, BLAST, and signal peptide results were piped into an Excel spreadsheet for manual annotation. Three categories of expressed genes derived from the manual annotation of the contigs (Table 1). The housekeeping (H) category had 24.4% and 17.8% of the clusters and sequences, respectively, and an average of 1.6 sequences per cluster. In contrast, the putatively secreted (S) category contained 47.1% of the clusters and 68.2% of the sequences, with an average number of 3.2 sequences per cluster. Similar results were observed in other transcriptome analysis of salivary glands in *Aedes aegypti*, *Anopheles gambiae*, *Ixodes scapularis* and *R. prolixus* (Francischetti et al., 2002; Ribeiro et al., 2004; Valenzuela et al., 2002). Finally, 28.5% of the clusters, containing 14% of all sequences, were classified as unknown (U) because no assignment for their putative function could be made; most of these consisted of singletons.

3.2. Housekeeping (H) genes

The 192 gene clusters (comprising 305 EST) attributed to H genes expressed in the salivary glands of *T. brasiliensis* were further characterized into 17 subgroups according to function (Table 2). Two of the largest sets were associated with protein synthesis machinery (117 EST in 53 clusters) and with energy metabolism (38 EST in 30 clusters). This is consistent with an organ specialized in secreting polypeptides. Also, the 40 EST (37 clusters),

representing conserved proteins of unknown function and indicated as 'Conserved' in Table 2, are presumably associated with cellular metabolism. This group also includes a homologue of the *Rhodnius prolixus* salivary protein MYS3 precursor (Ribeiro et al., 2004). Twelve clusters code for products associated with protein modification function, including two glutathione S-transferase proteins and a glutathione S-transferase-like protein and, also, two chaperones. Nine transporters/storage protein genes were also identified, including three coding for energy production and conversion, two for inorganic ion transport and metabolism, one involved with intracellular trafficking, secretion, and vesicular transport, and another two coding for carbohydrate or lipid transport and metabolism. Additional information on each of the 192 gene clusters is available online (Supplemental Table S1).

3.3. Possibly secreted proteins and peptides

Supplemental Table S1 indicates the presence of several gene families previously described in the salivary glands of *R. prolixus* (Ribeiro et al., 2004) and *T. infestans* (NCBI). Remarkably, of the 370 clusters of transcripts possibly associated with secretory products, 341 code for proteins of the lipocalin family (Flower et al., 2000). A summary of these transcripts organized by their abundance and protein family is shown in Table 3.

3.4. Preliminary characterization of the salivary transcriptome and proteome of *T. brasiliensis*

Several clusters of sequences coding for housekeeping and putative secreted polypeptides, indicated in Supplemental Table S1, are abundant and complete enough to extract consensus sequences of novel cDNA. Additionally, primer extension studies on several clones allowed us to obtain full- or near full-length sequences of products of interest. A total of 41 novel sequences, 36 of which code for putative secreted proteins, are grouped in Supplemental Table S2.

To obtain information on the most abundant proteins in the salivary glands of *T. brasiliensis*, 20 µL of saliva were collected from starved adult insects to perform chromatographic experiments. The ion-exchange chromatography is better suited than SDS-PAGE for this purpose because most *T. brasiliensis* salivary proteins are lipocalins with a molecular mass of 18–22 kDa. SDS-PAGE leads to poor separation of such proteins. Peaks of interest were collected and submitted to Edman degradation. These results are summarized in Fig. 1. Descriptions of the proteins identified are provided below.

3.4.1. Putative salivary-secreted proteins for which a function is known or presumed—From 370 clusters of transcripts possibly associated with secretory products, 341 (93.8%) are lipocalins. This family contains salivary proteins from triatomine insects and ticks that counteract vertebrate host haemostasis events such as coagulation, vasoconstriction and platelet aggregation (Ribeiro, 1995). All the results obtained in Edman degradation experiments were from proteins of this group (Supplemental Table S1).

Lipocalins: Among the lipocalin family, 218 sequences had similarity to Triabins from other triatomines. These include: (I) Triabin, a serine-protease inhibitor that forms a non-covalent complex with thrombin (Noeske-Jungblut et al., 1995); (II) Pallidipin, an anticollagen that prevents platelet aggregation (Noeske-Jungblut et al., 1994); (III) procalin, the major allergen of *T. protracta* saliva (Paddock et al., 2001); (IV) several triabin-like sequences and (V) others.

Five contigs of *T. brasiliensis* sialome were found matching the Nitrophorins. All of them matched the Nitrophorin 3B from *R. prolixus* with identity values that ranged from 31 to 42%. The findings are of interest, because Nitrophorins are hemeproteins found in the saliva

of blood-feeding insects that carry nitric oxide – typical of *Rhodnius* species. These proteins do not reveal amino acids sequence similar to lipocalins but their predicted crystal structure is typical of this large family of proteins (Andersen et al., 1997, 1998). Saliva of the blood-sucking bug *R. prolixus* contains four homologous nitrophorins, named NP1 to NP4, according to their relative abundance in the glands. Combined they represent near 50% of the salivary protein and confer a deep cherry color to the gland. However, heme-containing proteins have not been described in the saliva of *Triatoma* species, nor they appear to exist in large concentrations, as the *Triatoma* salivary glands are colorless.

