ELSEVIER

Insect Biochemistry and Molecular Biology 34 (2004) 61–79

Insect Biochemistry and Molecular Biology

www.elsevier.com/locate/ibmb

Exploring the sialome of the blood-sucking bug *Rhodnius prolixus*

J.M.C. Ribeiro [a,*], J. Andersen [a], M.A.C. Silva-Neto [b], V.M. Pham [a], M.K. Garfield [c], J.G. Valenzuela [a]

[a] *Medical Entomology Section, Laboratory of Parasitic Diseases, National Institute of Allergy and Infectious Diseases, National Institutes of Health, 4 Center Drive, Room 4/126, MSC 0425, Bethesda, MD 20892-0425, USA*

[b] *Departamento de Bioquímica Médica, Instituto de Ciências Biomédicas, Centro de Ciências da Saúde, Universidade Federal do Rio de Janeiro, UFRJ, P.O. Box 68041, Cidade Universitária, Ilha do Fundão, Rio de Janeiro CEP 21941-590, RJ, Brazil*

[c] *Biological Resources Branch, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Twinbrook I, Room 1013, Bethesda, MD 20892, USA*

Received 8 July 2003; received in revised form 29 August 2003; accepted 2 September 2003

Abstract

Rhodnius prolixus is a Hemiptera that feeds exclusively on vertebrate blood in all life stages. Its salivary glands produce potent pharmacological substances that counteract host hemostasis, including anti-clotting, anti-platelet, and vasodilatory substances. To obtain a further insight into the salivary biochemical and pharmacological complexity of this insect, a cDNA library was randomly sequenced, and salivary gland homogenates were fractionated by HPLC to obtain aminoterminal sequences of abundantly expressed proteins. Results indicate a remarkable expansion of the lipocalin family in *Rhodnius* salivary glands, among other protein sequences described. A summary of 31 new full length proteins deducted from their mRNA sequence is described, including several new members of the nitrophorin, triabin, and pallidipin families. The electronic version of the complete tables is available at http://www.ncbi.nlm.nih.gov/projects/vectors/rhodnius_prolixus.
Published by Elsevier Ltd.

Keywords: Nitrophorin; Triabin; Hemostasis; Pallidipin; Lipocalin; Proteome; Transcriptome

1. Introduction

Rhodnius prolixus is a hematophagous Hemiptera that feeds exclusively on blood throughout its entire life (Buxton, 1930). This bug is also a main vector of Chagas' disease in the northern part of South America and in some areas of Central America (Dias et al., 2002). *Rhodnius* is also a good laboratory model for both Chagas' disease and basic biological studies because of the ease with which it is reared, comparatively short life cycle, and remarkable biochemical and physiological changes following a blood meal (Wigglesworth, 1972).

Like most blood-sucking arthropods that have been studied to date (Ribeiro and Francischetti, 2002), *R. prolixus* is well equipped to disarm the host's hemostatic machinery, triggered to prevent blood loss following tissue injury. When probing in their victim's skin for blood, *Rhodnius* inject a salivary cocktail that contains apyrase, an enzyme that hydrolyses ADP, a nucleotide released by injured cells and a potent inducer of platelet aggregation (Sarkis et al., 1986). Additionally, a lipocalin with very high affinity for ADP mops up any residual nucleotide (Francischetti et al., 2000). Serotonin and thromboxane A_2, which are potent vasoconstrictors released by platelets, are also neutralized (Ribeiro, 1982; Ribeiro and Sarkis, 1982). Several distinct nitrophorins (NP), which are heme-containing lipocalins, carry nitric oxide (NO) from the salivary glands to the injured tissue, causing vasodilation and further inhibiting platelet aggregation (Ribeiro et al., 1993; Champagne et al., 1995). These NP also bind histamine with high affinity (Ribeiro and Walker, 1994; Andersen et al., 1998), and one, NP2,

* Corresponding author. Tel.: +1-301-496-3066; fax: +1-301-402-4941.
E-mail address: jribeiro@nih.gov (J.M.C. Ribeiro).

0965-1748/$ - see front matter Published by Elsevier Ltd.
doi:10.1016/j.ibmb.2003.09.004

also inhibits the Xase complex of the intrinsic pathway of the blood coagulation cascade (Ribeiro et al., 1995; Zhang et al., 1998; Isawa et al., 2000). An additional lipocalin, biogenic amine binding protein (BABP), removes serotonin and adrenergic mediators of vasoconstriction (Andersen et al., 2003). Although the saliva of most blood-sucking arthropods contains at least one anti-clotting, one anti-platelet, and one vasodilatory substance, the molecular nature of these compounds is very diverse, even in species belonging to the same family (Ribeiro and Francischetti, 2002).

The molecular diversity of hematophagous insect saliva, which may have arisen as a consequence of convergent evolution and/or a fast rate of evolution in saliva-expressed genes, represents a rich field for the discovery of novel pharmacologically active compounds and for understanding the evolutionary mechanisms leading to the insect's adaptation to this feeding habit. Toward this end, we have recently started to describe the sialome (=set of RNA message + set of proteins found in salivary glands) of hematophagous insects and ticks (Francischetti et al., 2002b; Valenzuela et al., 2002b, c). These studies demonstrate that the sialome from hematophagous insects and ticks is more complex than previously expected and contains many proteins to which we cannot yet ascribe a function. Presently, we initiate description of the sialome of *R. prolixus*, which is very rich in expressing lipocalins among other proteins. Lipocalins serve different roles, primarily as carriers of small ligands in vertebrates and invertebrates (Flower et al., 2000), although some have enzymatic activity, such as the glutathione-independent prostaglandin D_2 synthase (Urade and Hayaishi, 2000). Lipocalins were not found in the saliva of other blood-sucking insects, but they occur in tick saliva (Paesen et al., 2000) serving an anti-histaminic and anti-serotonin role. Other proteins are also described in this work, which should help in our understanding of the evolution to blood feeding by insects.

2. Materials and methods

2.1. Rhodnius prolixus

R. prolixus were reared in the laboratory at 27 °C, 70% relative humidity, and 16:8 h light cycle. They were fed at three-week intervals using a jacketed artificial feeder containing heparinized rabbit blood kept at 38 °C with a water bath circulator. The bugs were kept in cages containing vertical strips of coarse filter paper of chromatographic purity. A PCR-based salivary gland cDNA library was made as described before (Francischetti et al., 2002a, b). The library titer was 0.5×10^6/ml and it was not further amplified. Salivary glands were obtained from Vth instar nymphs dissected at days 5, 7, 10 and 14 following the blood meal to obtain an average representation of the messages expressed while the gland is replenishing its contents (Nussenzveig et al., 1995). This larval instar was chosen because their salivary glands contain the largest amount of protein of any instar, including adults (Ribeiro, unpublished), and because it has been shown to contain all nitrophorin forms expressed in nymphal instars (Moreira et al., 2003). For chromatographic experiments, glands from starved Vth instar nymphs were dissected between 5 and 30 days after the molt, because they contain the full set of proteins required for the salivary gland function during the feeding process (Nussenzveig et al., 1995). Previously, it had been estimated that ~90% of the salivary protein in starving *Rhodnius* salivary glands is of a secretory nature, and is contained within the large bladder-like cavity of *Rhodnius* glands (Nussenzveig et al., 1995; Ribeiro and Garcia, 1980).

2.2. Chromatography

Chromatographic 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, USA). A CM4100 pump and a SM4100 dual wavelength detector (both from ThermoSeparation Products, Rivera Beach, FL, USA) were used. To elute the proteins of interest, the ion-exchange columns were submitted to 2 ml gradients of NaCl from 0 (solution A) to 1 M (solution B) for 60 min at a flow rate of 50 μl/min. 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. To produce reliable gradients of this small nature with regular HPLC pumps, we introduced a 2 ml loop after the pump followed by a three-way valve (port A, pump; port B, waste; port C, sample injector) before the sample injector. With the three-way valve connecting the pump to the waste, a gradient from 0% to 100% solution B was made with the flow at 1 ml/min for 2 min + time of dead volume, after which the pump switches to 50 μl/min. The times of dead volumes were measured previously by observing the time it took for the change in UV absorbance when water is immediately switched to an organic buffer solution at 1 ml/min. When the pump switches to 50 μl/min, the three-way valve is manually switched to connect the gradient loop to the injector sample loop, which was previously loaded with sample. During the run, the pump delivers 100% solution B. The eluate was monitored for UV absorption (280 nm) as well as for absorbance at 404 nm to monitor the elution of NP. Fractions of 1 min intervals (50 μl) were collected into polypropylene 96 well plates using a 203B fraction collector (Gilson, Middleton, WI, USA). Fractions of

interest had 40 μl removed and diluted with an equal volume of 20% methanol containing 0.4% trifluoroacetic acid (TFA) and were applied to a ProSorb cartridge (Perkin Elmer, Foster City, CA, USA) previously treated with 10 μl 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. To further characterize the complex protein mixture eluting after the NP peak on the SCX experiment (Fig. 2), we submitted 100 pairs of salivary glands to SCX in a larger (25 × 1 cm) column, and the region of interest was mixed 1:1 with saturated NH_4SO_4 and applied to a Phenyl-TSK column from BioRad (Hercules, CA, USA) equilibrated with 2 M NH_4SO_4 + 20 mM Hepes buffer pH 7.4 at 0.5 ml/min. A gradient to 20 mM Hepes in 60 min eluted the proteins. The eluate was monitored at 280 nm, and peaks of interest were applied to ProSorb cartridges as above.

2.3. Sequencing

Bioinformatics procedures were as by Francischetti et al. (2002b) and Valenzuela et al. (2002c) 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). Trans-membrane 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). Bootstrapping 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/vectors/rhodnius_prolixus.

3. Results

To obtain an insight on the salivary transcriptome of *R. prolixus*, we randomly sequenced 539 cDNA clones from a salivary gland cDNA library from this insect and organized these into 252 clusters of related sequences after assembly. Using the BLAST package of programs (Altschul et al., 1997), we compared the sequence of each cluster in the database with the non-redundant protein and nucleotide sets of the NCBI and the gene ontology database (Ashburner et al., 2000; Lewis et al., 2000; Hvidsten et al., 2001). 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. With this information, the clustered database was annotated and classified into three categories of clusters: S, those associated with possibly secreted products, H, those possibly associated with housekeeping functions, and U, those of unknown function. Accordingly, 114 cDNA clusters containing a total of 127 sequences (23.5% of the transcriptome) were classified as being associated with products of the housekeeping class. These clusters have an average of 1.14 sequences per cluster. This contrasts with the 74 clusters containing 345 sequences (64.0% of the transcriptome) classified as associated with secreted products and providing for an average of 4.66 sequences per cluster. These results of average cluster size are very significant ($P< 0.01$ χ^2 test), and were observed also in other transcriptome analysis of salivary glands in *Aedes aegypti*, *An. gambiae* and *Ixodes scapularis* (Francischetti et al., 2002b; Valenzuela et al., 2002b, c), where clusters of putatively secreted proteins were abundantly expressed. Finally, 67 sequences (12.4% of the transcriptome) in 64 clusters were classified as being associated with products of unknown function.

3.1. Description of clusters associated with transcripts having probably a housekeeping function

From the 114 clusters of cDNA sequences associated with probably housekeeping-related products, only a few have more than one sequence per cluster, with a maximum size of three sequences per cluster. Twenty-nine of these clusters code for products associated with the protein synthetic machinery, including rRNA, and various ribosomal proteins. Four transcripts code for proteins associated with secretory pathways and three for proteins associated with protein modification products including a sulfotransferase and a mannosyltransferase. Seventeen transcripts coded for enzymes associated with energy metabolism, four coded for enzymes associated to lipid metabolism, and one transcript coded for uroporphyrinogen decarboxylase, an enzyme involved in heme synthesis. Seven clusters were associated with cytoskeleton proteins such as annexin and tubulin. Six transcripts were associated with the proteasome machinery, and three others with

transporters, including two for a UDP-Gal transporter. Seventeen transcripts were associated with signal transduction products and eight with nuclear regulation. When combined, these 25 transcripts include different members of intracellular signaling cascades ranging from surface receptors and adapter proteins to enzymes and transcriptional factors. It is noteworthy that among the transcripts related to signal transduction four display homology with members of the GTPase superfamily (ras, rab, sar1/arf, ran) which are involved in gene expression, vesicle trafficking, nucleocytoplasmic transport and cytoskeletal organization (Takai et al., 2001). Therefore, a sophisticated system of regulation of protein synthesis, sorting and exocytosis must work during refilling of glands as depicted by the number and the diversity of transcripts found. Finally, eight transcripts code for proteins that are well conserved among eukaryotes, suggesting a fundamental role in cellular processes, although their function is unknown.

