Kazal-type serine proteinase inhibitors in the midgut of *Phlebotomus papatasi*

Leah Theresa Sigle, Marcelo Ramalho-Ortigão/+

Department of Entomology, Kansas State University, Manhattan, KS, USA

Sandflies (Diptera: Psychodidae) are important disease vectors of parasites of the genus Leishmania*, as well as bacteria and viruses. Following studies of the midgut transcriptome of* Phlebotomus papatasi*, the principal vector of* Leishmania major*, two non-classical Kazal-type serine proteinase inhibitors were identified (*PpKzl1 *and* PpKzl2*). Analyses of expression profiles indicated that* PpKzl1 *and* PpKzl2 *transcripts are both regulated by blood-feeding in the midgut of* P. papatasi *and are also expressed in males, larva and pupa. We expressed a recombinant PpKzl2 in a mammalian expression system (CHO-S free style cells) that was applied to in vitro studies to assess serine proteinase inhibition. Recombinant PpKzl2 inhibited α-chymotrypsin to 9.4% residual activity and also inhibited α-thrombin and trypsin to 33.5% and 63.9% residual activity, suggesting that native PpKzl2 is an active serine proteinase inhibitor and likely involved in regulating digestive enzymes in the midgut. Early stages of* Leishmania *are susceptible to killing by digestive proteinases in the sandfly midgut. Thus, characterising serine proteinase inhibitors may provide new targets and strategies to prevent transmission of* Leishmania.

Key words: Diptera - sandflies - *Phlebotomus* - Kazal-type inhibitors - midgut - blood meal digestion

In arthropods, serine proteinases are involved in digestion, coagulation, phenoloxidase activation and other immune responses. Regulation of these enzymes by serine proteinase inhibitors is critical for maintaining homeostasis (Kanost 1999, Jiang & Kanost 2000, di Cera 2009). Several serine proteinase inhibitors have been identified in blood-feeding arthropods and linked to inhibition of thrombin and other components of the coagulation cascade to facilitate fluidity in the mouth parts and midgut following blood-feeding on a host [reviewed by Tanaka-Azevedo et al. (2010)]. Many of these thrombin inhibitors belong to the family of Kazal-type serine proteinase inhibitors.

The first Kazal-type thrombin inhibitor identified in a haematophagous insect was from *Rhodnius prolixus* (Friedrich et al. 1993). Since then, proteins containing Kazal-type domains have been identified in other triatomines as well as in many other blood-feeding arthropods including flies, mosquitoes and ticks (Mende et al. 1999, Campos et al. 2002, Takáč et al. 2006, Zhou et al. 2006, Araujo et al. 2007, Mulenga et al. 2007, Ribeiro et al. 2007, Meiser et al. 2010). Kazal-type inhibitors are known to inhibit a range of serine proteinases. Native Kazals from blood-feeding arthropods inhibit thrombin, trypsin, factor XIIa, subtilisin A, elastase, chymotrypsin and plasmin (Friedrich et al. 1993, Campos et al. 2002, 2004, Lovato et al. 2006, Meiser et al. 2010).

doi: 10.1590/0074-02761080620130001
Financial support: NIAID (R01AI074691)
+ Corresponding author: mortigao@ksu.edu
Received 27 February 2013
Accepted 2 July 2013

Kazal-type domains are characteristically 40-60 amino acids long and inhibitors may contain single or multiple active domains. Six cysteine residues forming three disulfide bridges, C_1:C_5, C_2:C_4, C_3:C_6, distinguish the conserved structure within classical and non-classical Kazal-type domains. The predicted reactive site, P1 amino acid residue, is located at position C_2-X-P1 and determines specificity within Kazal-type inhibitors (Kanost 1999). Within the domain, outside of the conserved cysteine residues, there are high amounts of variability in other amino acid residues (Rimphanitchayakit & Tassanakajon 2010).

Phlebotomine sandflies (Diptera: Psychodidae) are vectors of viruses, bacteria and parasites of the genus *Leishmania*. Transmission of *Leishmania* to suitable vertebrate hosts generally occurs during blood-feeding through the bite site of an infected sandfly vector [reviewed by Ramalho-Ortigão et al. (2010)].

Midgut transcriptome analyses of *Phlebotomus papatasi*, the principal vector of *Leishmania major*, revealed two Kazal-type serine proteinase inhibitors, *PpKzl1* and *PpKzl2* (Ramalho-Ortigão et al. 2007). These were the first Kazal-type serine proteinase inhibitors identified from sandflies. The mature *PpKzl1* cDNA is 231 base pairs (bp) encoding a 77 amino acid protein containing a single Kazal-type domain (GenBank ID: EU045342). The mature *PpKzl2* cDNA is 267 bp encoding an 89 amino acid protein (GenBank ID: JX171681). PpKzl1 and PpKzl2 have only 28% identity and 42% similarity in amino acid sequences (Ramalho-Ortigão et al. 2007). Both PpKzl1 and PpKzl2 have predicted signal peptides, suggesting that they are secreted in the midgut.

We are interested in the role of these proteins in *P. papatas*i as inhibitors of serine proteinases and their potential effects on blood digestion. We have analysed deduced sequences of the PpKzl1 and PpKzl2 for predicted activity and similarity, evaluated the expression

of *PpKzl1* and *PpKzl2* in developmental stages, adult female midguts and whole adult males and conducted in vitro analysis of inhibition activity of a recombinant PpKzl2 protein.

MATERIALS AND METHODS

Sandflies - *P. papatasi* Israel strain was reared in the Biology of Disease Vectors laboratory at the Department of Entomology, Kansas State University. Flies were maintained on 30% sucrose solution at 27ºC and 70% humidity with 12 h light and dark cycles. For blood feeding, sandflies were allowed to feed approximately 30 min on a BALB/c mouse anesthetised with 3 mg ketamine (Ketaset, Fort Dodge Animal Health, Fort Dodge, IA, USA) and 0.12 mg xylazine (AnaSed, Acorn Inc, Decatur, IL, USA) per mouse (100 mg/kg of ketamine and 4 mg/kg of xylazine). Use of animals was preapproved by the Kansas State University Institutional Animal Care and Use Committee under protocols 2747, 2748 and 2749. Infectious blood meals contained *L. major* amastigotes and were offered artificially, while simultaneously a control set of sandflies were fed on uninfected blood as previously described (Coutinho-Abreu et al. 2010a).

At 20 h post-blood meal (PBM) all blood-fed flies were briefly anesthetised with CO_2 and examined under a dissecting microscope. Fully fed flies (i.e., abdomen fully distended) of similar size were selected for dissection. Midguts were dissected in 30 µL 1X phosphate buffered saline RNAse free with ELIMINase (Fisher, Scientific, Pittsburgh, PA, USA) treated tools and equipment. Dissected midguts were then transferred to 50 µL of RNA later (Qiagen, Valencia, CA, USA), homogenised with a hand-held homogeniser for approximately 20 s and placed at -80ºC.

Sequence analysis - *PpKzl1* and *PpKzl2* were previously identified from *P. papatasi* cDNA midgut libraries (Ramalho-Ortigão et al. 2007). Molecular weights and isoelectric points (pI) were predicted using the Swiss Institute of Bioinformatics ExPASy tools (Gasteiger et al. 2003). Sequences similar to PpKzl1 and PpKzl2 were identified in National Center for Biotechnology Information using BLASTP for the non-redundant protein database (Altschul et al. 1997). The conserved six cysteine domain in PpKzl1 and PpKzl2 was used for multiple sequence alignments (MSA) with selected sequences from blast results. Protein sequence alignments were performed using CLUSTALW2 (Larkin et al. 2007) and manual edits were performed in Jalview version 2 (Waterhouse et al. 2009). A *Lutzomyia longipalpis* Kazal2 contig (69116) was identified using BLAST searching for homologs of PpKzl2 in the *L. longipalpis* Llon 0.1 preliminary Genome Assembly on the Baylor College of Medicine Human Genome Sequencing Center website (hgsc.bcm.tmc.edu/project-species-i-Lutzomyia_longipalpis.hgsc). The sequence was translated with Swiss Institute of Bioinformatics ExPASy (Gasteiger et al. 2003).