Lipocalins are remarkably diverse at the amino acid sequence level, often falling below 20% identity between members, yet have highly conserved structures (Flower et al., 2000). Functional genomic, proteomic, and functional studies have been performed to probe the role of salivary lipocalins in blood feeding arthropods. In the course of these investigations, anticoagulant, antiplatelet, anti-inflammatory, and vasodilatory molecules have been described (Andersen et al., 2005). The high number and variety of lipocalins found in *T. brasiliensis* has been reported previously in the saliva of *R. prolixus* and ticks. Ticks and kissing bugs evolved salivary lipocalins that act as efficient scavengers of biogenic amines, that is of strong adaptive value in the convergent evolution of arthropods to hematophagy (Calvo et al., 2006). As in *R. prolixus*, the proliferation of lipocalin genes from *Triatoma* species has probably occurred via gene duplication and subsequent divergence (Ribeiro et al., 2004).

A phylogenetic tree containing all published salivary lipocalins from triatomines including those described in this work reveals the large divergence of this group of proteins, indicated by the weak bootstrap support for many of its clades (Fig. 2). In particular, the tree clearly shows divergence between *Rhodnius* and *Triatoma* proteins, and, within *Triatoma*, it shows several instances of clades with strong bootstrap support containing solely members of one species (although clades with members deriving from more than one *Triatoma* species also exist). Some of these closely related proteins from the same species proteins could be alleles, but when the amino acid sequence diverge by more than 10% they are most probably derived from gene duplication events. The existence of mono-specific clades is indicative of gene duplication events that occurred after the ancestral *Triatoma* originated the present day *Triatoma* species, supporting the idea that gene duplication events must have been important in the evolution to blood feeding, as proposed before for ticks (Mans and Neitz, 2004).

Kazal peptides: Nine clusters of sequences displayed the Kazal domain signature. The similarity of the contigs with other Kazal inhibitors is shown in Fig. 3A. Among them, TBQ-contig 259 codes for a protein 53% identical to infestin, the intestinal thrombin inhibitor from *T. infestans* (Lovato et al., 2006), coherent to the neighbor-joining tree analysis that demonstrated higher similarity to triatomine intestinal inhibitors (Fig. 3B). Predicted amino acid sequence from TBQ-contig 54 (near full-length sequence) was similar to the vasodilator from horse flies (*Hybomitra bimaculata*, *Diptera*, *Tabanidae*), named vasotab precursor (Takac et al., 2006). Vasotab is a member of the Kazal-type protease inhibitor family. Physiologic tests with vasotab demonstrated positive inotropism in isolated rat hearts, vasodilatation of coronary and peripheral vessels, and Na, K-ATPase inhibition. Accordingly, we can speculate that the *Triatoma*'s vasodilator could be a Kazal peptide. The remaining six Kazal peptides showed homology to putative sequences of different species but with expected high blast values (> 0.01).

Secreted serine proteases: Five contigs had similarity to serine proteases. Among them, one (TBQ-contig 692) showed 39% identity to a trypsin-like serine protease from Zebrafish (gi|66911393). The four remaining matched trypsin-like salivary precursors previously identified in the plant-feeding bug *Lygus lineolaris* (Zhu et al., 2003), showing identity to

the amino acid level ranging from 31 to 67%. The most abundant contig (TBQ-40), with 24 sequences, codes for a protein 41% identical (BLAST E value $3E-37$) to the *L. lineolaris* trypsin-like precursor (LISgP4). The ClustalW alignment from *T. brasiliensis* and *L. lineolaris* trypsin-like sequences is shown in Fig. 4. *Lygus* bugs ingest plant liquids by inserting their sucking mouthparts into plant tissues where extra-oral digestion is facilitated by the secretion of digestive enzymes from the salivary glands (Cohen et al., 1998). Proteolytic activity has been detected in salivary glands from many mirids, including *L. rugulipennis* (Laurema et al., 1985) and *Creontiades dilutus* (Colebatch et al., 2001). In triatomines, a trypsin-like activity was described for the *T. infestans* triapsin (Amino et al., 2001). The protein was identified as an inactive precursor and a second trypsin-like protein may be responsible for the triapsin activation upon saliva release. Therefore, the trypsin-like sequences found in the *T. brasiliensis* sialome could be related to saliva processing upon the release on the host skin. Although they showed similarity to the typical trypsins with digestive function that are capable of cleaving a variety of proteins, the salivary trypsin-like from haematophagous hemiptera might be very specific in relation to their substrates. These proteases could participate in the processing of salivary proteins or act in specific targets from the host, once typical trypsins could elicit itch in the skin (Thomsen et al., 2002) and increase the possibility of the bug to be perceived by the host.