3.2. *Description of the clusters of transcripts in the salivary glands of* R. prolixus *probably associated with secretory products*

Remarkably, of the 74 clusters of transcripts possibly associated with secretory products, 62 code for proteins of the lipocalin family (Flower et al., 2000) (Table 1). Of these 62 lipocalin clusters, 27 code for proteins having similarity to NP. NP are NO-carrying proteins discovered in *Rhodnius* (Champagne et al., 1995) for which four protein sequences are known. Their amino acid sequences do not reveal similarity to lipocalins, but their crystal structure is typical of this large family of proteins (Andersen et al., 1997, 1998; Weichsel et al., 1998). NP account for the deep cherry color of *Rhodnius* salivary glands and may constitute half of the proteins in *Rhodnius* salivary glands (Wigglesworth, 1942; Champagne et al., 1995). We have also recently reported on a novel *Rhodnius* salivary protein similar to NP-2; this protein, however, does not bind heme and NO but rather biogenic amines such as adrenaline and serotonin (Andersen et al., 2003). This protein may be responsible for the anti-serotonin activity reported in *Rhodnius* saliva (Ribeiro, 1982).

Five of the 27 clusters of cDNA sequences shown in Table 1 represent the expected sequences for NP1, NP2, NP3, NP4, and BABP. The remaining 22 clusters code for proteins having diverse similarity to any of these five proteins. In a few instances, these clusters may represent truncated cDNA or alleles of the five known NP, but it appears that most are due to different gene products.

The remaining 35 clusters of lipocalin transcripts were found to be similar to salivary anti-hemostatic lipocalins from triatomine bugs previously described as RPAI (*Rhodnius* platelet aggregation inhibitor) and the *Triatoma* triabin (Noeske-Jungblut et al., 1995), procalin (Paddock et al., 2001), or pallidipin (Noeske-Jungblut et al., 1994). Thirty-four of the 35 clusters display the Pfam motif for triabin (Flower et al., 2000). These clusters include two perfect matches to anti-platelet salivary proteins previously described in *Rhodnius* as RPAI-1 and RPAI-2. The remaining clusters similar to RPAI produce various degrees of sequence similarities to RPAI-1 or RPAI-2 and may represent novel salivary proteins in *Rhodnius*. RPAI-1 and RPAI-2 inhibit platelet aggregation by strongly binding to adenosine nucleotides (Francischetti et al., 2000). ADP at the submicromolar concentrations normally found in the plasma decreases the threshold for collagen-induced platelet aggregation. At these ADP levels, salivary apyrase is not an efficient scavenger of ADP due to its K_m at >20 μM (Sarkis et al., 1986), but the salivary RPAI, with a nanomolar affinity for ADP, efficiently scavenges the nucleotide and raises the threshold concentration for collagen-induced aggregation (Francischetti et al., 2002a). Pallidipin produces a similar inhibition of platelets and may act by the same mechanism (Noeske-Jungblut et al., 1994; Haendler et al., 1996) despite claims of its activity on the collagen receptor. These novel proteins may, accordingly, act by removing mediators of hemostasis or as anti-clotting agents. Finally, a cluster with five sequences was not only found similar to RPPA-2 but also matching a lipocalin in the gene ontology database annotated as prostaglandin H_2 isomerase/PGD_2 synthase, suggesting that this protein product could act by transforming PGH_2 (normally produced by activated platelets) into the vasodilatory and anti-platelet PGD_2.

Several clusters matched *Triatoma pallidipennis* triabin. Triabin is a thrombin inhibitor, a function unrelated to its ability to bind small ligands (Noeske-Jungblut et al., 1995; Glusa et al., 1997), as is also the case with NP2, which, in addition to binding heme and NO, inhibits factor VIII in the clotting cascade. Although all anti-clotting activities from *Rhodnius* salivary glands may be explained by NP2 and no other inhibitor appears to block other sites of the clotting cascade other than in the extrinsic Xase complex (Ribeiro et al., 1995), the possibility of finding another clotting inhibitor in *Rhodnius* salivary glands remains a possibility. It is also interesting that several clusters of transcripts produce similar matches to both NP and triabin and have the NP, triabin, and lipocalin Pfam signatures. These sequences may represent evolutionary intermediate stages between the lipocalins and NP and between the lipocalins and triabins, most of which do not have the Pfam lipocalin signature.

Twelve of the 74 clusters associated with putative secreted proteins (Table 1) do not belong to the lipocalin family. Of these, eight represent sequences coding for proteins for which a protein family or function can

Table 1
Cluster of cDNA sequences coding for probably secreted proteins

Assembled contig	No. of sequences	Best match to NR protein database[a]	*E* value[b]	Best match to CDD database[c]	*E* value[b]	Comments	Edman sequence found[d]	Column-fraction[e]
Lipocalins of the nitrophorin (NP) family								
RP-contig_209	16	Biogenic amine-binding protein	1e−124	Nitrophorin	1e−017	BABP (6)	ASGCSTVDTVKD	C-8
RP-contig_113	1	Biogenic amine-binding protein	3e−009			BABP like-truncated clone		
RP-contig_181	2	Biogenic amine-binding protein	1e−011	Triabin	0.002	BABP-like		
RP-contig_89	1	Biogenic amine-binding protein	2e−031	Nitrophorin	2e−005	BABP-like		
RP-contig_248	1	Biogenic amine-binding protein	3e−039	Nitrophorin	2e−014	BABP-like 42% id (5)		
RP-contig_192	2	Biogenic amine-binding protein	2e−051	Nitrophorin	4e−009	BABP-like 60% id		
RP-contig_251	1	Biogenic amine-binding protein	1e−079	Nitrophorin	1e−011	BABP-like 69% id		
RP-contig_108	1	Biogenic amine-binding protein	2e−075	Nitrophorin	1e−007	BABP-like 75% id		
RP-contig_183	22	Biogenic amine-binding protein	1e−094	Nitrophorin	3e−013	BABP-like 78% id	ASGCLTVDTVKDFNKDNF	C-9
RP-contig_125	1	Biogenic amine-binding protein	7e−090	Nitrophorin	2e−013	BABP-like 78% id		
RP-contig_93	4	Biogenic amine-binding protein	1e−096	Nitrophorin	3e−015	BABP-like 81% id		
RP-contig_148	1	Biogenic amine-binding protein	3e−042	Nitrophorin	2e−004	BABP-like 85% id	ASGCLTVDTVKDFNKDNF	C-9
RP-contig_220	10	Nitrophorin 1	1e−117	Nitrophorin	4e−093	NP1	KCTKNALAQTGFNKDKYFNG	C-5
RP-contig_1	23	Nitrophorin 2 precursor	1e−113	Nitrophorin	1e−088	NP2		
RP-contig_174	2	Nitrophorin 2 precursor	1e−057	Nitrophorin	3e−034	NP2-like 59% id		
RP-contig_49	5	Nitrophorin 2 precursor (NP2)	1e−080	Nitrophorin	2e−066	NP2-like 71% id	EECSKNISPKSGLDKEKYYS	A-4
RP-contig_2	2	Nitrophorin 2 precursor (NP2)	1e−043	Nitrophorin	1e−031	NP2-like short		
RP-contig_234	1	Nitrophorin 3	4e−008	Triabin	1e−013	NP3-like		
RP-contig_82	4	Nitrophorin 3 precursor	4e−075	Nitrophorin	1e−056	NP3-like 67% id		
RP-contig_172	2	Nitrophorin-3	4e−030	Nitrophorin	4e−016	NP3-like 68% id		
RP-contig_242	8	Nitrophorin 3	3e−078	Nitrophorin	5e−044	NP3	DCSTNISPKKGLDKAKYFSG	C-2
RP-contig_171	31	Nitrophorin 4 precursor (NP4)	1e−115	Nitrophorin	4e−086	NP4	ACTKNAIAQTGFNKDKYFNG	C-5
RP-contig_4	1	Nitrophorin 4 precursor (NP4)	1e−097	Nitrophorin	5e−080	NP4-like 86% id	KCTQNAIAQTGFKKDQYFNG	C-4
RP-contig_85	1	Nitrophorin 4 precursor (NP4)	1e−097	Nitrophorin	5e−080	NP4-like 86% id		
RP-contig_115	52	Nitrophorin 4 precursor (NP4)	7e−097	Nitrophorin	1e−079	NP4-like 89% id		

(continued on next page)

Table 1 (*continued*)

Assembled contig	No. of sequences	Best match to NR protein database[a]	*E* value[b]	Best match to CDD database[c]	*E* value[b]	Comments	Edman sequence found[d]	Column-fraction[e]
RP-contig_3	30	Nitrophorin 4 precursor (NP4)	1e−104	Nitrophorin	2e−083	NP4-like 90% id	KCTQNAIAQTGFKKDQYFNG	C-4
RP-contig_170	5	Nitrophorin 4 precursor (NP4)	1e−106	Nitrophorin	1e−087	NP4-like 91% id	ACTKNAIAQTGFNKDKYFNG	C-5 C-7
Lipocalins of the Triabin family								
RP-contig_60	5	Pallidipin precursor	9e−017	Triabin	4e−015	Pallidipin-like		
RP-contig_168	2	Pallidipin precursor	1e−006			Pallidipin-like		
RP-contig_106	1	Pallidipin precursor	0.001	Triabin	0.005	Pallidipin-like		
RP-contig_210	1	Pallidipin-2 precursor	1e−006	Triabin	1e−006	Pallidipin-like		
RP-contig_81	1	Procalin	2e−008	Triabin	1e−014	Procalin like		
RP-contig_5	7	RPAI	6e−098	Triabin	5e−053	RPPAI1	ANPPKMPTGCKDLNSKAV	A-8
RP-contig_147	1	RPAI	4e−006	Triabin	2e−008	RPAI1-like	IQCDCESVEAAGNDAGFFKG	A-8
RP-contig_117	1	RPAI	2e−006	Triabin	9e−006	RPAI1-like		
RP-contig_180	2	RPAI	4e−010	Triabin	3e−014	RPAI1-like 28%id	MQCDCESVEAAGNDGEFFKG	A-7 A-9
RP-contig_56	1	RPAI	4e−010	Triabin	2e−013	RPAI1-like 28%id	VQYDCESVEAAGNDEGF	A-5
RP-contig_38	5	RPAI	2e−011	Triabin	8e−015	RPAI1-like 29%id	MQCDCESVEAAGNDGEFFKG	A-7 A-10
RP-contig_198	1	RPAI	1e−011	Triabin	7e−015	RPAI1-like 29%id		
RP-contig_199	1	RPAI	3e−016	Triabin	7e−015	RPAI1-like 30%id		
RP-contig_6	1	RPAI	3e−016	Triabin	3e−016	RPAI1-like 32%id		
RP-contig_137	1	RPAI	1e−017	Triabin	1e−017	RPAI1-like 32%id		
RP-contig_245	1	RPAI	6e−021	Triabin	2e−019	RPAI1-like 35%id		
RP-contig_176	2	RPAI	6e−050	Triabin	2e−035	RPAI1-like 54%id		
RP-contig_71	5	RPAI	2e−050	Triabin	7e−035	RPAI1-like 56%id	ATVPKMPQGCADVHNKAVSD	C-12
RP-contig_28	1	RPAI	6e−049	Triabin	5e−026	RPAI1-like 57%id		
RP-contig_138	3	RPAI	6e−044	Triabin	7e−031	RPAI1-like 58%id	ATVPKMPDGCADVHNKAVSD	C-12
RP-contig_9	1	RPAI	8e−035	Triabin	2e−023	RPAI1-like 58%id		
RP-contig_177	2	RPAI	2e−040	Triabin	5e−020	RPAI1-like 67%id	ATVPKMPDGCADVHNKAVSD	C-12
RP-contig_149	3	RPAI	1e−102	Triabin	1e−055	RPAI2		
RP-contig_160	2	RPAI	7e−015	Triabin	7e−008	RPAI2-like		
RP-contig_231	8	RPAI	9e−058	Triabin	1e−040	RPAI2-like 64%id		
RP-contig_27	5	RPAI	4e−068	Triabin	4e−045	RPAI2-like 71%id—PGD_2		
RP-contig_165	3	RPAI	2e−069	Triabin	3e−039	RPAI2-like 73%id		
RP-contig_16	6	RPAI	3e−064	Triabin	3e−048	RPAI2-like 74%id		
RP-contig_98	1	Triabin precursor	2e−007	Triabin	5e−011	Triabin-like		
RP-contig_57	1	Triabin precursor	7e−011	Triabin	2e−015	Triabin-like		
RP-contig_249	1	Triabin precursor	2e−011	Triabin	5e−017	Triabin-like		
RP-contig_185	1	Triabin precursor	6e−011	Triabin	8e−017	Triabin-like		
RP-contig_196	1	Triabin precursor	6e−011	Triabin	8e−017	Triabin-like		
RP-contig_184	1	Triabin precursor	1e−010	Triabin	2e−014	Triabin-like		
RP-contig_145	1	Triabin precursor	3e−012	Triabin	9e−018	Triabin-like	ITNGECDAVTAQENIDEFFT	A-3