RNA extraction and cDNA synthesis - Total RNA was extracted from whole sample pools or individual dissected midguts using the RNeasy Mini Kit (Qiagen) and eluted in 40 µL of RNase-free water. Three RNAs were obtained for each developmental stage from pools of 20 eggs, 10 L_1 larvae and five each for stages L_2, L_3, L_4 and pupae. Extracted RNA was treated with TURBO DNase (Ambion, Austin, TX, USA) to eliminate any residual genomic DNA. Up to 100 ng of each RNA was used for first strand cDNA synthesis and was added to 3.3 µM oligo-dT_{20} primer, 0.67 mM deoxynucleotide triphosphates and RNase-free water to total volume of 15 µL. Samples were incubated at 65ºC for 5 min and then placed on ice for 1 min. Addition of 4 µL of 5X SuperScript III Reverse Transcriptase First-Strand Buffer, 5 mM DTT, 0.5 µL RNaseOUT (40 units/µL) and 1 µL of SuperScript III Reverse Transcriptase (200 units/µL) (Invitrogen, Carlsbad, CA, USA) was followed with 1 h incubation at 50ºC. All cDNA was stored at -20ºC.

Real-time polymerase chain reaction (RT-PCR) - *PpKzl1* and *PpKzl2* relative expression was analysed in non-blood-fed and blood-fed adult female sandflies. Individual midguts were dissected from non-blood-fed flies (0 h) and blood-fed flies at 24 h, 48 h and 72 h PBM. Total RNA was extracted from individual midguts and used for first-strand cDNA synthesis. RT-PCR was carried out on an Eppendorf Mastercycler ep Realplex[4] in 8 µL reactions. Forward and reverse 0.3 µM primers (Table) were mixed with 4 µL iQ SYBR green Super-

TABLE

Complete list of primers

Primer	Primer sequence 5'-3' forward	Primer sequence 5'-3' reverse	Annealing (°C)	PCR
PpKzl859	GCACCAGCCCAAAAGACC	TCACTGCAATCTGATGGCGC	56.5	PCR
VR1020	ACAGGAGTCCAGGGCTGGAGAGAA	AGTGGCACCTTCCAGGGTCAAGGA	49	PCR
PpKzl2-R-His	GCACCAGCCCAAAAGACC	His tag[a]-CTGCAATCTGATGGCGC	60	PCR[b]
PpKzl1_137	AGAGCGTTACCTGTCCTTG	CCAGCGAATACTGAGGTTC	58	RT-PCR
PpKzl2_152	AATGAATGTCTGAAGGCCTG	CCTTGGGATTTCACCTCCC	58	RT-PCR
Pp40S_S3_136	GGACAGAAATCATCATCATG	CCTTTTCAGCGTACAGCTC	58	RT-PCR

a: His tag-TCAGTGGTGATGGTGATGATG; *b*: touchdown polymerase chain reaction (PCR); RT: real-time.

mix (BioRad, Hercules, CA, USA) and added to 0.2 µL cDNA and 3.32 µL molecular grade water (Invitrogen). All cDNA samples were run in duplicate for *PpKzl1* and *PpKzl2* and in parallel for 40S ribosomal protein S3 (GenBank accession FG113203). Reactions were carried out 40 cycles of 95ºC/30 s, 58ºC/1 min and 72ºC/30 s, followed by 95ºC/15 s, 60ºC/15 s and a melt curve up to 95ºC/20 min. C_T values from the Realplex Software were used for expression analysis.

Expression levels of mRNA were calculated with the comparative C_T method as previously described (Coutinho-Abreu et al. 2010b). Briefly, C_T values were normalised to the expression of a non-regulated internal control gene, 40S ribosomal protein S3 and then normalised to a calibrator. Calibrators for analysis of temporal, developmental and infected expression were mean averages of expression in 0 h, eggs and non-infected blood-fed samples respectively. Comparative C_T method: $\Delta\Delta C_T$ = [ΔC_T Variant X Sample] – [average (ΔC_T Calibrator Samples)], where variant X equals time points or tissue type. Fold change was calculated by $2^{-\Delta\Delta C}{}_C$ (Livak & Schmittgen 2001). Mean fold change of at least five individual samples or three pools were graphed for each time point or tissue. Distribution of the data was tested with the Kolmogorov-Smirnov test for normality and Levene's test for equality of variance. Nonparametric data was logarithmically transformed for statistical analysis. Data was evaluated with one-way analysis of variance and a parametric t test with the Bonferroni correction for multiple comparisons. For temporal expression profiles of *L. major* infected sandflies, statistical analysis used two-tailed unpaired t tests for parametric analysis and the two-tailed Mann-Whitney *U* test for nonparametric statistical comparisons. Prism 5 Software (GraphPad, La Jolla, CA, USA) was used for all graphing and statistical analysis.

Recombinant protein expression and purification - The mature (minus signal peptide) *PpKzl2* cDNA was amplified using the forward primer PpKzl859 and the reverse primer PpKzl2-R-His containing a 6X-His tag on its 3' end (Table), touchdown reverse transcriptase PCR was performed as follows, 95ºC/3 min, three cycles of 94ºC/1 min, 72ºC/1 min, three cycles of 94ºC/1 min, 68ºC/1 min, 72ºC/1 min, five cycles of 94ºC/1 min, 62ºC/1 min, 72ºC/1 min, 25 cycles of 94ºC/1 min, 60ºC 1 min, 72ºC 1 min, finished with 72ºC 5 min. Two microlitres of the PCR product was separated on an agarose gel for analysis and to assess concentration. The mature *PpKzl2* was cloned into VR1020-TOPO vector as described previously (Ramalho-Ortigão et al. 2005, Oliveira et al. 2006). Insert-containing clones were screened by PCR (Table) and orientation was confirmed by sequencing. Plasmid purification was as described by Oliveira et al. (2006). Final concentration was 2.5 mg/mL and plasmid sequence was confirmed by sequencing.

The recombinant rPpKzl2 was expressed in CHO-S free style cells, following transfection using 37.5 µg of purified plasmid following the manufacturer's protocol (Invitrogen). Transfected CHO supernatant was collected after 72 h of culture, concentrated using a 3 kDa cut-off Centricon filter (Milipore, Billerica, MA, USA) and purified by nickel-nitrilotriacetic acid chromatography with a gravity flow column. The column was washed with 15 mL of 20 mM sodium phosphate buffer-300 mM sodium chloride-20 mM imidazole, eluted with 5 mL 20 mM sodium phosphate buffer-300 mM sodium chloride-300 mM imidazole and the eluted rPpKzl2 was concentrated to 1.5 µg/µL. Two hundred and fifty nanograms of protein were analysed by sodium dodecyl sulfate polyacrylamide gel electrophoresis using 4-12% reducing Bis-Tris NuPAGE pre-cast gel purchased from Invitrogen. The protein was transferred to nitrocellulose and incubated with anti-His antibody (Santa Cruz, Santa Cruz, CA, USA) overnight at 4ºC and followed by three washes of 10 min each in tris buffered saline buffer with 0.1% Tween-20 (TBS-T). The blot was incubated with anti-mouse antibody conjugated to alkaline phosphatase (Promega, Madison, WI, USA) diluted 1:10,000 in TBS-T for 1 h at room temperature and washed in TBS-T. The protein bands were visualised using the Western Blue substrate (Promega).