Apyrase: Apyrases are enzymes ubiquitously found in the salivary glands of blood-feeding insects and ticks (Valenzuela et al., 2003). These enzymes, belonging to different protein families, degrade the neutrophil-activating substance ATP and the platelet-aggregating nucleotide ADP to AMP, presumably facilitating blood feeding. *T. infestans* apyrase is a member of the 5'-nucleotidase family (Champagne et al., 1995) and was reported as a set of five different molecular weight proteins (Faudry et al., 2004). The contigs showed 43% similarity to an *Aedes aegypti* apyrase. The fact that only one contig with only one EST was found suggests that the level of expression in *T. brasiliensis* is low, and indeed *T. infestans* apyrase activity has been recorded to be about 10-fold higher than that in *T. brasiliensis* (Ribeiro et al., 1998).

Inositol phosphatase: Two contigs produced similarity to inositol phosphatases from *R. prolixus*. The inositol polyphosphate 5-phosphatase (IPP) enzymes act on both soluble inositol phosphate and phosphoinositide substrates and are involved in many cellular processes related to signal transduction, secretion, and cytoskeletal structure. In *R. prolixus*, it was previously thought to be responsible for the salivary apyrase activity, but later identified with no apyrase activity (Ribeiro et al., 2004). *R. prolixus* inositol polyphosphate 5-phosphatase exists as an isolated IPP domain, which is secreted into the saliva of this blood-feeding insect. It shows selectivity for soluble and lipid substrates having a 1,4,5-trisphosphate substitution pattern while only poorly hydrolyzing substrates containing a D3 phosphate (Andersen and Ribeiro, 2006). The role of salivary Inositol phosphatases for triatomine bugs remains unclear.

Salivary protein MYS2 precursor: One contig showed similarity to the *R. prolixus* MYS-2. The MYS proteins have no homology to sequences with known functions. Two MYS proteins were identified in the *T. brasiliensis* sialome, MYS-2 and MYS-3. Each of these showed high homology to *R. prolixus*, probably secreted MYS protein, but while MYS-2 was classified as a secreted protein, the MYS-3 was classified as housekeeping. The *R. prolixus* sialome also identified a third member of the family, the MYS-1, which was also predicted as probably secreted and showed no homology to the ESTs sequenced here (Ribeiro et al., 2004).

Antigen-5-like protein precursor: One cDNA contig indicated similarity to the antigen-5-protein family from *R. prolixus* (Ribeiro et al., 2004). These are a widespread extracellular family of proteins ubiquitously found in animals and plants with mostly unknown functions (Schreiber et al., 1997; Valenzuela et al., 2003). Closely related proteins from this family have been reported in the salivary glands of Hymenoptera and Diptera, such as sand flies (Charlab et al., 1999), tsetse (Li et al., 2001), mosquitoes (Francischetti et al., 2002; Valenzuela et al., 2002) and *Culicoides sonorensis* (Campbell et al., 2005).

Other proteins: Other contigs were identified with homology to proteins deposited in the GenBank. They are a contigs with similarity to sulfatase, similar to *C. sonorensis* protein, similar to major epididymal secreted protein and two contigs with similarity to heme-binding proteins from *R. prolixus*.

3.5. Concluding remarks

The sialoma of *T. brasiliensis* showed a high number and variety of lipocalins. A similar situation has been reported previously in the saliva of *R. prolixus* and ticks (Ribeiro et al., 2004;). Remarkably, the sialoma showed the presence of a great number of serine proteases that were not observed in previous sialotranscriptomes. Although mosquitoes have a chymotrypsin-like enzyme, the number of transcripts is low. Ticks have a metalloprotease that breaks fibrin into fibrinogen, but it's a completely different family of proteins.

The presence of the Kazal peptide with high homology to the tabanid vasodilator is also interesting, suggesting that the *Triatoma* vasodilator could be a Kazal protein.

Acknowledgments

We thank Fabrício Rodrigues dos Santos and Paula Lara Ruiz from the Laboratório de Biodiversidade e Evolução Molecular (ICB/UFMG), and Cristiane Quimelli Snoeijer and Carlos Rodrigo Bueno from the Laboratório de Protozoologia, (MIP - CCB/ UFSC) for helping in the mass sequencing reactions. This work was supported by the Wellcome Trust Fundation, CNPq, CAPES, Fapemig, ECLAT and the Intramural Research Program of the National Institute of Allergy and Infectious Diseases.