Table 1 (*continued*)

Assembled contig	No. of sequences	Best match to NR protein database[a]	*E* value[b]	Best match to CDD database[c]	*E* value[b]	Comments	Edman sequence found[d]	Column-fraction[e]
SCP/antigen 5 family								
RP-contig_188	2	Venom allergen 5 (Antigen 5)	5e−015	SCP	9e−012	Antigen 5	WSDSDQNLRVVRNSCNSP	C-15
RP-contig_173	2	CG8483-PA (*Drosophila*)	3e−023	SCP	1e−012	Antigen 5		
Secreted basic protein								
RP-contig_224	1	gene_id:F1D9.26	9e−026	Protamine	2e−020	Basic secreted protein		
Inositol phosphatase								
RP-contig_197	2	Apyrase	1e−085	IPPc	4e−031	Inositol phosphatase		
RP-contig_230	1	Apyrase	2e−043	IPPc	9e−016	Inositol phosphatase		
Cytochrome P450								
RP-contig_104	3	Cytochrome P450	4e−020	p450	0.002	Cytochrome p450		
Mucin								
RP-contig_159	1	CG16707-PA (*Drosophila*)	2e−020	Tryp_mucin	4e−008	Mucin		
Lectin								
RP-contig_76	1	AgCP14122 (*An gambiae*)	9e−063	Lectin_leg-like	1e−052	Mannose-binding lectin		
Melibiase								
RP-contig_103	1	CG7997-PA (*Drosophila*)	5e−057	Melibiase	1e−034	Melibiase		
Unknown family								
RP-contig_186	2	Hypothetical protein XP_207739	1e−011	Caldesmon	2e−006	EK rich protein	DDANEEGAEDGTQ	H-8
RP_contig_21	1	Hypothetical protein Y50E8A.i	0.026			Low complexity		
RP-contig_195	18					Unknown	LDEEEIDNCEDGPGYRT	H-7

[a] Blastx performed against the non-redundant (NR) protein database of the NCBI as of 2/12/2003.
[b] Probability of the match occuring by chance.
[c] RPSblast performed against the Conserved Domains Database (CDD) from the NCBI.
[d] Sequence found by Edman degradation (underlined) matching predicted sequence in cluster.
[e] Column fractions of Edman degradation results. A, strong anion exchange chromatography (Fig. 1); C, cation exchange chromatography (Fig. 2); H, hydrophobic interaction chromatography (Fig. 3). Numbers refer to region of the chromatogram, as indicated in the figure bars.

be tentatively assigned. These include two clusters coding for proteins with high similarity to the antigen 5 family, which is a widespread extracellular family of proteins found in animals and plants (Schreiber et al., 1997), most of unknown function. Members of this protein family are found in salivary glands of sand flies, mosquitoes, and tsetse (Charlab et al., 1999; Li et al., 2001; Valenzuela et al., 2001, 2002c; Francischetti et al., 2002b) and are here described for the first time in the salivary glands of a Hemiptera. One cluster contains a sequence coding for a low complexity basic protein with a clear signal peptide, indicative of secretion. This basic putative protein yields a Pfam motif of protamines and histones. Two clusters produce substantial similarities to a protein previously deposited in the NCBI database as *Rhodnius* apyrase, because it purified together with two other lipocalins when attempts were made to isolate *Rhodnius* salivary apyrase (Champagne and Ribeiro, unpublished). The two putative proteins have strong similarities to inositol polyphosphate 5-phosphatase and were thought to be responsible for the apyrase activity of saliva. We have now expressed one of these clones and will report elsewhere that it indeed contains inositol phosphatase, but no apyrase activity (Andersen, Francischetti and Ribeiro, manuscript in preparation). Another cluster codes for a putative protein with strong similarities to insect cytochrome P450 proteins, which are enzymes that oxidize a very broad range of substrates including insecticides, allochemicals, and hormones. It is possible that this secreted P450 enzyme participates in the oxidation of arachidonic acid metabolites or in oxidizing serotonin, as does the salivary peroxidase of anophelines (Ribeiro and Nussenzveig, 1993). A low complexity protein with a Pfam match to mucins was coded by another cluster. This protein could help to lubricate the feeding canal of *Rhodnius* or interact with host matrix proteins. Finally, two sequences coding for carbohydrate-related functions were found. One sequence coded for a protein with very strong similarity to mannose-binding lectins of insects and mammals. Hemagglutinins in the salivary glands of anophelines are common (Metcalf, 1945; Gooding, 1972) but have not been found previously in *Rhodnius* (Gregorio and Ratcliffe, 1991). Their role in feeding is not clear, but they may play a role in insect immunity, as was proposed for several immune-related salivary proteins found in mosquitoes (Francischetti et al., 2002b; Valenzuela et al., 2002c). Alternatively, this lectin could be confined to the endoplasmic reticulum or the Golgi apparatus, and indeed, it has similarities to proteins annotated as such. A second sequence, with substantial similarity to melibiase, a sugar hydrolyzing enzyme, was also found.

Finally, three clusters have transcripts coding for possibly secreted proteins, but their family or function is unknown. These clusters code for proteins of low complexity, and include an abundant cluster having 18 transcripts (Table 1).

3.3. *Preliminary characterization of the salivary proteome of* R. prolixus

In our previous work (Francischetti et al., 2002b; Ribeiro and Francischetti, 2002; Valenzuela et al., 2002b, c), we have resorted to 1D sodium dodecyl sulfate polyacrylamide electrophoresis (SDS-PAGE) to separate the salivary proteins of hematophagous insects and ticks. Most *Rhodnius* salivary proteins, however, are lipocalins having 18–22 kDa in molecular mass, leading to poor separation of the proteins. We now utilize ion-exchange chromatography to separate the proteins. Five pairs of salivary glands from Vth instar nymphs (~0.5 mg protein) were submitted to anion exchange chromatography, and selected UV absorbing peaks were applied to PVDF membranes using a Prosorb cartridge (see Materials and methods for details) (Fig. 1). The non-retained portion of this anion exchange chromatogram was acidified to pH 5.0 and submitted to cation exchange column chromatography. Selected protein peaks were applied to PVDF membranes to obtain amino terminal amino acid sequence by Edman degradation (Fig. 2). Because we observed considerable protein complexity of the post-NP peak in the cation exchange chromatogram (NP absorb at 404 nm), we submitted this region (indicated by a bar in Fig. 3A) to hydrophobic interaction chromatography (HIC) and adsorbed selected peaks to a PVDF membrane to obtain Edman degradation information. In most cases, Edman degradation results yielded multiple amino acids per sequence cycle, but these could be deconvoluted using an in-house program that compared all possible sequences obtained by Edman degradation with all possible protein translation products deduced from the cDNA library. This program has previously been described in more detail (Valenzuela et al., 2002c). Table 1 summarizes the results obtained from these experiments when separating *R. prolixus* proteins into 26 fractions by ion-exchange chromatography (Figs. 1 and 2) and into an additional 9 fractions by HIC (Fig. 3). These 35 Edman degradation experiments allowed identification of 22 cDNA clusters matching the observed sequences. In some cases, no unambiguous assignment to a cDNA cluster could be made. For example, the sequence KXTQNAIAQTGF-KKDQYFNG could be derived either from NP4 (where X matches a Cys) or from a cluster coding for a protein with at least 86% identity to NP4 (Table 1). The amino terminal sequences of most of the already known salivary proteins from *R. prolixus* were found in these chromatographic experiments as follows: NP1, NP3, NP4, BABP (which was actually identified during

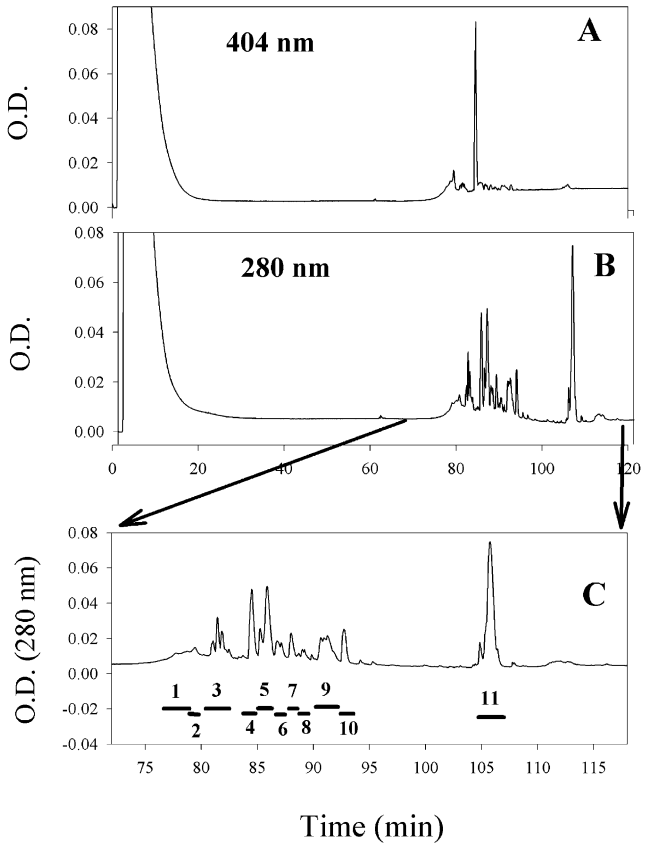

Fig. 1. Strong anion exchange chromatography of homogenized *R. prolixus* salivary glands. Five pairs of salivary glands were applied to the column, which was eluted as indicated in Materials and methods. (A) The column effluent was monitored at 404 nm to show retention of heme proteins. (B) Column effluent monitored at 280 nm. (C) Expansion of the chromatogram in (B), indicated by the arrows. The bars with numbers above or below indicate the fractions used to obtain amino terminal sequence information by Edman's degradation.

