Savicalin, a lipocalin from hemocytes of the soft tick, *Ornithodoros savignyi*

Paul H. Cheng[a], Ben J. Mans[b], Albert W.H. Neitz[a] and Anabella R.M. Gaspar[a]*

[a]Department of Biochemistry, University of Pretoria, Pretoria 0002, South Africa
[b]Parasites, Vectors and Vector-Borne Diseases, Agricultural Research Council-Onderstepoort Veterinary Institute, Onderstepoort 0110, South Africa
*Corresponding author: anabella.gaspar@up.ac.za

Abstract

Savicalin, is a lipocalin found in the hemocytes of the soft tick, *Ornithodoros savignyi*. It could be assigned to the tick lipocalin family based on BLAST analysis. Savicalin is the first non-salivary gland lipocalin described in ticks. The mature sequence is composed of 188 amino acids with a molecular mass of 21481.9 Da. A homolog for savicalin was found in a whole body EST-library from a related soft tick *O. porcinus*, while other tick salivary gland derived lipocalins retrieved from the non-redundant sequence database are more distantly related. Homology modeling supports the inclusion of savicalin into the lipocalin family. The model as well as multiple alignments suggests the presence of five disulphide bonds. Two conserved disulphide bonds are found in hard and soft tick lipocalins. A third disulphide bond is shared with the TSGP4-clade of leukotriene C4 binding soft tick lipocalins and a fourth is shared with a lipocalin from the hard tick *Ixodes scapularis*. The fifth disulphide bond is unique and links strands D-E. Phylogenetic analysis showed that savicalin is a distant relative of salivary gland derived lipocalins, but groups within a clade that is possibly non-salivary gland derived. It lacks the biogenic amine-binding motif associated with tick histamine and serotonin binding proteins. Expression profiles indicate that savicalin is found in hemocytes, midgut and ovaries, but not in the salivary glands. Up-regulation occurs in hemocytes after bacterial challenge and in midguts and ovaries after feeding. Given its tissue distribution and up-regulation of expression, it is possible that this lipocalin functions in tick development after feeding or in an anti-microbial capacity.

Key words: Savicalin, lipocalin-like molecule, hemocytes, homology modeling, phylogenetic analysis, *Ornithodoros savignyi*

Introduction

Lipocalins play important roles in immune responses, transport of hydrophobic molecules (such as pheromones, steroids, bilins, retinoids and lipids), cancer cell interactions and allergies (Peitsch et al. 1991; Nagata et al. 1992; Cowan et al. 1990; Flower et al. 1991; Pervaiz and Brew, 1987). These molecules have been detected across all living organisms and exhibit three characteristic features,

namely, an unusually low amino acid sequence similarity (typically 15-25% between paralogs), a highly conserved protein tertiary structure, and a similar arrangement of exons and introns in the coding sequence of their genes (Mans and Neitz, 2004a; Mans et al. 2008b). Lipocalins are single modular proteins of around 150-200 amino acids that fold tightly in a β-barrel that wind around a central axis with potential for binding small hydrophobic molecules in a central pocket (Skerra 2000; Flower 2000; Flower et al. 2000). These proteins usually have two helices at the N- and C-terminal ends, the N-terminal helix closing of one side of the barrel and the C-terminal α-helix packed against the outer surface of the barrel. Most lipocalins classified are based on variations observed in the length of the N- and C- terminal segments. These changes are not expected to significantly alter the β-barrel core, and could provide a substrate of variation for potential functional diversification and specialization (Monfort et al. 2000).

In ticks, there are more than 300 lipocalins known to this date (Mans et al. 2008b). Tick lipocalins identified from saliva or salivary glands are distinct from other arthropod lipocalins and could only be assigned to the lipocalin family based on structural similarity (Paesen et al. 2000; Mans et al. 2003). The crystal structure of Histamine-Binding Protein (HBP) from the hard tick, *Rhipicephalus appendiculatus,* established the first functional relationship of tick lipocalins and their ligands and indicated that tick lipocalins could function in a immunomodulatory capacity by scavenging histamine (Paesen et al. 1999, 2000). Since then, both hard and soft tick lipocalins have been implicated in the binding of a variety of bio-active ligands such as histamine, serotonin, leukotriene B_4, leukotriene C_4, and thromboxane A_2 involved in immuno-modulation and platelet aggregation (Sangamnatdej et al. 2002; Mans et al. 2008a,b; Mans and Ribeiro, 2008a,b). Targeting of complement C5 has also been indicated for soft tick lipocalins as well as being toxic (Nunn et al. 2005; Mans et al. 2002; Mans et al. 2003; Mans 2005; Mans et al. 2008b; Mans and Ribeiro, 2008a,b).

All tick lipocalins described thus far are salivary gland derived and presumed to be involved in tick feeding (Mans et al. 2008b, Valenzuela et al. 2002, Francischetti et al. 2008a; Francischetti et al. 2008b; Ribeiro et al. 2006). However, lipocalins have been described in other arthropods that are not involved in blood-feeding, but in processes such as development, coloration, defense mechanisms, transport of ligands (Kayser et al. 2005; Sánchez et al. 1995, 2000, 2006; Mauchamp et al. 2006; Andersen et al. 2005; Weichsel et al. 1998). The possibility thus exists that lipocalins might play a much larger role in tick biology that is not limited to the feeding process alone. The current study describes such a lipocalin from the hemocytes of *O. savignyi*, named savicalin that is not associated with feeding.

Materials and methods

Cloning and sequencing

Ticks obtained from the Upington region, Northern Cape, South Africa, were dorsally immobilized on double-sided tape. A 30 gauge needle was used to puncture the first pair of coxae at the base of the trochanter followed by gentle pressure on abdomen (Johns et al. 1998). Hemolymph was collected from 20 ticks using a glass capillary and immediately added to Tri-Reagent (Sigma). RNA was isolated according to the manufacturer's instructions.

Single stranded cDNA was prepared from total RNA (500 ng) using a 5'Smart IIA anchor primer AAGCAGTGGTATCAACGCAGAGTACGCGGG, a poly-T anchor primer GCTATCATTACCACAAACCACTCTTTTTT and Superscript III reverse transcriptase (Invitrogen, Carlsbad, CA) according to the manufacturer's instructions. Purification of cDNA was performed using Nucleospin Extract II PCR clean up and gel extraction kit (Macherey Nagel, Duren, Germany).

To obtain the coding gene and 3'-untranslated region, a degenerate primer TGGACGGA(T/C)TA(T/C)TA(T/C)GA(T/C)(A/C)G (Integrated DNA Technologies, Coralville, USA) was designed from a *de novo* sequence WTDYYDRM. Single-stranded cDNA, degenerate primer and poly-T anchor primer were used to optimize 3' RACE using exTaq polymerase (Takara, Japan). Optimized conditions consisted of an initial cDNA denaturation (94^{o}C, 3min), hot start addition of enzyme (80^{o}C), followed by 35 cycles of DNA denaturation (94^{o}C, 30s), annealing (52^{o}C, 25s) and extension (72^{o}C, 2 min) followed by a final extension (72^{o}C, 7 min).