References

- Alencar, JE. História natural da doença de Chagas no Estado do Ceará. Fortaleza: Imprensa Universitária da UFC; 1987. 339 p.
- Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. Gapped BLAST and PSIBLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 1997; 25:3389–3402. [PubMed: 9254694]
- Amino R, Porto RM, Chammas R, Egami MI, Schenkman S. Identification and characterization of a sialidase released by the salivary gland of the hematophagous insect *Triatoma infestans*. *J. Biol. Chem.* 1998; (38):24575–24582. [PubMed: 9733752]
- Amino R, Tanaka AS, Schenkman S. Triapsin, an unusual activatable serine protease from the saliva of the hematophagous vector of Chagas' disease *Triatoma infestans* (Hemiptera: Reduviidae). *Insect Biochem. Mol. Biol.* 2001; 31:465–472. [PubMed: 11222956]
- Amino R, Martins RM, Procópio J, Hirata IY, Juliano MA, Schenkman S. Trialysin, a novel pore-protein from saliva of hematophagous insects activated limited proteolysis. *J. Biol. Chem.* 2002; 277:6207–6213. [PubMed: 11751887]
- Andersen JF, Champagne DE, Weichsel A, Ribeiro JM, Balfour CA, Dress V, Montford WR. Nitric oxide binding and crystallization of recombinant nitrophorin I, a nitric oxide transport protein from blood-sucking bug *Rhodnius prolixus*. *Biochemistry.* 1997; 36:4423–4428. [PubMed: 9109649]
- Andersen JF, Weichsel A, Balfour CA, Champagne DE, Montford WR. The crystal strycture of nitrophorin 4 at 1.5 Å resolution: transport of nitric oxide by a lipocalin-based heme protein. *Structure.* 1998; 6:1315–1327. [PubMed: 9782054]

- Andersen JF, Gudderra NP, Ivo, Francischetti MB, Ribeiro JMC. The role of salivary lipocalins in blood feeding by *Rhodnius prolixus*. *Arch. Insect Biochem. Physiol.* 2005; 58:97–105. [PubMed: 15660358]
- Andersen JF, Ribeiro JMC. A secreted salivary inositol polyphosphate 5-phosphatase from a blood-feeding insect: allosteric activation by soluble phosphoinositides and phosphatidylserine. *Biochemistry.* 2006; 45:5450–5457. [PubMed: 16634626]
- Andrade BB, Teixeira CR, Barral A, Barral-Netto M. Haematophagous arthropod saliva and host defense system: a tale of tear and blood. *An. Acad. Bras. Cienc.* 2005; 77:665–693. [PubMed: 16341443]
- Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, et al. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nature Genetics.* 2000; 25:25–29. [PubMed: 10802651]
- Bargues MD, Marcilla A, Ramsey JM, Durjardin JP, Schofield CJ, Mas-Coma S. Nuclear DNA-based molecular clock of the evolution of Triatominae (Hemiptera: Reduviidae), vectors of Chagas' disease. *Mem. Inst. Oswaldo Cruz.* 2000; 95:567–573. [PubMed: 10904416]
- Bateman A, Birney E, Durbin R, Eddy SR, Howe KL, Sonnhammer EL. The Pfam protein families database. *Nucleic Acids Res.* 2000; 28:263–266. [PubMed: 10592242]
- Calvo E, Mans BJ, Andersen JF, Ribeiro JMC. Function and Evolution of a Mosquito Salivary Protein Family. *J. Biol. Chem.* 2006; 281:1935–1942. [PubMed: 16301315]
- Campbell CL, Vandyke KA, Letchworth GJ, Drolet BS, Hanekamp T, Wilson WC. Midgut and salivary gland transcriptomes of the arbovirus vector *Culicoides sonorensis* (Diptera: Ceratopogonidae). *Insect Mol. Biol.* 2005; 14:121–136. [PubMed: 15796745]
- Catalá S. Antennal sensilla id Triatominae (Hemiptera, Reduviidae): A comparative study over five genera. *Int. J. Insect Morphol. Embr.* 1997; 26:67–73.
- Cavalcante RR, Pereira MH, Gontijo NF. Anti-complement activity in the saliva of phlebotomine sand flies and other haematophagous insects. *Parasitology.* 2003; 127:87–93. [PubMed: 12885192]
- Champagne D, Nussenzeig RH, Ribeiro JMC. Purification, characterization, and cloning of nitric oxide-carrying heme proteins (nitrophorins) from salivary glands of the blood-sucking insect *Rhodnius prolixus*. *J. Biol.Chem.* 1995; 270:8691–8695. [PubMed: 7721773]
- Charlab R, Valenzuela JG, Rowton ED, Ribeiro JMC. Toward an understanding of the biochemical and pharmacological complexity of the saliva of a hematophagous sand fly *Lutzomyia longipalpis*. *Proc. Natl. Acad. Sci. USA.* 1999; 96:15155–15160. [PubMed: 10611354]
- Cohen AC. Solid-to-liquid feeding: the inside(s) story of extra-oral digestion in predaceous Arthropoda. *Am. Entomologist.* 1998; 44:103–117.
- Colebatch G, East P, Cooper P. Preliminary characterisation of digestive proteases of the green mirid, *Creontiades dilutus* (Hemiptera: Miridae). *Insect Biochem. Mol. Biol.* 2001; 31:415–423. [PubMed: 11222951]
- Costa J, Almeida CE, Dotson E, Lins A, Vinhaes M, Silveira AC, Beard CB. The epidemiologic importance of *Triatoma brasiliensis* as a Chagas disease vector in Brazil: a revision of domiciliary captures between 1993–1999. *Mem. Inst. Oswaldo Cruz.* 2003; 98:443–449. [PubMed: 12937751]
- Dan A, Pereira MH, Pesquero JL, Diotaiuti L, Beirao PS. Action of the saliva of *Triatoma infestans* (Heteroptera: Reduviidae) on sodium channels. *J. Med. Entomol.* 1999; 36:875–879. [PubMed: 10593094]
- Dias JCP, Machado EMM, Fernandes AL, Vinhaes M. General situation and perspectives of Chagas disease in Northeastern Region, Brazil. *Cad. Saúde Publ.* 2000; 16:13–34.
- Diotaiuti L, Faria Filho OF, Carneiro FCF, Dias JCP, Pires HHR, Schofield CJ. Operational aspects of *Triatoma brasiliensis* control. *Cad. Saúde Pública.* 2000; 16:61–67.
- Dujardin JP, Steinel M, Chavez T, Machyane M, Schofield CJ. Changes in the sexual dimorphism of Triatominae in the transition from natural to artificial habitats. *Mem. Inst. Oswaldo Cruz.* 1999; 94:565–569. [PubMed: 10446020]
- Faudry E, Lozzi SP, Santana JM, D'Souza-Ault M, Kieffer S, Felix CR, Ricart CA, Sousa MV, Vernet T, Teixeira AR. *Triatoma infestans* apyrases belong to the 5'-nucleotidase family. *J. Biol. Chem.* 2004; 279:19607–19613. [PubMed: 14985353]