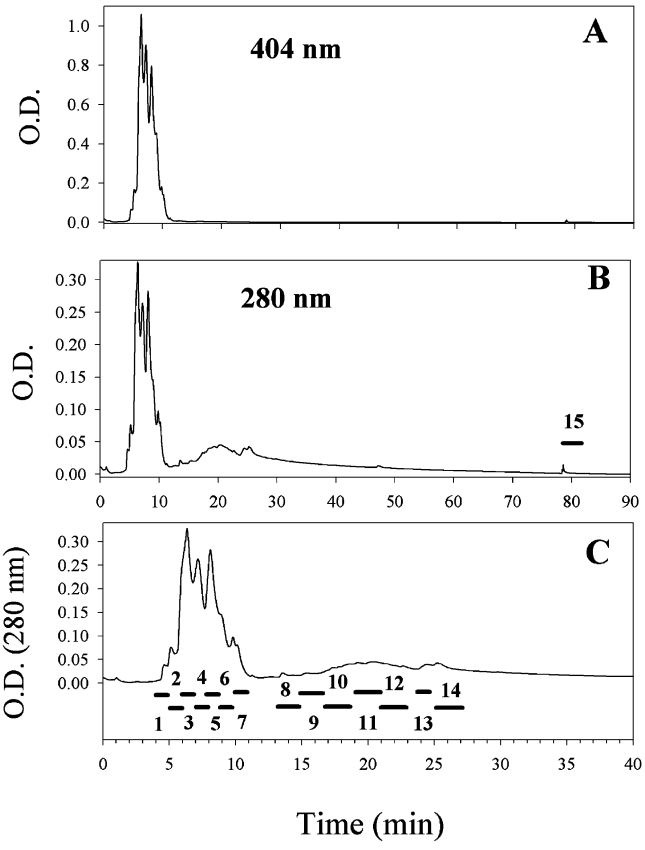

Fig. 2. Strong cation exchange from the void volume (first 20 min) of the strong anion exchange chromatogram shown in Fig. 1. (A) The column effluent was monitored at 404 nm to show retention of heme proteins. (B) Column effluent monitored at 280 nm. (C) Expansion of the chromatogram in (B). (B,C) The bars with numbers above or below indicate the fractions used to obtain amino terminal sequence information by Edman's degradation.

the course of this work), and RPAI-1. Among the novel sequences identified by Edman degradation are novel NP, triabin-like molecules, RPAI-1-like proteins, a protein member of the antigen 5 family, and two other protein sequences of unknown function. Full length sequence of the cDNA from these clusters may help to identify these proteins. We accordingly proceeded to obtain full length sequence information for the clusters possibly associated with the Edman sequences, as well as of other clusters that might be involved in blood feeding.

3.4. *Characterization of 31 novel full length cDNA clones from* R. prolixus *salivary glands*

3.4.1. *Lipocalin peptides*

Six of the 30 novel *Rhodnius* full length sequences code for proteins with a Pfam NP domain (Table 2) with significant similarities to NP and BABP. These predicted polypeptides have all a signal peptide indicative of secretion. Alignment of these novel proteins with the four NP and the BABP is shown in Fig. 4A. Notice that two of the NP are shorter than the other NP proteins, one of which has a polylysine tail that is not an artifact due to a missing stop codon; four independent sequences confirmed the final sequence. A phylogenetic tree (Fig. 4B) indicates four robust clades for these proteins consisting of the sole short NP protein containing the polylysine tail (RPSNP3A); the clade containing BABP with three novel putative proteins (RPNP3B, RPNP4B, and RPNP1A); the clade containing NP1, NP4, and the novel RPNP4A; and the clade containing NP2, NP3, and the short NP2-like protein named RPSNP2A. Notice that most of this family has four conserved cysteines, with the exception of the short NP (Fig. 4A). Edman degradation products of the chromatographic experiments (shown in Figs. 1–3) had sequences matching three of the six novel NP, as indicated in Table 2.

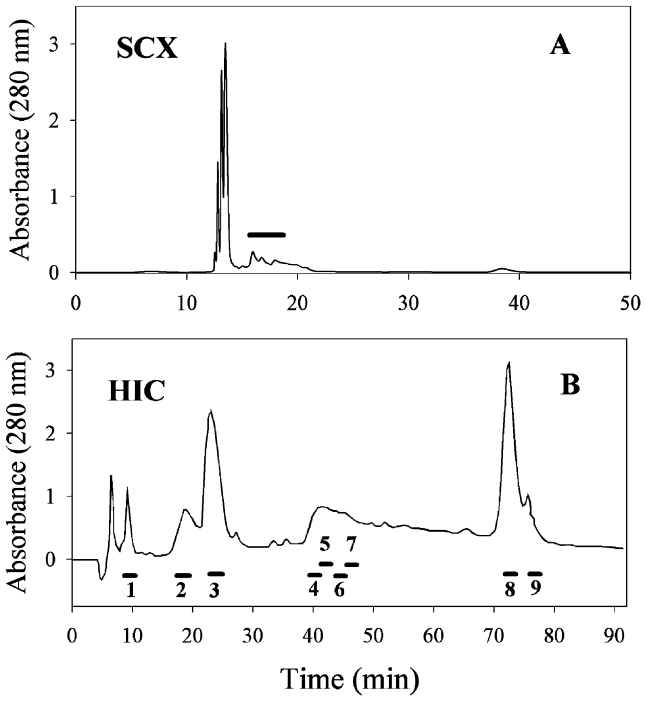

Fig. 3. (A) Strong cation exchange chromatography of 100 pairs of homogenized salivary glands of *R. prolixus*. The bar in (A) represents the fractions submitted to hydrophobic interaction chromatography shown in (B). The bars in (B) represent fractions submitted to Edman degradation to obtain amino terminal sequence information.

Four novel cDNA sequences match proteins previously described as *Rhodnius* platelet aggregation inhibitor proteins (RPAI) (Francischetti et al., 2000). These four sequences have a clear signal peptide indicative of secretion. RPAI-1 has been shown (Francischetti et al., 2000, 2002a) to bind adenosine derivatives and inhibit platelet aggregation. Except for RPAI-6, which has five cysteines, the remaining sequences have six conserved cysteines indicative of three disulfide bonds (Fig. 5A). The phylogenetic tree in Fig. 5B indicates that this family of proteins probably derives from ancient gene duplication events followed by divergence. The role of these novel proteins is not known. They could bind adenosine, as previously demonstrated for RPAI-1, or bind to other pro-hemostatic/inflammatory substances.

Triabin is an anti-thrombin peptide obtained from *T. pallidipennis* salivary glands (Noeske-Jungblut et al., 1995; Glusa et al., 1997). Four full length cDNA clones have similarity to triabin (25–28% identity, 40–50% similarity) and weaker similarity to RPAI and NP (Fig. 6A) (Table 2). These proteins have a signal peptide indicative of secretion. The amino terminal sequence of Rptriab2 was found in one of the anion exchange chromatographic fractions (Table 2). These protein sequences have six conserved cysteines except for RPTriab3 and RPTriab4, which have only five of the conserved cysteines; however, these last two proteins have in common a significant lysine repeat in their carboxyterminal region, as did one of the short NP (Fig. 4A). Other conserved domains are shown in Fig. 6A. Although the anterior midgut of *Rhodnius* has significant anti-thrombin activity (Hellmann and Hawkins, 1965), no such activity is found in pure salivary gland homogenates. The function of *Rhodnius* triabins remains to be investigated.

Two full length cDNA sequences predict proteins with similarities to a salivary *Triatoma* allergen named procalin (Fig. 7). The amino terminal sequence of RPPROCAL1 was found by Edman degradation in fractions of the anion exchange column experiment (Table 2). The role of these proteins in feeding is unknown, but they may be secreted, as indicated by the presence of a signal peptide.

Pallidipin is an inhibitor of collagen-induced platelet aggregation found in the salivary glands of *T. pallidipennis* (Noeske-Jungblut et al., 1994). Two full length cDNA clones from *Rhodnius* with similarities to pallidipin were accordingly named RPPAL-1 and RPPAL2. They have six conserved cysteines and a few other conserved amino acid residues (Fig. 8). Both have a signal peptide indicative of secretion. The role of these putative *Rhodnius* proteins in feeding is unknown.

3.4.2. *Low complexity proteins*

Two predicted protein sequences are almost identical in their first 34 amino acids but diverge abruptly into low complexity sequences containing repeats of glutamic acid and glycine or lysine repeats yielding either acidic or basic polypeptides named RPGE and RPPK (Fig. 9A). The nucleotide sequences reveal that RPGE has AGG repeats, while RPPK has AAG repeats in its message (Fig. 9B). These types of nucleotide repeats have been found previously to be associated with high mutation rates, possibly associated with polymerase slippage (Debrauwere et al., 1997; Wilson et al., 1998). The amino terminal sequence DDANEEGAEDGTQG was found by Edman degradation of one of the fractions in the anion exchange column experiment (Table 2), but it cannot distinguish between the two predicted polypeptides. This amino terminal sequence starts after the predicted cleavage of the signal peptide. Finding this amino terminal sequence in a well retained region of an anion exchange column suggests that the actual sequence derived from RPGE, which is quite acidic (Table 2). The function of these proteins is unknown, but they might play a role in adhesion to extracellular matrix or to cell membranes, as postulated for the low complexity proteins of the Stevor, Rifin and Var families from *Plasmodium falciparum* (Gardner et al., 2002). Glycine and glutamic acid-rich proteins have also been described in *Aedes* (Simons and Peng, 2001; Valenzuela et al., 2002c) and *Anopheles*

Table 2
Summary of 31 novel full length cDNA sequences from *Rhodnius prolixus* salivary glands

Name	Best match to NR protein database[a]	*E* values[b]	Best match to the CDD[c]	*E* values[b]	Comments	MW[d]	pI[e]	SigP[f]	MW mat[g]	pI-mat[h]	Edman degradation product[i]	Column fraction[j]
Sequences similar to previously described triatome salivary proteins of the lipocalin family												
Similar to Rhodnius nitrophorins												
RPNP3B	Nitrophorin-3	4e−008	Nitrophorin	3e−010	Similar to NP3	20.888	8.92	20–21	18.75	8.9		
RPNP4A	Nitrophorin-4	1e−104	Nitrophorin	1e−074	Similar to NP4	22.528	8.64	23–24	20.28	6.38	ACTKNAIAQTGFNKDKYFNG	C-3 C-4
RPSNP2A	Nitrophorin-2	1e−062	Nitrophorin	2e−038	Short NP2-K rich carboxyterminus	14.876	9.73	23–24	12.5	9.6	DCSTNISPKQGLDKAKYFSG	C-2
RPSNP3A	Nitrophorin-3	2e−029	Nitrophorin	9e−015	Short NP3	9.5	8.81	22–23	7.201	7.52		
RPNP4B	Biogenic amine-binding protein	8e−043	Nitrophorin	6e−016	Intermediate between NP3 and NP1	27.666	6.2	19–20	25.66	5.3		
RPNP1A	Biogenic amine-binding protein	9e−095	Nitrophorin	4e−011	Similar to NP1	24.45	5.05	21–22	22.23	7.33	ASGCLTVDTVKDFNKDNF	C-9
Similar to Rhodnius platelet aggregation inhibitor proteins (RPPA)												
RPAI-3	RPAL-1	4e−049			Similar to RPAI-1	19.468	9.56	16–17	17.71	9.64		
RPAI-4	RPAL-2	8e−059			Similar to RPAI-2	18.692	6.15	17–18	16.89	8.96		
RPAI-5	RPAL-1	9e−012	Nitrophorin	0.007	Similar to RPAI-1	20.249	6.03	20–21	17.89	5.22		
RPAI-6	RPAL-1	1e−020	Lipocalin	0.013	Simillar to RPAI-1	18.825	9.77	18–19	16.84	9.85		
RPAI-7	RPAL-1	2e−44	Triabin	8e−30	Simillar to RPAI-1	16.856	9.04	18–19	15.05	8.98		
Similar to Triatoma triabin												
RpTriab1	Triabin	3e−013			Similar to triabin	20.459	9.42	21–22	18.25	9.49		
RpTriab2	Triabin	2e−012	Nitrophorin	6e−009	Similar to triabin	19.451	5.95	20–21	17.15	5.22	ITNGECDAVTAQENIDEFFT	A-3
RpTriab3	Triabin	4e−011	Nitrophorin	0.003	Similar to triabin-K rich carboxyterminus	18.701	9.76	21–22	16.48	9.82		
RpTriab4	Triabin	1e−011	Nitrophorin	1e−006	Similar to triabin-K rich carboxyterminus	18.4	9.52	18–19	16.49	9.64		
Similar to Triatoma pallidipin												
RPPAL1	Pallidipin precursor	5e−017			Similar to pallidipin	21.472	8.95	18–19	19.57	8.95		
RPPAL2	Pallidipin-2	2e−009	Nitrophorin	0.002	Similar to pallidipin 2	26.977	6.16	18–19	24.93	7.54		
Similar to Triatoma procalin												
RPPROCAL1	Procallin (*Triatoma* protracta)	6e−004			Similar to procalin	12.019	5.46	20–21	9.654	4.88	MQCDCESVEAAGND	A-10
RPPROCAL2	Procallin (*Triatoma* protracta)	3e−009	Nitrophorin	2e−007	Similar to procalin	21.984	6.64	18–19	20.07	7.47		
Other probably secreted proteins												
RPGE	Retinitis pigmentosa GTPase	6e−014	Daxx	4e−008	GE rich protein	12.26	3.71	19–20	10.29	3.71	DDANEEGAEDGTQG	A-11
RPPK	gene_id:F1D9.26-unknown	2e−027	P53	2e−004	Lysine rich protein	16.242	10.51	19–20	14.22	10.54	DDANEEGAEDGTQG	A-11
RPMBL	agCP14122	2e−085			Mannose-binding lactin maybe endosomal	28.598	5.67	19–20	26.51	5.57		