One-tenth of amplified products were analyzed by agarose gel electrophoresis and the rest was ligated into the pGEM T easy vector system (Promega, Madison, WI, USA) and transformed into electrocompetent BL21 cells. Positive clones were identified by colony PCR using Kapa ready mix (Kapa Bioscience, South Africa).

Plasmid was recovered using the NucleoSpin® Ready-to-use system for fast purification of nucleic acids (Machery-Nagel, Germany) as described in the manual. Positive clones (20) were sequenced using both upstream T7 (TAATACGACTCACTATAGGG) and downstream SP6 (TATTTAGGTGACACTATAG) primers with the ABI Big Dye solution kit on an automated Applied Biosystems 3130 DNA sequencer.

DNA sequences were analyzed for similarities with known sequences using the BLAST (Basic Local Alignment Search Tool) algorithm (www.ncbi.nlm.nih.gov). The BLAST algorithm (Altschul et al. 1990, 1997) searches for local (as opposed to global) alignments and reports the significance of the search results as an

expect value. Generally, an e-value of <0.0001 (1e-04) is considered highly significant.

The amino acid sequence of savicalin was deduced using Bioedit, while Expasy was used to predict its amino acid composition, hydrophobicity profile, pI and functional family.

Multiple sequence alignments and phylogenetic analysis

Lipocalin sequences from both hard and soft ticks were retrieved from the National Center for Biotechnology Information (NCBI) Genebank database, by BLASTP, TBLASTN and PSI-BLAST analysis using the sequence of savicalin (Altschul et al. 1990; 1997). Multiple sequence alignments of the tick lipocalins were performed using ClustalW with default parameters (Jeanmougin et al. 1998; Larkin et al. 2007). Sequences were manually checked and adjusted accordingly. Neighbor joining (NJ) analysis was conducted using MEGA version 4.0 (Saitou and Nei 1987; Tamura et al. 2007). Gapped positions were completely deleted so that 55 informative sites were used for analysis. Reliability of the inferred tree was evaluated by bootstrap analysis (100 000 replicates).

Homology modeling of savicalin, quality assessment of the modeled structure and fold recognition

For homology modeling of savicalin, female specific histamine binding protein (PDB ID: 1QFT, 1QFV; Paesen et al. 1999) was used as a template using SWISS-PdbViewer (Guex and Peitsch 1997). The initial model was submitted to the SWISS-MODEL automated comparative protein modeling server (Guex et al. 1999). Savicalin's sequence was also submitted to the Phyre fold recognition server (Kelley and Sternberg, 2009) and analyzed using the Conserved Domain Database (CDD) (Marchler-Bauer et al. 2009).

Tissue distribution pattern of savicalin

RT-PCR was carried out to analyze gene expression of savicalin. *Bacillus subtilis* (ATCC: 13933) cells were resuspended in physiological saline to a final concentration of the 2.5 x 10^6 cells/ml and 1 μl of the suspension was heat-inactivated and injected into 20 unfed adult female ticks. Saline only was injected into the same number ticks as a control. Total RNA was isolated as described above from the midguts, ovaries, salivary glands and hemolymph obtained from both groups of ticks 24 hours post injection.

In the second part of the experiment, 20 ticks were fed artificially on heparinized cattle blood (Experimental Farm, University of Pretoria, SA) infected with *B.*

subtilis (2.5 x 10^6 cells/ml blood) as described by Schwan et al. (1991). Native blood (no bacteria added) was used for the control group. At 1 day and 10 days post feeding, total RNA was isolated from the same tissues as described for unfed, hemocoelic injected ticks.

First strand synthesis was performed as described above. To obtain the transcript of savicalin a first strand cDNA synthesis for each tissue (500 ng), a gene specific primer TGGACGGATTACTACGACCG and a poly-T anchor primer were used. The primers for the housekeeping gene, actin, were CAGATCATGTTTGAGACCTTCAAC (forward primer) and G(C/G)CCATCTC(T/C)TGCTCGAA(A/G)TC (reverse primer). PCR reactions were performed using a denaturation step (94^{o}C, 3min), hot start addition of enzyme (80^{o}C) followed by 35 cycles of DNA denaturation (94^{o}C, 30s), annealing (54^{o}C, 25s) and extension (72^{o}C, 2 min) followed by a final extension (72^{o}C, 7 min).

Results

In a previous study undertaken by our group, *O. savignyi* hemolymph proteins that recognize and bind to Gram-negative *Escherichia coli* bacteria were subjected to tandem mass spectrometry (MS/MS) analysis for identification. None of the proteins analyzed could be identified by searching the current databases with both the MS/MS ion spectra as well as the derived *de novo* sequences obtained for these proteins. In this study, a degenerate primer derived from a *de novo* sequence for one of these hemolymph proteins using 3'-RACE resulted in the amplification of a single 900 bp fragment. Sequencing revealed the full gene sequence of a non-related hemocyte protein (Fig 1). This sequence contains the 5'UTR, open reading frame, stop codon, poly-adenylation site and 3'UTR. The translated protein sequence has a signal peptide indicating that this protein is targeted to the secretory pathway in hemocytes and other tissues. The mature processed protein has a calculated pI of 4.37 and molecular mass of 21481.9 Da, that includes 10 cysteine residues predicted to be involved in disulphide bonds.

BLASTP analysis indicated similarity to tick lipocalins. The three best hits include lipocalins from *Rhipicephalus* (*Boophilus) microplus* (E-0.004), *I. scapularis* (E-0.004) and *Argas monolakensis* (E-0.003). In addition, a TBLASTN query of the non-redundant EST database retrieved EST sequences (5E-14) from a whole body cDNA library of the closely related tick, *Ornithodoros porcinus*. The translated EST sequences showed 22% sequence identity to savicalin. In addition, savicalin was submitted to the CDD and the Phyre servers in order to confirm that it belongs to the lipocalin fold. In the case of the CDD analysis savicalin was assigned to the His_binding superfamily, which essentially describes all tick lipocalins. The top 4 hits obtained with the Phyre server were all tick lipocalins for which structures were previously solved and in all cases the

estimated precision was greater than 95%, while all other hits corresponded to lipocalins from other organisms. As such, savicalin was assigned to the lipocalin family using three different algorithms that preferentially selected the lipocalin fold from a variety of known sequences and folds. This increased the confidence that savicalin belongs to the lipocalin fold, even if it is divergent.