Inhibition assays - The inhibition activity of rPpKzl2 was tested against human α-thrombin and trypsin and bovine α-chymotrypsin. Increasing concentrations of rPpKzl2 were pre-incubated with 0.05 µM human α-thrombin (Calbiochem, EMD Chemicals Inc, Gibbstown, NJ, USA), 2 µM trypsin (Sigma, St. Louis, MO, USA) or 0.25 µM α-chymotrypsin (Calbiochem, EMD Chemicals Inc) in 50 mM Hepes-0.5% BSA, pH 7.3 for thrombin and in 50 mM Tris-HCl, pH 8.0 for trypsin and α-chymotrypsin. Each enzyme and rPpKzl2 combination was incubated for 15 min at 37ºC in a 96-well non-binding microtitre plate. Chromogenic peptide substrate H-D-Phenylalanyl-L-pipecolyl-Larginine-p-nitroaniline dihydrochloride (S-2238) (Chromogenix, diaPharma, West Chester Township, OH, USA), Na-Benzoyl-D,L-arginine 4-nitroanilide hydrochloride (BAPNA) (Sigma) or N-Succinyl-L-alanyl-L-alanyl-L-prolyl-L-phenylalanine 4-nitroanilide (Suc-AAPF-pNA) (Sigma) was added at increasing concentrations for α-thrombin, trypsin or α-chymotrypsin respectively for a total reaction volume of 100 µL. Inhibiton of trypsin activity was measured for 3 nM, 30 nM and 300 nM rPpKzl2 at increasing concentrations of BAPNA (25 µM, 125 µM, 250 µM, 500 µM and 1000 µM). Inhibition of α-chymotrypsin activity was measured for 0.0005 nM, 0.005 nM and 0.05 nM rPpKzl2 and inhibition of α-thrombin was measured at 0.5 nM, 3 nM and 300 nM rPpKzl2 at increasing concentrations 250 µM, 500 µM and 1000 µM of Suc-AAPF-pNA or S-2238, respectively. The rate of proteinase hydrolysis of the chromogenic substrate was measured at 405 nm every 35 s during the reaction with a Biotek Synergy HT microplate reader (Biotek, Winooski, VT, USA). Each reaction was run in triplicate and each assay was repeated at least twice.

Graphs of initial velocity (V) vs. substrate concentration [S] were fit with the Michaelis-Menten equation to obtain the kinetic constant (K_m) and maximum velocity (V_{max}), $v = \frac{V\,\max\,[S]}{Km + [S]}$ (Copeland 2000). Residual activity in the presence of different concentrations of rPpKzl2 was calculated with apparent V_{max} values, residual activity = $\frac{V\,\max}{V\,\max,\,0}$ x 100 (Copeland 2005).

RESULTS

Sequence analysis - Both *PpKzl1* and *PpKzl2* code for six cysteine residues in a conserved arrangement characterised as a non-classical Kazal-type domain. Predicted molecular weights and isoelectric points for *PpKzl1*and *PpKzl2* are estimated to be 6.4 kDa and 5.22 pI and 7.6 kDa and 6.10 pI respectively. In PpKzl1 an arginine residue is in the deduced P1 site, the predicted active site for Kazal-type inhibitors (Fig. 1A) and PpKzl2 contains a tyrosine in the P1 site (Fig. 1B). Arginine has been shown to confer thrombin and trypsin inhibitory activities and tyrosine in the P1 commonly shows chymotrypsin inhibitory activity (Kanost 1999).

PpKzl1 has 81% similarity and 73% identity to a putative protein identified in the New World sandfly *L. longipalpis*, vector of *Leishmania infantum chagasi* (Jochim et al. 2008, Pitaluga et al. 2009) (Fig. 1A). PpKzl1 has conserved sequence features previously described in non-classical Kazal-type domains in blood-feeding and non-blooding insects such as P-X-C_3-G-X_4-T-Y-X-N-X-C_4 and G-X-C_6, with (X) representing various residues (Augustin et al. 2009). A MSA with the top blast results for PpKzl1 was assembled as described in Materials and Methods section and displayed high conservation of arginine in the P1 site for this group of Kazals (Supplementary data). PpKzl2 is also similar to another predicted protein in *L. longipalpis* (Ramalho-Ortigão et al. 2007), but to a lesser degree with only 44% identity and 53% similarity (Fig. 1B). While the amino acids in the predicted P1 site in the *P. papatasi* and *L. longipalpis* proteins differ, tyrosine and phenylalanine do share similar structural and chemical properties and have both been shown to inhibit chymotrypsin. The Kazal-type domains in the PpKzl2 MSA displayed large diversity in P1 residues (Supplementary data).

Non-classical Kazal-type domain patterns are partially conserved in PpKzl2 including regions P-X-C_3 and G-X-C_6, (Fig. 1B, Supplementary data). PpKzl2 also has more residues between C_3 and C_4 shifting the location of the fifth cysteine closer to the C-terminus, which has been seen in other non-classical Kazal-type domains (Hemmi et al. 2005, Rimphanitchayakit & Tassanakajon 2010). Conserved residues specific to the PpKzl2 MSA include N-C_5-E/Q and a phenylalanine located four residues upstream of the fourth cysteine (Supplementary data).

Expression profiles - Expression of *PpKzl1* and *PpKzl2* in the female midgut increased after blood feeding. Temporal expression was analysed 0 h, 24 h, 48 h and 72 h PBM. *PpKzl1* transcript expression was up-regulated at 24 h and 48 h PBM ($p < 0.05$, $p < 0.001$) (Fig. 2A). After a significant increase in expression at 48 h PBM, *PpKzl1* expression decreased to pre-blood feeding levels (0 h) between 48-72 h PBM ($p < 0.01$). Expression of *PpKzl2* was up-regulated 24 h, 48 h and 72 h PBM ($p < 0.01$, $p < 0.001$, $p < 0.05$) (Fig. 2B). Transcript levels were up-regulated at 24 h and continued to increase significantly at 48 h PBM ($p < 0.01$). *PpKzl2* expression was then down-regulated by 72 h ($p < 0.001$) with expression at 72 h decreasing to levels similar to 24 h expression.

Following results indicating that *PpKzl*1 and *PpKzl2* expression is regulated following a blood meal, we then investigated if these transcripts are expressed in developmental (non-blood feeding) stages. Expression profiles of developmental stages for *PpKzl1* and *PpKzl2* show both transcripts expressed during early development (Fig. 3). Both *PpKzl1* and *PpKzl2* are expressed in larval stages L_1, L_2, L_3, L_4 and pupa at constant levels showing no significant differential regulation in expression between developmental stages. Both *PpKzl1* and *PpKzl2* are expressed in whole male tissues, but expression was not detected in eggs for either Kazal transcript (Supplementary data).

PpKzl1 and *PpKzl2* expression was further analysed at 24 h, 48 h and 72 h following an infective blood feeding with 5 x 10^6 *L. major* amastigotes per mL of blood. No significant difference in the mRNA expression levels of *PpKzl1* and *PpKzl2* between non-infected vs. *L. major* infected flies were detected in these three time points (Fig. 4).

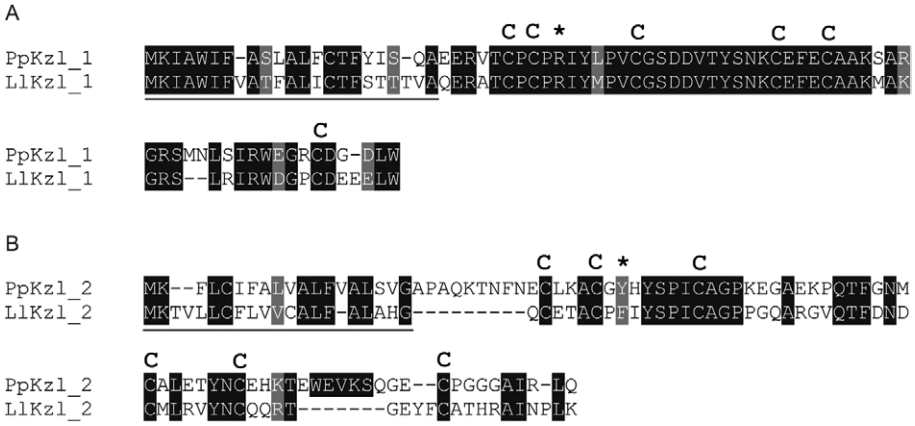


Fig. 1: PpKzl1 and PpKzl2 alignments with *Lutzomyia longipalpis* Kazal domains. PpKzl1 (GenBank ID: EU045342) (A) and PpKzl2 (GenBank ID: JX171681) (B) are both similar to putative proteins in *L. longipalpis* with Kazal-type domains: LlKzl1 (GenBank ID: ABV60319) and LlKzl2 (contig 69116). Conserved residues are in black and similar residues are in grey. Predicted signal peptides are underlined, asterisks mark predicted P1 residues, conserved cysteines are marked (C) and gaps are indicated by dashes.