(continued on next page)

Table 2 (*continued*)

Name	Best match to NR protein database[a]	*E* values[b]	Best match to the CDD[c]	*E* values[b]	Comments	MW[d]	pI[e]	SigP[f]	MW mat[g]	pI-mat[h]	Edman degradation product[i]	Column fraction[j]
RPSAG5	LD39025p (*D. melanogaster*)	5e−032	SCP	1e−030	Antigen 5 related protein	28.267	9.43	26–27	25.43	9.3		
RPSOBP	Odorant-binding protein	1e−006		Weak similarity to odorant binding	Weak similarity to odorant binding proteins	15.946	9.23	19–20	13.92	9.38		
RPSP450	Cytochrome P450 enzyme	5e−098	p450	2e−078	Salivary p450 enzyme	59.4	9.16	20–21	57.26	9.18		
RPMYS1					Novel protein	15.975	5.84	16–17	14.3	8.6	LDEEEIDNCEDGPGYRT	H-7
RPMYS2	RE71014p	4e−006			Weak similarity to *Drosophila*CG7781	18.756	5.02	21–22	16.5	4.64		
RPMYS3	Toxin secretion ATP-binding	0.099			Weak similarity to *Vibrio cholera* ATP binding protein	22.126	4.94	21–22	19.77	4.73		
Probably housekeeping proteins												
RpUGALT	CG5802-PA	1e−070			UDP-Gal transporter	34.722	9.54					
RpMapmod	CG5784-PA	2e−073			Homologue of Mapmodulin	27.812	4.23					
RPLET	Unknown (MGC:53950)	1e−031			Similar to leptin receptor overlapping transcript	14.75	4.75	52–53	9.141	4.44		

[a] Blasix performed against the non-redundant (NR) protein database of the NCBI as of 2/12/2003.
[b] Probability of the match occurring by chance.
[c] RPSblast performed against the Conserved Domains Database (CDD) from the NCBI.
[d] Predicted molecular weight of the protein.
[e] Predicted pI of the protein.
[f] Predicted cleavage site of the signal peptide by the SignalP server.
[g] Predicted molecular weight of the mature protein.
[h] Predicted pI of the mature protein.
[i] Sequence found by Edman degradation (underlined) matching predicted sequence in cluster.
[j] Column fractions of Edman degradation results. A, strong anion exchange chromatography (Fig. 1). C, cation exchange chromatography (Fig. 2). H, hydrophobic interaction chromatography (Fig. 3). Numbers refer to region of the chromatogram, as indicated in the figure bars.

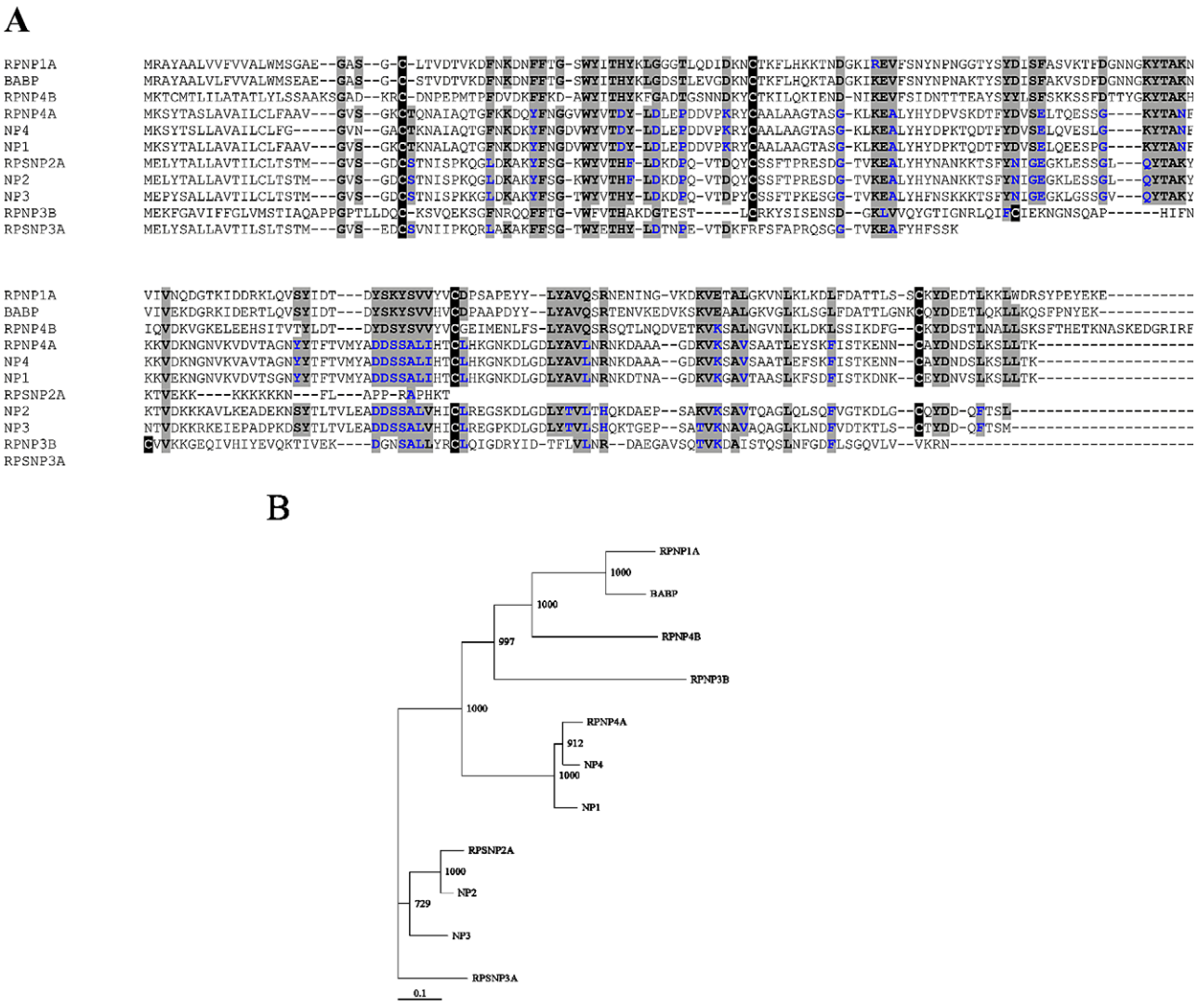

Fig. 4. The salivary nitrophorin (NP) family of *R. prolixus*. (A) Clustal alignment of novel NP with known NP (NP1, NP2, NP3 and NP4), and with BABP (*b*iogenic *a*mine-*b*inding *p*rotein). The signal peptide indicative of secretion is shown in gray background at the beginning of the sequences. Cysteines are shown in white with black background. Conserved residues are shown in black with gray background or in blue with gray background. The signal peptide is indicated in gray background at the beginning of sequences. (B) Phylogenetic tree showing the sequence distance relationships between members of the family. The bar shows 10% divergence at the amino acid level. The numbers indicate the bootstrap value from 1000 iterations (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.).

(Francischetti et al., 2002b) salivary glands, but their function is unknown.

3.4.3. *Other putatively secreted proteins*

Eight additional full length cDNA clones code for various proteins that have a clear signal peptide indicative of secretion (Table 2). The predicted protein sequences include RPMBL, which has similarities to mannose-binding lectins including endoplasmic reticulum-located lectins. In this latter case, this is indicative of a protein with a housekeeping function. Previously, we have not found hemagglutinating activity in *Rhodnius* salivary homogenates (unpublished), indicating either that this protein is expressed in low amounts or that it has no sugar-binding activity or that it is monovalent and thus does not cross-link with red blood cells. RPSAG5 is a protein with similarities to extracellular proteins of the antigen 5 family, ubiquitously found in the salivary glands of hematophagous Diptera (Valenzuela et al., 2002c) and in other insects (Schreiber et al., 1997). It has no known function. It is here described for the first time in a Hemiptera. RPSOBP has similarities to several proteins of the odorant-binding protein (OBP) family found in the mosquito *An. gambiae* and in the moth *Manduca sexta*. Of interest, the D7 family of proteins found in blood feeding Diptera (Valenzuela et al., 2002a) is a distant member of the OBP superfamily (Hekmat-Scafe et al., 2000). The

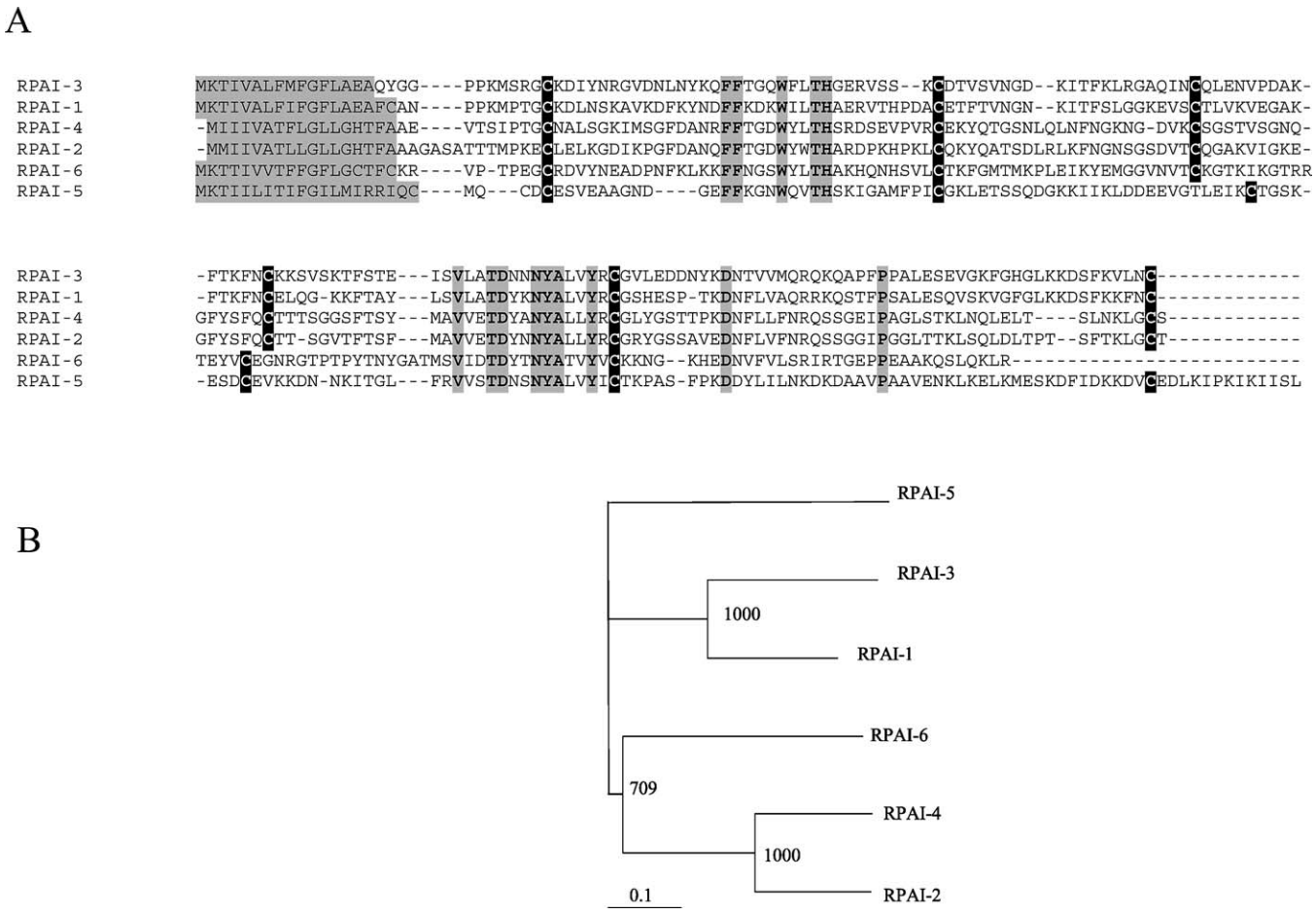

Fig. 5. The salivary RPAI family of proteins in *Rhodnius*. RPAI-1 and RPAI-2 were previously reported *Rhodnius* platelet *a*ggregation *i*nhibitors. For other details, see Fig. 4A.