Multiple sequence alignments with these proteins as well as tick lipocalins that have been functionally characterized indicated that savicalin shows overall less than 20% sequence identity to these tick lipocalins (Fig. 2). Conserved features include two disulphide bonds found in hard and soft ticks (Cys59-Cys180; Cys132-Cys159) (Mans et al. 2003). A third disulphide bond (Cys145-Cys167) is shared with TSGP4, the serotonin and histamine binding protein from *Dermacentor reticulatus*, and an *I. scapularis* sequence, that is characteristic of the leukotriene C4 binding clade of soft ticks (Mans and Ribeiro 2008a). A fourth is shared with the same *I. scapularis* sequence (Cys35-Cys129). The fifth disulphide bond is unique and links the β strands DE (Cys96-Cys106) (Fig. 2). Savicalin lacks the biogenic amine-binding (BAB) motif (CL[L]X(11)VL[G]X(10)C vs. CD[VIL]X(7,17)EL[WY]X(11,30)C), and would therefore not bind biogenic amines (Mans et al. 2008b). In addition, the residues proposed to be involved in leukotriene binding and complement C5 interaction (Mans and Ribeiro 2008b), are not conserved in savicalin either, predicting that it will lack these functions as well (results not shown).

The molecular model obtained presents all secondary structure features associated with lipocalins (Fig. 3). This includes the eight stranded anti-parallel +1 beta-barrel, the N-terminal helix that closes off the barrel and the C-terminal alpha-helix that packs against the barrel (Skerra 2000; Flower 2000; Flower et al. 2000). This indicates that the overall features of the model fits well with the proposed lipocalin structure and supports the inclusion of savicalin into this family. All cysteine residues are spatially organized to form intact disulphide bonds in the model, supporting the proposed disulphide bond pattern for savicalin (Fig. 2; Fig. 3).

Phylogenetic analysis using the sequence set from the multiple sequence alignments indicated that savicalin does not group with any of the known functionally characterized clades, implying that it will lack these functions (Fig. 4).

Expression profiling by mRNA level detection showed that savicalin was up-regulated in hemocytes of unfed ticks upon hemocoelic bacterial challenge as well as ten days after feeding (Fig. 5). Down-regulation occurred, however, 1 day after feeding. In contrast, savicalin was not up-regulated in midguts and ovaries irrespective of bacterial challenge or 1 day after feeding, but seems to be constitutively expressed. It is, however, down-regulated 10 days after feeding.

Savicalin was also up-regulated in midguts of unfed ticks. In contrast, no transcription was detected in salivary glands.

Discussion

The limited number of tick protein sequences available in databases is a drawback to the identification of tick proteins using a proteomics approach, if species or organs specific EST libraries are not available (Madden et al. 2002; Oleaga et al. 2007). The number of tick sequences in Genbank is rapidly growing and MS/MS ion spectra data as well as *de novo* sequences may be archived and used to search databases in the future (Blackburn and Goshe 2009; Shevchenko et al. 2009). However until then, *de novo* sequences may be employed for the design of degenerate primers for cloning of the genes of the corresponding proteins by a PCR-based approach (Lingner et al. 1997, Shevchenko et al. 2001). Using this approach we failed to identify the original tick hemolymph *E. coli* binding protein, but discovered a novel lipocalin-like molecule in hemocytes.

Ticks are obligate blood-feeding parasites that interact with their hosts mainly at the feeding site. Soft ticks such as *O. savignyi* feed multiple times imbibing relatively small amounts of blood that is utilized for laying small batches of eggs and for development and molting (Mans et al. 2004a, Sonenshine 1991). During feeding a salivary gland derived cocktail of bio-active proteins is secreted into the feeding site that suppresses the host's immune and hemostatic defenses. Soft ticks may secrete up to 200 different proteins that belong to various protein families during the feeding process (Mans et al. 2008a; Francischetti et al. 2008a; b). Characterization of the protein families found in salivary glands by proteomic methods have become a standard method to identify potential proteins secreted during feeding that are involved in the regulation of the host's defense mechanisms (Ribeiro and Francischetti 2003; Francischetti et al. 2009).

The presence of a signal peptide is indicative of the secretory nature of a salivary gland protein and implicates it as being functional at the feeding site in the host (Valenzuela et al. 2002; Mans et al. 2008a; Ribeiro et al. 2006). Signal peptides also target secretory proteins to the secretory granules where proteins are stored as aggregated masses until secretion (Mans et al. 2001). These proteins will as such not have a house-keeping function (functions not involved with tick-host interaction) in the tick. A basic assumption is thus that salivary gland derived proteins are limited to the salivary glands and will not be present in other tissues. The absence of proteins in tissues other than the salivary gland is thus also used as evidence for potential roles in feeding (Stutzer et al. 2008). It was, however, indicated that most protein families present in salivary glands are related to much larger families generally found in arthropods and that salivary gland proteins probably derived from these families with house-keeping function during the evolution of hematophagy in ticks (Mans and Neitz 2004a; Mans et al. 2008a). As

such, protein family members should also exist in non-salivary gland tissues. For example, exon-intron gene structure similarities suggested that tick salivary gland derived lipocalins are evolutionary related to lazarillo sequences from insects and it was suggested that such non-salivary gland derived lipocalins should also exist in ticks (Mans et al. 2004a).

Lipocalins are the most abundant proteins found in tick salivary glands, in terms of transcript numbers, protein expression levels and numbers of family members (Mans et al. 2008a,b). To date more than 300 lipocalin sequences have been deposited in the sequence databases and most have been indicated to derive from salivary glands (Mans et al. 2008 a,b). This could suggest, perhaps erroneously that tick lipocalins are limited to salivary glands and the tick feeding site. However, a number of lipocalins have been identified in whole-body derived EST libraries from *B. microplus* and *O. porcinus*, and has the potential to function in other facets of tick biology (Mans et al. 2008a). In addition, salivary gland-derived lipocalin transcripts were also detected in midguts from the hard tick *Ixodes ricinus* (Beaufays et al. 2008). However, this study is the first to assign a lipocalin-like molecule, savicalin, uniquely to non-salivary gland derived tissues that includes hemocytes, midgut and ovaries of the soft tick *O. savignyi* to the exclusion of the salivary gland. This is the first concrete proof that tick lipocalins can also function in ticks in a non-feeding capacity.

Savicalin groups within a clade formed by the three best hits obtained by BLASTP analysis and the translated EST sequence from *O. porcinus*. The support for this clade is quite high, but does not necessarily imply that these proteins are orthologous, as the expected species relationships for *Ornithodoros*, *Argas*, *Ixodes* and *Boophilus* are not recapitulated. It is of interest though that *Ixodes*, *Boophilus* and possibly the *O. porcinus* sequences derive from non-salivary gland tissues and could suggest that their orthologous/ paralogous relationships date back to a split between salivary and non-salivary gland derived sequences.

The fact that savicalin is absent from salivary glands, suggests that it does not function at the tick-host interface, as found for salivary gland-derived lipocalins. Expression patterns do, however, indicate that its expression is up-regulated during feeding and this could suggest a role in the post-feeding development of the tick, which could include processes such as blood-digestion, molting and embryogenesis (Sonenshine, 1991). Orthologs for this lipocalin could also be discovered once more lipocalins are described that are not expressed in the salivary glands. The function for savicalin has not yet been determined, but given its tissue distribution, expression patterns and being a lipocalin, it would be likely that it can act as a scavenger or transporter of bio-active molecules involved in post-feeding development of soft ticks or as an anti-microbial.