- Flower DR, North AC, Sansom CE. The lipocalin protein family: structural and sequence overview. *Biochim. Biophys. Acta.* 2000; 1482:9–24. [PubMed: 11058743]
- Francischetti IM, Ribeiro JM, Champagne D, Andersen J. Purification, cloning, expression, and mechanism of action of a novel platelet aggregation inhibitor from the salivary gland of the blood-sucking bug, *Rhodnius prolixus*. *J. Biol. Chem.* 2000; 275:12639–12650. [PubMed: 10777556]
- Francischetti IM, Valenzuela JG, Pham VM, Garfield MK, Ribeiro JMC. Toward a catalog for the transcripts and proteins (sialome) from the salivary gland of the malaria vector *Anopheles gambiae*. *J. Exp. Biol.* 2002; 205:2429–2451. [PubMed: 12124367]
- Galvão C, Carcavallo R, Rocha DS, Jubert J. A checklist of the current valid species of the subfamily Triatominae Jeannel, 1919 (Hemiptera, Reduviidae) and their geographical distribution, with nomenclatural and taxonomic notes. *Zootaxa.* 2003; 202:1–36.
- Garcia BA, Powell JR. Phylogeny of species of *Triatoma* (hemiptera: Reduviidae) based on mitochondrial DNA sequences. *J. Med. Entomol.* 1998; 35:232–238. [PubMed: 9615540]
- Golodne DM, Monteiro RQ, Graca-Souza AV, Silva-Neto MA, Atella GC. Lysophosphatidylcholine acts as an anti-hemostatic molecule in the saliva of the blood-sucking bug *Rhodnius prolixus*. *J. Biol. Chem.* 2003; 278:27766–27771. [PubMed: 12740385]
- Hellmann K, Hawkins RI. Anticoagulants and fibrinolytic activities from *Rhodnius prolixus*. *Stal. Nature.* 1964; 201:1008–1009. [PubMed: 14191567]
- Hellmann K, Hawkins RI. Prolixin-S and Prolixin-G; two anticoagulants from *Rhodnius prolixus* Stal. *Nature.* 1965; 207:265–267. [PubMed: 5887330]
- Hvidsten TR, Komorowski J, Sandvik AK, Laegreid A. Predicting gene function from gene expressions and ontologies. *Pac. Symp. Biocomput.* 2001:299–310. [PubMed: 11262949]
- Huang X. A contig assembly program based on sensitive detection of fragment overlaps. *Genomics.* 1992; 14:18–25. [PubMed: 1427824]
- Jurberg, J. Uma abordagem filogenética entre os Triatomíneos baseada nas estruturas fállicas. In: Schofield, CJ.; Dujardin, JP.; Jurberg, J., editors. *Proceedings of the International Workshop on Population Genetics and Control of Triatominae*, Santo Domingo de los Colorados, Ecuador; Mexico City: INDRE; 1996. p. 51-52.
- Kalvachova P, Hribalova V, Kodym P, Volf P. Modulation of murine lymphocyte responsiveness by the saliva of *Rhodnius prolixus* (Hemiptera: Reduviidae). *J. Med. Entomol.* 1999; 36:341–344. [PubMed: 10337105]
- Laurema S, Varis AL, Miettinen H. Studies on enzymes in the salivary glands of *Lygus rugulipennis* (Hemiptera Miridae). *Insect Biochem.* 1985; 15:211–224.
- Lewis S, Ashburner M, Reese MG. Annotating eukaryote genomes. *Current Opinions in Structural Biology.* 2000; 10:349–354.
- Li S, Kwon J, Aksoy S. Characterization of genes expressed in the salivary glands of the tsetse fly, *Glossina morsitans morsitans*. *Insect Mol. Biol.* 2001; 10:69–76. [PubMed: 11240638]
- Lovato DV, Nicolau de Campos IT, Amino R, Tanaka AS. The full-length cDNA of anticoagulant protein infestin revealed a novel releasable Kazal domain, a neutrophil elastase inhibitor lacking anticoagulant activity. *Biochimie.* 2006; 88:673–681. [PubMed: 16469426]
- Mans BJ, Neitz AW. Adaptation of ticks to a blood-feeding environment: evolution from a functional perspective. *Insect Biochem. Mol. Biol.* 2004; 34:1–17. [PubMed: 14723893]
- Marcilla A, Bargues MD, Ramsey JM, Magallon-Gastelum E, Salazar-Schettino PM, Abad-Franch F, Dujardin JP, Schofield CJ, Mas-Coma S. The ITS-2 of the nuclear DNAr as a molecular marker for populations, species and phylogenetic relationships in Triatominae (Hemiptera: Reduviidae), vectors of Chagas' disease. *Mol. Phylogenet. Evol.* 2001; 18(1):136–142. [PubMed: 11161750]
- Nielsen H, Engelbrecht J, Brunak S, von Heijne G. Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. *Protein Eng.* 1997; 10:1–6. [PubMed: 9051728]
- Noeske-Jungblut C, Kratzchmar J, Haendler B, Alagon A, Possani L, Verhallen P, Donner P, Schleuning WD. An inhibitor of collagen-induced platelet aggregation from the saliva of *Triatoma pallidipennis*. *J. Biol. Chem.* 1994; 269:5050–5053. [PubMed: 8106481]
- Noeske-Jungblut C, Haendler B, Donner P, Alagon A, Possani L, Schleuning WD. Triabin, a highly potent exosite inhibitor of thrombin. *J. Biol. Chem.* 1995; 270:28629–28634. [PubMed: 7499380]