role of these proteins in blood feeding is unknown except for Hamadarin, which inhibits factor XII of the clotting/kininase cascade (Isawa et al., 2002). RPSP450 is a cytochrome p450 of the CYP4 family, which includes enzyme-hydroxylating arachidonic acid derivatives having vasodilatory and immunomodulatory function (Simpson, 1997). The secretory fate, substrate specificity, and function of this cytochrome remain to be investigated. Finally, three other protein sequences were found with no significant similarities to known proteins. These were named RPMYS-1, RPMYS-2, and RPMYS-3. The amino terminal sequence predicted for the mature RPMYS-1 protein was found in one of the fractions of the HIC experiment (Table 2).

3.4.4. Protein sequences associated with probable housekeeping functions

Three protein sequences predicted by full length cDNA clones are indicative of a housekeeping function: RPUGALT, an UDP-galactose transporter without a signal peptide and having eight predicted intra-membrane helices by the TMHMM program (Sonnhammer et al., 1998); RPMAPMOD, the homologue of *Drosophila* mapmodulin; and RPLET, the *Rhodnius* homologue of the leptin receptor overlapping transcript (Huang et al., 2001). RPLET has a signal peptide indicative of secretion; it has four predicted intra-membrane helices, indicating it to be a membrane protein probably associated with the leptin receptor. Leptin is a peptide hormone associated with feeding behavior, and the finding of the receptor for this peptide in *Rhodnius* salivary glands is of interest (Saper et al., 2002), as it may be related to regulation of saliva secretion.

4. Discussion

A remarkable finding in the present work was the large expansion of the lipocalin family of proteins observed to be transcribed and, in several cases, demonstrated to be expressed (Tables 1 and 2), in the salivary glands of *Rhodnius*. Except for the NP, which appear to be unique to the *Rhodnius* genus, lipocalins similar to salivary proteins of North American species

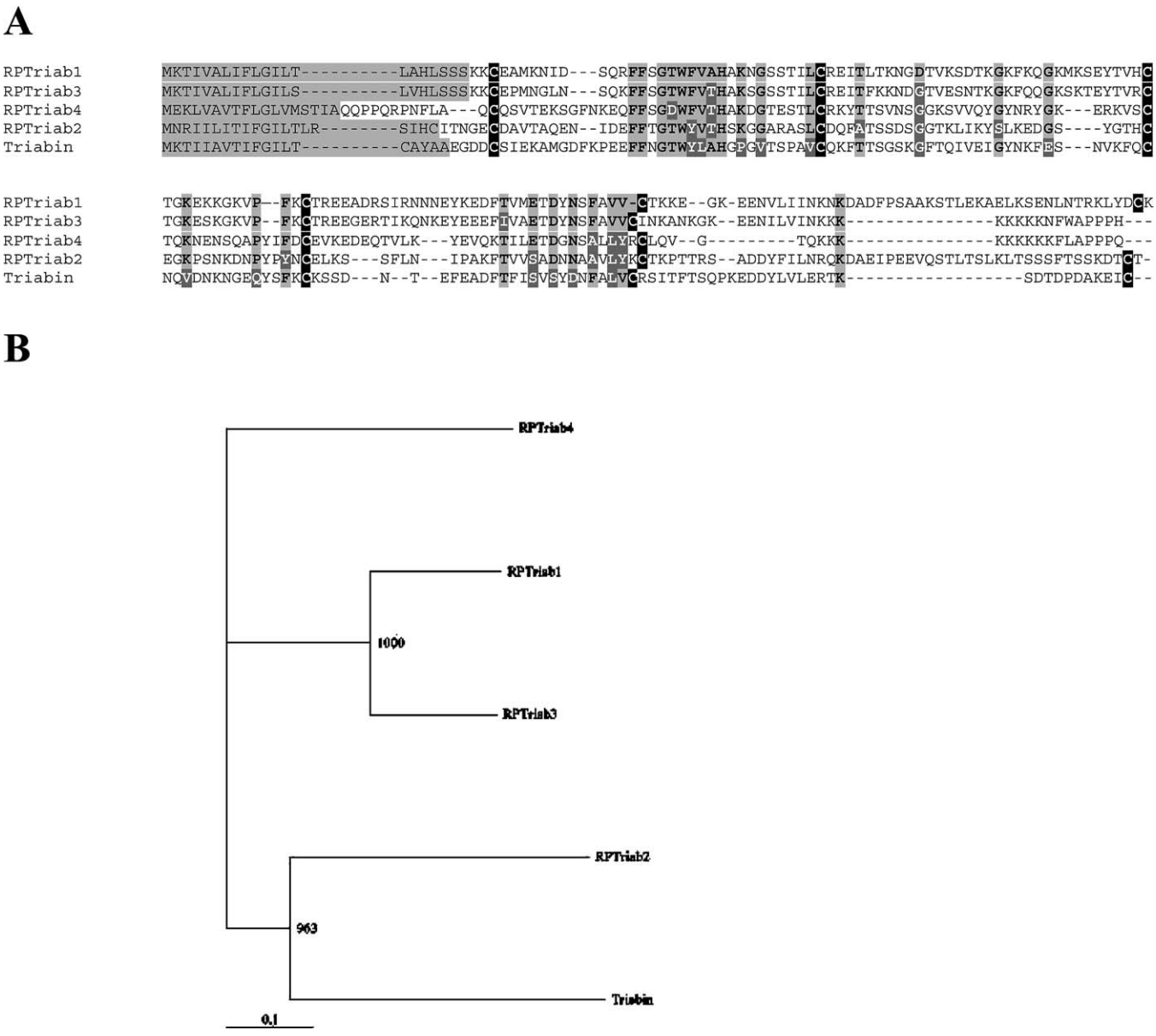

Fig. 6. The salivary triabin family in *R. prolixus*. Four novel *Rhodnius* triabins are compared with the previously reported *Triatoma* protein. For other details, see Fig. 4A.

of the genus *Triatoma* such as triabin (Noeske-Jungblut et al., 1995), pallidipin (Noeske-Jungblut et al., 1994), and procalins (Paddock et al., 2001) have also been found in *Rhodnius*. More recently, the South American *T. brasiliensis* was shown to also contain salivary cDNA sequences similar to these three previously described *Triatoma* lipocalins (Sant Anna et al., 2002). Although we know the function for some of these proteins, as indicated above, most of these novel proteins have unknown function. The probable scenario for the

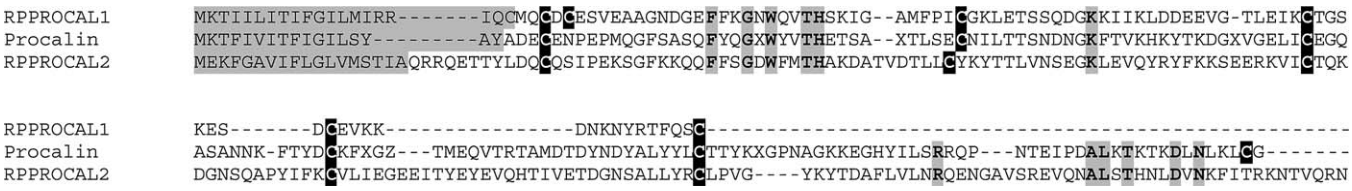

Fig. 7. The salivary procalin family in *R. prolixus*. Two novel *Rhodnius* procalins are compared with the previously reported *Triatoma* protein. For other details, see Fig. 4A.

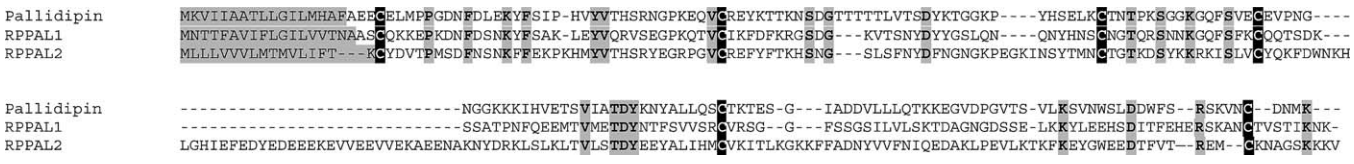

Fig. 8. The salivary pallidipin family in *R. prolixus*. Two novel *Rhodnius* pallidipins are compared with the previously reported *Triatoma* protein. For other details, see Fig. 4A.

evolution of these lipocalins, as proposed for the evolution of BPTI salivary proteins in *Ornithodorus* ticks (Mans et al., 2002), must have been several events of gene duplication and divergence of function. While lipocalins have also been found in tick saliva performing similar functions as in *Rhodnius*, such as histamine and serotonin binding (Paesen et al., 2000; Sangamnatdej et al., 2002), this contrasts with mosquitoes and sand flies, where no salivary lipocalins have been described to date, although another family of small ligand-binding proteins, the OBP (Hekmat-Scafe et al., 2000), have evolved into the D7 subfamily found in blood-sucking Diptera (Valenzuela et al., 2002a). Of interest, a putative OBP with a clear signal peptide indicative of secretion was also found in *Rhodnius*.

When all known triatomine lipocalin sequences were aligned and an NJ phylogenetic tree constructed, bootstrap values indicated that these sequences have evolved beyond recognition of a common ancestor (Fig. 10), although three clades were generated. Of these, a single robust NP/BABP clade was observed; however, the remaining sequences only produced robust associations with two or three other protein sequences at most, despite containing similarities to the Pfam triabin or NP motifs. These results may indicate a long evolutionary history for these proteins, or, alternatively a fast rate of evolution, as has been suggested by comparing the differences in salivary and housekeeping protein sequences between *Anopheles stephensi* and *An. gambiae* (Valenzuela et al., 2003). The diverse nature of lipocalins has been noted before (Flower et al., 2000), where a single Trp is found conservatively at the amino terminal end of the superfamily. The superfamily is clearly recognizable by the fold structure verified by crystallization of its several family members, which share primary sequence similarities within each family.