There are similarities between hemocytes, midguts and ovaries that might explain the distribution of savicalin in these tissues. Note that antibodies cross-react with hemocytes, midguts and ovaries (Magnarelli et al. 1991, Cristofoletti et al. 2005, Wang et al. 2007). This implies that hemocytes, midguts and ovaries share common antigenic determinants or that the same or similar proteins are expressed in these organs. It has also been shown that the number of hemocytes (nongranular cells and granule-scant basophilic granular cells) increases during ecdysis of fed nymphs and that basophilic granular cells increase in fed adult ticks and nymphs (Kadota et al. 2003). Up and down-regulation of savicalin could thus be closely associated by hemocyte numbers during ecdysis. Molecules known to influence tick development and ecdysis, include ecdysone and juvenile hormone.

Savicalin could also act as an anti-microbial by scavenging siderophores, as found for the mammalian lipocalin NGAL (Alpízar-Alpízar et al. 2009). Given its up-regulation in hemocytes upon bacterial challenge, savicalin might also play a critical role in the defense against harmful pathogens. In this regard, it should be noted that ecdysone and juvenile hormone influence the innate immune system of insects and could be potential ligands for savicalin (Meister and Richards 1996; Flatt et al. 2008; Figueiredo et al. 2006). Investigation into the ligand binding properties of savicalin is currently being pursued.

In conclusion, savicalin is widely expressed in tissues not directly involved in tick-host interactions. Expression does, however, relate to post-feeding events such as development and ecdysis or bacterial challenge. This would implicate savicalin in one of these crucial and inter-related processes and indicate that the role of lipocalins in tick biology is not limited to tick-host interaction.

Acknowledgements

We wish to thank the National Research Foundation and University of Pretoria for financial support.

References

Andersen JF, Gudderra NP, Francischetti IM, Ribeiro JM (2005) The role of salivary lipocalins in blood feeding by *Rhodnius prolixus*. Arch Insect Biochem Physiol 58: 97-105

Alpízar-Alpízar W, Laerum OD, Illemann M, Ramírez JA, Arias A, Malespín-Bendaña W, Ramírez V, Lund LR, Borregaard N, Nielsen BS (2009) Neutrophil gelatinase-associated lipocalin (NGAL/Lcn2) is upregulated in gastric mucosa infected with *Helicobacter pylori*. Virchows Arch 455: 225-233.

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol 215: 403-410.

Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 25: 3389-402

Beaufays J, Adam B, Decrem Y, Prévôt PP, Santini S, Brasseur R, Brossard M, Lins L, Vanhamme L, Godfroid E. (2008) Ixodes ricinus tick lipocalins: identification, cloning, phylogenetic analysis and biochemical characterization. PLoS One. 3: e3941.

Blackburn K, Goshe MB (2009) Mass spectrometry bioinformatics: Tools for navigating the proteomics landscape. Curr. Anal. Chem 5: 131-143.

Cowan SW, Jones TA, Newcomer ME (1990) Crystallographic refinement of human serum retinol binding protein at 2A resolution Proteins 8: 44–61.

Cristofoletti PT, Ribeiro AF, Terra WR (2005) The cathepsin L-like proteinases from the midgut of Tenebrio molitor larvae: sequence, properties, immunocytochemical localization and function. Insect Biochem Mol Biol 35: 883-901.

Figueiredo MB, Castro DP, S Nogueira NF, Garcia ES, Azambuja P (2006) Cellular immune response in *Rhodnius prolixus*: role of ecdysone in hemocyte phagocytosis. J Insect Physiol 52: 711-716.

Flatt T, Heyland A, Rus F, Porpiglia E, Sherlock C, Yamamoto R, Garbuzov A, Palli SR, Tatar M, Silverman N (2008) Hormonal regulation of the humoral innate immune response in *Drosophila melanogaster*. J Exp Biol 211: 2712-2724.

Flower DR, Attwood TK, North AC (1991) Mouse oncogene protein 24p3 is a member of the lipocalin protein family. Biochem Biophys Res Commun 180: 69–74.

Flower DR (2000) Experimentally determined lipocalin structures. Biochim Biophys Acta 1482: 46-56.

Flower DR, North AC, Sansom CE (2000) The lipocalin protein family: structural and sequence overview. Biochim Biophys Acta 1482: 9-24.

Francischetti IM, Mans BJ, Meng Z, Gudderra N, Veenstra TD, Pham VM, Ribeiro JM (2008a). An insight into the sialome of the soft tick, *Ornithodorus parkeri.* Insect Biochem Mol Biol 38: 1-21.

Francischetti IM, Meng Z, Mans BJ, Gudderra N, Hall M, Veenstra TD, Pham VM, Kotsyfakis M, Ribeiro JM (2008b). An insight into the salivary transcriptome and proteome of the soft tick and vector of epizootic bovine abortion, *Ornithodoros coriaceus*. J Proteomics 71: 493-512

Francischetti, IMB, Sa-Nunes A, Mans BJ, Santos IM, Ribeiro JMC (2009) The role of saliva in tick feeding. Front. Biosci 14: 2051-2088.

Guex N, Peitsch MC (1997) SWISS-MODEL and the Swiss-PdbViewer: An environment for comparative protein modeling. Electrophoresis 18: 2714-2723.

Guex N, Diemand A, Peitsch MC (1999) Protein modelling for all. Trends Biochem Sci 24: 364-367

Jeanmougin F, Thompson JD, Gouy M, Higgins DG, Gibson TJ (1998) Multiple sequence alignment with Clustal X. Trends Biochem Sci 23: 403-405

Johns R, Sonenshine DE, Hynes WL (1998) Control of bacterial infections in the hard tick Dermacentor variabilis (Acari: Ixodidae): evidence for the existence of antimicrobial proteins in tick hemolymph. J Med Entomol 35: 458-464

Kadota K, Walter S, Claveria FG, Igarashi I, Taylor D, Fujisaki K (2003) Morphological and populational characteristics of hemocytes of Ornithodoros moubata nymphs during the ecdysial phase. J Med Entomol 2003; 40: 770-776

Kayser H, Krull-Savage U, Rilk-van Gessel R (2005) Developmental profiles of 5-aminolevulinate, porphobilinogen and porphobilinogen synthase activity in Pieris brassicae related to the synthesis of the bilin-binding protein. Insect Biochem Mol Biol 35:165-174.

Kelley LA and Sternberg MJE (2009) Protein structure prediction on the web: a case study using the Phyre server. Nature Protocols 4: 363 - 371.

Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG (2007) Clustal W and Clustal X version 2.0. Bioinformatics 23: 2947-2948.

Lingner J, Hughes TR, Shevchenko A, Mann M, Lundblad V, Cech TR (1997) Reverse transcriptase motifs in the catalytic subunit of telomerase. Science 276: 561-567.

Madden RD, Sauer JR, Dillwith JW (2002) A proteomics approach to characterizing tick salivary secretions. Exp Appl Acarol 28: 77-87.