Inhibition assays - Inhibition activity of rPpKzl2 was tested for α-thrombin, trypsin and α-chymotrypsin enzymes. Residual activity of enzymes in the presence of rPpKzl2 was reduced to 9.4% for α-chymotrypsin, 33.5% for α-thrombin and 63.9% for trypsin (Fig. 5). Both V_{max} and K_m decreased in all inhibition assays with increasing concentrations of rPpKzl2 (Supplementary data). Recombinant PpKzl2 inhibited α-chymotrypsin at the nanomolar level and inhibited α-thrombin and trypsin at micromolar levels.

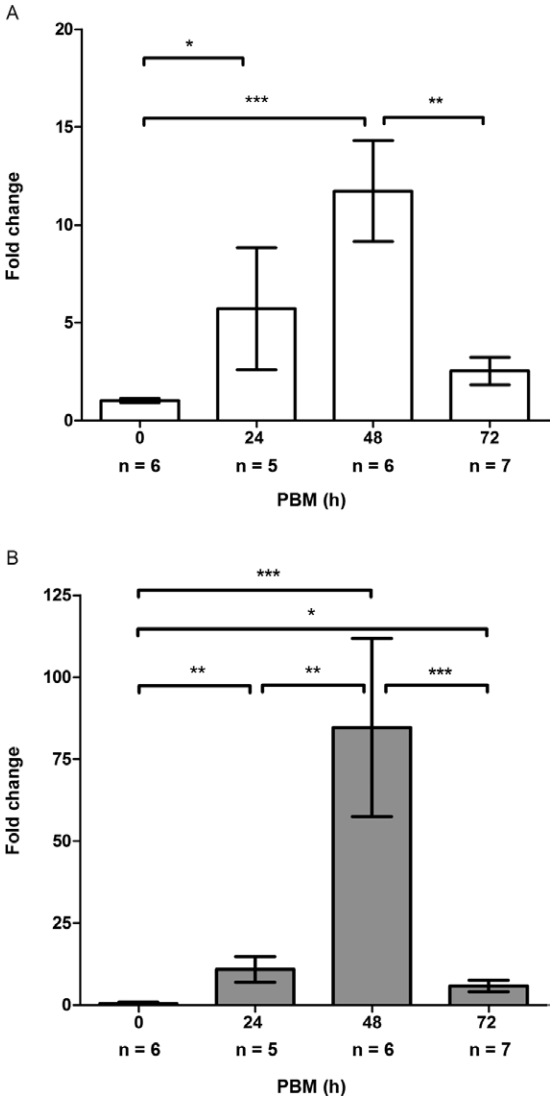


Fig. 2: *PpKzl1* and *PpKzl2* expression in adult females post-blood meal (PBM). *PpKzl1* and *PpKzl2* mRNA expression levels are regulated after a blood meal. A: *PpKzl1* is up-regulated 24 h and 48 h PBM with highest expression at 48 h PBM. By 72 h expression is down-regulated to levels similar to 0 h; B: *PpKzl2* is up-regulated 24 h, 48 h and 72 h PBM. Expression is highest at 48 h and decreases between 48-72 h PBM. Values are the mean fold change of five or more individual midguts with standard error of the mean. Expression was calibrated to 0 h expression levels. Analysis used ANOVA t test with the Bonferroni correction for multi-comparisons. *: $p < 0.05$; **: $p < 0.01$; ***: $p < 0.001$.

DISCUSSION

Kazal-type inhibitors are a diverse group of serine proteinase inhibitors with a wide range of roles in invertebrates. In blood-feeding triatomines, Kazal-type inhibitors in the midgut prevent coagulation of the blood meal (Friedrich et al. 1993, Mende et al. 1999, Campos et al. 2002, 2004, Araujo et al. 2007, Meiser et al. 2010).

Here, we characterised two single domain non-classical Kazal-type inhibitors from the sandfly *P. papatasi*. *PpKzl1* and *PpKzl2* mRNA transcripts are expressed in non-blood-fed and blood-fed female midguts and expression is regulated by the blood meal with up-regulation at 24 h and 48 h PBM. The decrease in *PpKzl1* and *PpKzl2* expression detected around 72 h PBM correlates with the completion of blood meal digestion, which culminates with the midgut emptying between 72-144 h PBM.

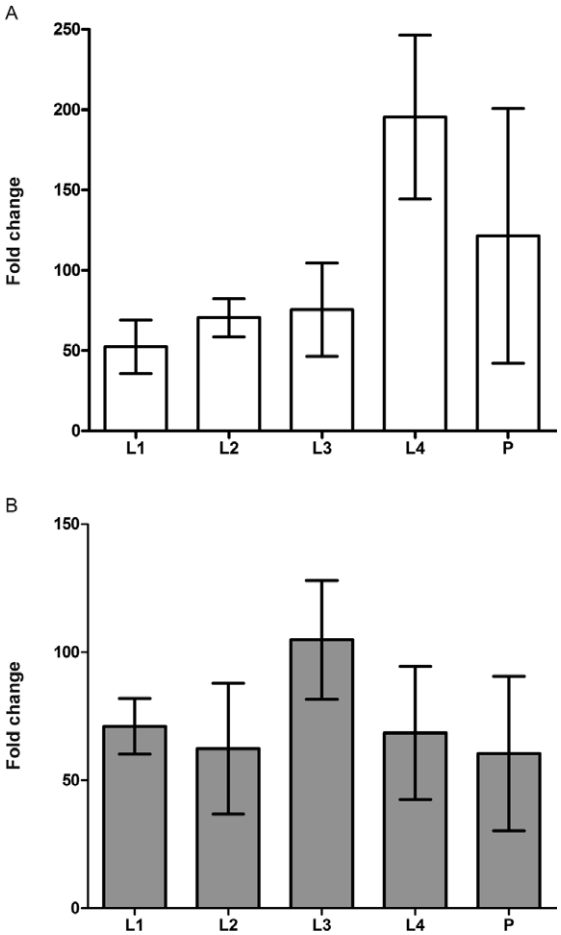


Fig. 3: *PpKzl1* and *PpKzl2* expression in larval stages and pupa. A: *PpKzl1* was expressed in all larval stages and pupae. *PpKzl1* expression was not significantly different when compared between larval stages; B: *PpKzl2* was also expressed in all larval stages and pupae at similar expression levels. Five or more individuals were pooled for each developmental stage and this was repeated for a total of three replicates. Values are the mean fold change with standard error of the mean. Expression was calibrated to expression levels in eggs. ANOVA t test with the Bonferroni correction for multi-comparisons was used for statistical analysis. L: larval stage; P: pupa.

Furthermore, the expression levels of both *PpKzl1* and *PpKzl2* remain constant between 72-144 h PBM (Supplementary data). Such expression profiles of *PpKzl1* and *PpKzl2* are suggestive of a role in digestion for their respective proteins. In addition, as *PpKzl1* and *PpKzl2* also are expressed in all larval stages, pupae and males, inhibition during digestion is likely not specific to serine proteinases involved in the coagulation cascade, but rather serine proteinases engaged across life stages and sexes.

The predicted PpKzl1 is similar to a single domain non-classical Kazal-type inhibitor from *Aedes aegypti*, AaTI (Ribeiro et al. 2007). Interestingly, a recombinant AaTI was shown to inhibit trypsin and plasmin, with weak inhibition of thrombin activity; the AaTI transcript also was shown to be expressed in larva, pupa, male and female tissues (Watanabe et al. 2010, 2011). PpKzl1 is also similar to the multi-domain Kazal-type inhibitors infestin and dipetalogastin, identified in *T. infestans* (Campos et al. 2002), and *Dipetalogaster maximus* (Mende et al. 2004), respectively, but with the highest identity to infestin's domain-4. This domain was found to strongly inhibit factor XIIa, plasmin and trypsin, with no activity for thrombin (Campos et al. 2002, 2004). Consistent with previous findings, PpKzl1 as a non-classical Kazal-type domain displays a predicted active site residue that suggests it likely possess inhibitory activity for trypsin-like serine proteinases.