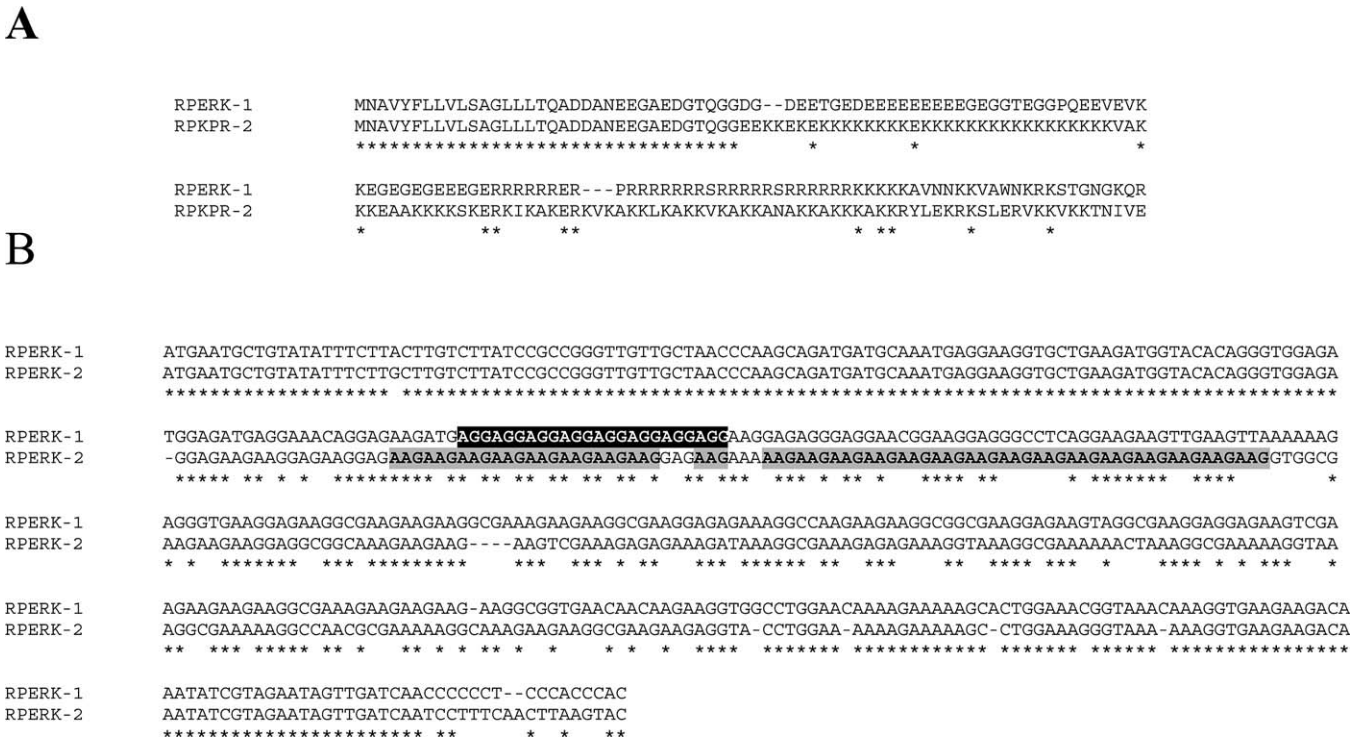

Fig. 9. The salivary RPERK family of *R. prolixus*. (A) Alignments of predicted protein sequences RPGE and RPPK. Other details as in the legend for Fig. 4A. (B) Alignment of the nucleotide sequences coding for RPGE and RPPK. The identical nucleotides of the 5′ region are shown in gray background. The GGA repeat region of RPGE and the AAG repeat region of RPPK are shown in reverse color.

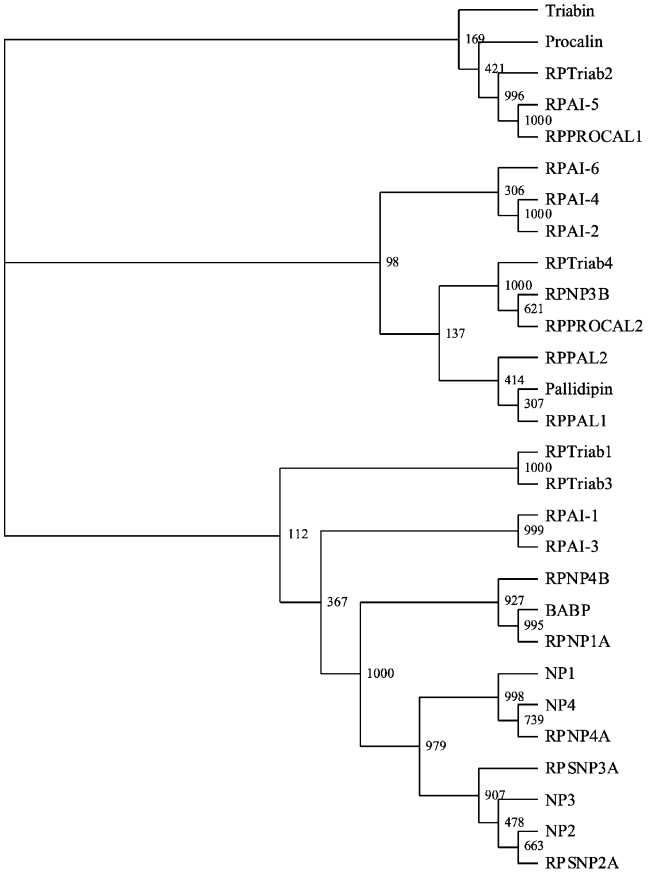

Fig. 10. Phylogenetic tree showing the sequence distance relationships between the salivary lipocalins of triatomines. The bar shows 10% divergence at the amino acid level. The numbers indicate the bootstrap value from 1000 iterations.

It was reported recently that expression of the salivary NP of *Rhodnius* is differentially regulated during insect development (Moreira et al., 2003). NP2 is the only NP found in the salivary glands of Ist instar insects; the number of NP expressed increases with each insect stage. In the same work, two novel amino terminal sequences from NP were presented and have been named NP5 and NP6; however, the sequence for NP6 in the paper actually represents the known sequence for NP2, while the sequence shown therein for NP2 has a single amino acid difference from that of NP2, where a KK substitutes for a KW. NP5 differs on a single amino acid from NP4, where a Q substitutes a K. It is possible that the KK observed in one of the novel sequences results from the normally low signal produced by W (tryptophane) and a misreading of K (lysine) that "bled" from the previous Edman cycle. In any case, neither of these two sequences was found in our library.

It is interesting to observe that several of the novel proteins described in Table 2 have amino terminal sequences containing stretches of basic amino acids. These are reminiscent of sequences found in proteins interacting with plasma membranes and in the so-called signal transduction sequences, which direct peptide transport through cell membranes. Farnesylated and myristoylated proteins are normally directed to the surface of the plasma membrane, where they become anchored by insertion of the lipid moiety into the membrane bilayer (Murray et al., 1997, 1999; Leventis and Silvius, 1998; Macia et al., 2000). Membrane targeting for these proteins is mediated by C-terminal polybasic sequences that interact non-specifically with negatively charged phospholipid head groups. Experimental and theoretical studies have shown that as few as five basic residues (either arginine or lysine) clustered near the C-terminal of a protein suffice for adsorption to a negatively charged membrane. This is true even in the absence of a protein-bound lipid moiety. Significantly, the outer leaflet of the platelet membrane becomes negatively charged following activation, due to exposure of phosphatidylserine, and serves as a procoagulant surface by promoting the assembly of coagulation factor complexes. The polybasic sequences found in salivary proteins may direct these proteins to the surface of activated platelets, where they could bind and/or release biologically active ligands, block the interaction of coagulation factors with the membrane surface, or interact with other proteins on the platelet surface.

It is remarkable to observe the molecular complexity and redundancy of the *Rhodnius* salivary transcriptome. Expression and bioassay of the novel proteins will ultimately characterize the salivary pharmacological complexity and redundancy resulting from *Rhodnius*' evolution to blood feeding.

Acknowledgements

We are thankful to Dr. Robert Gwadz, Dr. Thomas Wellems, and Dr. Thomas Kindt for support and encouragement and to Ms. Brenda Rae Marshall and Ms. Nancy Schulman for editorial assistance.

References

Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W., Lipman, D.J., 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 25, 3389–3402.

Andersen, J.F., Champagne, D.E., Weichsel, A., Ribeiro, J.M., Balfour, C.A., Dress, V., Montfort, W.R., 1997. Nitric oxide binding and crystallization of recombinant nitrophorin I, a nitric oxide transport protein from the blood-sucking bug *Rhodnius prolixus*. Biochemistry 36, 4423–4428.

Andersen, J.F., Francischetti, I.M., Valenzuela, J.G., Schuck, P., Ribeiro, J.M., 2003. Inhibition of hemostasis by a high affinity biogenic amine-binding protein from the saliva of a blood-feeding insect. J. Biol. Chem. 278, 4611–4617.

Andersen, J.F., Weichsel, A., Balfour, C.A., Champagne, D.E., Montfort, W.R., 1998. The crystal structure of nitrophorin 4 at 1.5 A resolution: transport of nitric oxide by a lipocalin-based heme protein. Structure 6, 1315–1327.

Ashburner, M., Ball, C.A., Blake, J.A., Botstein, D., Butler, H., Cherry, J.M., Davis, A.P., Dolinski, K., Dwight, S.S., Eppig, J.T., Harris, M.A., Hill, D.P., Issel-Tarver, L., Kasarskis, A., Lewis, S., Matese, J.C., Richardson, J.E., Ringwald, M., Rubin, G.M., Sherlock, G., 2000. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nat. Genet. 25, 25–29.

Bateman, A., Birney, E., Durbin, R., Eddy, S.R., Howe, K.L., Sonnhammer, E.L., 2000. The Pfam protein families database. Nucleic Acids Res. 28, 263–266.

Buxton, P.A., 1930. The biology of a blood sucking bug, *Rhodnius prolixus* Stahl. Trans. R. Entomol. Soc. Lond. 78, 227–236.

Champagne, D., Nussenzveig, R.H., Ribeiro, J.M.C., 1995. Purification, characterization, and cloning of nitric oxide-carrying heme proteins (nitrophorins) from salivary glands of the blood sucking insect *Rhodnius prolixus*. J. Biol. Chem. 270, 8691–8695.

Charlab, R., Valenzuela, J.G., Rowton, E.D., Ribeiro, J.M., 1999. Toward an understanding of the biochemical and pharmacological complexity of the saliva of a hematophagous sand fly *Lutzomyia longipalpis*. Proc. Natl. Acad. Sci. USA 96, 15155–15160.

Debrauwere, H., Gendrel, C.G., Lechat, S., Dutreix, M., 1997. Differences and similarities between various tandem repeat sequences: minisatellites and microsatellites. Biochimie 79, 577–586.

Dias, J.C., Silveira, A.C., Schofield, C.J., 2002. The impact of Chagas disease control in Latin America: a review. Mem. Inst. Oswaldo Cruz 97, 603–612.

Flower, D.R., North, A.C., Sansom, C.E., 2000. The lipocalin protein family: structural and sequence overview. Biochim. Biophys. Acta 1482, 9–24.

Francischetti, I.M., Andersen, J.F., Ribeiro, J.M., 2002a. Biochemical and functional characterization of recombinant *Rhodnius prolixus* platelet aggregation inhibitor 1 as a novel lipocalin with high affinity for adenosine diphosphate and other adenine nucleotides. Biochemistry 41, 3810–3818.

Francischetti, I.M., Ribeiro, J.M., Champagne, D., Andersen, J., 2000. 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. 275, 12639–12650.

Francischetti, I.M., Valenzuela, J.G., Pham, V.M., Garfield, M.K., Ribeiro, J.M., 2002b. Toward a catalog for the transcripts and proteins (sialome) from the salivary gland of the malaria vector *Anopheles gambiae*. J. Exp. Biol. 205, 2429–2451.

Gardner, M.J., Hall, N., Fung, E., White, O., Berriman, M., Hyman, R.W., Carlton, J.M., Pain, A., Nelson, K.E., Bowman, S., Paulsen, I.T., James, K., Eisen, J.A., Rutherford, K., Salzberg, S.L., Craig, A., Kyes, S., Chan, M.S., Nene, V., Shallom, S.J., et al., 2002. Genome sequence of the human malaria parasite *Plasmodium falciparum*. Nature 419, 498–511.

Glusa, E., Bretschneider, E., Daum, J., Noeske-Jungblut, C., 1997. Inhibition of thrombin-mediated cellular effects by triabin, a highly potent anion-binding exosite thrombin inhibitor. Thromb. Haemost. 77, 1196–1200.

Gooding, R.H., 1972. Digestive process of haematophagous insects. I. A literature review. Quaest. Ent. 8, 5–60.