- Paddock CD, McKerrow JH, Hansell E, Foreman KW, Hsieh I, Marshall N. Identification, cloning, and recombinant expression of procalin, a major triatomine allergen. *J Immunol.* 2001; 167:2694–2699. [PubMed: 11509613]
- Page RD. TreeView: an application to display phylogenetic trees on personal computers. *Comput. Appl. Biosci.* 1996; 12:357–358. [PubMed: 8902363]
- Pereira MH, Souza MEL, Vargas AP, Penido CM, Martins MS, Diotaiuti L. Anticoagulant activity of *Triatoma infestans* and *Panstrongylus megistus* saliva (Hemiptera/ Triatominae). *Acta Tropica.* 1996; 61:255–264. [PubMed: 8790775]
- Ribeiro JMC, Garcia ES. The salivary and crop apyrase activity of *Rhodnius prolixus*. *J. Insect Physiol.* 1980; 26:303–307.
- Ribeiro JMC, Garcia ES. Platelet antiaggregating activity in the salivary secretion of the blood-sucking bug *Rhodnius prolixus*. *Experientia.* 1981; 37:384–386. [PubMed: 7016577]
- Ribeiro JMC, Sarkis JJF. Anti-thromboxane activity in *Rhodnius prolixus* salivary secretion. *J. Insect Physiol.* 1982; 28:655–660.
- Ribeiro JMC, Marinotti O, Gonzles R. A salivary vasodilator in the blood-sucking bug, *Rhodnius prolixus*. *Br. J. Pharmacol.* 1990; 101:932–936. [PubMed: 2085715]
- Ribeiro JMC, Hazzard JMH, Nussenzveig RH, Champagne DE, Walker FA. Reversible binding of nitric oxide by a salivary heme protein from a bloodsucking insect. *Science.* 1993; 260:539–541. [PubMed: 8386393]
- Ribeiro JMC, Nussenzveig RH. Nitric oxide synthase activity from a hematophagous insect salivary gland. *FEBS Lett.* 1993; 330:165–168. [PubMed: 7689981]
- Ribeiro JMC, Walker FA. High affinity histamine-binding and antihistaminic activity of the salivary nitric oxidecarrying heme protein (nitrophorin) of *Rhodnius prolixus*. *J. Exp. Med.* 1994; 180:2251–2257. [PubMed: 7964498]
- Ribeiro JMC. Blood-feeding arthropods: live syringes or invertebrate pharmacologists? *Infect. Agents Dis.* 1995; 4:143–152.
- Ribeiro JMC, Schneider M, Isaias T, Jurberg J, Galvo CC, Guimares JA. Role of salivary antihemostatic components in blood feeding by triatomine bugs (Heteroptera). *J. Med. Entomol.* 1998; 35:599–610. [PubMed: 9701952]
- Ribeiro JM, Francischetti IM. Role of arthropod saliva in blood feeding: sialome and post-sialome perspectives. *Annu. Rev. Entomol.* 2003; 48:73–88. [PubMed: 12194906]
- Ribeiro JMC, Andersen J, Silva-Neto MAC, Pham VM, Garfield MK, Valenzuela JG. Exploring the sialome of the blood-sucking bug *Rhodnius prolixus*. *Insect Biochem. Mol. Biol.* 2004; 34:61–79. [PubMed: 14976983]
- Saitou N, Nei M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* 1987; 4:406–425. [PubMed: 3447015]
- Sant'Anna MR, Araujo JG, Pereira MH, Pesquero JL, Diotaiuti L, Lehane SM, Lehane MJ. Molecular cloning and sequencing of salivary gland-specific cDNAs of the blood-sucking bug *Triatoma brasiliensis* (Hemiptera: Reduviidae). *Insect Biochem. Mol. Biol.* 2002; 11:585–593. [PubMed: 12421416]
- Sarkis JJ, Guimares JA, Ribeiro JMC. Salivary apyrase of *Rhodnius prolixus*. Kinetics and purification. *Biochem. J.* 1986; 233:885–891. [PubMed: 3010945]
- Schofield, CJ. Biosystematics of the triatominae. In: Service, MW., editor. *Biosystematics of Haematophagous Insects*. Vol. 37. Oxford: Clarendon Press; 1988. p. 284–312. Systematics Association Special
- Schreiber MC, Karlo JC, Kovalick GE. A novel cDNA from *Drosophila* encoding a protein with similarity to mammalian cysteine-rich secretory proteins, wasp venom antigen 5, and plant group 1 pathogenesis-related proteins. *Gene.* 1997; 191:135–141. [PubMed: 9218711]
- Schultz J, Copley RR, Doerks T, Ponting CP, Bork P. SMART: a web-based tool for the study of genetically mobile domains. *Nucleic Acids Res.* 2000; 28:231–234. [PubMed: 10592234]
- Silveira AC, Feitosa VR, Borges R. Distribuio de triatomneos capturados no ambiente domiciliar, no perodo de 1975/83, Brasil. *Rev. Bras. Malariol. Doenas Trop.* 1984; 36:15–312.