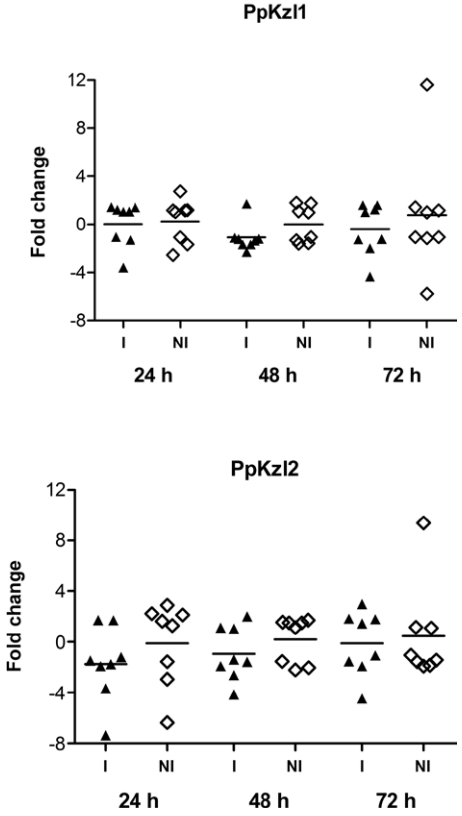


Fig. 4: *PpKzl1* and *PpKzl2* expression in adult females infected with *Leishmania major*. Temporal expression profiles 24 h, 48 h and 72 h post-infective blood meal (I) (▲) and post-non-infected blood meal (NI) (◇). Eight individual midguts were assayed for each infected and non-infected time point. *PpKzl1* and *PpKzl2* expression was not significantly different 24 h, 48 h and 72 h I when compared to NI control groups. Bars are the mean fold change of eight individual midguts. Expression was calibrated to expression in NI controls. Statistical analysis used two-tailed unpaired t tests and two-tailed Mann-Whitney *U* tests for parametric and nonparametric comparisons respectively ($p < 0.05$).

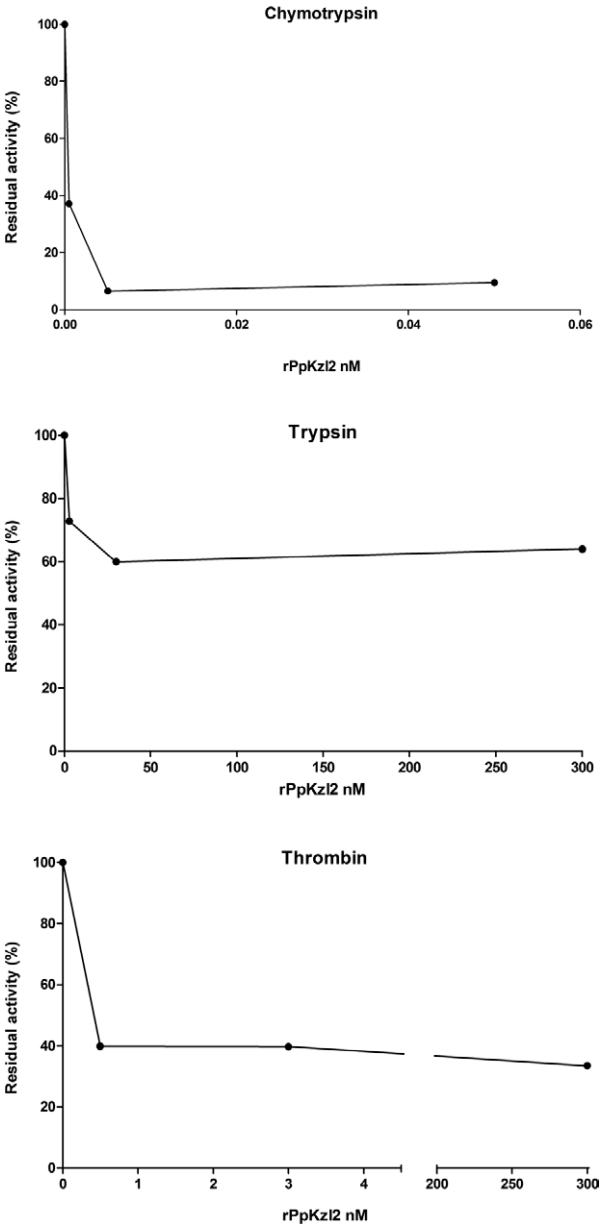


Fig. 5: rPpKzl2 enzyme inhibition activity. Activity was measured at increasing concentrations of both rPpKzl2 and substrate. Reactions were fit with Michaelis-Menten non-linear regression and apparent maximum velocity (V_{max}) values were used to calculate residual activity. Inhibition of α-chymotrypsin activity was observed with decreasing V_{max}. Activity of 0.25 μM α-chymotrypsin was reduced to 9.4%. Residual activity of 2 μM trypsin was reduced to 63.9%. Activity of 0.05 μM α-thrombin in the presence of rPpKzl2 was reduced to 33.5%. Reactions were run in triplicate and each graph represents one of two replicates of each experiment.

PpKzl2 on the other hand is similar to Kazal-type domains from dipteran, lepidopteran and hymenopteran species. Though no functional characterisation for these Kazal domains have been described, putative proteins were identified in expressed sequence tag and cDNA libraries of immune-challenged insects (Bartholomay et al. 2004, Gandhe et al. 2006).

A recombinant PpKzl2 was obtained and tested against various substrates. Inhibition activity of rPpKzl2 was observed for α-chymotrypsin, α-thrombin and trypsin, in agreement with previous reports on single-domain Kazal-type inhibitors having activity against multiple serine proteinases (Nirmala et al. 2001, Watanabe et al. 2010). The ability of PpKzl2 to inhibit serine proteinases in *P. papatasi* midgut is dependent upon the rate of inhibition and concentrations present in the midgut (Kanost & Jiang 1996) and therefore in vivo activity may be enzyme specific. Whereas rPpKzl2 inhibited α-thrombin, the inhibition activity for α-chymotrypsin was the strongest. We previously characterised two chymotrypsin-like and four trypsin-like proteases from *P. papatasi* and demonstrated that chymotrypsin and trypsin activities in the midgut of this sandfly peak between 27-48 h PBM and by 72 h PBM no such activities were detected (Ramalho-Ortigão et al. 2003). Also, as our results indicate, the peak in RNA abundance for Kazals in *P. papatasi* is 48 h PBM. These data, together with the observations that rPpKzl2 inhibited both chymotrypsin and trypsin and expression of the mRNA was also observed in non-blood-feeding life stages, suggest to us that PpKzl2 is more likely involved in regulating digestive proteases than blood fluidity within the midgut. Knock down by injection of 127 ng of double stranded RNA produced against each target did not affect mRNA expression levels of *PpKzl2* and *PpKzl1* in the midgut of *P. papatasi* and therefore analysis of effects on blood meal digestion rate via haemoglobin levels in female midguts were not informative (data unpublished).

Some Kazals have been shown to have immune-like activity; however there was no response in transcript expression of *PpKzl1* and *PpKzl2* during *L. major* infection. No effects were observed on *PpKzl1* and *PpKzl2* expression during *L. major* infection in the midgut at 24 h, 48 h or 72 h post-infective-blood meal. It has been described in sandflies that infection with *Leishmania* leads to modulation of trypsin-like activity in the midgut during digestion, suggesting that modulation of trypsin activity allows the parasites to survive (Borovsky & Schlein 1987, Sant'Anna et al. 2009, Telleria et al. 2010). This has been supported with data showing that RNAi of a trypsin gene increased parasite numbers during infection (Sant'Anna et al. 2009). The dynamics of serine proteinases and serine proteinase inhibitors in the midgut are not only crucial to sandfly metabolism and digestion, but may also affect *Leishmania* development. Further characterisation of the serine proteinase cascades and their inhibitors in *P. papatasi* may provide insight into the complex interactions that constitute vector competence.