Magnarelli LA, Andreadis TG, Stafford KC 3rd, Holland CJ (1991) *Rickettsiae* and *Borrelia burgdorferi* in ixodid ticks. J Clin Microbiol 29: 2798-2804.

Mans BJ, Venter JD, Very PJ, Louw AI, Neitz AW (2001) Identification of putative proteins involved in granule biogenesis of tick salivary glands. Electrophoresis 22: 1739-1746.

Mans BJ, Steinmann CM, Venter JD, Louw AI, Neitz AW (2002) Pathogenic mechanisms of sand tampan toxicoses induced by the tick, *Ornithodoros savignyi.* Toxicon 40: 1007-1016.

Mans BJ, Louw AI, Neitz AW (2003) The major tick salivary gland proteins and toxins from the soft tick, *Ornithodoros savignyi*, are part of the tick Lipocalin family: implications for the origins of tick toxicoses. Mol Biol Evol 20: 1158-1167.

Mans, BJ, Neitz AW (2004a) Exon-intron structure of outlier tick lipocalins indicate a monophyletic origin within the larger lipocalin family. Insect Biochem Mol Biol 34: 585-594.

Mans BJ, Neitz AW (2004b) Molecular crowding as a mechanism for tick secretory granule biogenesis. Insect Biochem Mol Biol 34: 1187-1193.

Mans BJ (2005) Tick histamine-binding proteins and related lipocalins: potential as therapeutic agents. Curr Opin Investig Drugs 6: 1131-1135.

Mans BJ, Andersen JF, Francischetti IM, Valenzuela JG, Schwan TG, Pham VM, Garfield MK, Hammer CH, Ribeiro JM (2008a) Comparative sialomics between hard and soft ticks: implications for the evolution of blood-feeding behavior. Insect Biochem Mol Biol 38: 42-58.

Mans BJ, Ribeiro JM, Andersen JF (2008b) Structure, function, and evolution of biogenic amine-binding proteins in soft ticks. J Biol Chem 283: 18721-18733.

Mans BJ, Ribeiro JM (2008a) A novel clade of cysteinyl leukotriene scavengers in soft ticks. Insect Biochem Mol Biol 38: 862-870.

Mans BJ, Ribeiro JM (2008b) Function, mechanism and evolution of the moubatin-clade of soft tick lipocalins. Insect Biochem Mol Biol 38: 841-852

Marchler-Bauer A, Anderson JB, Chitsaz F, Derbyshire MK, DeWeese-Scott C, Fong JH, Geer LY, Geer RC, Gonzales NR, Gwadz M, He S, Hurwitz DI, Jackson JD, Ke Z, Lanczycki CJ, Liebert CA, Liu C, Lu F, Lu S, Marchler GH, Mullokandov M, Song JS, Tasneem A, Thanki N, Yamashita RA, Zhang D, Zhang N, Bryant SH. (2009) CDD: specific functional annotation with the Conserved Domain Database. Nucleic Acids Res. 37: D205-D210.

Mauchamp B, Royer C, Garel A, Jalabert A, Da Rocha M, Grenier AM, Labas V, Vinh J, Mita K, Kadono K, Chavancy G. (2006) Polycalin (chlorophyllid A binding protein): a novel, very large fluorescent lipocalin from the midgut of the domestic silkworm *Bombyx mori L*. Insect Biochem Mol Biol 36: 623-633.
Meister M, Richards G. (1996) Ecdysone and insect immunity: the maturation of the inducibility of the diptericin gene in *Drosophila* larvae. Insect Biochem Mol Biol 26: 155-160.

Montfort WR, Weichsel A, Andersen JF (2000) Nitrophorins and related antihemostatic lipocalins from *Rhodnius prolixus* and other blood-sucking arthropods. Biochim Biophys Acta 1482: 110-118.

Nagata A, Igarashi M, Toh H, Urade Y, Hayaishi O (1992) Structural organization of the gene for prostaglandin D synthase in the rat brain. Proc Natl Acad Sci USA 89: 5376–5380.

Nunn MA, Sharma A, Paesen GC, Adamson S, Lissina O, Willis AC, Nuttall PA (2005) Complement inhibitor of C5 activation from the soft tick *Ornithodoros moubata*. J Immunol 174: 2084-2091.

Oleaga A, Escudero-Poblacion A, Camafeita E, Perez-Sanchez R (2007) A proteomic approach to the identification of salivary proteins from the argasid ticks *Ornithodoros moubata* and *Ornithodoros erraticus*. Insect Biochem. Mol Biol 37: 1149-1159.

Paesen GC, Adams PL, Harlos K, Nuttall PA, Stuart DI (1999) Tick histamine-binding proteins: isolation, cloning, and three-dimensional structure. Mol Cell 3: 661-671.

Paesen GC, Adams PL, Nuttall PA, Stuart DL (2000) Tick histamine-binding proteins: lipocalins with a second binding cavity. Biochim Biophys Acta 1482: 92-101.

Peitsch MC, Tschopp J, Jenne DE, Haefliger JA (1991) Structural and functional characterization of complement C8 gamma, a member of the lipocalin protein family. Mol Immunol 28: 123–131.

Pervaiz S, Brew K (1987) Homology and structure-function correlations between alpha 1-acid glycoprotein and serum retinol-binding protein and its relatives. FASEB J 1: 209–214.

Ribeiro, JM, Francischetti, IM (2003) Role of arthropod saliva in blood feeding: sialome and post-sialome perspectives. Annu. Rev. Entomol 48, 73-88.

Ribeiro JM, Alarcon-Chaidez F, Francischetti IM, Mans BJ, Mather TN, Valenzuela JG, Wikel SK (2006) An annotated catalog of salivary gland transcripts from *Ixodes scapularis* ticks. Insect Biochem Mol Biol 36: 111-129

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

Sánchez D, Ganfornina MD, Bastiani MJ (1995) Developmental expression of the lipocalin Lazarillo and its role in axonal pathfinding in the grasshopper embryo. Development 121: 135-147.

Sánchez D, Ganfornina MD, Bastiani MJ (2000) Lazarillo, a neuronal lipocalin in grasshoppers with a role in axon guidance. Biochim Biophys Acta 1482: 102-109.

Sánchez D, López-Arias B, Torroja L, Canal I, Wang X, Bastiani MJ, Ganfornina MD. (2006) Loss of glial lazarillo, a homolog of apolipoprotein D, reduces lifespan and stress resistance in Drosophila. Curr Biol 16: 680-686.

Sangamnatdej S, Paesen GC, Slovak M, Nuttall PA (2002) A high affinity serotonin- and histamine-binding lipocalin from tick saliva. Insect Mol Biol 11: 79-86.

Schwan EV, Hutton D, Shields KJ, Townson S (1991) Artificial feeding and successful reproduction in Ornithodoros moubata moubata (Murray, 1877) (Acarina: Argasidae). Exp Appl Acarol 13: 107-115

Shevchenko A, Sunyaev S, Loboda A, Shevchenko A, Bork P, Ens W, Standing KG (2001) Charting the proteomes of organisms with unsequenced genomes by MALDI-quadrupole time-of-flight mass spectrometry and BLAST homology searching. Anal Chem 73: 1917-1926.