- Sonnhammer EL, von Heijne G, Krogh A. A hidden Markov model for predicting transmembrane helices in protein sequences. *Proc. Int. Conf. Intell. Syst. Mol. Biol.* 1998; 6:175–182. [PubMed: 9783223]
- Stothard JR, Yamamoto Y, Cherchi A, Garcia AL, Valente SAS, Schofield CJ, Miles MA. A preliminary survey of mitochondrial sequence variation within triatomine bugs (Hemiptera: Reduviidae) using polymerase chain reaction-based strand conformational polymorphism (SSCP) analysis and direct sequencing. *Bull. Entomol. Res.* 1998; 88:553–560.
- Takac P, Nunn MA, Meszaros J, Pechanova O, Vrbjar N, Vlasakova P, Kozanek M, Kazimirova M, Hart G, Nuttall PA, Labuda M. Vasotab, a vasoactive peptide from horse fly *Hybomitra bimaculata* (Diptera, Tabanidae) salivary glands. *J. Exp. Biol.* 2006; 209:343–352. [PubMed: 16391356]
- Thompson JD, Higgins DG, Gibson TJ. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* 1994; 22:4673–4680. [PubMed: 7984417]
- Thomsen JS, Sonne M, Benfeldt E, Jensen SB, Serup J, Menne T. Experimental itch in sodium lauryl sulphate-inflamed and normal skin in humans: a randomized, double-blind, placebo-controlled study of histamine and other inducers of itch. *Br J Dermatol.* 2002; 146:792–800. [PubMed: 12000375]
- Valenzuela JG, Pham VM, Garfield MK, Francischetti IM, Ribeiro JMC. Toward a description of the sialome of the adult female mosquito *Aedes aegypti*. *Insect Biochem. Mol. Biol.* 2002; 32:1101–1122. [PubMed: 12213246]
- Valenzuela JG, Francischetti IM, Pham VM, Garfield MK, Ribeiro JM. Exploring the salivary gland transcriptome and proteome of the *Anopheles stephensi* mosquito. *Insect Biochem. Mol. Biol.* 2003; 33:717–732. [PubMed: 12826099]
- Zhu YC, Zeng F, Oppert B. Molecular cloning of trypsin-like cDNAs and comparison of proteinase activities in the salivary glands and gut of the tarnished plant bug *Lygus lineolaris* (Heteroptera: Miridae). *Insect Biochem. Mol. Biol.* 2003; 33:889–899. [PubMed: 12915180]

Fig.1.

(A) Strong cation exchange and (B) Strong anion exchange chromatography of 20 μ L of saliva from *T. brasiliensis* starving adult insects. The bars with number represent the fractions submitted to Edman degradation to obtain amino terminal sequence information.

Fig. 2. Phylogenetic tree of lipocalin sequences among triatomine species with bootstrap concordance higher than 85%, obtained by the Neighbour Joining method. The bar shows 20% divergence at the amino acid level. The numbers in the tree branches indicate the bootstrap value from 10,000 interactions. Figure legend: Red braches – lipocalins from *R. prolixus*; Green branches – lipocalins from only one *Triatoma* species; Blue branches - lipocalins from other *Triatoma* species; Black branches – lipocalins from *R. prolixus* and *Triatoma* species. *T. brasiliensis* sequences are indicated by TB. Other sequences are indicated by the first 3 letter of the genus followed by the first 2 letters of the species name followed by the Genbank accession number.