REFERENCES

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*: 3389-3402.

Araujo N, Campos ITN, Tanaka AS, Santos A, Gontijo NF, Lehane MJ, Pereira MH 2007. Brasiliensin: a novel intestinal thrombin inhibitor from *Triatoma brasiliensis* (Hemiptera: Reduviidae) with an important role in blood intake. *Int J Parasitol 37*: 1351-1358.

Augustin R, Siebert S, Bosch TCG 2009. Identification of a kazal-type serine protease inhibitor with potent anti-staphylococcal activity as part of Hydra's innate immune system. *Dev Comp Immunol 33*: 830-837.

Bartholomay LC, Cho W, Rocheleau TA, Boyle JP, Beck ET, Fuchs JF, Liss P, Rusch M, Butler KM, Wu RC, Lin S, Kuo H, Tsao I, Huang C, Liu T, Hsiao K, Tsai S, Yang U, Nappi AJ, Perna NT, Chen C, Christensen BM 2004. Description of the transcriptomes of immune response-activated hemocytes from the mosquito vectors *Aedes aegypti* and *Armigeres subalbatus*. *Infect Immun 72*: 4114-4126.

Borovsky D, Schlein Y 1987. Trypsin and chymotrypsin-like enzymes of the sandfly *Phlebotomus papatasi* infected with *Leishmania* and their possible role in vector competence. *Med Vet Entomol 1*: 235-242.

Campos IT, Amino R, Sampaio CA, Auerswald EA, Friedrich T, Lemaire HG, Schenkman S, Tanaka AS 2002. Infestin, a thrombin inhibitor presents in *Triatoma infestans* midgut, a Chagas disease vector: gene cloning, expression and characterization of the inhibitor. *Insect Biochem Mol Biol 32*: 991-997.

Campos IT, Tanaka-Azevedo AM, Tanaka AS 2004. Identification and characterization of a novel factor XIIa inhibitor in the hematophagous insect, *Triatoma infestans* (Hemiptera: Reduviidae). *FEBS Lett 577*: 512-516.

Copeland RA 2000. *Enzymes: a practical introduction to structure, mechanism and data analysis*, 2nd ed., John Wiley & Sons, New Jersey, 390 pp.

Copeland RA 2005. Evaluation of enzyme inhibitors in drug discovery: a guide for medicinal chemists and pharmacologists. *Methods Biochem Anal 46*: 1-265.

Coutinho-Abreu IV, Sharma NK, Robles-Murguia M, Ramalho-Ortigão M 2010a. Targeting the midgut secreted PpChit1 reduces *Leishmania major* development in its natural vector, the sandfly *Phlebotomus papatasi*. *PLoS Negl Trop Dis 4*: e901.

Coutinho-Abreu IV, Wadsworth M, Stayback G, Ramalho-Ortigão M, McDowell MA 2010b. Differential expression of salivary gland genes in the female sandfly *Phlebotomus papatasi* (Diptera: Psychodidae). *J Med Entomol 47*: 1146-1155.

di Cera E 2009. Serine proteases. *IUBMB Life 61*: 510-515.

Friedrich T, Kröger B, Bialojan S, Lemaire HG, Höffken HW, Reuschenbach P, Otte M, Dodt J 1993. A Kazal-type inhibitor with thrombin specificity from *Rhodnius prolixus*. *J Biol Chem 268*: 16216.

Gandhe A, Arunkumar K, John S, Nagaraju J 2006. Analysis of bacteria-challenged wild silkmoth, *Antheraea mylitta* (Lepidoptera) transcriptome reveals potential immune genes. *BMC Genomics 7*: 184.

Gasteiger E, Gattiker A, Hoogland C, Ivanyi I, Appel RD, Bairoch A 2003. ExPASy: the proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Res 31*: 3784-3788.

Hemmi H, Kumazaki T, Yoshizawa-Kumagaye K, Nishiuchi Y, Yoshida T, Ohkubo T, Kobayashi Y 2005. Structural and functional study of an anemonia elastase inhibitor, a "nonclassical" kazal-type inhibitor from *Anemonia sulcata*. *Biochemistry 44*: 9626-9636.

Jiang H, Kanost MR 2000. The clip-domain family of serine proteinases in arthropods. *Insect Biochem Mol Biol 30*: 95-105.

Jochim RC, Teixeira CR, Laughinghouse A, Mu J, Oliveira F, Gomes RB, Elnaiem D, Valenzuela JG 2008. The midgut transcriptome of *Lutzomyia longipalpis*: comparative analysis of cDNA libraries from sugar-fed, blood-fed, post-digested and *Leishmania infantum chagasi*-infected sandflies. *BMC Genomics 9*: 15.

Kanost MR 1999. Serine proteinase inhibitors in arthropod immunity. *Dev Comp Immunol 23*: 291-301.

Kanost MR, Jiang H 1996. Proteinase inhibitors in invertebrate immunity. In K Söderhäll, S Iwanaga, G Vanta, *New directions in invertebrate immunology*, SOS Publications, New Jersey, p. 155-174.

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. CLUSTALW and CLUSTALX version 2.0. *Bioinformatics 23*: 2947-2948.

Livak KJ, Schmittgen TD 2001. Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta CT}$ Method. *Methods 25*: 402-408.

Lovato DV, de Campos ITN, Amino R, Tanaka AS 2006. The full-length cDNA of anticoagulant protein infestin revealed a novel releasable Kazal domain, a neutrophil elastase inhibitor lacking anticoagulant activity. *Biochimie 88*: 673-681.

Meiser CK, Piechura H, Werner T, Dittmeyer-Schäfer S, Meyer HE, Warscheid B, Schaub GA, Balczun C 2010. Kazal-type inhibitors in the stomach of *Panstrongylus megistus* (Triatominae, Reduviidae). *Insect Biochem Mol Biol 40*: 345-353.

Mende K, Lange U, Nowak G 2004. Three recombinant serine proteinase inhibitors expressed from the coding region of the thrombin inhibitor dipetalogastin. *Insect Biochem Mol Biol 34*: 971-979.

Mende K, Petoukhova O, Koulitchkova V, Schaub GA, Lange U, Kaufmann R, Nowak G 1999. Dipetalogastin, a potent thrombin inhibitor from the bloodsucking insect *Dipetalogaster maximus*: cDNA cloning, expression and characterization. *Eur J Biochem 266*: 583-590.

Mulenga A, Blandon M, Khumthong R 2007. The molecular basis of the *Amblyomma americanum* tick attachment phase. *Exp Appl Acarol 41*: 267-287.

Nirmala X, Kodrik D, Zurovec M, Sehnal F 2001. Insect silk contains both a kunitz-type and a unique Kazal-type proteinase inhibitor. *Eur J Biochem 268*: 2064-2073.

Oliveira F, Kamhawi S, Seitz AE, Pham VM, Guigal PM, Fischer L, Ward J, Valenzuela JG 2006. From transcriptome to immunome: identification of DTH inducing proteins from a *Phlebotomus ariasi* salivary gland cDNA library. *Vaccine 24*: 374-390.

Pitaluga AN, Beteille V, Lobo AR, Ortigão-Farias JR, Dávila AM, Souza AA, Ramalho-Ortigão JM, Traub-Cseko YM 2009. EST sequencing of blood-fed and *Leishmania*-infected midgut of *Lutzomyia longipalpis*, the principal visceral leishmaniasis vector in the Americas. *Mol Genet Genomics 282*: 307-317.

Ramalho-Ortigão JM, Jochim R, Anderson J, Lawyer P, Pham V, Kamhawi S, Valenzuela J 2007. Exploring the midgut transcriptome of *Phlebotomus papatasi*: comparative analysis of expression profiles of sugar-fed, blood-fed and *Leishmania major*-infected sandflies. *BMC Genomics 8*: 300.