Shevchenko A, Valcu CM, Junqueira M (2009) Tools for exploring the proteomosphere. J Proteomics 72, 137-144.

Skerra A (2000) Lipocalins as a scaffold. Biochim Biophys Acta 1482: 337-350.

Sonenshine DE (1991) Biology of Ticks. Oxford University Press, New York/ Oxford :3-12

Stutzer C, Mans BJ, Gaspar AR, Neitz AW, Maritz-Olivier C (2009) *Ornithodoros savignyi*: soft tick apyrase belongs to the 5'-nucleotidase family. Exp Parasitol 122: 318-27

Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Mol Biol Evol 24:1596-1599.

Valenzuela JG, Francischetti IM, Pham VM, Garfield MK, Mather TN, Ribeiro JM (2002) Exploring the sialome of the tick *Ixodes scapularis*. J Exp Biol 205 2843-2864.

Wang L, Li Z, Du C, Chen W, Pang Y (2007) Characterization and expression of a cecropin-like gene from *Helicoverpa armigera*. Comp Biochem Physiol B Biochem Mol Biol 148: 417-425.

Weichsel A, Andersen JF, Champagne DE, Walker FA, Montfort WR (1998) Crystal structures of a nitric oxide transport protein from a blood-sucking insect. Nat Struct Biol 5:304-309.

List of figures

Fig 1: Nucleotide and deduced amino acid sequence of the cDNA encoding savicalin. The mature peptide is indicated in bold type. The arrow indicates the putative cleavage site of the signal peptide. The stop codon is marked with an astrerisk. The polyadenylation signal is underlined. Nucleotide sequence data has been submitted to the database (accession number: GU733379)

Fig 2: Multiple sequence alignments of tick lipocalins. Alignment of savicalin with the following molecules and their accession codes: Monomine (114152936), Monotonin (114152976) and Am-33 (114152974) from the soft tick, *Argas monolakensis;* SHBP (18032205) from the hard tick*, Dermacentor reticulatus*; Ir-LBP (219935277) from the hard tick, *Ixodes ricinus;* M.like3 (Moubatin like 3, 149287030) from the soft tick, *Ornithodoros parkeri;* HBP (7767032) from the hard tick, *Rhipicephalus appendiculatus;* Moubatin (159945) and OMCI (49409517) from the soft tick, *Ornithodoros moubata;* TSGPs (TSGP1: 25991387, TSGP2: 25991389, TSGP3: 25991391, TSGP4: 25991438) from the soft tick*, Ornithodoros savignyi;* O.porc.lip *(Ornithodoros* porcinus lipocalin, 29779506) from the soft tick, *Ornithodoros porcinus*; I.scap.lip *(Ixodes scapularis*

lipocalin, 241679301) from the hard tick, *Ixodes scapularis*. Secondary structures based on the SWISS model of savicalin is boxed in black as α-helices (a) and β-sheets (A~H). Conserved cysteine residues found in both hard and soft ticks with predicted disulphide bonds are indicated with solid black line. Solid grey line indicates share TSGP 4 fold disulphide bond. Dotted grey line indicates disulphide bond share *I. scapularis* lipocalin. Light grey line indicates disulphide bond unique only to savicalin. Red rectangular box indicates biogenic amine binding motif compared to CL-VLG-C sequence obtained from savicalin.

Fig 3: Structural modeling of savicalin. Cysteine residues are indicated with spacefill spheres with their corresponding disulphide bonds. The RMSD value compared with the modeling template HBP-2 is also indicated.

Fig 4: Phylogenetic analysis of the tick lipocalin family. Lipocalins were retrieved from non-redundant database by BLAST analysis of savicalin. Neighbor joining (NJ) analysis was conducted using MEGA version 4.0. Reliability of the inferred tree was evaluated by bootstrap analysis (100 000 replicates). Sequences are described by a species designation (A. mon: *Argas monolakensis*; Am. Amer: *Amblyomma americanum*; B. micro: *Boophilus microplus*; D. recti: *Dermacentor reticulatus*; I. paci: *Ixodes pacificus*; I. scap: *Ixodes scapularis*; I. rici: *Ixodes ricinus*; O. cori: *Ornithodoros coriaceus*; O. park: *Ornithodoros parkeri*; O. moub: *Ornithodoros moubata*; O. porc: *Ornithodoros porcinus*; O. sav: *Ornithodoros savignyi*; R. app: *Rhipicephalus appendiculatus*). Savicalin is shown in bold red. Red circle indicate lipocalins that have more than 8 cysteine residues. Grey dotted line divides classes as of ticks. Bold black: Functions of lipocalins have been determined experimentally with ligands indicated in different colours (Blue: Serotonin; Green: Histamine; Light purple: Leukotriene C4, Purple: Leukotriene B4; Light blue: Complement C5; Light Green: Thromboxane A2).

Fig 5: Transcriptional profile of savicalin. Tissue distribution was analyzed by RT-PCR in the hemolymph, salivary glands, midguts and ovaries of the ticks (a) 1day after hemocoelic inoculation with heat killed *B. subtilis* or saline and (b) 1 day/ 10 days after artificial feeding of ticks with either native or *B. subtilis* infected blood. Actin is shown as the internal standard.

tggacggattactacgaccgcatgaaccttcttcggcaacgccccgaaaggattacttcttgc
C
gaccagtcgcaaagtgcacaaagtctgacggacatcagcgccaaacgaacgagaatgaag
M K
gatcttttttgtgatcggcactcttttggcgtgcattttatgcattcctgctgaccctctg
D L F V I G T L L A C I L C I P A D P L
aagcgcgatctttcgcgtcgagagacagatgaagtttcaggaaccaccgcaggaagcaca
K R D L S R R E T D E V S G T T A G S T
gtcgacttcacgaagtttcaggatgcgggcaagtgcttcagtactccaacaacgtactat
V D F T K F Q D A G K C F S T P T T Y Y
gggtactacaggtcttatctagatgaccctgtgttggccgacatgtgtatcgctgcgaag
G Y Y R S Y L D D P V L A D M C I A A K
cacatcggcgtccctgagaatggaacacttacgttcgatttcgtatacgagaggaacgga
H I G V P E N G T L T F D F V Y E R N G
gaaaccatttttagaccctcagttcatatgtacctctgttctactccggactcagatgtt
E T I F R P S V H M Y L C S T P D S D V
caagactgcttccggatgccagttcttgatgttccaggcgctgaatatgtgcagtatgac
Q D C F R M P V L D V P G A E Y V Q Y D
ttcctctacttcaaatgtggcgactgcctattaatgtggatgccacagtattctgagggt
F L Y F K C G D C L L M W M P Q Y S E G
tcttgcgttttaggagtgacggacctgcacaacattgatgaatcatgtcaccaggcttat
S C V L G V T D L H N I D E S C H Q A Y
gattactactgctccccgaacaagttctttattcacgatgaagacatgtgtccggaagtc
D Y Y C S P N K F F I H D E D M C P E V
tacagtgctcttcagtaattcccagcgttacaaatgtgctggtcatttggtcatttaga
Y S A L Q *
gggattcaataaaacgcacacgtcaaataaggtaaaaaaaaaaaaaaaaa