Gregorio, E.A., Ratcliffe, N.A., 1991. The distribution of agglutinins and lytic activity against *Trypanosoma rangeli* and erythrocytes in *Rhodnius prolixus* and *Triatoma infestans* tissue extracts and haemolymph. Mem. Inst. Oswaldo Cruz 86, 181–186.

Haendler, B., Becker, A., Noeske-Jungblut, C., Kratzschmar, J., Donner, P., Schleuning, W.D., 1996. Expression, purification and characterisation of recombinant pallidipin, a novel platelet aggregation inhibitor from the haematophageous triatomine bug *Triatoma pallidipennis*. Blood Coagul. Fibrinolysis 7, 183–186.

Hekmat-Scafe, D.S., Dorit, R.L., Carlson, J.R., 2000. Molecular evolution of odorant-binding protein genes OS-E and OS-F in *Drosophila*. Genetics 155, 117–127.

Hellmann, K., Hawkins, R.I., 1965. Prolixin-S and Prolixin-G: two anticoagulants from *Rhodnius prolixus* Stahl. Nature 207, 265–267.

Huang, X., 1992. A contig assembly program based on sensitive detection of fragment overlaps. Genomics 14, 18–25.

Huang, Y., Ying, K., Xie, Y., Zhou, Z., Wang, W., Tang, R., Zhao, W., Zhao, S., Wu, H., Gu, S., Mao, Y., 2001. Cloning and characterization of a novel human leptin receptor overlapping transcript-like 1 gene (LEPROTL1). Biochim. Biophys. Acta 1517, 327–331.

Hvidsten, T.R., Komorowski, J., Sandvik, A.K., Laegreid, A., 2001. Predicting gene function from gene expressions and ontologies. Pac. Symp. Biocomput. 299–310.

Isawa, H., Yuda, M., Orito, Y., Chinzei, Y., 2002. A mosquito salivary protein inhibits activation of the plasma contact system by binding to factor XII and high molecular weight kininogen. J. Biol. Chem. 13, 13.

Isawa, H., Yuda, M., Yoneda, K., Chinzei, Y., 2000. The insect salivary protein, prolixin-S, inhibits factor IXa generation and Xase complex formation in the blood coagulation pathway. J. Biol. Chem. 275, 6636–6641.

Leventis, R., Silvius, J.R., 1998. Lipid-binding characteristics of the polybasic carboxy-terminal sequence of K-ras4B. Biochemistry 37, 7640–7648.

Lewis, S., Ashburner, M., Reese, M.G., 2000. Annotating eukaryote genomes. Curr. Opin. Struct. Biol. 10, 349–354.

Li, S., Kwon, J., Aksoy, S., 2001. Characterization of genes expressed in the salivary glands of the tsetse fly, *Glossina morsitans morsitans*. Insect Mol. Biol. 10, 69–76.

Macia, E., Paris, S., Chabre, M., 2000. Binding of the PH and polybasic C-terminal domains of ARNO to phosphoinositides and to acidic lipids. Biochemistry 39, 5893–5901.

Mans, B.J., Louw, A.I., Neitz, A.W., 2002. Evolution of hematophagy in ticks: common origins for blood coagulation and platelet aggregation inhibitors from soft ticks of the genus *Ornithodoros*. Mol. Biol. Evol. 19, 1695–1705.

Metcalf, R.L., 1945. The physiology of the salivary glands of *Anopheles quadrimaculatus*. J. Nat. Malaria Soc. 4, 271.

Moreira, M.F., Coelho, H.S., Zingali, R.B., Oliveira, P.L., Masuda, H., 2003. Changes in salivary nitrophorin profile during the life cycle of the blood-sucking bug *Rhodnius prolixus*. Insect Biochem. Mol. Biol. 33, 23–28.

Murray, D., Arbuzova, A., Hangyas-Mihalyne, G., Gambhir, A., Ben-Tal, N., Honig, B., McLaughlin, S., 1999. Electrostatic properties of membranes containing acidic lipids and adsorbed basic peptides: theory and experiment. Biophys. J 77, 3176–3188.

Murray, D., Ben-Tal, N., Honig, B., McLaughlin, S., 1997. Electrostatic interaction of myristoylated proteins with membranes: simple physics, complicated biology. Structure 5, 985–989.

Nielsen, H., Engelbrecht, J., Brunak, S., von Heijne, G., 1997. Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. Protein Eng 10, 1–6.

Noeske-Jungblut, C., Haendler, B., Donner, P., Alagon, A., Possani, L., Scheleuning, W.-D., 1995. Triabin, a highly potent exosite inhibitor of thrombin. J. Biol. Chem. 270, 28629–28634.

Noeske-Jungblut, C., Krätzschmar, J., Haendler, B., Alagon, A., Possani, L., Verhallen, P., Donner, P., Schleuning, W.-D., 1994. An inhibitor of collagen-induced platelet aggregation from the saliva of *Triatoma pallidipennis*. J. Biol. Chem. 269, 5050–5053.

Nussenzveig, R.H., Bentley, D.L., Ribeiro, J.M.C., 1995. Nitric oxide loading of the salivary NO-heme carrying proteins (nitrophorins) in the blood sucking bug, *Rhodnius prolixus*. J. Exp. Biol. 198, 1093–1098.

Paddock, C.D., McKerrow, J.H., Hansell, E., Foreman, K.W., Hsieh, I., Marshall, N., 2001. Identification, cloning, and recom-

binant expression of procalin, a major triatomine allergen. J. Immunol. 167, 2694–2699.

Paesen, G.C., Adams, P.L., Nuttall, P.A., Stuart, D.L., 2000. Tick histamine-binding proteins: lipocalins with a second binding cavity. Biochim. Biophys. Acta 1482, 92–101.

Page, R.D., 1996. TreeView: an application to display phylogenetic trees on personal computers. Comput. Appl. Biosci. 12, 357–358.

Ribeiro, J.M., Francischetti, I.M., 2002. Role of arthropod saliva in blood feeding: Sialome and post-sialome perspectives. Annu. Rev. Entomol. 19, 19.

Ribeiro, J.M.C., 1982. The antiserotonin and antihistamine activities of salivary secretion of *Rhodnius prolixus*. J. Insect Physiol. 28, 69–75.

Ribeiro, J.M.C., Garcia, E.S., 1980. The salivary and crop apyrase activity of *Rhodnius prolixus*. J. Insect Physiol. 26, 303–307.

Ribeiro, J.M.C., Hazzard, J.M.H., Nussenzveig, R.H., Champagne, D., Walker, F.A., 1993. Reversible binding of nitric oxide by a salivary nitrosylhemeprotein from the blood sucking bug, *Rhodnius prolixus*. Science 260, 539–541.

Ribeiro, J.M.C., Nussenzveig, R.H., 1993. The salivary catechol oxidase/peroxidase activities of the mosquito, *Anopheles albimanus*. J. Exp. Biol. 179, 273–287.

Ribeiro, J.M.C., Sarkis, J.J.F., 1982. Anti-thromboxane activity in *Rhodnius prolixus* salivary secretion. J. Insect Physiol. 28, 655–660.

Ribeiro, J.M.C., Schneider, M., Guimaraes, J.A., 1995. Purification and characterization of Prolixin S (nitrophorin 2), the salivary anticoagulant of the blood sucking bug, *Rhodnius prolixus*. Biochem. J. 308, 243–249.

Ribeiro, J.M.C., Walker, F.A., 1994. High affinity histamine-binding and anti-histaminic activity of the salivary NO-carrying heme protein (nitrophorin) of *Rhodnius prolixus*. J. Exp. Med. 180, 2251–2257.

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

Sangamnatdej, S., Paesen, G.C., Slovak, M., Nuttall, P.A., 2002. A high affinity serotonin- and histamine-binding lipocalin from tick saliva. Insect Mol. Biol. 11, 79–86.

Sant Anna, M.R., Araujo, J.G., Pereira, M.H., Pesquero, J.L., Diotaiuti, L., Lehane, S.M., Lehane, M.J., 2002. Molecular cloning and sequencing of salivary gland-specific cDNAs of the blood-sucking bug *Triatoma brasiliensis* (Hemiptera: Reduviidae). Insect Mol. Biol. 11, 585–593.

Saper, C.B., Chou, T.C., Elmquist, J.K., 2002. The need to feed: homeostatic and hedonic control of eating. Neuron 36, 199–211.

Sarkis, J.J.F., Guimaraes, J.A., Ribeiro, J.M.C., 1986. Salivary apyrase of *Rhodnius prolixus*: kinetics and purification. Biochem. J. 233, 885–891.

Schreiber, M.C., Karlo, J.C., Kovalick, G.E., 1997. 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 191, 135–141.

Schultz, J., Copley, R.R., Doerks, T., Ponting, C.P., Bork, P., 2000. SMART: a web-based tool for the study of genetically mobile domains. Nucleic Acids Res. 28, 231–234.

Simons, F.E., Peng, Z., 2001. Mosquito allergy: recombinant mosquito salivary antigens for new diagnostic tests. Int. Arch. Allergy Immunol. 124, 403–405.

Simpson, A.E., 1997. The cytochrome P450 4 (CYP4) family. Gen. Pharmacol. 28, 351–359.

Sonnhammer, E.L., von Heijne, G., Krogh, A., 1998. A hidden Markov model for predicting transmembrane helices in protein sequences. Proc. Int. Conf. Intell. Syst. Mol. Biol. 6, 175–182.

Takai, Y., Sasaki, T., Matozaki, T., 2001. Small GTP-binding proteins. Physiol. Rev. 81, 153–208.

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

Urade, Y., Hayaishi, O., 2000. Prostaglandin D synthase: structure and function. Vitam. Horm. 58, 89–120.

Valenzuela, J.G., Belkaid, Y., Garfield, M.K., Mendez, S., Kamhawi, S., Rowton, E.D., Sacks, D.L., Ribeiro, J.M., 2001. Toward a defined anti-*Leishmania* vaccine targeting vector antigens: characterization of a protective salivary protein. J. Exp. Med. 194, 331–342.

Valenzuela, J.G., Charlab, R., Gonzalez, E.C., Miranda-Santos, I.K.F., Marinotti, O., Francischetti, I.M., Ribeiro, J.M.C., 2002a. The D7 family of salivary proteins in blood sucking Diptera. Insect Mol. Biol. 11, 149–155.

Valenzuela, J.G., Francischetti, I.M.B., Pham, V.M., Garfield, M.K., Mather, T.N., Ribeiro, J.M.C., 2002b. Exploring the sialome of the tick, *Ixodes scapularis*. J. Exp. Biol. 205, 2843–2864.

Valenzuela, J.G., Pham, V.M., Garfield, M.K., Francischetti, I.M., Ribeiro, J.M.C., 2002c. Toward a description of the sialome of the adult female mosquito *Aedes aegypti*. Insect Biochem. Mol. Biol. 32, 1101–1122.

Valenzuela, J.G., Francischetti, I.M., Pham, V.M., Garfield, M.K., Ribeiro, J.M., 2003. Exploring the salivary gland transcriptome and proteome of the *Anopheles stephensi* mosquito. Insect Biochem. Mol. Biol. 33, 717–732.

Weichsel, A., Andersen, J.F., Champagne, D.E., Walker, F.A., Montfort, W.R., 1998. Crystal structures of a nitric oxide transport protein from a blood-sucking insect. Nat. Struct. Biol. 5, 304–309.

Wigglesworth, V.B., 1942. The fate of haemoglobin in *Rhodnius prolixus* (Hemiptera) and other blood-sucking arthropods. Proc. R. Soc. (Lond.) 131B, 313–339.

Wigglesworth, V.B., 1972. The principles of insect physiology. Chapman and Hall, New York.

Wilson, P., Liu, Y.J., Banchereau, J., Capra, J.D., Pascual, V., 1998. Amino acid insertions and deletions contribute to diversify the human Ig repertoire. Immunol. Rev. 162, 143–151.

Zhang, Y., Ribeiro, J.M., Guimaraes, J.A., Walsh, P.N., 1998. Nitrophorin-2: a novel mixed-type reversible specific inhibitor of the intrinsic factor-X activating complex. Biochemistry 37, 10681–10690.