Ramalho-Ortigão JM, Kamhawi S, Joshi MB, Reynoso D, Lawyer PG, Dwyer DM, Sacks DL, Valenzuela JG 2005. Characterization of a blood activated chitinolytic system in the midgut of the sand fly vectors *Lutzomyia longipalpis* and *Phlebotomus papatasi*. *Insect Mol Biol 14*: 703-712.

Ramalho-Ortigão JM, Kamhawi S, Rowton ED, Ribeiro JMC, Valenzuela JG 2003. Cloning and characterization of trypsin and chymotrypsin-like proteases from the midgut of the sandfly vector *Phlebotomus papatasi*. *Insect Biochem Mol Biol 33*: 163-171.

Ramalho-Ortigão JM, Saraiva EM, Traub-Csekö YM 2010. Sandfly-*Leishmania* interactions: long relationships are not necessarily easy. *Open Parasitol J 4*: 195-204.

Ribeiro JM, Arca B, Lombardo F, Calvo E, Phan VM, Chandra PK, Wikel SK 2007. An annotated catalogue of salivary gland transcripts in the adult female mosquito, *Aedes aegypti*. *BMC Genomics 8*: 6.

Rimphanitchayakit V, Tassanakajon A 2010. Structure and function of invertebrate Kazal-type serine proteinase inhibitors. *Dev Comp Immunol 34*: 377-386.

Sant'Anna MRV, Diaz-Albiter H, Mubaraki M, Dillon RJ, Bates PA 2009. Inhibition of trypsin expression in *Lutzomyia longipalpis* using RNAi enhances the survival of *Leishmania*. *Parasit Vectors 2*: 62.

Takáč P, Nunn MA, Mészáros J, Pecháňová O, Vrbjar N, Vlasáková P, Kozánek M, Kazimírová M, Hart G, Nuttall PA, Labuda M 2006. Vasotab, a vasoactive peptide from horse fly *Hybomitra bimaculata* (Diptera, Tabanidae) salivary glands. *J Exp Biol 209*: 343-352.

Tanaka-Azevedo AM, Morais-Zani K, Torquato RJS, Tanaka AS 2010. Thrombin inhibitors from different animals. *J Biomed Biotechnol 2010*: 641025.

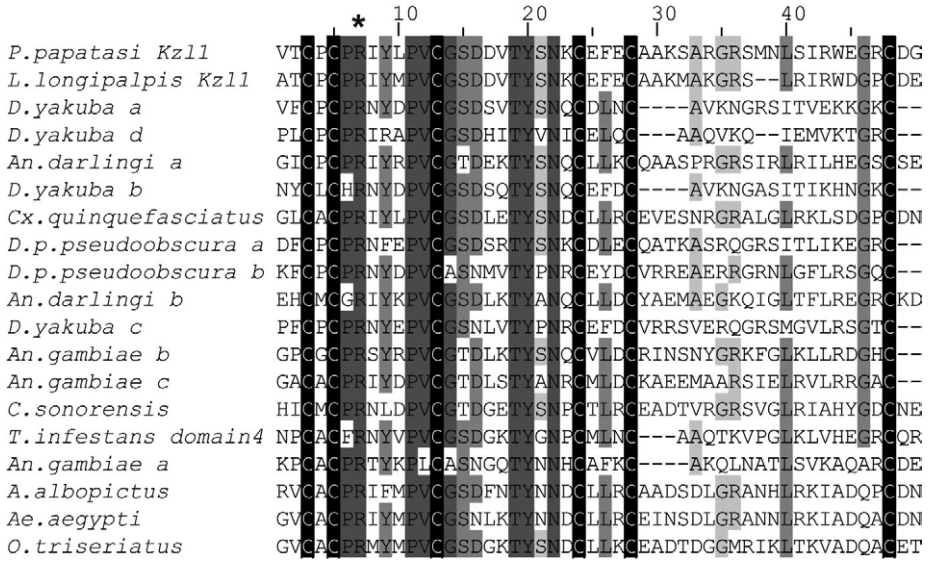
Telleria EL, Araújo A, Secundino NF, d'Avila-Levy CM, Traub-Csekö YM 2010. Trypsin-like serine proteases in *Lutzomyia longipalpis* - expression, activity and possible modulation by *Leishmania infantum chagasi*. *PLoS ONE*: e10697.

Watanabe RMO, Soares TS, Morais-Zani K, Tanaka-Azevedo AM, Maciel C, Capurro ML, Torquato RJS, Tanaka AS 2010. A novel trypsin Kazal-type inhibitor from *Aedes aegypti* with thrombin coagulant inhibitory activity. *Biochimie 92*: 933-939.

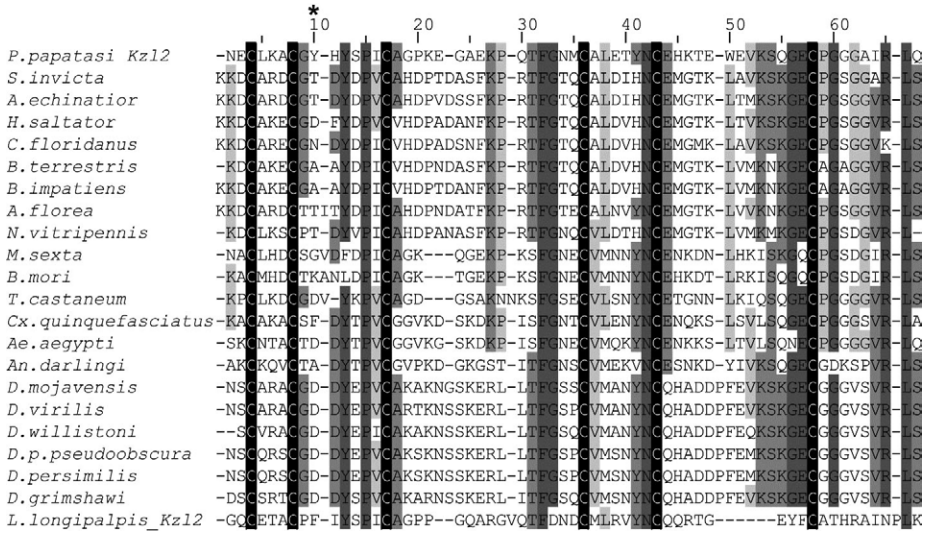
Watanabe RMO, Tanaka-Azevedo AM, Araujo MS, Juliano MA, Tanaka AS 2011. Characterization of thrombin inhibitory mechanism of rAaTI, a Kazal-type inhibitor from *Aedes aegypti* with anticoagulant activity. *Biochimie 93*: 618-623.

Waterhouse AM, Procter JB, Martin DMA, Clamp M, Barton GJ 2009. Jalview version 2 - A multiple sequence alignment editor and analysis workbench. *Bioinformatics 25*: 1189-1191.

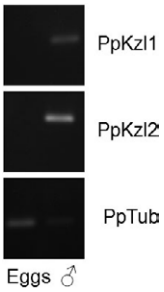
Zhou J, Liao M, Hatta T, Tanaka M, Xuan X, Fujisaki K 2006. Identification of a follistatin-related protein from the tick *Haemaphysalis longicornis* and its effect on tick oviposition. *Gene 372*: 191-198.