	10	20	30	40	50	60
TBQ-contig_558	CSDSCPLSYDPVCG	---	YNGEEHVRFNTMCELEKSE	---	---	---
TBQ-contig_789	CRDSCPLSYDPVRG	---	YNGEEYVRFNSKCELEKSE	---	---	---
TBQ-contig_132	CSSSCGFDEPVCG	---	YNGQEHVRFNNECELEKSE	---	---	---
TBQ-contig_262	CNTNCPDTYEPVCTQP	---	KEKEEFQTFKNPCELNRALQCNGGRNDDYT	---	---	---
TBQ-contig_394	CPRTCPSSTNYVCG	---	QKLNHQKTFDSDCEMEKDNLCGSG	---	---	---
TBQ-contig_54	CPFVCSNDYRPVCG	---	GRDTERRTFDNQCKIDEENDCGRS	---	---	---
Vasotab	CPRICTADYRPVCGT	---	PSGGRRSANRTFGNQCSLNAHN	---	---	---
Dipetalin	C	---	ECPRALHRVCG	---	---	---
INF2R	C	---	ECPRALHRVCG	---	---	---
rhodniin1	C	---	ACPHALHRVCG	---	---	---
TBQ-contig_259	C	---	DCDRMLFRVCG	---	---	---
	C	C	C	Y	N	C

Fig. 3.

(A) ClustalW alignment of predicted sequences from seven contigs that showed homology to kazal peptides with other kazal inhibitors - vasotab from *Hybomitra bimaculata* (gi94730670), INF2R from *T. infestans* (gi14211145), dipetalin from *Dipetalogaster maximus* (gi21465882), rhodniin1 from *R. prolixus* (gi1827578). Residues identical in more than four sequences are boxed and gaps are indicated by dashes. Amino acids that define the kazal family protease inhibitor domain are indicated below the sequences. (B) NJ phylogenetic tree showing the sequence distance between members of the group. The two most divergent contigs (TBQ-531 and 155) were excluded from the analysis.

Fig. 4.

ClustalW alignment of predicted sequences from four contigs that showed similarity to trypsin-like precursors with five trypsin-like sequences from *L. lineolaris* (Zhu et al., 2003). Amino acid gaps are indicated by dashes. Identical or highly similar residues among sequences are highlighted with a black or gray background. Functionally important residues are boxed. The most divergent contig (TBQ-692) was excluded from the analysis.

Table 1

Types of transcripts found in *Triatoma brasiliensis* salivary glands.

Types of transcripts	Clusters	Sequences	Sequences/Cluster	% Cluster	% Sequence
Secreted	370	1168	3.2	47.1	68.2
Housekeeping	192	305	1.6	24.4	17.8
Unknown	224	239	1.1	28.5	14.0
Total	786	1712			

Table 2
Functional classification of the housekeeping genes expressed in *Tritoma brasiliensis* salivary glands.

Types of transcripts	Clusters	Sequences	Sequences/Cluster	% Cluster	% Sequence
Protein synthesis machinery	53	117	2.2	27.6	38.4
Conserved	37	40	1.1	19.3	13.1
Metabolism, energy	30	38	1.3	15.6	12.5
Protein modification	12	29	2.4	6.3	9.5
Transport/Storage	9	18	2.0	4.7	5.9
Signal transduction	8	11	1.4	4.2	3.6
Cytoskeletal	9	10	1.1	4.7	3.3
Metabolism, oxidant	5	9	1.8	2.6	3.0
Transcription machinery	7	7	1.0	3.6	2.3
Protein export machinery	5	6	1.2	2.6	2.0
Metabolism, nucleotides	2	4	2.0	1.0	1.3
Proteasome machinery	3	4	1.3	1.6	1.3
Transcription factors	3	3	1.0	1.6	1.0
Metabolism, lipid	3	3	1.0	1.6	1.0
Nuclear regulation	2	2	1.0	1.0	0.7
Metabolism, amino acid	2	2	1.0	1.0	0.7
Metabolism, carbohydrate	2	2	1.0	1.0	0.7
Total	192	305			

Table 3

Classification of transcripts coding for putative secreted proteins in *Triatoma brasiliensis* salivary glands.

Types of transcripts	Clusters	Sequences	% Sequences
Lipocalins	341	1096	93.8
Kazal domain containing	9	26	2.2
Secreted serine protease	5	29	2.5
Others proteins	5	5	0.4
Similar to <i>Rhodnius</i> binding protein	2	2	0.2
Inositol phosphatase	2	4	0.3
Similar to <i>Rhodnius</i> MYS protein	1	1	0.1
Sulfatase - secreted?	1	1	0.1
Similar to <i>Culicoides</i> protein	1	1	0.1
Similar to major epididymal secreted protein	1	1	0.1
5'-nucleotide/apyrase	1	1	0.1
Antigen 5 protein	1	1	0.1
Total	370	1168	