Fig. 1

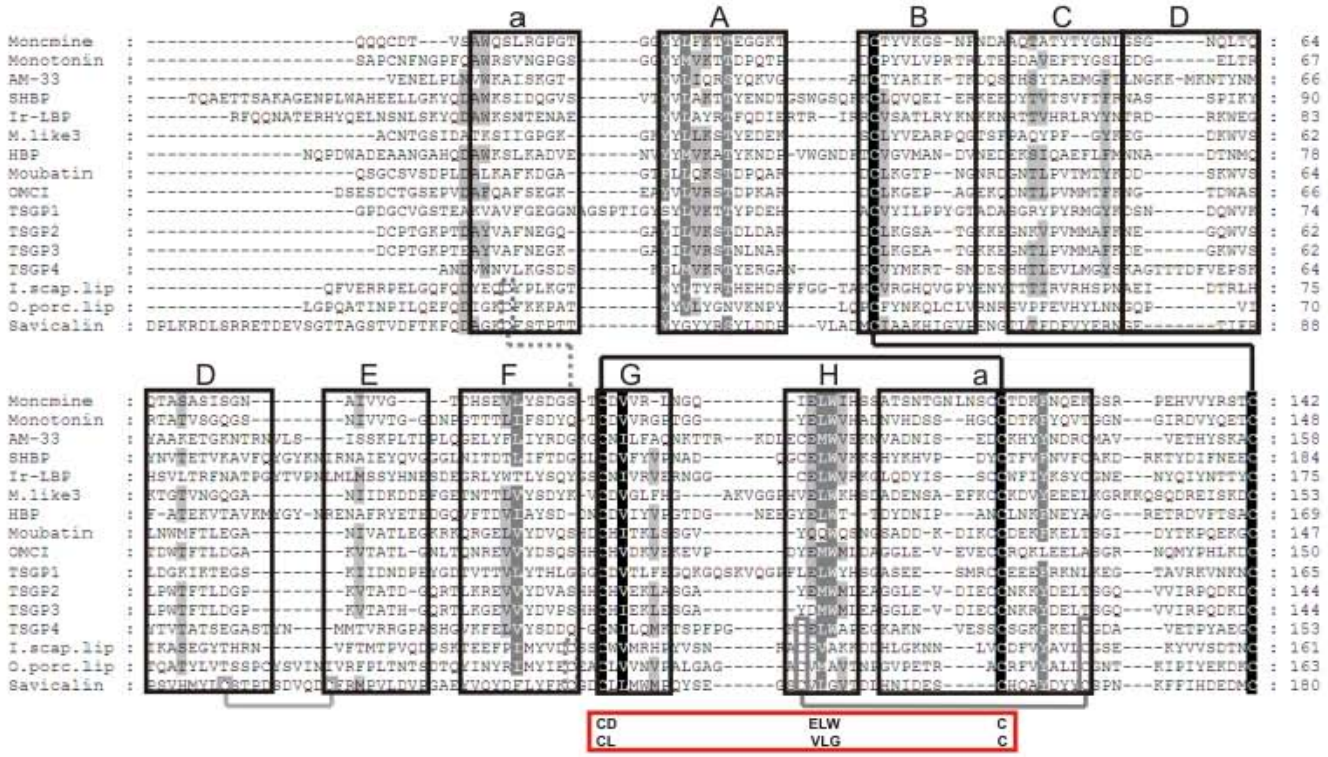
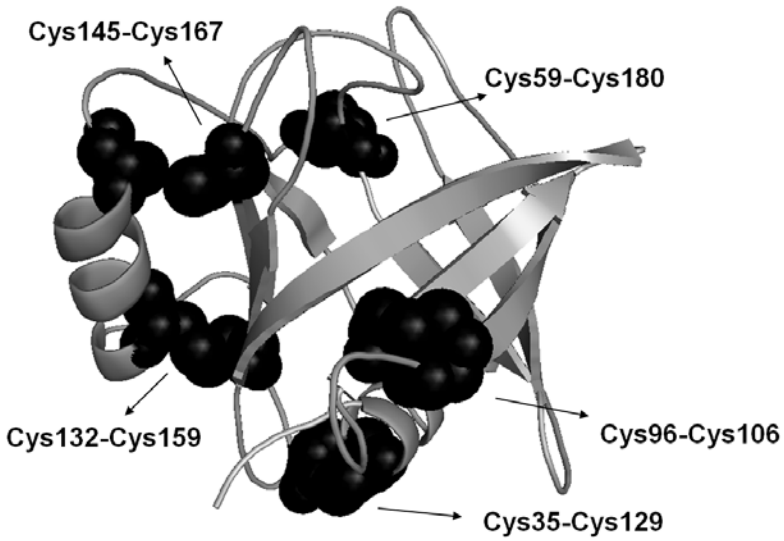

a
A
B
C
D
D
E
F
G
H
a
Moncmine
Monotonin
AM-33
SHBP
Ir-LBP
M.like3
HBP
Moubatin
OMCI
TSGP1
TSGP2
TSGP3
TSGP4
I.scap.lip
O.porc.lip
Savicalin
CD ELW C
CL VLG C