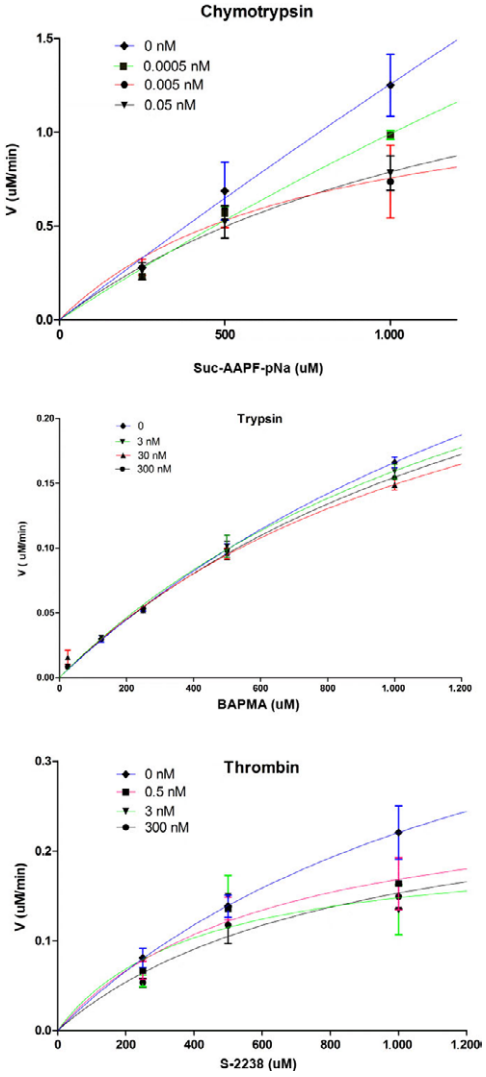
PpKzl1 multiple sequence alignment. *Phlebotomus papatasi* PpKzl1 (GenBank ID: EU045342), *Culicoides sonorensis* (GenBank ID: AAV84258), *Drosophila yakuba*_a (GenBank ID: XP_002088400), *Anopheles darlingi*_a (GenBank ID: ACI30143), *Lutzomyia longipalpis* (GenBank ID: ABV60319), *Aedes aegypti* AaTi (GenBank ID: ABF18209), *Culex quinquefasciatus* (GenBank ID: XP_001868221), *Ochlerotatus triseriatus* salivary (GenBank ID: ACU30983), *Aedes albopictu*s (GenBank ID: AAV90671), *Triatoma infestans* infestin domain 4 (GenBank ID: AAK57342), *Anopheles gambiae*_a (GenBank ID: XP_001230687), *An. gambiae*_b (GenBank ID: XP_317819), *An. gambiae*_c (GenBank ID: EAA12788), *Drosophila yakuba*_b (GenBank ID: XP_002088399), *D. yakuba*_c (GenBank ID: XP_002088401), *D. yakuba*_d (GenBank ID: XP_002088397), *Drosophila pseudoobscura pseudoobscura*_a (GenBank ID: XP_001356962), *D. pseudoobscura pseudoobscura*_b (GenBank ID: XP_001356963) and *An. darlingi*_b (GenBank ID: ACI30165). Asterisk means the predicted P1 residue, gaps are indicated by dashes, conserved cysteines are in black and residues with more than 50% conserved identity are in shades of grey.



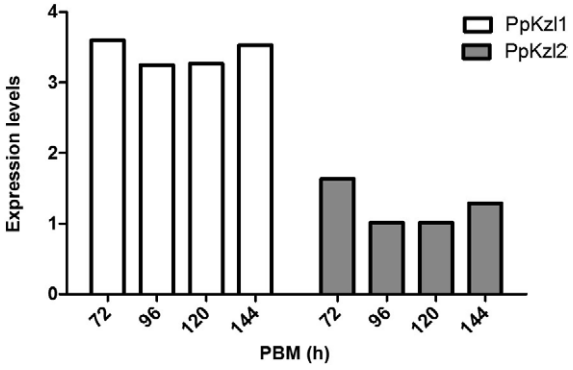
PpKzl2 multiple sequence alignment. *Phlebotomus papatasi* PpKzl2 (GenBank ID: JX171681), *Anopheles darlingi* (GenBank ID: ACI30205), *Drosophila mojavensis* (GenBank ID: XP_002000106), *Culex quinquefasciatus* (GenBank ID: XP_001842298), *Aedes aegypti* (GenBank ID: XP_001658905), *Manduca sexta* (GenBank ID: AAF16698), *Nasonia vitripennis* (GenBank ID: XP_001600330), *Bombyx mori* (GenBank ID: NP_001040250), *Bombus terrestris* (GenBank ID: XP_003401213), *Drosophila pseudoobscura pseudoobscura* (GenBank ID: XP_001359513), *Drosophila persimilis* (GenBank ID: XP_002017393), *Tribolium castaneum* (GenBank ID: XP_974370), *Drosophila willistoni* (GenBank ID: XP_002070657), *Drosophila grimshawi* (GenBank ID: XP_001994337)*, Solenopsis invicta* (GenBank ID: ADC34234), *Drosophila virilis* (GenBank ID: XP_002053264), *Acromyrmex echinatior* (GenBank ID: EGI69242), *Harpegnathos saltator* (GenBank ID: EFN89909), *H. saltator* 2 (GenBank ID: EFN81812)*, Bombus impatiens* (GenBank ID: XP_003486913), *Apis florae* (GenBank ID: XP_003692060), *Camponotus floridanus* (GenBank ID: EFN62548) and *Lutzomyia longipalpis* (contig 69116). Asterisk means the predicted P1 residue, gaps are indicated by dashes, conserved cysteines are in black and residues with more than 50% conserved identity are in shades of grey.



PpKzl1 and *PpKzl2* expression in *Phlebotomus papatasi* males and not in eggs. Reverse-transcriptase polymerase chain reaction (PCR) was carried out in 20 µL reactions with cDNA from one whole male and a pool of 10 eggs on an Eppendorf Mastercycler gradient. Reactions were prepared with 10 µL of 2X GoTaq PCR master mix (Promega, Madison, WI, USA), 0.2 µM forward and reverse primers, 1 µL cDNA and molecular grade water (Invitrogen, Carlsbad, CA, USA) to a total volume of 20 µL. Reactions were carried out as follows: 95ºC/1 min followed by 26 cycles of 94ºC/30 s, 56.5ºC/30 s and 72ºC/1 min and a final step at 72ºC/5 min. The primers PpKzl111, PpKzl859 and PpTub148 specific for *P. papatasi* Pp*Kzl1*, *PpKzl2* and β-tubulin were used for PCR. PCR fragments (10 µL) were separated on a 1.5% agarose gel stained with ethidium bromide. Primers for PpTub148 forward: GCGATGACTCCTTCAACAC and reverse: GTGATCAATTGTTCGGGATG.



Michaelis-Menten non-linear regression of rPpKzl2 inhibition. Initial velocity (V) over substrate concentration (S) was fit with Michaelis-Menten non-linear regression for each concentration of rPpKzl2. A reduction in maximum velocity and kinetic constant values was observed with increasing rPpKzl2 when compared to the fit of 0 nM rPpKzl2. BAPNA: Na-Benzoyl-D,L-arginine 4-nitroanilide hydrochloride; S-2238: H-D-Phenylalanyl-L-pipecolyl-Larginine-p-nitroaniline dihydrochloride; Suc-AAPF-pNA: N-Succinyl-L-alanyl-L-alanyl-L-prolyl-L-phenylalanine 4-nitroanilide.



PpKzl1 and *PpKzl2* expression 72-144 h post-blood meal (PBM). Temporal analysis with semi-quantitative reverse transcriptase-polymerase chain reaction (PCR) indicated that *PpKzl1* and *PpKzl2* transcript expression remains constant 72-144 h PBM. Time points were pools of five midguts from female sandflies. PCR was carried out in 20 µL reactions on an Eppendorf Mastercycler gradient. Reactions were prepared with 10 µL of 2X GoTaq PCR master mix (Promega, Madison, WI, USA), 0.2 µM forward and reverse primers, 1 µL cDNA and molecular grade water (Invitrogen, Carlsbad, CA, USA) to a total volume of 20 µL. Reactions were carried out as follows: 95ºC/1 min followed by 26 cycles of 94ºC/30 s, 56.5ºC/30 s and 72ºC/1 min and a final step at 72ºC/5 min. The primers PpKzl111, PpKzl859 and PpTub148 specific for *Phlebotomus papatasi* Pp*Kzl1*, *PpKzl2* and β-tubulin were used for PCR. PCR fragments (10 µL) were separated on a 1.5% agarose gel stained with ethidium bromide alongside 0.5 µg of exACTGene cloning DNA Ladder (Fisher, Scientific, Pittsburgh, PA, USA). Intensities of PCR fragments were standardised to β-tubulin and compared with known quantities of the reference ladder using Total Lab 100 software (BioSystematica, Sarnau, UK).