Fig. 2



RMSD~0.45Å
153 amino acid residues
Alpha carbon

Fig. 3

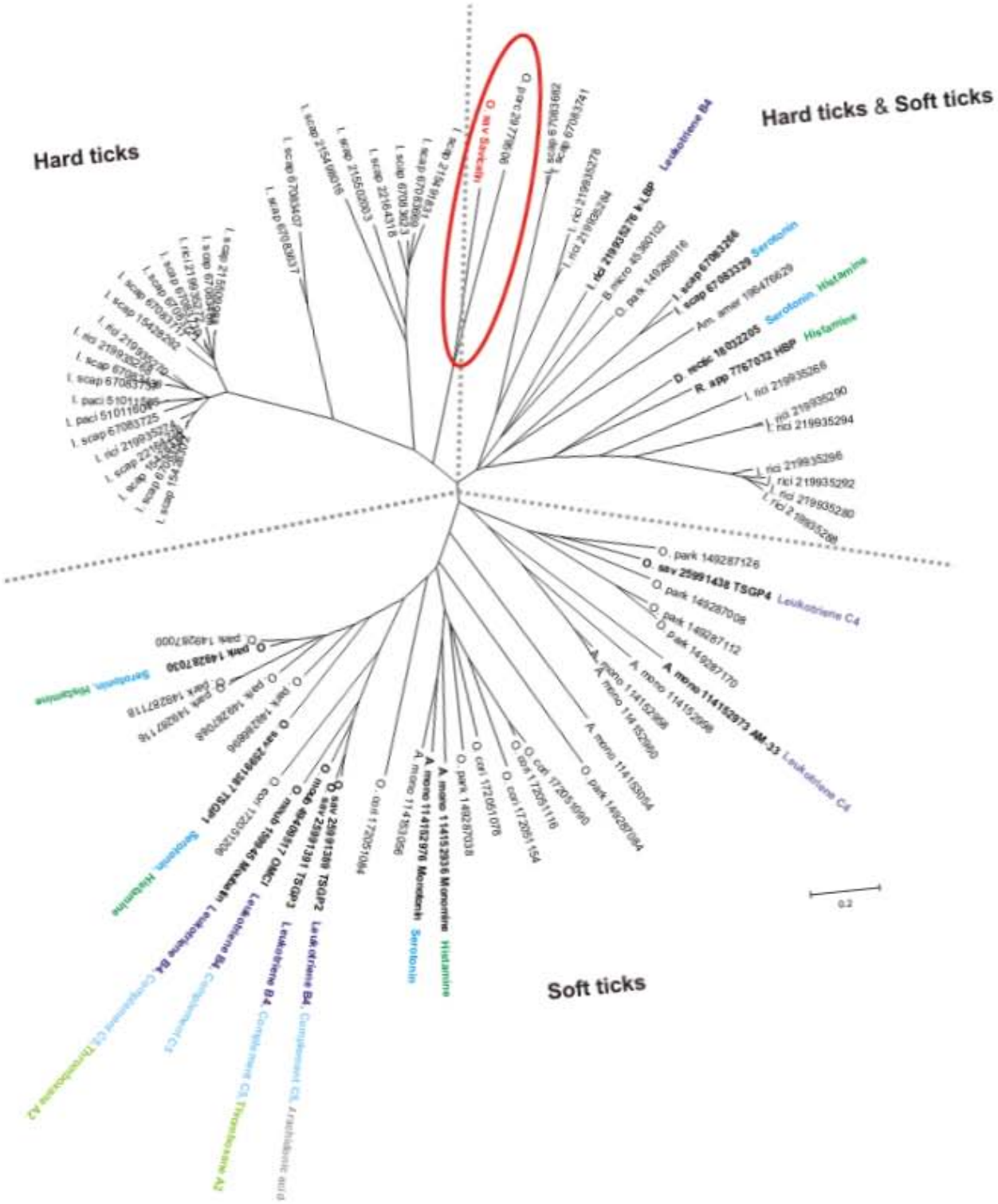

Hard ticks
Hard ticks & Soft ticks
Soft ticks
O. sav Salivary
Leukotriene B4
Serotonin
Serotonin, Histamine
Histamine
Leukotriene C4
Leukotriene B4, Complement C3
Leukotriene B4, Complement C3, Thromboxane A2
Complement C3
Complement C3, Thromboxane A2
Leukotriene B4, Complement C3, Arachidonic acid
Serotonin, Histamine
0.2

Fig. 4

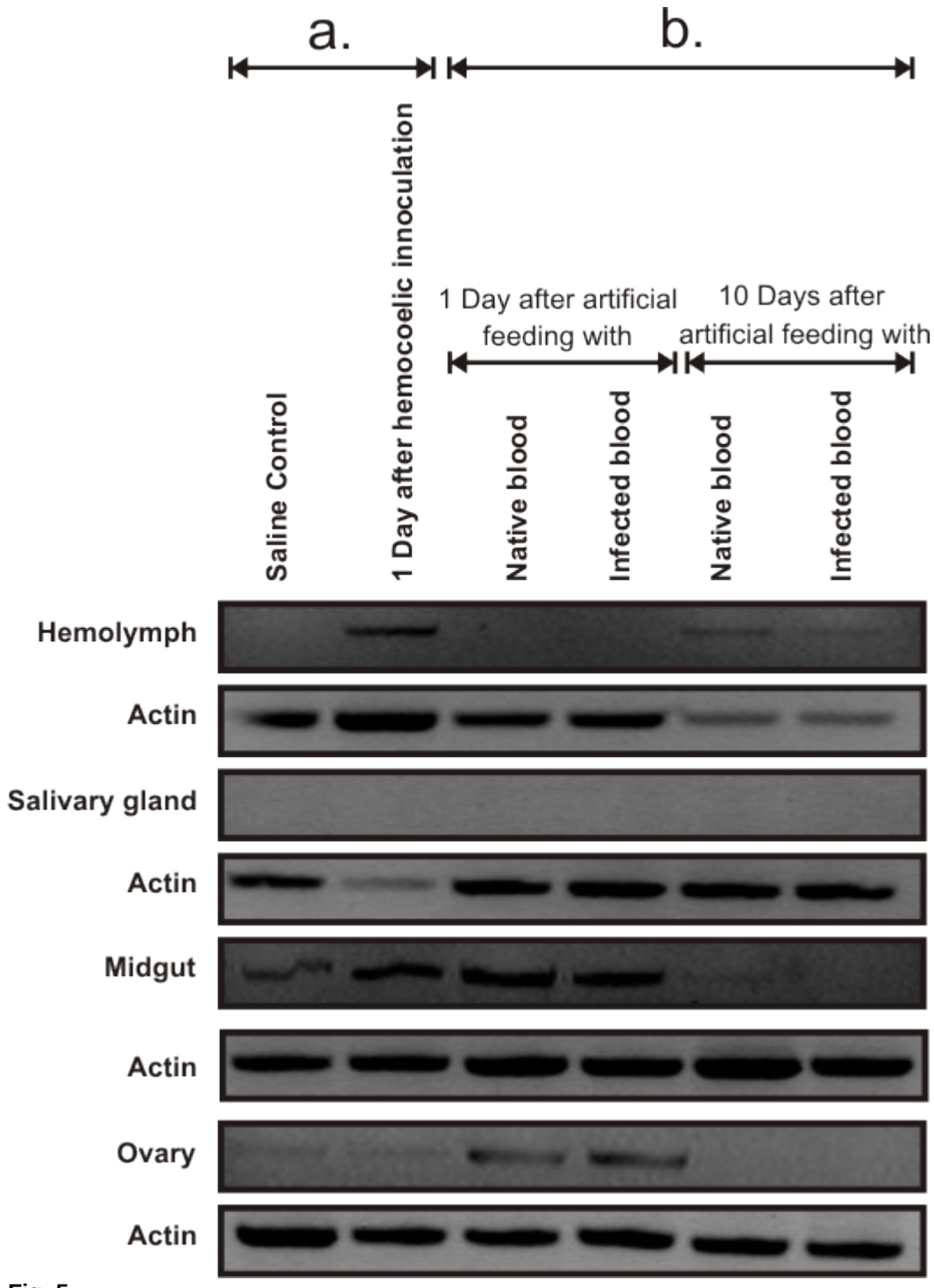

a.
b.
1 Day after artificial feeding with
10 Days after artificial feeding with
Saline Control
1 Day after hemocoelic innoculation
Native blood
Infected blood
Native blood
Infected blood
Hemolymph
Actin
Salivary gland
Actin
Midgut
Actin
Ovary
Actin

Fig. 5